Sun ZG, Chen SC, Qiao LS. A two-step regularization framework for non-local means. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 29(6): 1026–1037 Nov. 2014. DOI 10.1007/s11390-014-1487-9

A Two-Step Regularization Framework for Non-Local Means

Zhong-Gui Sun[1,2] (孙忠贵), Song-Can Chen[2,*] (陈松灿), and Li-Shan Qiao[1] (乔立山)

[1] *Department of Mathematics Science, Liaocheng University, Liaocheng 252000, China*

[2] *College of Computer Science and Technology, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China*

E-mail: {altlp, s.chen, qiaolishan}@nuaa.edu.cn

Received January 28, 2014; revised July 8, 2014.

Abstract As an effective patch-based denoising method, non-local means (NLM) method achieves favorable denoising performance over its local counterparts and has drawn wide attention in image processing community. The implementation of NLM can formally be decomposed into two sequential steps, i.e., computing the weights and using the weights to compute the weighted means. In the first step, the weights can be obtained by solving a regularized optimization. And in the second step, the means can be obtained by solving a weighted least squares problem. Motivated by such observations, we establish a two-step regularization framework for NLM in this paper. Meanwhile, using the framework, we reinterpret several non-local filters in the unified view. Further, taking the framework as a design platform, we develop a novel non-local median filter for removing salt-pepper noise with encouraging experimental results.

Keywords non-local means, non-local median, framework, image denoising, regularization

1 Introduction

Denoising is an important task in the image processing area. Up to date, a variety of denoising methods have been developed in different forms. Because the existing methods mostly assume that the images are relatively smooth, their denoising results are usually achieved by a weighted averaging[1]. The typical schemes, such as the mean or Gaussian filter, are inefficient in preserving details due to their isotropy characteristic. In order to better preserve the details, several edge-aware filtering schemes, e.g., anisotropic filter[2], Yaroslavsky's filter[3], Susan filter[4], bilateral filter[5], have been invented respectively. Though improving the preservation of details in different degrees, these filters still have two defects[1]: 1) the weights involved in averaging only depend on the single pixels, which is not robust when these pixels are contaminated with noise; 2) there exist artificial shocks in the denoised results. And, a common point in the filters mentioned above is that they all operate in a local neighborhood.

Motivated by the existence of high redundancy in natural images[6], Buades *et al.* proposed a non-local means filter (NLM)[1]. Unlike its precursors typically operating in a local neighborhood, NLM operates in a non-local area (even the whole image) by using a dissimilarity measure between patches[7-8]. Despite its simplicity and intuition, NLM has been empirically validated to clearly outperform other classic filters[7]. Similar to NLM, the UINTA filter proposed by Awate and Whitaker[8] also restores the pixels by non-local and path-based strategy which, in fact, has become a core of most state-of-the-art filters including BM3D[9] and K-SVD[10]. However, the main difference between UINTA and NLM is that the former is an iterative algorithm, while the latter is based on a non-iterative scheme.

As reviewed in [7], all the follow-up researches of NLM can roughly be divided into 3 lines: 1) improving NLM itself, 2) combining it with other denoising methods, and 3) adapting it to other image processing tasks. Following line 1, we develop a two-step regularization framework for NLM, whose idea is mainly from the observations of the NLM's implementation. Though such a two-step viewpoint on NLM has appeared in several literatures such as [11-14], our framework is more general than the existing ones. For the existing models, despite their two-step solving procedures for filter designs, their objective functions are single-step. Thus essentially, the definitions for the existing models are single-step. By contrast, in our

Regular Paper

The research was partially supported by the National Natural Science Foundation of China under Grant No. 61300154, the Natural Science Foundations of Shandong Province of China under Grant Nos. NZR2010FL011, ZR2012FQ005, Jiangsu Qing Lan Projects, the Fundamental Research Funds for the Central Universities of China under Grant No. NZ2013306, and the Natural Science Foundation of Liaocheng University under Grant No. 318011408.

*Corresponding Author

framework, not only the solving procedure but also the objective definition is two-step. As a result, all components (terms) involved in our framework can be selected more flexibly. For example, one can naturally use different measures in the two steps. As illustrated shortly in Subsection 2.3, some non-local filters can be interpreted by the proposed framework but cannot be done by the existing models[11-14].

Finally, we summarize the contributions of this paper as follows.

1) Inspired by the observations mentioned above, we propose a two-step regularization framework for NLM.

2) To illustrate the effectiveness of the framework, based on it, we reinterpret several non-local filters such as NLM and NLEM[15] in a unified view.

3) The proposed framework provides a platform for developing new non-local filters. With its help, we develop a novel non-local median filter (NLMED) with encouraging experimental results.

The rest of this paper is structured as follows. In Section 2, we propose our framework and reinterpret two classical non-local filters (NLM and NLEM) based on it. In Section 3, based on the framework, we develop a novel non-local filter (NLMED) to remove salt-pepper noise. Finally, we give the concluding remarks in Section 4.

2 Proposed Two-Step Regularization Framework

2.1 Proposition of the Framework

Consider the following image denoising problem,

$$Y = X + N,$$

where Y is the observed (noisy) image, X is the ideal (noise-free) image to be recovered, N is Gaussian noise with mean zero, and $\Omega = [1, \ldots, m] \times [1, \ldots, n]$ is a bounded domain on which the image is defined.

Let $\widehat{X}$ denote the denoised results, which is an estimate of X. For any pixels $i, j \in \Omega$, $w_{ij} = -\frac{\|Y_i - Y_j\|^2_{2,\alpha}}{h^2}$ is the similarity weight between i and j, $C(i)$ is the normalized factor of w_{ij} defined as

$$C(i) = \sum_{j \in \Omega} \exp\Big(-\frac{\|Y_i - Y_j\|^2_{2,\alpha}}{h^2}\Big), \tag{1}$$

and the gray value $\widehat{X}(i)$ is computed by NLM[1] in terms of

$$\widehat{X}(i) = \frac{1}{C(i)} \sum_{j \in \Omega} w_{ij} Y(j), \tag{2}$$

where Y_i and Y_j denote image patches of size $P \times P$ centered at pixels i and j. h acts as a filtering parameter.

In (1), $\|\cdot\|_{2,\alpha}$ is the Gaussian weighted L_2 norm, and α is the standard deviation of Gaussian kernel. That means $\|Y_i - Y_j\|_{2,\alpha}$ is the weighted Euclidean distance measure between Y_i and Y_j. For the sake of computation, a non-local search window of size $S \times S$ is generally recommended to replace the whole image in (2) to restore gray value of each pixel i. Thus, the complexity of NLM is $O(P^2S^2)$ per pixel. More details of NLM can be found in [1, 7].

