International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 8, No. 4 (2015), pp. 257-268
http://dx.doi.org/10.14257/ijsip.2015.8.4.23

Image Processing Model with K-support Norm

Junli Fan and *Xiaowei He

College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, P.R. China
E-mail: fjl9008@163.com
E-mail: jhhxw@zjnu.edu.cn

Abstract

In recent years, l_1 norm is usually considered as the regularization term in the field of sparse representation. However, the non-zero entries obtained by the l_1 regularization term always neglect the correlations with each other. In fact, different relationships or structures among non-zero entries are necessary in many applications. K-support norm is firstly proposed in the field of sparse prediction. The most important property of the k-support norm is grouping feature of the largest entries in the obtained solution. In this paper, we present a new image processing model by introducing the k-support norm to image gradient domain. The proposed model can be applied to image denoising and edge detection simultaneously. Some examples demonstrate the effectiveness of the novel model and its improvements.

Keywords: *k-support norm, image denoising, edge detection, sparse representation*

1. Introduction

Image denoising is a fundamental task in image science. The goal of denoising is to remove noise from a given noisy measurement:

$$b = Au + w \tag{1.1}$$

where $b \in R^m$ is the observed image, $A \in R^{m*n}$ is a linear operator and often an identity operator in classic denoising problems, $u \in R^n$ and $w \in R^m$ denote the true image and the noise respectively. In the past decades, a vast variety of denoising methods have been proposed, such as the ROF model [1], the wavelet-based approach [2], the non-local means filter [3] and the sparse and redundant representation model [4, 5]. The famous ROF model proposed by Rudin, Osher and Fatemi in [1] is demonstrated to be very successful in image denoising and is formulated as:

$$\min_u \{\frac{1}{2} \| Au - b \|^2 + \lambda \| |\nabla u| \|_1\} \tag{1.2}$$

where $\| |\nabla u| \|_1 = \int \sqrt{\nabla u_x^2 + \nabla u_y^2}$ and $A = I$, λ is the regularization parameter that should be positive. The model (1.2) is also called the isotropic ROF model. There are some efficient methods proposed in [6-9] to solve the model (1.2). We can see in (1.2) that the ROF model takes $\| |\nabla u| \|_1$ as its regularization term. It is very popular to obtain a sparse solution by a l_1 regularization term in various optimization problems including image processing [10], pattern recognition [11] and compressed sensing [12]. The

* Corresponding author: jhhxw@zjnu.edu.cn

ISSN: 2005-4254 IJSIP

methods with the l_1 regularization term have achieved great success in relevant applications.

However, the sparse representation method with l_1 regularization term is just minimizing the number of non-zero entries while neglecting the correlations among them. In fact, different relationships and structures among non-zero entries are often necessary in many applications [14, 15]. Therefore, a new group-variable selection method named elastic net [16] is proposed. The elastic net penalizes with both l_1 and l_2 norms and encourages a grouping effect. That means the elastic net pays more attention to the correlations among non-zero entries and often outperforms LASSO [13] that regularized with l_1 norm.

Similar to the elastic net, a novel norm named k-support norm $\| \cdot \|_k^{sp}$ is proposed in [17] for sparse prediction. According to its definition, it is a norm between l_1 and l_2 norms. But we should point out that this new norm is not equal to the $l_p (1 < p < 2)$ norm. Actually, the k-support norm chooses the largest k entries of a vector. The unit ball of the k-support norm is more "rounder" than that of LASSO penalty and the elastic net penalty. Namely, the tops of the unit balls of LASSO penalty and the elastic net penalty are not as smooth as the unit ball of the k-support norm. This property results the k-support norm less biased towards sparse vectors and can enhance the correlations among the non-zero entries of a solution. Besides, according to the designed method of the k-support norm, there are no more than k non-zero entries of every atom in the dictionary. That means the obtained solution is composed by the atoms which have no more than k non-zero entries.

As mentioned before, the solution obtained by (1.2) makes those non-zero entries individually. Moreover, a large proportion of $|\nabla u_x|$ and $|\nabla u_y|$ are close to zero. Our goal is to seek a sparse solution and take the correlations between non-zero entries into account. If we pay more attentions to the largest non-zero entries of $|\nabla u_x|$ and $|\nabla u_y|$, a natural choice for us is to choose the k-support norm $\| \nabla u \|_k^{sp}$ as the regularization term.

In this paper, we propose a new image processing model that regularized with $\| \nabla u_x \|_k^{sp}$ and $\| \nabla u_y \|_k^{sp}$. The new model can be applied to image denoising and edge detection simultaneously. We demonstrate the performance of the proposed model and show our model often outperforms the isotropic ROF model [1] and the anisotropic ROF model [18]. More details of our method are described in Section 3.

In Section 2, the k-support norm, its dual norm and the proximal method are stated. An image processing model with the k-support norm is proposed in Section 3. In Section 4, some examples are given to illustrate the effectiveness of the new model. In the last section, we draw some conclusions for correlated sparse representation.

