OPTIMIZED PARAMETERS FOR BELL-SHAPED ERROR FUNCTION IN IMAGE DENOISING

Kemal ÖZKAN [1, *], Erol SEKE [2]

[1] Department of Computer Engineering, Faculty of Engineering and Architecture, Eskişehir Osmangazi University, Eskişehir, Turkey

[2] Department of Electrical Electronics Engineering, Faculty of Engineering and Architecture, Eskişehir Osmangazi University, Eskişehir, Turkey

ABSTRACT

Adaptive image denoising algorithms rely on an error function that measure the distance between an estimated result and expectations. Selection of the error function and its parameters are crucial for a successful denoising implementation. In this paper, a method for determining close-to-optimal parameters for a bell-shaped error function is evaluated. The function with calculated parameters is employed within a gradient optimization algorithm and tested using test images with varying noise types and levels. The restoration results of the denoising test runs that use the proposed parameters are compared against the results of algorithms that employ well-known least squares and sum of absolute differences methods along with a method that combines both. The clear superiority of the bell-shaped error function for the proposed parameters is shown by the test results.

Keywords: Adaptive iterative restoration, Bell-shaped error function, Data fidelity, Geometric tight framelet, Image denoising

1. INTRODUCTION

Digital images unsurprisingly contain noise introduced during almost all stages of handling it, including acquisition, transmission and compression. Denoising (reduction of noise), consequently is a fundamental and widely studied problem in all fields of digital image processing. Preservation of original information and not introducing additional artifacts while reducing noise component is one of the main challenges in image denoising. A wide variety of algorithms have been proposed over the past few decades, including various filtering-based spatial methods [28-29], transform domain methods [25-27], wavelet thresholding-based approaches[30] and total-variation (TV)-based approaches [31] . State-of-the-art is represented by BM3D [32], centralized sparse representation (CSR) [33] and learned simultaneous sparse coding (LSSC) [34]. In all those methods, selection of error measure is of critical importance. Equally, results of the spatial methods also affected by the choice. L1 and L2, described in the following paragraphs, are two of the well-known measures used in image restoration problems. In this paper, we proposed a method for calculating parameters of the influence function which is used to switch between these two measures. We employed this adaptive schema to test images and obtained better results than the methods that use these individual measures and another method that use both.

This section introduces the problem formulation through widely used models and terms. The following section describes the influence functions and motive for the proposed approach. The remaining sections present the experimental setup and give the results and comparisons, concluding at the end.
Many signal/image restoration problems are modeled as inverse formulations:

$$Y = HX + \eta \tag{1}$$

---

*Corresponding Author: kozkan@ogu.edu.tr

where X is the desired solution, Y is known and some inaccurate knowledge on transformation H and noise η exist. In addition to the difficulties caused by noise, H is an ill-conditioned sparse matrix, and X is usually larger than Y in size, so a rank deficiency exists, making (1) impossible to solve for an exact X without some assumptions. For example, Y is a set of images arranged to have a $MNK \times 1$ matrix, where K is the number of $M \times N$ images and X is a $r^2 MN \times 1$ matrix representing the desired image pixels, where r ($r \geq 1$) is the ratio of upsizing. In the special case of $K = 1$ and $r = 1$, the problem can be viewed as single-image deblurring when H is a blur matrix and as denoising when H is a unity matrix. When $r > 1$, the problem becomes a super-resolution (SR) restoration, and when $K > 1$, it is called multiframe SR [1]. Usually, both noise and blur effects exist in observation images Y. When a direct solution is not possible, one resorts to optimization techniques to determine a solution, among infinite number of solutions, which is as consistent as possible with all available data and prior knowledge/assumption/expectation. Consistency is usually measured by an objective function whose inputs are observed data, estimated solution and/or data generated from the estimation. The indifference or correlation between observed data and data generated from the estimation may well be a consistency measure. It is very common to have an objective function to be minimized in the form of $\left\| \varepsilon \right\|_p := (\sum_{i=1}^{n} |\varepsilon_i|^p)^{1/p}$ for an N-dimensional difference vector

$$\varepsilon = Y - HX \tag{2}$$

When $p = 1$ or $p = 2$, this measure is called L_1 [2-5] or L_2 [6-8] norm, respectively, in image restoration. However, in both cases, the solution is very sensitive to noise in the observed data, forcing the researchers to limit the solution space using prior knowledge and expectations. Such applications of limitations are called regularizations. Total variation (TV) [9-10] is the most used regularization term, and it is the L_1 norm of the first-order derivative of error. Although TV aims to protect edges in images, smoothing in relatively edge-free regions locally limits flexibility. In the research described in this paper, we used L_1 norm of the framelet coefficients of the estimated high-resolution image [11-12]. Using the L_2 norm, (2) becomes

$$\varphi(Y, X) = \left\| Y - HX \right\|^2 = \varepsilon^T \varepsilon \tag{3}$$

To minimize (3), the 1st order derivative with respect to X is set to zero to obtain the least squares solution:

$$\Delta X = \frac{\partial \varphi(Y, X)}{\partial X} = (H^T H)^{-1} H^T Y = 0 \tag{4}$$