The implementation of NLM in (2) can be divided into two steps, i.e., obtaining the weights and computing the weighted means. Though superficially the weights are directly defined in NLM, inspired by the work in [12-13, 16], we know that they can be obtained through optimizing some objective with Shannon entropy regularizers. Meanwhile, the means also result from optimizing a weighted least squares objective. It is well known that different objectives can lead to different optimization results. And at the same time, different regularizers can reflect different problem priors. Inspired by the observations above, we propose a two-step regularization framework in (3)∼(4):

$$\text{step 1}: \quad \hat{w} = \arg\min_{w} \sum_{i,j \in D} w_{ij}^{p_1} M_1(Y_i, Y_j) + \lambda_1 J_1(w), \tag{3}$$

$$\text{step 2}: \quad \widehat{X} = \arg\min_{X} \sum_{i,j \in D} \hat{w}_{ij}^{p_2} M_2(X_i, Y_j) + \lambda_2 J_2(X), \tag{4}$$

where M_k $(k = 1, 2)$ are dissimilarity measures between patches or points. Step 1 is to obtain the weights w_{ij} by minimizing an objective regularized by $J_1(w)$. And the objective depends on the patch dissimilarities, i.e., $M_1(Y_i, Y_j)$. Meanwhile, step 2 is to recover the ideal image X by minimizing another regularized objective. Both steps are described as corresponding optimization problems. Some constraints associated with specific problem could also be imposed. The two steps can work not only independently but also iteratively if necessary. The patches could be reduced to single pixels if needed. Instead of being directly defined in many other non-local algorithms such as [1, 17-18], the weights w_{ij} and the image X both have clear optimization backgrounds in our framework.

For convenience of the problem formulation, we specially call $\sum_{i,j \in D} w_{ij}^{p_1} M_1(\cdot, \cdot)$ and $\sum_{i,j \in D} \hat{w}_{ij}^{p_2} M_2(\cdot, \cdot)$ in the framework as harmonious terms. And $J_k(\cdot)$ $(k = 1, 2)$ are regularization terms associated with some known priori knowledge. $\lambda_k \geqslant 0$ $(k = 1, 2)$ are trade-off parameters used to balance the two terms. D is the working range of pixels i and j. Before introducing these components of the framework in detail, we list key

Table 1. Key Notations

Notation	Description
X	Ideal image
Y	Noisy image
$\Omega = [1, \ldots, m] \times [1, \ldots, n]$	A bounded domain on which the image is defined
$\Omega' \subset \Omega$	Noise-free candidate index
$D \subset \Omega$	Working range of pixels
$i = (i_1, i_2) \in \Omega$, $j = (j_1, j_2) \in \Omega$	Pixels
$X(i)$	Gray value of pixel i
X_i	Image patch centered at pixel i
Y_i^t	A patch whose size is varied and limited by threshold t
w_{ij}	Similar weight between i and j
h	Filtering parameter of NLM[1]
α	Standard deviation of Gaussian kernel
u	Meaning the operation is based on unpolluted (noise-free) pixels
t	Threshold parameter in UNLM[19]
p_1, p_2	Nonnegative real number
$P \times P$	Size of image patch
$S \times S$	Size of search window
M_k $(k = 1, 2)$	Dissimilarity measures between patches or points
J_k $(k = 1, 2)$	Regularization terms
λ_k $(k = 1, 2)$	Trade-off parameters
$\|\cdot\|_1$	L_1 norm
$\|\cdot\|_2$	L_2 norm
$\|\cdot\|_{2,\alpha}$	Gaussian weighted L_2 norm where α is the standard deviation of Gaussian kernel
$\|\cdot\|_{2,\alpha,u}$	Measure based on unpolluted pixels
$WMed\{\cdot,\cdot\}$	Weighted median operator

notations involved in this paper in Table 1 for convenient reading.

2.2 Components of the Framework

2.2.1 Harmonious Terms

In our framework, the harmonious terms aim to keep consistence between the weights w_{ij} and the dissimilarities M_k. That is, large weight should correspond to small dissimilarity (vice verse). In order to facilitate solving, we restrict $p_k \geqslant 1$. The dissimilarity measures M_k satisfy the following definition[20]:

Definition 1. *The dissimilarity measure M on a set of vectors $\boldsymbol{V}$ is a real valued function on $\boldsymbol{V} \cdot \boldsymbol{V}$ satisfying*

$$\begin{aligned} M_a^* &= M(a, a) \leqslant M(a, b) \\ &= M(b, a) < \infty, \quad \forall a, b \in \boldsymbol{V}. \end{aligned} \tag{5}$$

Compared with the definition of distance, the dissimilarity measure is not necessarily to be restrained by triangle inequality. Thus it is generally more convenient for us to define a dissimilarity measure than a distance. Some often-used dissimilarity measures are listed and roughly classified into three categories in [21]. And according to specific conditions, different dissimilarities adopted will lead to different optimized results.

2.2.2 Regularization Terms

In the framework, the regularization terms J_k $(k = 1, 2)$ are generally used to impose problem priori knowledge on the corresponding objective variables. For example, in step 1, one often expects the weights w_{ij} to have clear probability meaning. And thus we need to impose the constraints: $\sum_j w_{ij} = 1$, $w_{ij} \geqslant 0$. In fact, there exists a class of (discrete) entropy measures (called as generalized entropy) to describe the uncertainty within a probability distribution, i.e., $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)^{\mathrm{T}} \in \Delta = \{\boldsymbol{v} \in R^n | \sum_{i=1}^n v_i = 1, v_i \geqslant 0\}$[22]. And hence, any one of them can be used in the framework according to problem prior. In what follows, we simply introduce the definition of generalized entropy[23]:

Definition 2. *Generalized entropy is a mapping $H : \Delta \to R_+$ (R_+ denotes the set of nonnegative real numbers) that satisfies the following two criteria (symmetry and concavity):*

1) $H(\boldsymbol{v}_a) = H(\boldsymbol{v}_b)$, for any $\boldsymbol{v}_a \in \Delta$ and any $\boldsymbol{v}_b \in \Delta$ whose elements are a permutation of the elements of $\boldsymbol{v}_a$;

2) $H(\cdot)$ is a concave function.

We refer the reader to [24] for the details of some typical examples of generalized entropy. Replacing J_1 with $-H$ in (3) means applying the maximum entropy

priori knowledge in step 1 in our framework. Further, due to the concavity of $H(\cdot)$ and linear constraints of w_{ij}, the step becomes a convex subproblem and thus can produce a global optimization solution.

In contrast to J_1 which usually adopts generalized entropy regularization for the desired property of the weights w_{ij}, J_2 plays the same regularization role but respects a different desired property of the ideal image X. For example, the image is usually expected to have smoothness, sparsity, fidelity or other characteristics. And some corresponding regularizations for the properties of the ideal image have been discussed deeply in [11].

2.3 Reinterpretations of NLM and NLEM

As mentioned before, in NLM, the weights can be obtained through optimizing an objective regularized by Shannon entropy, and the means are also derived from optimizing a weighted least squares problem. Furthermore, with the help of the framework, we formulate NLM as a two-step optimization procedure as follows:

$$\begin{aligned}\text{step 1}: \quad \hat{w} =& \arg\min_{w} \sum_{i,j\in\Omega} w_{ij}\|Y_i - Y_j\|_{2,\alpha}^2 + \\ & h^2 \sum_{i,j\in\Omega} w_{ij}\ln w_{ij}, \\ & \text{s.t.} \sum_{j\in\Omega} w_{ij} = 1, \quad w_{ij} \geqslant 0,\ i,j \in \Omega, \end{aligned} \tag{6}$$

$$\text{step 2}: \quad \widehat{X} = \arg\min_{X} \sum_{i,j\in\Omega} \hat{w}_{ij}\|X(i) - Y(j)\|_2^2, \tag{7}$$

where solution $\widehat{X}(i)$ exactly equals that in (2). And the detailed formulation can be found in Appendix A.1. Different from the direct definition of NLM in [1], the reinterpretation of the same filter provided in (6)$\sim$(7) is a specific case of our two-step framework, which is beneficial for us to survey NLM in an alternative (optimization) view:

1) The filter is a weighted least square regression process.