2. K-support Norm and Related Notions

In this section, we introduce the k-support norm briefly and more details are in [17] and references therein. The k-support norm is given by the following formulation:

$$\| z \|_k^{sp} = \min\{ \sum_{I \in G_k} \| v_I \|_2 : support \quad (v_I) \subseteq I, \sum_{I \in G_k} v_I = z \} \tag{2.1}$$

where $z \in R^d$, $k \in \{1, \ldots, d\}$ and G_k denotes the set of all subsets of $\{1, \ldots, d\}$ of cardinality at most k. It is already mentioned in [17] that $\| \cdot \|_1^{sp} = \| \cdot \|_1$ and $\| \cdot \|_d^{sp} = \| \cdot \|_2$. Actually, the k-support norm is a norm between l_1 and l_2 norms with regard to $1 < k < d$. It is known to all that the unit ball of the l_1 norm is a square with four corners located at coordinate axes. This property leads to obtain a strictly sparse solution

for the relevant optimization problem. The unit ball of the k-support norm is much rounder than that of the l_1 norm. Therefore, the k-support norm tends to obtain edge-tangency with constraint hyper-plane when the unit ball is inflated gradually. It makes the solution of relevant optimization problem combined by some k-support atoms. Naturally, the obtained solution also has the features like all k-support atoms have. Besides, the dual norm of the k-support norm is:

$$\{\| z \|_k^{sp}\}^* = \max\{ < w, z >: \| w \|_k^{sp} \le 1\} = \max\{(\sum_{i \in I} z_i^2)^{\frac{1}{2}} : I \in G_k\} = (\sum_{i=1}^{k} (| z |_i^{\downarrow})^2)^{\frac{1}{2}} \quad (2.2)$$

Actually the dual norm is the l_2 norm of the largest k entries in vector z and interpolates between the l_∞ norm and the l_2 norm. If $k = 1$, we have $\| z \|_k^{sp} = \| z \|_1$ and its dual norm $\{\| z \|_k^{sp}\}^* = \| z \|_1^* = \| z \|_\infty$; when $k = d$, the k-support norm $\| z \|_k^{sp} = \| z \|_2$ and its dual norm $\{\| z \|_k^{sp}\}^* = \| z \|_2^* = \| z \|_2$. Moreover, when solving the optimization problem with the k-support regularization term, a proximal algorithm is given in [17]. Here, we describe it briefly in Algorithm 1.

According to the elaboration in [17], the k-support norm is tighter than the elastic net by a factor of $\sqrt{2}$ and can also encourage a grouping effect. When regularized with the k-support norm, we can obtain a sparse solution that not neglects the correlations among non-zero entries. Therefore, we introduce the k-support norm in image processing. However, if we apply the k-support norm to the image processing problem directly, the obtained solution will not be good enough. Take a 256*256 image as an example. In such an image, there are 65536 pixels. When the k-support norm is chosen to sparse represent the image directly, k pixels with a big gray value are retained while the rest pixels are disregarded. Since the gray value of a pixel distributes randomly in [0, 255], the value of k should set to be very large so that the k-support norm can represent the image precisely. However, it is contrary to our original intention and also impracticable.

Considering the problem in the image gradient domain, the points with a big gradient value are always located around the edges and points with a gradient value close to zero are located in smooth areas. In an image, the most parts are smooth areas and only a small portion of it is edges. That is to say, in the gradient domain, the number of the points with a big gradient value is small and a lot of points with a gradient value close to zero. Moreover, the gradient values of points in the same edge are almost the same with each other. Therefore when one of them has been selected, then the rest may also be picked out. These characters are well in line with the properties of the k-support norm. Naturally, we introduce the k-support norm to our model in the following section.

Input: $v \in R^d$

Output: $q = prox_{\frac{1}{2L}(\|\cdot\|_k^{sp})^2}(v)$

Find $r \in \{0, 1, ..., k-1\}$, $l \in \{k, ..., d\}$, such that

$$\frac{1}{L+1} z_{k-r-1} > \frac{T_{r,l}}{l-k+(L+1)r+L+1} \geq \frac{1}{L+1} z_{k-r}$$

$$z_l > \frac{T_{r,l}}{l-k+(L+1)r+L+1} \geq z_{l+1}$$

where $z := |v|^{\downarrow}$, $z_0 := +\infty$, $z_{d+1} := -\infty$, $T_{r,l} := \sum_{i=k-r}^{l} z_i$

$$q = \begin{cases} \frac{L}{L+1} z_i & if \quad i = 1, ..., k-r-1 \\ z_i - \frac{T_{r,l}}{l-k+(L+1)r+L+1} & if \quad i = k-r, ..., l \\ 0 & if \quad i = l+1, ..., d \end{cases}$$

Algorithm 1. Computation of the Proximal Operator for $\|\cdot\|_k^{sp}$ Norm

3. Image Processing Model and Optimization

In this section, an image processing model with the k-support norm is proposed. We formulate the model as:

$$\min_u \{\frac{1}{2}\|u-f\|^2 + \frac{\alpha_1}{2}(\|\nabla u_x\|_k^{sp})^2 + \frac{\alpha_2}{2}(\|\nabla u_y\|_k^{sp})^2\} \tag{3.1}$$

where f denotes the observed image, $\alpha_1, \alpha_2 > 0$ are regularization parameters, ∇u_x and ∇u_y are the gradient in x and y direction respectively. To settle (3.1), we introduce two auxiliary variables z_x and z_y to separate the calculation of the non-differentiable terms and the fidelity term. The model (3.1) is thus equivalent to the following formulation:

$$\min\{\frac{1}{2}\|u-f\|^2 + \frac{\alpha_1}{2}(\|z_x\|_k^{sp})^2 + \frac{\alpha_2}{2}(\|z_y\|_k^{sp})^2 \quad s.t.\ z_x = \nabla u_x \quad z_y = \nabla u_y\} \tag{3.2}$$

The constrained optimization problem (3.2) can be solved by the Augmented Lagrangian method [19, 20].

$$\min_{u,z_x,z_y} \max_{\mu_x,\mu_y} L_c(u, z_x, z_y, \mu_x, \mu_y) = \frac{1}{2}\|u-f\|^2 + \frac{\alpha_1}{2}(\|z_x\|_k^{sp})^2 + \frac{\alpha_2}{2}(\|z_y\|_k^{sp})^2 +$$

$$< z_x - \nabla u_x, \mu_x > + \frac{r_1}{2}\|z_x - \nabla u_x\|^2 + < z_y - \nabla u_y, \mu_y > + \frac{r_2}{2}\|z_y - \nabla u_y\|^2$$

(3.3)

where μ_1, μ_2 are the Lagrangian multipliers and r_1, r_2 are positive constants. To settle the problem (3.3), we should seek a saddle point of the augmented Lagrangian function $L_c(u, z_x, z_y, \mu_x, \mu_y)$ [20]. So we solve the problem (3.3) as following.

Fix μ and z, the first is for u :

$$\min_{u} \{\frac{1}{2}\| u - f \|^2 + < z - \nabla u, \mu > + \frac{r}{2}\| z - \nabla u \|^2\} \tag{3.4}$$

for the sake of simplicity, we use $< z - \nabla u, \mu >$ denotes $\{< z_x - \nabla u_x, \mu_x > + < z_y - \nabla u_y, \mu_y >\}$ and $\frac{r}{2}\| z - \nabla u \|^2$ denotes $\{\frac{r_1}{2}\| z_x - \nabla u_x \|^2 + \frac{r_2}{2}\| z_y - \nabla u_y \|^2\}$. The optimality condition for (3.4) is:

$$\frac{\partial L_c}{\partial u} = (u - f) + \nabla \cdot \mu + r\nabla \cdot (z - \nabla u) = 0 \tag{3.5}$$

The second is to obtain the solutions of z_x and z_y by fixing μ and u. To settle this problem, we need to address the following two sub-problems:

$$\min_{z_x} \{\frac{\alpha_1}{2}(\| z_x \|_k^{sp})^2 + < z_x - \nabla u_x, \mu_x > + \frac{r_1}{2}\| z_x - \nabla u_x \|^2\} \tag{3.6}$$

$$\min_{z_y} \{\frac{\alpha_1}{2}(\| z_y \|_k^{sp})^2 + < z_y - \nabla u_y, \mu_y > + \frac{r_2}{2}\| z_y - \nabla u_y \|^2\} \tag{3.7}$$

(3.6) and (3.7) are also optimization problems. We firstly solve the problem (3.6) as following:

$$\begin{aligned}
&\min_{z_x} \{\frac{\alpha_1}{2}(\| z_x \|_k^{sp})^2 + \frac{r_1}{2}(\| z_x - \nabla u_x \|^2 + \frac{2}{r_1} < z_x - \nabla u_x, \mu_x > + \frac{1}{r_1^2}\mu_x^2 - \frac{1}{r_1^2}\mu_x^2)\} \\
&= \min_{z_x} \{\frac{\alpha_1}{2}(\| z_x \|_k^{sp})^2 + \frac{r_1}{2}(\| z_x - \nabla u_x \|^2 + \frac{2}{r_1} < z_x - \nabla u_x, \mu_x > + \frac{1}{r_1^2}\mu_x^2) - \frac{1}{2r_1}\mu_x^2\} \\
&= \min_{z_x} \{\frac{\alpha_1}{2}(\| z_x \|_k^{sp})^2 + \frac{r_1}{2}\| z_x - \nabla u_x + \frac{1}{r_1}\mu_x \|^2 - \frac{1}{2r_1}\mu_x^2\}
\end{aligned} \tag{3.8}$$

Now do the same steps to (3.7), we can get

$$\min_{z_y} \{\frac{\alpha_2}{2}(\| z_y \|_k^{sp})^2 + \frac{r_2}{2}\| z_y - \nabla u_y + \frac{1}{r_2}\mu_y \|^2 - \frac{1}{2r_2}\mu_y^2\} \tag{3.9}$$

after ignoring constant terms, (3.8) and (3.9) can be rewritten as:

$$\min_{z_x} \{\frac{\alpha_1}{2}(\| z_x \|_k^{sp})^2 + \frac{r_1}{2}\| z_x - (\nabla u_x - \frac{1}{r_1}\mu_x) \|^2\} \tag{3.10}$$

$$\min_{z_y} \{\frac{\alpha_2}{2}(\| z_y \|_k^{sp})^2 + \frac{r_2}{2}\| z_y - (\nabla u_y - \frac{1}{r_2}\mu_y) \|^2\} \tag{3.11}$$