2) The weights w_{ij} result from optimizing a Shannon entropy regularized objective.

3) More interestingly, h^2 in NLM exactly plays a role of regularization parameter. And the larger h^2, the larger the penalty imposed.

It should be noted that the above understanding for NLM does not seem so apparent in its original definition[1].

To further indicate the efficiency of the framework, based on it, we reinterpret the robust version of NLM (NLEM) too, which is recently proposed in [15]. Similar to that of NLM, the reinterpretation is directly given as follows:

$$\begin{aligned}\text{step 1}: \quad \hat{w} =& \arg\min_{w} \sum_{i,j\in\Omega} w_{ij}\|Y_i - Y_j\|_2^2 + \\ & h^2 \sum_{i,j\in\Omega} w_{ij}\ln w_{ij}, \\ & \text{s.t.} \sum_{j\in\Omega} w_{ij} = 1, \quad w_{ij} \geqslant 0, \quad i,j \in \Omega, \\ \text{step 2}: \quad \widehat{X} =& \arg\min_{X} \sum_{i,j\in\Omega} \hat{w}_{ij}\|X_i - Y_j\|_2. \end{aligned}$$

Step 1 can be solved in the same way as that in NLM in (6). And the objective in step 2 is called the Euclidean median (or geometric median)[25], which is computed by iterations in [15]. The overall computational complexity of NLEM is $O(P^2S^2N_{\text{iter}})$ per pixel, where N_{iter} is the average number of iterations.

When using the framework to interpret NLEM, we find that the dissimilarity measures M_k $(k = 1, 2)$ in the two steps are different. Concretely, compared with $M_1 = \|\cdot\|_2^2$ (the squares of the L_2 norm), $M_2 = \|\cdot\|_2$ is more robust to heavy noise[25] and thus leads to better performance in denoising. However, the inconsistence between M_1 and M_2 in NLEM also leads to another consequence, i.e., this filter cannot be interpreted by the existing models[11-14]. Recently, a new non-local median filter (INLEM) is developed in [26], which not only compensates such inconsistence but also improves the denoising performance of NLEM. Meanwhile, the filter shares the same computational complexity of $O(P^2S^2N_{\text{iter}})$ per pixel as NLEM. Based on our framework, an interpretation of INLEM will be tabulated in the following section.

In addition to being able to reinterpret (analyze) some existing non-local methods, the two-step framework also provides us a platform to develop novel algorithms. And, in the following section, a novel non-local median filter for removing salt-pepper noise is developed based on it.

3 Non-Local Median Filter (NLMED)

3.1 Salt-Pepper Noise and Median Filters

It has been proved that NLM is a state-of-the-art filter for Gaussian noise[7]. However image noise is not only limited to this kind. For example, salt-pepper noise exists pervasively in real world too, which can usually be caused by analog-to-digital converter errors, bit errors in transmission, faulty memory locations, etc.[27] Denote the dynamic range of noisy image Y to be $[l_{\min}, l_{\max}]$, the noise model is defined as

$$Y(i)=\begin{cases} l_{\min}, & \text{with probability } p/2,\\ l_{\max}, & \text{with probability } p/2,\\ X(i), & \text{with probability } 1-p,\end{cases}$$

where p is the noise level.

Up to date, researchers have developed various filters to remove salt-pepper noise. For the typical median filter (MF), the fixed window and the equally-weighted nature limit its performance to some extent. As a result, different improved methods, including weighted median filter (WMF)[28], adaptive median filter (AMF)[29], etc., are successively developed. Among them, AMF is still widely used for its significantly ability in removing salt-pepper noise. In its implementation, AMF first manages to distinguish whether the pixel is polluted or not, and if true, then replace it by the median. Such an "identification-based" idea is a highlight of AMF. An overview of median-type filters can be found in [30].

3.2 Motivations

There are two possible factors which lead to the fact that NLM is not so effective in removing salt-pepper noise.

1) According to the reinterpretation of NLM in (6), the weight w_{ij} relies on the dissimilarity of two corresponding patches Y_i, Y_j, and the patch-based dissimilarities can preserve the order for Gaussian noise, i.e., the following equality (relation) holds[1].

$$E\|Y_i-Y_j\|_{2,\alpha}^2=\|X_i-X_j\|_{2,\alpha}^2+2\sigma^2.$$

However, this equality is not true anymore for salt-pepper noise.

2) An important characteristic of salt-pepper noise is that only some of the pixels are corrupted and the rest are all still noise free[29-30]. However, NLM manages to restore all pixels regardless of whether the pixels have been corrupted or not.

3.3 Implementations

According to the first factor mentioned above, an intuitive idea is to redefine the dissimilarity measures M_k only based on the noise-free candidates.

In the framework described in (3)∼(4), let $[p_{\min},p_{\max}]$ be the dynamic range of patches Y_i and Y_j, adopt a similar strategy to that of AMF, the noise-free candidate index set of the two patches is then defined by

$$\begin{aligned}K_u=\{&k|Y(i-k)\in Y_i,Y(j-k)\in Y_j,Y(i-k)\notin\\ &\{p_{\min},p_{\max}\},Y(j-k)\notin\{p_{\min},p_{\max}\},\\ &Y(j-k)\notin\{p_{\min},p_{\max}\}\}.\end{aligned}$$

Next, specially taking M_1 as the form of (8) as

$$\begin{aligned}M_1(Y_i,Y_j)&=\|Y_i-Y_j\|_{2,\alpha,u}^2\\ &=\sum_{k\in K_u}G_\alpha(k)(Y(i-k)-Y(j-k))^2.\end{aligned}\tag{8}$$

Then, for a fixed K_u, we have

$$\|Y_i-Y_j\|_{2,\alpha,u}^2=\|X_i-X_j\|_{2,\alpha,n}^2,\tag{9}$$

i.e., $M_1(Y_i,Y_j)=M_1(X_i,X_j)$. (See Appendix A.2 for the characteristics of M_1 in detail). (9) states that the novel M_1 can preserve the dissimilarity based on the noise-free candidates.

Considering the fact pointed out in Subsection 3.2, only some of the pixels are corrupted and the rests are all noise free, and we divide all the pixels into two classes, i.e., noise-free ones ($\Omega'=\cup K_u$) and noisy ones ($\Omega\backslash\Omega'$). And for the noise-free ones, we regularize them with the fidelity regularization term, i.e., $\sum_i\|X(i)-Y(i)\|_1$.

Further, considering that L_1 norm is more reasonable than L_2 norm for designing salt-pepper noise filters[31], we replace L_2 norm with L_1 norm to define the dissimilarity measure M_2. Except for adjusting M_k ($k=1,2$) and the working range, the other components of the framework in (3)∼(4) are set to the same as those for NLM in (6)∼(7). Now substituting the above settings to the framework, we get

$$\begin{aligned}\text{step 1}:\quad \hat{w}=&\arg\min_{w}\sum_{i\in\Omega\backslash\Omega',j\in\Omega'}w_{ij}\|Y_i-Y_j\|_{2,\alpha,u}^2+\\ &\lambda_1\sum_{i\in\Omega\backslash\Omega',j\in\Omega'}w_{ij}\ln w_{ij},\\ \text{s.t.}\quad &w_{ij}\geqslant 0,\quad \sum_j w_{ij}=1,\\ &i\in\Omega\backslash\Omega',\ j\in\Omega',\end{aligned}\tag{10}$$

$$\begin{aligned}\text{step 2}:\quad \widehat{X}=&\arg\min_{X}\sum\nolimits_{i\in\Omega\backslash\Omega',j\in\Omega'}\hat{w}_{ij}\|X(i)-\\ &Y(j)\|_1+\lambda_2\sum_{i\in\Omega'}\|X(i)-Y(i)\|_1,\end{aligned}\tag{11}$$

where $\Omega'=\cup K_u$ is the noise-free candidate index set of the whole image.

Solving the above optimization problem (Appendix A.3) turns out

$$\widehat{X}(i)=\begin{cases}Y(i), & \text{if } i\in\Omega',\\ WMed\{w_{ij},Y(j)\}, & \text{if } i\in\Omega\backslash\Omega',j\in\Omega',\end{cases}\tag{12}$$

where $WMed\{\cdot,\cdot\}$ is the weighted median operator[28] and $w_{ij}=\exp(-\frac{\|Y_i-Y_j\|_{2,\alpha,u}^2}{\lambda_1})$. The new non-local median filter is termed as NLMED for short. Similar to NLM, the complexity of NLMED is mainly from

computing patch-based dissimilarities in the non-local search window, and thus is also $O(P^2S^2)$ per pixel.

Although NLMED can be developed and interpreted based on our framework, similar to NLEM, it also cannot be interpreted based on the existing models[11-14] due to the difference between M_1 and M_2.

Here, a universal version of NLM (UNLM[19]) should be noted, which also borrows the "identification-based" idea from AMF to compute the patch-based dissimilarities. However, UNLM is significantly different from our NLMED.

1) Similar to NLM, UNLM is still a mean filter.

2) In UNLM, the pixels are recovered without fidelity regularization.

3) The patch size in UNLM is adaptive and limited by a threshold. As a result, the corresponding dissimilarity measure in this filter is different from that in NLMED.

4) The computational complexity of UNLM per pixel can be up to $O(P_{\max}^2 S^2)$, where $P_{\max} \times P_{\max}$ is the largest patch size adopted in UNLM. And our experiments will corroborate that UNLM generally consumes more denoising time than NLMED.

Here, we denote the dissimilarity measure in UNLM by $\|Y_i^t - Y_j^t\|_{2,\alpha,u}^2$ to distinguish from that in NLMED (t is the threshold). For more details of UNLM filter, we refer the reader to [19].

For more clarity, we summarize the non-local filters involved above in Table 2, in terms of the framework proposed in (3)$\sim$(4). The corresponding two-step optimization problems can be solved in the similar ways which have been proposed in Appendix A.1 and Appendix A.3 for reinterpreting NLM and NLMED respectively.

From Table 2, we can observe that our framework is very general and flexible. In addition, it leaves some blanks which are able to be filled to develop novel filters. And we can design other type filters by modifying the components in the framework. Naturally this work is just an initial realization.

Furthermore, what needs to be stressed is the difference between the existing models[11-14] and this proposed framework. For the former, despite adopting implementation in a two-step way, i.e., alternation between the computation of the non-local means and re-estimating the weights, their objective functions involved are all still single-step. Therefore, the operations in the two steps of the existing models should be consistent. And as have been illustrated, some non-local filters cannot be interpreted via the existing models because of the inconsistency of the dissimilarity measures employed respectively in the two steps. On the other hand, for this proposed framework, not only the implementation but also the objective definition is two-step, and thus it is more general and flexible than the existing ones.

3.4 Experiments

3.4.1 Experimental Settings

Similar to NLM which generalizes the mean filter from local to non-local, for removing salt-pepper noise, NLMED generalizes the median filter (MF) from local to non-local too. In the following experiments, we aim to validate the effects of our generalization by comparing it with its local counterparts, i.e., MF and WMF. Since the "identification-based" idea is borrowed in our generalization, we also compare it with AMF. Meanwhile, we take the universal version of NLM (UNLM) into our experiments, due to its activity in removing salt-pepper noise.

In the following experiments, we adopt 10 popularly-used images with 256×256 resolutions as listed in Fig.1.

For testing their denoising performances, we corrupt all the images using salt-pepper noises with different levels and then evaluate both visual effects and quantitative index (PSNR) in denoising. The PSNR[30] is defined as $PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right)$ with $MSE = \frac{1}{|\Omega|}\sum_{i\in\Omega}(\widehat{X}(i) - X(i))^2$. For MF, WMF and AMF, we conduct six rounds of experiments with 3×3, $5\times 5, \ldots,$

Table 2. Interpretations of Non-Local Filters Based on Our Framework

	Step 1 (Optimizing w)				Step 2 (Optimizing X)			
	Harmonious Term			Regularization Term	Harmonious Term			Regularization Term
	$\sum_{i,j\in D} w_{ij}^{p_1} M_1(Y_i, Y_j)$			$\lambda_1 J_1(w)$	$\sum_{i,j\in D} \hat{w}_{ij}^{p_2} M_2(X_i, Y_j)$			$\lambda_2 J_2(X)$
	D	p_1	M_1		D	p_2	M_2	
NLM[1]	$i,j \in \Omega$	1	$\|Y_i - Y_j\|_{2,\alpha}^2$	$h^2 \sum_{i,j\in\Omega} w_{ij} \ln w_{ij}$	$i,j \in \Omega$	1	$\|X(i) - Y(j)\|_2^2$	
UINTA[8]	$i \in \Omega$, $j \in A_i$	1	$\|Y_i^m - Y_j^m\|_{2,\alpha}^2$	$h^2 \sum_{i,j\in\Omega} w_{ij} \ln w_{ij}$	$i,j \in \Omega$	1	$\|X^{m+1}(i) - Y^m(j)\|_2^2$	
NLEM[15]	$i,j \in \Omega$	1	$\|Y_i - Y_j\|_2^2$	$h^2 \sum_{i,j\in\Omega} w_{ij} \ln w_{ij}$	$i,j \in \Omega$	1	$\|X_i - Y_j\|_2$	
INLEM[26]	$i,j \in \Omega$	1	$\|Y_i - Y_j\|_2$	$h^2 \sum_{i,j\in\Omega} w_{ij} \ln w_{ij}$	$i,j \in \Omega$	1	$\|X_i - Y_j\|_2$	
UNLM[19]	$i,j \in \Omega$	1	$\|Y_i^t - Y_j^t\|_{2,\alpha,u}^2$	$h^2 \sum_{i,j\in\Omega} w_{ij} \ln w_{ij}$	$i,j \in \Omega$	1	$\|X(i) - Y(j)\|_2^2$	
NLMED	$i \in \Omega\backslash\Omega'$, $j \in \Omega'$	1	$\|Y_i - Y_j\|_{2,\alpha,u}^2$	$\lambda_1 \sum_{i\in\Omega\backslash\Omega', j\in\Omega'} w_{ij} \times \ln w_{ij}$	$i \in \Omega\backslash\Omega'$, $j \in \Omega'$	1	$\|X(i) - Y(j)\|_1$	$\lambda_2 \sum_{i\in\Omega'} \|X(i) - Y(i)\|_1$

Note: UINTA is an iterative algorithm, m is the iteration flag, and A_i is a subset of Ω, chosen at random for each i.