The Lagrangian multipliers μ_1 and μ_2 are updated by (3.12) and (3.13) respectively:

$$\mu_x^{k+1} = \mu_x^k + r_1(z_x^{k+1} - \nabla u_x^{k+1}) \tag{3.12}$$

$$\mu_y^{k+1} = \mu_y^k + r_2(z_y^{k+1} - \nabla u_y^{k+1}) \tag{3.13}$$

Deducing from (3.5), (3.10), (3.11), (3.12) and (3.13), the solution for (3.3) can be obtained by iterating the followings:

$$\begin{cases} u^{k+1} = (I - r\Delta)^{-1}(f - div(\mu^k) - r * div(z^k)) \\ z_x^{k+1} = \arg\min_{z_x} \frac{r_1}{2} \| z_x^k - (\nabla u_x^{k+1} - \frac{\mu_x^k}{r_1}) \|^2 + \frac{\alpha_1}{2} (\| z_x \|_k^{sp})^2 \\ z_y^{k+1} = \arg\min_{z_y} \frac{r_2}{2} \| z_y^k - (\nabla u_y^{k+1} - \frac{\mu_y^k}{r_2}) \|^2 + \frac{\alpha_2}{2} (\| z_y \|_k^{sp})^2 \\ \mu_x^{k+1} = \mu_x^k + r_1 (z_x^{k+1} - \nabla u_x^{k+1}) \\ \mu_y^{k+1} = \mu_y^k + r_2 (z_x^{k+1} - \nabla u_y^{k+1}) \end{cases} \tag{3.14}$$

where $\mu^k = (\mu_x^k, \mu_y^k)^T$, $z^k = (z_x^k, z_y^k)^T$. The first line in (3.14) can be solved easily. Here, we use the function handle and the PCG function in Matlab to solve it and the solution achieves very high precision. The second and third lines are solved by the proximal method:

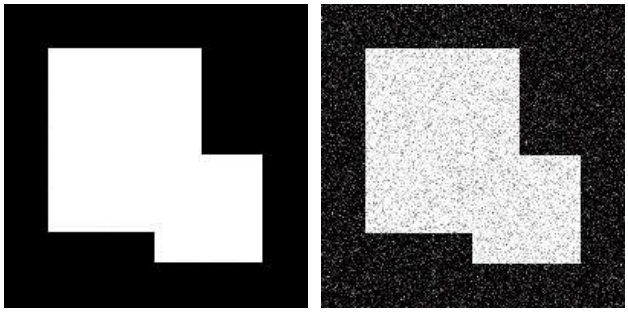
$$z_x^{k+1} = Prox_{\frac{\alpha_1}{2}(\|\cdot\|_k^{sp})^2, r_1}(z_x^k - \theta\delta_1(z_x^k - \nabla u_x + \frac{\mu_x^k}{\alpha_1})) \tag{3.15}$$

$$z_y^{k+1} = Prox_{\frac{\alpha_2}{2}(\|\cdot\|_k^{sp})^2, r_2}(z_y^k - \theta\delta_2(z_y^k - \nabla u_y + \frac{\mu_y^k}{\alpha_2})) \tag{3.16}$$

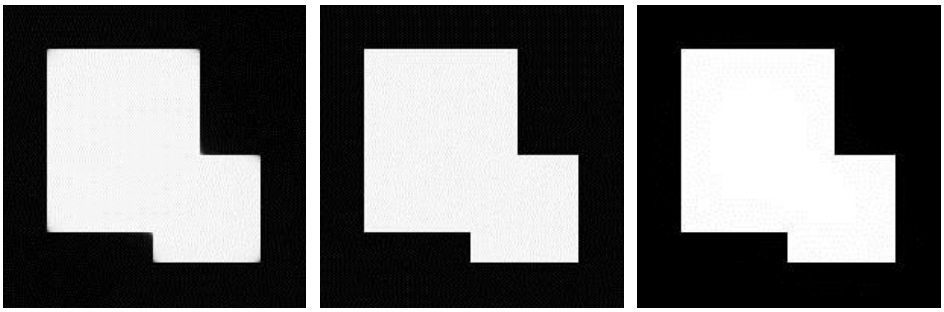
where δ_1 and δ_2 are the Lipschitz constants of the convex functions $\frac{r_1}{2} \| z_x^k - (\nabla u_x^{k+1} - \frac{\mu_x^k}{r_1}) \|^2$ and $\frac{r_2}{2} \| z_y^k - (\nabla u_y^{k+1} - \frac{\mu_y^k}{r_2}) \|^2$, $\theta \in (0, 2)$.

4. Examples

In this section, some experimental results are given to demonstrate the effectiveness of our model. We compare our model with the isotropic ROF model [1] and anisotropic ROF model [18]. Two simple geometry examples are firstly designed for comparing our new model with the two ROF models in terms of the visual effects, corner protecting, edge-preserving and the brightness level. In all experiments, we set the most suitable parameters and do enough iteration when applying the two ROF models.