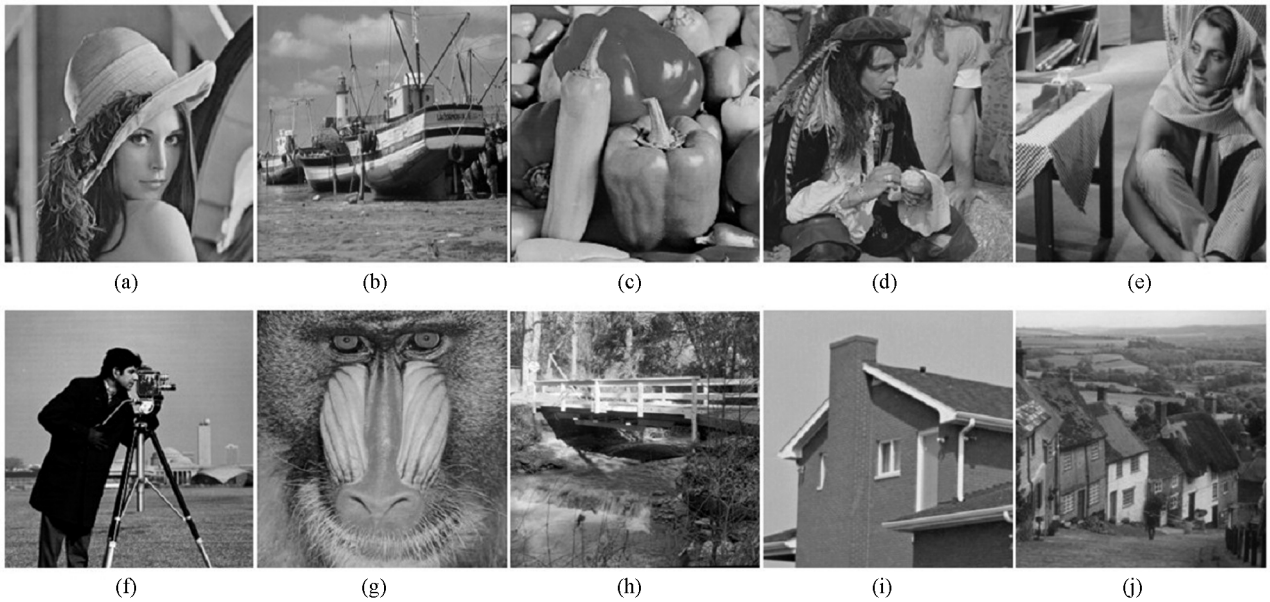

Fig.1. 10 popularly-used images on which we perform denoising experiments. (a) Lena. (b) Boats. (c) Peppers. (d) Man. (e) Barbara. (f) Cameraman. (g) Mandrill. (h) Stream. (i) House. (j) Hill.

13×13 windows respectively, and report their largest average PSNR on the 10 images respectively. In the experiments, we take the bell-shaped WMF[32] concretely. For UNLM, we set the parameters to the values recommended in the original references[19]. For our NLEM, the 21×21 search window is adopted, as that recommended in [1, 7] for NLM. And in order to sufficiently guarantee many uncorrupted pixels, the patch size in NLMED is set to 13×13, which is slightly larger than that in NLM. From step 2, we know that different $\lambda_2 > 0$ will lead to the same result for X. Meanwhile, as shown in Fig.2, by searching a large range of candidates on two tested images (Lena and Boats), we empirically set $\lambda_1 = 30$ in NLMED.

3.4.2 *Experimental Results*

Firstly, we show the CPU time of the filters involved in the experiments, where the test image "House" is corrupted with salt-pepper noise (30% noise level) and all the filters are implemented by MATLAB 7.11 codes on a PC equipped with INTEL 2.5-GHz CPU and 2.00 GB RAM. For the three local filters, the typical run time of MF is about 3 seconds, and that for WMF is about 24 seconds. AMF is the fastest, which consumes less than 1 second. Compared with the local filters, the non-local ones all need more time in denoising. The run time of NLMED is about 200 seconds. Especially for UNLM, the implementation consumes up to 1 500 seconds.

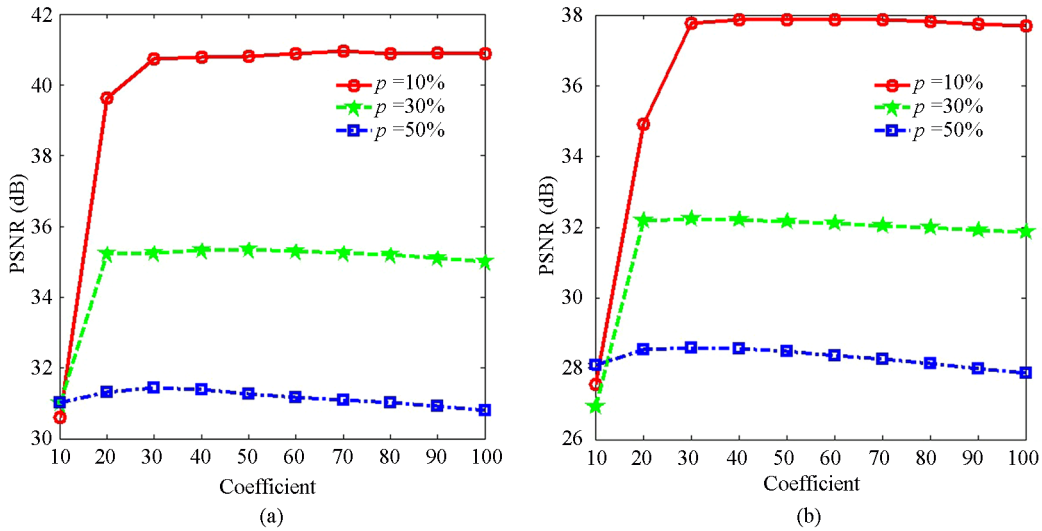

Fig.2. PSNR results filtered by NLMED versus different coefficients on two tested images. (a) Results on Lena. (b) Results on Boats.

Then, we summarize and report the average PSNR of the compared filters on all the 10 images respectively in Fig.3 and Table 3.

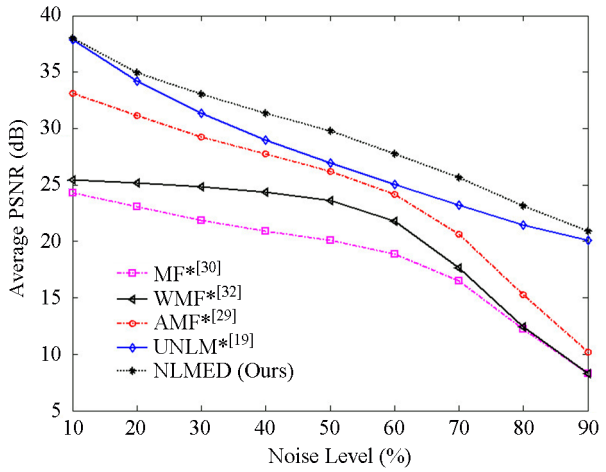

Fig.3. Results of the average PSNR on the 10 images (* stands for experiments conducted with 7×7 window).

Furthermore, we give the comparison of visual effects on the image "Barbara" with 30% noise level in Fig.4 as well.

From these results, we can obtain the following observations.

1) For the three local filters involved, WMF and AMF both more effectively than MF in removing salt-pepper noise. Especially AMF is significantly more powerful than the other two in all noise levels.