(a) (b)



(c) (d) (e)

Figure 1. (a) Original Image (b) Noisy Image with $\sigma = 50$, PSNR=14.1 (c) Denoising Result by the Isotropic ROF Model, PSNR=29.8 (d) Denoising Result by the Anisotropic ROF Model, PSNR=30.79 (e) Denosing Result by Our Model, PSNR=39.45

The first example is shown in Figure 1(a). It is a clean 200*200 pixel image with two rectangles located in (30:150, 30:130) and (100:170, 100:170) respectively. We add the zero-mean additive Gaussian noise with $\sigma = 50$ to the image and get the corrupted version that is shown Figure 1(b). Figure 1(c) and (d) are the denoising results of the isotropy ROF model [1] and the anisotropy ROF model [18] respectively. The denoising result of our method with $k = 20$ is shown in Figure 1(e). We can see directly that Figure 1(e) looks much better than Figure 1(c). Obviously, one shortcoming of Figure 1(c) is that the corners are seriously defiled while they are well preserved in Figure 1(e). Looking carefully, we can find that Figure 1(e) is brighter than Figure 1(d) in terms of visual effects. Moreover, the PSNR value of Figure 1(e) is much better than that of Figure 1(d).

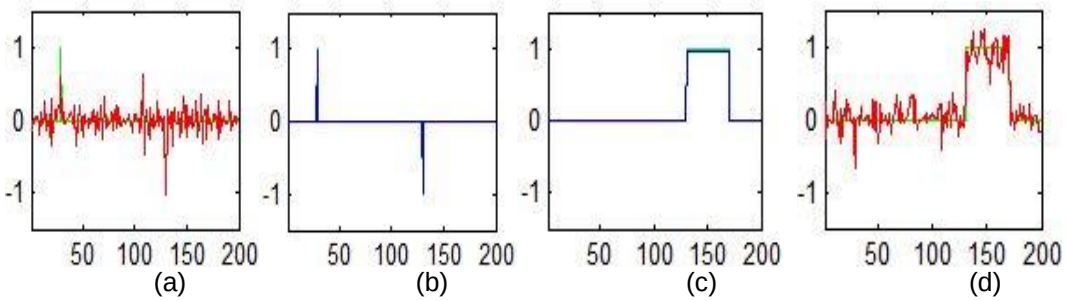


Figure 2. The Gradient of 99th Row for Figure 1(a), Figure 1(d) and Figure 1(e)

Figure 2 shows the gradient value of the 99th row (the top edge of the small rectangle) for Figure 1(a), Figure 1(d) and Figure 1(e). The green line in Figure 2 is for the original image Figure 1(a). The red line and the blue line are for Figure 1(d) and Figure 1(e) respectively. Figure 2(a) and Figure 2(b) are the gradient in x direction and the rest of Figure 2 is for the gradient in y direction. Comparing with the red line in Figure 2, the blue line fits the green line much better. We can see that in Figure 2(a) and Figure 2(c), the green line and the blue line are almost overlapping. It means our new method can provide a well approximation for the gradient of the original image. This property can also be used to edge detection and its applications will be given latter.

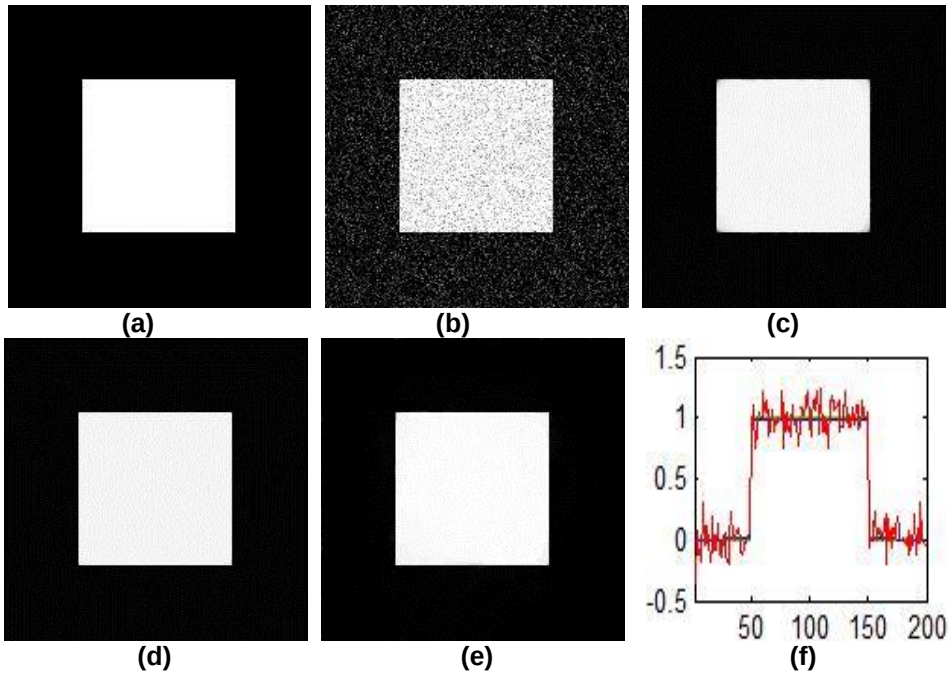


Figure 3. (a) Original Image (b) Noising Image with $\sigma = 50$, PSNR=14.1 (c) Denoising Result by Isotropic ROF Model, PSNR=31.7 (d) Denoising Result by Anisotropic ROF Model, PSNR=32.6 (e) Denoising Result by Our Model, PSNR=44.30 (f) Intensity of the 100th Row Pixels

The second example is shown in Figure 3. Figure 3(a) is the original image of size 200*200 and the white rectangle is located in (50:150, 50:150). After adding the zero-mean additive Gaussian noise with $\sigma = 50$, we get the noisy image Figure 3(b) (PSNR=14.1). Figure 3(c) and Figure 3(d) are the denoising results of the isotropic ROF model and the anisotropic ROF model respectively. Figure 3(e) is the denoising result by (3.1) with $k = 20$. We see that four corners in Figure 3(c) are defiled. It means that the isotropic ROF model fails to protect corners. Similar to the first example, the denoising result of our model is brighter than that of the anisotropic ROF model in terms of visual effects. Moreover, we find the PSNR value of our method is much better than the anisotropic ROF model. Looking locally, Figure 3(f) shows the grey value of the 100th row pixels for Figure 3(a), Figure 3(d) and Figure 3(e). We use the blue line and the green line to denote the denoising result of our model and the original image respectively. The red line is for the anisotropic ROF model. In Figure 3(f), we see that the blue line fits the green line much better than the red line. This means that our method outperforms the anisotropic ROF model.

The above examples illustrate our model performs much better than the two ROF models when denoising simple geometry images. Now we give two natural image examples to compare our model with the two ROF models. In the experiments, we set the most suitable parameters and do enough iteration to the two ROF models. The value of k in our model is set to 28 in both real image examples. The results are shown in Figure 4 and Figure 5. As we see, Figure 4(a) is the original Hepburn image and the noisy version of it is shown in Figure 4(b). Here we add the zero-mean additive Gaussian noise with $\sigma = 25$ to the original image. The denoising results of the isotropic ROF model and the anisotropic ROF model are shown in Figure 4(c) and Figure 4(d) respectively. Figure 4(e) is the denoising result of our method. Figure 4(f) is the edge detection result by our method and more details are shown in Figure 4. In the second real image experiment, we use the Pepper image that is shown in Figure 5(a). The same noise is added to the original image and the corrupted version is shown in Figure 5(b). All the results obtained in the

second experiment are shown in the rest of Figure 5. The PSNR values of denoising results demonstrate our method performs better than two ROF models. Moreover, we can obtain the edge detection result simultaneously. Figure 4(f) and Figure 5(f) show the edge detection results of our method. We can see the edge detection results fit the edges of the original image very well. After doing these two experiments, we demonstrate the model (3.1) can be applied to image denoising and edge detection simultaneously.

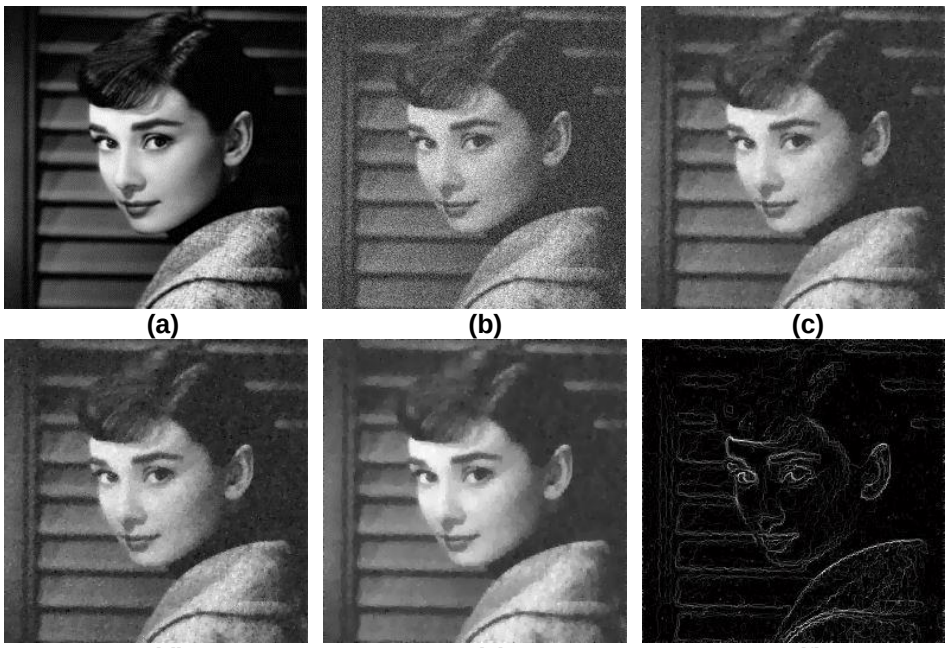
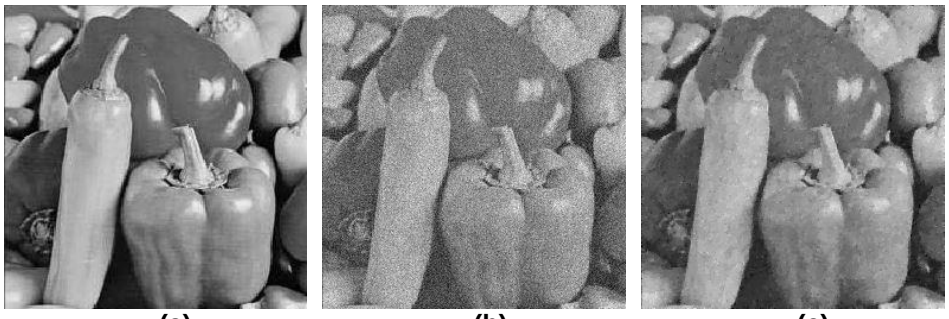