2) The two non-local filters (UNLM, NLMED) both obtain significantly larger PSNR and better visual effects than the local ones involved.

3) Compared with the visual effects of NLMED, the denoised result of UNLM shows over-smoothness on the texture areas to a certain degree. And the proposed NLMED obtains the best performance among the filters involved.

In addition to the observations above, a fact should be noted that considering computational complexity, NLMED just adopts the fixed rather than adaptive patch size as used in UNLM. Undoubtedly, the adoption of the adaptive technique can further improve the performance of our NLMED.

4 Conclusions

In this paper, we proposed a general two-step regularization framework for NLM. This framework is instinctive and flexible, from which we can analyze some non-local filters in a unified view. Moreover, this framework provides a platform to improve existing filters or design new ones. For instance, based on the framework, a non-local median filter (NLMED) was developed just through selecting dissimilarity measure, regularization term, and so on. Our final experiments exhibit encouraging results.

It must be noted that, similar to NLM, NLMED also has high computational complexity. Fortunately there are some speed up schemes that have been studied and designed[17-18]. Undoubtedly, they can be borrowed to speed up our filtering process.

To conclude, we noted some recent interesting studies[33-34], which generalize non-local schemes from denoising and deblurring to more general inverse problems including positron emission tomography (PET) image reconstruction, compressed sensing, etc. However, the objective functions involved in these schemes are still one-step, and thus not so flexible as the separate objective functions in the two-step framework proposed in this paper, which naturally provides a room for further improvement.

References

[1] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising. In *Proc. IEEE Computer Society Conference Computer Vision and Pattern Recognition* (*CVPR*), June 2005, Vol.2, pp.60-65.

Table 3. Average PSNR (dB) on the 10 Images for Salt-Pepper Noise

	Noise Level (%)								
	10	20	30	40	50	60	70	80	90
Noisy images	15.37	12.37	10.60	9.37	8.41	7.63	6.96	6.36	5.85
MF*[30]	24.28	23.10	21.87	20.92	20.10	18.90	16.48	12.21	8.29
WMF*[32]	25.43	25.18	24.85	24.37	23.61	21.79	17.62	12.43	8.30
AMF*[29]	33.13	31.18	29.26	27.73	26.17	24.19	20.63	15.25	10.17
MF**[30]	28.68	26.04	23.23	21.89	20.22	18.90	17.25	15.02	10.96
WMF**[32]	29.05	26.64	25.11	24.37	23.73	22.86	21.14	16.12	9.85
AMF**[29]	33.13	31.19	29.26	27.74	26.23	24.78	23.20	21.11	15.65
UNLM[19]	37.89	34.17	31.33	29.01	26.94	25.03	23.19	21.45	20.12
NLMED	**37.97**	**34.92**	**33.03**	**31.38**	**29.76**	**27.77**	**25.65**	**23.15**	**20.89**

Note: *: the average PSNR corresponding to 7×7 window. **: the largest average PSNR among the six rounds with different windows.

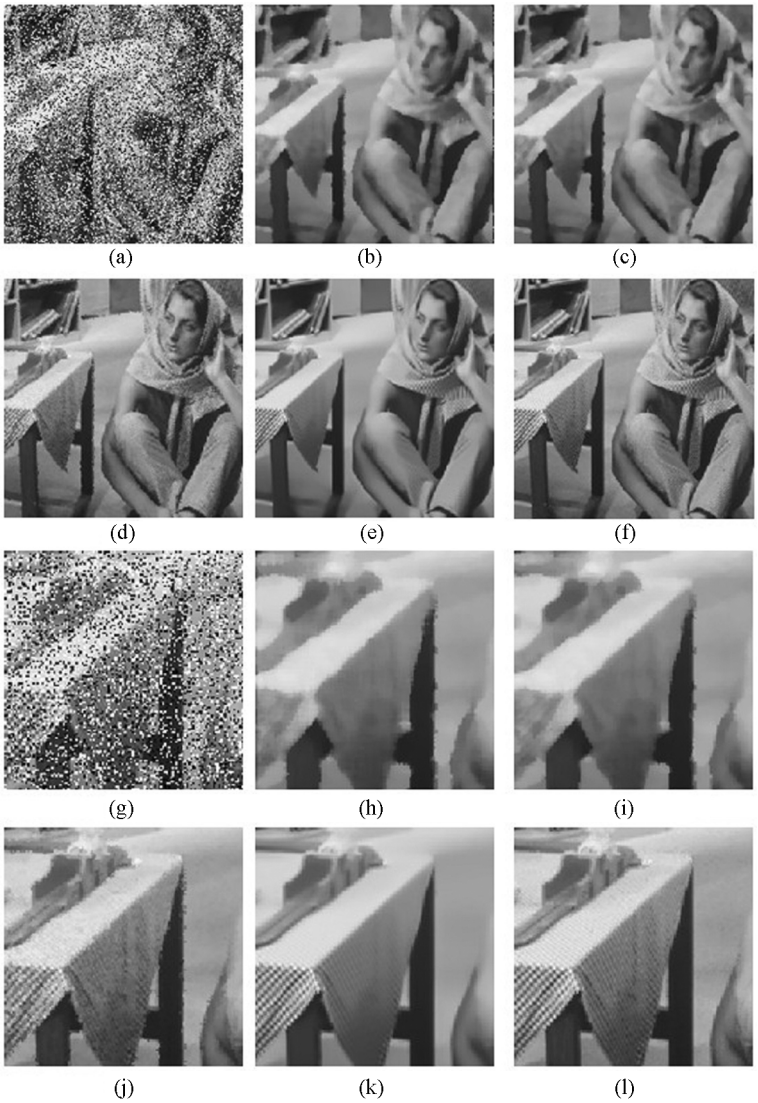

Fig.4. (a) Image "Barbara" corrupted by salt-pepper noise with 30% level. Restored images by (b) MF (7×7) with PSNR = 22.28 dB. (c) WMF (7×7) with PSNR = 25.65 dB. (d) AMF (7×7) with PSNR = 29.46 dB. (e) UNLM with PSNR = 32.81 dB. (f) NLMED with PSNR = 35.29 dB. (g)∼(l) are their corresponding closeups.

[2] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(7): 629-639.

[3] Yaroslavsky L P. Digital Picture Processing: An Introduction (1st edition). Springer-Verlag, 1985.

[4] Smith S M, Brady J M. SUSAN — A new approach to low level image processing. *Int. J. Computer Vision*, 1997, 23(1): 45-78.

[5] Tomasi C, Manduch R. Bilateral filtering for gray and color images. In *Proc. the 6th International Conference on Computer Vision* (*ICCV*), Jan. 1998, pp.839-846.

[6] Efros A A, Leung T K. Texture synthesis by non-parametric sampling. In *Proc. the 7th International Conference on Computer Vision* (*ICCV*), Sept. 1999, pp.1033-1038.

[7] Buades A, Coll B, Morel J M. Image denoising methods. A new nonlocal principle. *SIAM Review: Multiscale Modeling and Simulation*, 2010, 52(1): 113-147.

[8] Awate S P, Whitaker R T. Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(3): 364-376.

[9] Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095.

[10] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Processing*, 2006, 54(11): 4311-4322.

[11] Peyré G, Bougleux S, Cohen L. Non-local regularization of inverse problems. *Inverse Problems and Imaging*, 2011, 5(2): 511-530.

[12] Facciolo G, Arias P, Caselles V, Sapiro G. Exemplar-based interpolation of sparsely sampled images. In *Proc. the 7th Int. Conf. Energy Minimization Methods in Computer Vision and Pattern Recognition*, Aug. 2009, pp.331-344.

[13] Arias P, Caselles V, Sapiro G. A variational framework for non-local image inpainting. In *Proc. the 7th Int. Conf. Energy Minimization Methods in Computer Vision and Pattern*

Recognition, Aug. 2009, pp.345-358.

[14] Peyré G, Bougleux S, Cohen L. Non-local regularization of inverse problems. In *Lecture Notes in Computer Science 5304*, Forsyth D, Torr P, Zisserman A (eds.), Springer-Verlag, 2008, pp.57-68.

[15] Chaudhury K N, Singer A. Non-local Euclidean medians. *IEEE Signal Processing Letters*, 2012, 19(11): 745-748.

[16] Zhang L, Qiao L, Chen S. Graph-optimized locality preserving projections. *Pattern Recognition, 2010*, 43(6): 1993-2002.

[17] Dowson N, Salvado O. Hashed non-local means for rapid image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(3): 485-499.

[18] Vignesh R, Byung T O, Kuo C C J. Fast non-local means (NLM) computation with probabilistic early termination. *IEEE Signal Processing Letters*, 2010, 17(3): 277-280.

[19] Sun Z, Chen S. Modifying NL-means to a universal filter. *Optics Communications*, 2012, 285(24): 4918-4926.

[20] Malerba D, Esposito F, Gioviale V, Tamma V. Comparing dissimilarity measures for symbolic data analysis. In *Proc. Techniques and Technologies for Statistics - Exchange of Technology and Know-How*, June 2001, pp.473-481.

[21] Liu H, Song D, Stefan R, Hu R, Victoria U. Comparing dissimilarity measures for content-based image retrieval. In *Lecture Notes in Computer Science* 4993, Li H, Liu T, Ma W Y *et al.* (eds.), Springer-Verlag, 2008, pp.44-50.

[22] Sun J, Zhao W, Xue J, Shen Z, Shen Y. Clustering with feature order preferences. In *Lecture Notes in Computer Science 5351*, Ho T B, Zhou Z H (eds.), Springer-Verlag, 2008, pp.382-393.

[23] Luo P, Zhan G, He Q, Shi Z, Lu K. On defining partition entropy by inequalities. *IEEE Transactions on Information Theory*, 2007, 53(9): 3233-3239.

[24] Sun J, Zhao W, Xue J, Shen Z, Shen Y. Clustering with feature order preferences. *Intelligent Data Analysis*, 2010, 14: 479-495.

[25] Huber P J, Ronchetti E M. Robust Statistics (2nd edition). New Jersey: John Wiley & Sons, 2009.

[26] Sun Z, Chen S. Analysis of non-local Euclidean medians and its improvement. *IEEE Signal Processing Letters*, 2013, 20(4): 303-306.

[27] Cai J F, Chan R H, Nikolova M. Fast two-phase image deblurring under impulse noise. *Journal of Mathematical Imaging and Vision*, 2010, 36(1): 46-53.

[28] Brownrigg D R K. The weighted median filter. *Communications of the ACM*, 1984, 27(8): 807-818.

[29] Hwang H, Haddad R A. Adaptive median filters: New algorithms and results. *IEEE Transactions on Image Processing*, 1995, 4(4): 499-502.

[30] Bovik A. Handbook of Image and Video Processing. Academic Press, 2000.

[31] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, 1992, 60(1/2/3/4): 259-268.

[32] Yang R, Yin L, Gabbouj M, Astola J, Neuvo T. Optimal weighted median filtering under structural constraints. *IEEE Transactions on Signal Processing*, 1995, 43(3): 591-604.

[33] Wang G, Qi J. Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization. *IEEE Transactions on Medical Imaging*, 2012, 31(12): 2194-2204.

[34] Yang Z, Jacob M. Nonlocal regularization of inverse problems: A unified variational framework. *IEEE Transactions on Image Processing*, 2013, 22(8): 3192-3203.

Zhong-Gui Sun received his B.S. degree in mathematics from Shandong Normal University in 1996, and M.S. degree in computer software and theory from Sichuan Normal University in 2006. Currently, he is an associate professor at Department of Mathematics Science at Liaocheng University as well as a Ph.D. student at the Department of Computer Science and Engineering at Nanjing University of Aeronautics & Astronautics (NUAA). His research interests include image processing, pattern recognition, and machine learning.

Song-Can Chen received his B.S. degree in mathematics from Hangzhou University (merged into Zhejiang University in 1998) in 1983. In December 1985, he got his M.S. degree in computer applications at Shanghai Jiao Tong University and then worked at NUAA in January 1986 as an assistant lecturer. There he received his Ph.D. degree in communication and information systems in 1997. Since 1998, as a full professor, he has been with the College of Computer Science and Technology at NUAA. His research interests include pattern recognition, image processing, machine learning and neural computing.

Li-Shan Qiao received his B.S. degree in mathematics from Liaocheng University (LU) in 2001. He received his M.S. degree in applied mathematics from Chengdu University of Technology in 2004, and then worked in LU till today. In 2010, he received his Ph.D. degree in computer applications from the Department of Computer Science and Engineering, NUAA. Currently, he is an associate professor at the Department of Mathematics Science, LU, and his research interests focus on image processing, pattern recognition, and machine learning.

Appendix A.1 Solving the Optimization Problem in (6) and (7)

The optimization problem is described as follows:

$$\begin{aligned}\text{step 1}: \quad \hat{w} = &\arg\min_{w} \sum_{i,j\in\Omega} w_{ij}\|Y_i - Y_j\|_{2,\alpha}^2 + \\ &h^2 \sum_{i,j\in\Omega} w_{ij}\ln w_{ij}, \\ &\text{s.t.} \sum_{j\in\Omega} w_{ij} = 1, \quad w_{ij} \geqslant 0,\ i,j\in\Omega, \end{aligned} \tag{6}$$

$$\text{step 2}: \quad \widehat{X} = \arg\min_{X} \sum_{i,j\in\Omega} \hat{w}_{ij}\|X(i) - Y(j)\|_2^2. \tag{7}$$

For optimizing w_{ij} in (6), the corresponding Lagrangian function is constructed as follows:

$$L(w_{ij}) := \sum_{i,j\in\Omega} w_{ij}\|Y_i - Y_j\|_{2,\alpha}^2 + h^2 \sum_{i,j\in\Omega} w_{ij}\ln w_{ij} - \eta\Big(\sum_{i,j\in\Omega} w_{ij} - 1\Big).$$

Let $\frac{\partial L}{\partial w_{ij}} = \|Y_i - Y_j\|_{2,\alpha}^2 + h^2(\ln w_{ij} + 1) - \eta = 0$, then

$$w_{ij} = \exp\Big(\frac{\eta}{h^2} - 1\Big)\exp\Big(-\frac{1}{h^2}\|Y_i - Y_j\|_{2,\alpha}^2\Big).$$

Since $\sum_{j\in\Omega} w_{ij} = 1$, we get

$$\exp\Big(\frac{\eta}{h^2} - 1\Big) = \frac{1}{\sum\limits_{j\in\Omega}\exp\Big(-\frac{1}{h^2}\|Y_i - Y_j\|_{2,\alpha}^2\Big)}.$$

Thus, we obtain

$$\hat{w}_{ij} = \frac{1}{C(i)}\exp\Big(-\frac{\|Y_i - Y_j\|_{2,\alpha}^2}{h^2}\Big),$$

where, $C(i) = \sum_{j\in\Omega}\exp\big(-\frac{\|Y_i - Y_j\|_{2,\alpha}^2}{h^2}\big)$.