Figure 4. (a) Original Image (b) Noisy Image with $\sigma = 25$, PSNR=20.2 (c) Denoising Result by Isotropic ROF Model, PSNR=27.44 (d) Denoising Result by Anisotropic ROF Model, PSNR=27.59 (e) Denoising Result by Our Model, PSNR=28.47 (f) Edge Detection Result by Our Method



(d) (e) (f)

Figure 5. (a) Original Image (b) Noisy Image with $\sigma = 25$, PSNR=20.15 (c) Denoising Result by Isotropic ROF Model, PSNR=27.51 (d) Denoising Result by Anisotropic ROF Model, PSNR=27.64 (e) Denoising Result by Our Method, PSNR=28.79 (f) Edge Detection Result by Our Model

Besides, a series of experiments on the two natural images have been done. We use different parameters and compare our method with the two ROF models. The results of experiments are given in Table 1 and Table 2. In experiments, we add different noise to original images to get corrupted version, then use our method and the two ROF models to deal with the noisy images. When applying our model, different k are set to obtain different denoising results. As we can see, under a little noisy situation, the denoising results by our method are slightly better than the two ROF models. However, with the increase of the noise level, our method performs much better than the latter two. There exist significant differences between our method and the two ROF models. In addition, the value of k can also influence the result of our method. Under the heavy noisy situation, the value of k has a great influence on PSNR and should not be chosen too large. We find that with the increasing of k , our method takes more time. It is meaningful for us to choose an appropriate k for our model.

Table 1. PSNR Values of Denoisied Hepburn Image with Different Parameters

Hepburn		$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$	$\sigma = 35$
ours	k=10	32.52	30.69	29.56	28.56	27.72	27.09
	k=20	32.72	30.83	29.63	28.55	27.60	26.83
	k=30	32.82	30.87	29.60	28.36	27.44	26.53
	k=40	32.89	30.82	29.46	28.09	27.08	25.96
	k=50	32.95	30.78	29.53	27.67	26.91	25.74
ROF [1]		31.53	30.70	29.44	27.39	25.17	23.22
ani-ROF [18]		31.51	30.75	29.56	27.53	25.29	23.32

Table 2. PSNR Values of Denoisied Pepper Image with Different Parameters

Pepper		$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$	$\sigma = 35$
ours	k=10	32.71	30.98	29.71	28.80	27.92	27.12
	k=20	33.01	31.26	29.84	28.86	27.90	27.06
	k=30	33.26	31.35	29.91	28.73	27.87	26.78
	k=40	33.43	31.33	29.73	28.51	27.35	26.18
	k=50	33.46	31.32	29.71	28.16	27.01	25.83
ROF [1]		32.03	31.20	29.63	27.55	25.28	23.16
ani-ROF [18]		32.09	31.28	29.73	27.67	25.40	23.25

Although many experiments have been done, they are on the same test images. In order to verify the practicability of our model, we also do some experiments on other images. In these experiments, we set the most suitable value to k and add the zero-mean additive Gaussian noise with different σ to the test images. Table 3 shows the PSNR values of the denoising images. The results demonstrate that our method is well performed with different pictures.

Table 3. PSNR Values of Different Denoising Images by Our Method

Image	$\sigma = 10$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 30$	$\sigma = 35$
barbara	31.15	29.10	27.66	26.59	25.75	24.95
bowl	32.50	30.49	29.03	27.91	27.04	26.31
chart	29.30	27.53	27.20	26.79	26.31	25.53
clock	32.91	31.31	30.05	28.76	27.44	26.47
bartexture	29.78	27.55	26.20	25.21	24.58	24.08
lena	32.05	30.07	28.79	27.74	26.80	25.75
cameraman	31.13	29.46	28.31	27.42	26.45	25.62
plane	36.41	33.78	31.75	30.61	28.87	27.53

After doing a series of experiments on different examples, the results indicate that our model (3.1) outperforms the isotropic ROF model and the anisotropic ROF model. Moreover, the edge detection result can be obtained simultaneously since we use the k-support norm to sparse represent the gradient of the image.