Similarly, for (7), let

$$\frac{\partial \sum\limits_{i,j\in\Omega} \hat{w}_{ij}\|X(i) - Y(j)\|_2^2}{\partial X(i)} = 2w_{ij}(X(i) - Y(j)) = 0,$$

we get

$$\begin{aligned}\widehat{X}(i) &= \sum_{j\in\Omega} \hat{w}_{ij}Y(j)\\ &= \frac{1}{C(i)}\sum_{j\in\Omega}\exp\Big(-\frac{\|Y_i - Y_j\|_{2,\alpha}^2}{h^2}\Big)Y(j).\end{aligned}$$

Appendix A.2 Characteristics of M_1 in (8)

In (8), the dissimilarity measure M_1 is defined as

$$\begin{aligned}M_1(Y_i, Y_j) &= \|Y_i - Y_j\|_{2,\alpha,u}^2\\ &= \sum_{k\in K_u} G_\alpha(k)(Y(i-k) - Y(j-k))^2,\end{aligned} \tag{8}$$

where

$$\begin{aligned}K_u = \{&k | Y(i-k)\in Y_i,\ Y(j-k)\in Y_j,\\ &Y(i-k)\notin\{p_{\min}, p_{\max}\}, Y(j-k)\notin\{p_{\min}, p_{\max}\},\\ &Y(j-k)\notin\{p_{\min}, p_{\max}\}\}.\end{aligned}$$

For a fixed K_u, we have

$$\begin{aligned}M_1(Y_i, Y_j) &= \sum_{k\in K_u} G_\alpha(k)(Y(i-k) - Y(j-k))^2\\ &= \sum_{k\in K_u} G_\alpha(k)(X(i-k) - X(j-k))^2\\ &= M_1(X_i, X_j),\end{aligned}$$

which can preserve the dissimilarity based on the noise-free candidates. Further, for any two image patches A and B, M_1 satisfies the following constraints clearly:

$$0 = M_1(A, A) \leqslant M_1(A, B) = M_1(B, A) < \infty.$$

That is, it is a dissimilarity measure defined in (5). However, M_1 does not satisfy the triangle inequality, which can be illustrated by the following three specific patches:

255	0	0
255	0	255
0	255	0

Y_1

255	60	0
255	80	100
200	130	120

Y_2

60	60	0
60	80	100
150	160	255

Y_3

The three patches are all 3×3 size, the maximum and the minimum gray levels of them are $\{p_{\min}, p_{\max}\} = \{0, 255\}$. From the definition in (8), it is clear that

$$M_1(Y_1, Y_2) = 0; \quad M_1(Y_1, Y_3) = 0; \quad M_1(Y_2, Y_3) > 0.$$

And thus

$$M_1(Y_1, Y_2) + M_1(Y_1, Y_3) < M_1(Y_2, Y_3),$$

i.e., it is not a distance metric.

Appendix A.3 Solving the Optimization Problem in (10) and (11)

The optimization problem is described as follows:

$$\begin{aligned}\text{step 1}: \quad \hat{w} =& \arg\min_{w} \sum_{i\in\Omega\backslash\Omega', j\in\Omega'} w_{ij}\|Y_i - Y_j\|_{2,\alpha,u}^2 + \\ &\lambda_1 \sum_{i\in\Omega\backslash\Omega', j\in\Omega'} w_{ij}\ln w_{ij},\\ \text{s.t.} \quad & w_{ij}\geqslant 0, \quad \sum_j w_{ij} = 1,\\ & i\in\Omega\backslash\Omega',\ j\in\Omega',\end{aligned} \tag{10}$$

$$\begin{aligned}\text{step 2}: \quad \widehat{X} =& \arg\min_{X} \sum_{i\in\Omega\backslash\Omega', j\in\Omega'} \hat{w}_{ij}\|X(i) - \\ &Y(j)\|_1 + \lambda_2 \sum_{i\in\Omega'} \|X(i) - Y(i)\|_1.\end{aligned} \tag{11}$$

Similar to optimizing w_{ij} in Appendix A.1, we first construct Lagrangian function for (10) as:

$$L(w_{ij}) = \sum_{i\in\Omega\backslash\Omega',j\in\Omega'} w_{ij}\|Y_i - Y_j\|^2_{2,\alpha,u} + \lambda_1 \sum_{i\in\Omega\backslash\Omega',j\in\Omega'} w_{ij}\ln w_{ij} - \eta\Big(\sum_{i\in\Omega\backslash\Omega',j\in\Omega'} w_{ij} - 1\Big).$$

Then let $\frac{\partial L}{\partial w_{ij}} = 0$, we obtain

$$\hat{w}_{ij} = \frac{1}{C(i)} \sum_{j\in\Omega'} \exp\Big(-\frac{\|Y_i - Y_j\|^2_{2,\alpha,u}}{\lambda_1}\Big) Y(j),$$

where $C(i) = \sum_{j\in\Omega'} \exp\big(-\frac{\|Y_i - Y_j\|^2_{2,\alpha,u}}{\lambda_1}\big)$ is a normalizing factor.

To optimize (11), we note that the pixels to be recovered in both the harmonious term and the regularization term are disjoint and thus can be optimized separately. Specifically, for any $i \in \Omega'$ and $\lambda_2 > 0$, we have $\widehat{X}(i) = Y(i)$ from the second term. Meanwhile, for any $i \in \Omega\backslash\Omega'$, we formally differentiate $\sum_{i\in\Omega\backslash\Omega',j\in\Omega'} w_{ij}\|X(i) - Y(j)\|_1$ with respect to $X(i)$ to get

$$\frac{\partial \sum_{i\in\Omega\backslash\Omega',j\in\Omega'} \hat{w}_{ij}\|X(i) - Y(j)\|_1}{\partial X(i)} = \sum_{i\in\Omega\backslash\Omega',j\in\Omega'} w_{ij}\mathrm{sign}(X(i) - Y(j)).$$

Zeroing the above derivative, we obtain

$$\begin{aligned}\widehat{X}(i) &= \mathit{WMed}\{w_{ij}, Y(j)\} \\ &= \mathit{WMed}\Big\{\Big(\frac{1}{C(i)}\exp\Big(-\frac{\|Y_i - Y_j\|^2_{2,\alpha,u}}{\lambda_1}\Big)\Big), Y(j)\Big\} \\ &= \mathit{WMed}\Big\{\exp\Big(-\frac{\|Y_i - Y_j\|^2_{2,\alpha,u}}{\lambda_1}\Big), Y(j)\Big\},\end{aligned}$$

where $\mathit{WMed}\{\cdot,\cdot\}$ is the weighted median operator. That is

$$\widehat{X}(i) = \begin{cases} Y(i), & \text{if } i \in \Omega', \\ \mathit{WMed}\{w_{ij}, Y(j)\}, & \text{if } i \in \Omega\backslash\Omega',\ j \in \Omega'. \end{cases} \tag{12}$$

In this way, we solve the optimization problem completely.