5. Conclusions

Considering the largest k non-zero entries of the sparse vectors, we propose a new model with the k-support norm for image processing. In this model, we use the k-support norm to sparse represent the gradient of the image. Our model can be applied to image denoising and edge detection simultaneously. In fact, how to organize atoms in dictionary for correlated sparse norm in image processing is worth to be studied and is easily extensible according to different tasks.

Acknowledgements

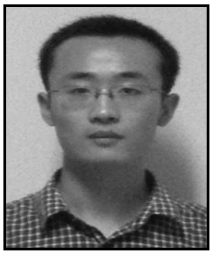
This research is supported by Zhejiang Provincial Natural Science Foundation of China (Grant No: LY14F010008).

References

[1] L. I. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms", Physica D: Nonlinear Phenomena, vol. 60, no. 1, **(1992)**, pp. 259-268.
[2] J. Portilla, V. Strela, M. Wainwright and E. Simoncclli, "Image denoising using scale mixtures of guassians in the wavelet domain", IEEE Transactions on Image Processing, vol. 12, no. 11, **(2003)**, pp. 1338-1351.
[3] A. Buades, B. Coll and J. M. Morel, "A review of image denoising algorithms with a new one", Multiscale Modeling and Simulation, vol. 4, no. 2, **(2005)**, pp. 490–530.
[4] A. M. Bruchstein, D. L. Donoho and M. Elad, "From sparse solutions of systems of equaitons to sparse modeling of signals and images", SIAM review, vol. 51, no. 1, **(2009)**, pp. 34-81.
[5] M. Elad, "Sparse and redundant representation: from theory to applications in signal and image processing", Springer, **(2010)**.
[6] A. Chambolle, "An algorithm for total variation minimization and applications", Journal of Mathematical imaging and vision, vol. 20, no. 1-2, **(2004)**, pp. 89-97.
[7] T. F. Chan, G. H. Golub and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration", SIMA Journal on Scientific Computing, vol. 20, no. 6, **(1999)**, pp. 1964-1977.
[8] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin, "An iterative regularization method for total nariation-based image restoration", Multiscale Modeling and Simulation, vol. 4, no. 2, **(2005)**, pp. 460-489.

[9] W. Yin, S. Osher, D. Goldfarb and J. Darbon, "Bregman iterative algorithms for l1-minimization with applications to compressed sensing", SIAM Journal on Imaging Sciences, vol. 1, no. 1, **(2008)**, pp. 143-168.
[10] M. Elad and M. Aharon, "Image squence denoising via sparse and redundant representations", IEEE Trans on Image Processing, vol. 18, no. 1, **(2009)**, pp. 27-35.
[11] J. Wright, Y. Ma and J. Maival J, G. Sapiro, T. Huang and S. Yan, "Sparse representation for computer vision and pattern recognition", Preceedings of the IEEE, vol. 98, no. 6, **(2010)**, pp. 1031-1044.
[12] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling", Inverse Problems, vol. 23, **(2007)**, pp. 969-985.
[13] R. Tibshirani, "Regression shrinkage and selection via the lasso", Journal of the Royal Statistical Society. Series B (Methodological), vol. 58, no. 1, **(1996)**, pp. 267-288.
[14] M. Elad, "Sparse and Redundant Representation Modeling — What Next?", IEEE Signal Processing Letters, vol. 19, no. 12, **(2012)**, pp. 922-928.
[15] J. Friedman, T. Hastie, H. Hofling and R. Tibshirani, "Pathways coordinate optimization", The Annals of Applied Statistics, vol. 1, no. 2, **(2007)**, pp. 302–332.
[16] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net", Journal of the Royal Statistical Society. Series B (Statistical Methodology), vol. 67, no. 2, **(2005)**, pp. 301–320.
[17] A. Argyriou, R. Foygel and N. Srebro, "Sparse Prediction with the k-Support Norm", **(2012)**, arXiv:1204.5043v2 [stat.ML],
[18] S. Esedoglu and S. Osher, "Decomposition of images by the anisotropic Rudin-Osher-Fatemi model", Communications on pure and applied mathematics, vol. 57, no. 12, **(2004),** pp. 1609-1626.
[19] J. Nocedal and S.J. Wright, "Penalty and Augmented Lagrangian Methods", Springer, **(2006)**.
[20] X. C. Tai and C. Wu, "Augmented lagrangian method, dual methods and split Bregman iteration for ROF model", Scale Space and Variational Methods in Computer Vision, **(2009)**, pp. 502-513.

Authors



Junli Fan, was born in China, in August 1990. He received the B.S. degree in Telecommunication Engineering from Hangzhou Dianzi University in 2012, China. He is currently studying for a M.S. degree in Computer Science in Zhejiang Normal University. His main research interest is image processing including image denoising, deblurring and reconstruction.

E-mail: fjl9008@163.com

Xiaowei He, was born in China, in April 1968. He received the B.S degree in Mathematics form Zhejiang Normal University, and the M.S degree in Computer Application from Zhejiang University, Jinhua, P.R.China in 1992 and Hangzhou in 2000 respectively. Now, he is an associate professor and also a Master Tutor in Computer Software and Theory in Zhejiang Normal University. His research interests include image processing, signal processing and inverse problems.

E-mail: jhhxw@zjnu.edu.cn