Nonlinear problems are usually solved iteratively through $X_{k+1} = X_k + \Delta X_k$ update formulation. In addition, a damping factor called the Lagrange multiplier is usually added to diagonal elements of $H^T H$ because singular values may prevent inversion. The equation

$$\Delta X = \left(H^T H + \lambda I \right)^{-1} H^T Y \tag{5}$$

called the Levenberg-Marquard solution of nonlinear problems [13-14] is obtained. However, the stability of (5) in ill-conditioned problems is questionable and may cause unrealistic solutions. To overcome this, the general functional

$$P_\lambda(Y, X) = \varphi(Y, X) + \lambda \phi(X) \tag{6}$$

where λ and $\phi(X)$ are called the regularization parameter and the stabilizing functional, respectively,

is proposed by Tikhonov and Arsenin [15], so that the solution is found by

$$X = \arg\min_X \{\varphi(Y, X) + \lambda\phi(Y, X)\} \tag{7}$$

minimization [14]. In practice, when the $\phi(X)$ functional is selected to be the energy function $\phi(X) = \|X\|^2$, as Tikhonov and Arsenin suggested, λ needs to be large to eliminate noise, causing overly smoothed solutions. Although the motive is different, the cost function is very similar, if not the same. Generally, the functional is defined as $\phi(X) = \sum_{i=1}^{r} \phi(|DX|)$, where $\{D_i : 1 \le i \le r\}$ is a group difference operator. Two of the popular designs for $\phi(X)$, among several others in the literature, are $\phi(X) = |X|^l$ [16] and $\phi(X) = \sum\sum \alpha^{|l|+|m|} \|X - S_{x_1}^l S_{x_2}^m X\|^l$ which is used in a generalization of total variation [2] and called bilateral-TV. Here $S_{x_1}^l$ and $S_{x_2}^m$ are horizontal and vertical shift operators and α is a scalar between 0 and 1 for spatial decay. Bilateral-TV considers total variation in horizontal and vertical directions. In addition, one can also use variations in the angled directions of $\pm\pi/4$ radians. In the experimental part of the proposed method, a regularization term involving the following 18 operators comprising these angled variations and their 1st and 2nd order derivatives, called geometric tight framelets [11-12] is used;

$$\lambda_0 = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad \lambda_1 = \frac{1}{16}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \lambda_2 = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad \lambda_3 = \frac{\sqrt{2}}{16}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$\lambda_4 = \frac{\sqrt{2}}{16}\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad \lambda_5 = \frac{\sqrt{7}}{24}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \lambda_6 = \frac{1}{48}\begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix} \quad \lambda_7 = \frac{1}{16}\begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\lambda_8 = \frac{1}{12}\begin{bmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \lambda_9 = \frac{1}{12}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \lambda_{10} = \frac{\sqrt{2}}{12}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \lambda_{11} = \frac{\sqrt{2}}{16}\begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\lambda_{12} = \frac{\sqrt{2}}{16}\begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{bmatrix} \quad \lambda_{13} = \frac{1}{48}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \quad \lambda_{14} = \frac{\sqrt{2}}{12}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$
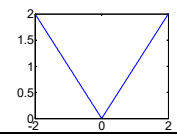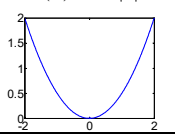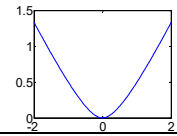
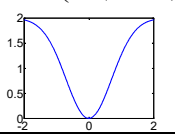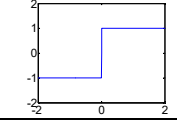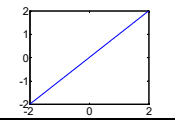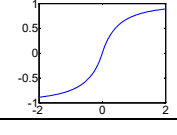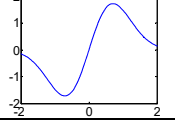Discussion of the properties of these operators is somewhat lengthy and found in [11, 12] for interested readers.

2. DATA FIDELITY TERM and THE PROPOSED METHOD

The least squares error measure is the most used measure for generating approximate solutions to linear systems. It seems that L_2 is a decent choice when the noise has a Gaussian distribution. The sum of the absolute differences (L_1) is preferred when it is suspected that the data contain outliers. The use of absolute values limits the sensitivity to outliers, whereas squaring is expected to increase errors in the neighborhood, which is the motive for using L_1. The behavioral difference between L_1 and L_2 is associated with their derivatives. The derivative of L_2, $\partial\varphi/\partial X = -2H(Y - HX)$, linearly depends on X

, making it sensitive to outliers. The derivative of L_1, $\partial\varphi/\partial X = H\,sign(Y - HX)$, in contrast, is bounded, even though it is undefined at zero. Therefore, it is reasonable to employ L_2 when noise is low and to use L_1 otherwise. Li et al. [11,12] suggested an even, strictly convex, twice differentiable function bounded by $\eta = \|Y - HX\|^1\big/M$ and adaptive selection for the norms. Such a function with these and additional desired features is proposed by Pham et al. [17]. It is infinitely differentiable, adaptive to noise variance and has an asymptotically zero derivative for higher error values. Although it is desired to not influence the local average and the data structure with outliers, instead of completely eliminating their influence as Li et al. [11-12] do, it is more logical to have their contribution lessen and diminish gradually until they become completely irrelevant. The derivative of the function proposed in [17] asymptotically approaches zero and reduces the contribution of outliers. The L_1, L_2, Li et al. [11-12] and bell-shaped [17] functions are given in Table 1, along with their first derivatives (influence functions).

Table 1. Examined error functions, their derivatives and parameters

	ABSOLUTE	LEAST-SQUARES	Y.R. LI	BELL-SHAPED
Function	$\varphi(x) = |x|$	$\varphi(x) = 1/2|x|^2$	$\varphi(x) = \eta x^2/(\eta + |x|)$	$\varphi(x) = \dfrac{a^2}{b}\left(1 - \dfrac{1}{\left(1 + \dfrac{x^2}{2a^2}\right)^b}\right)$
Derivative	$\varphi'(x) = sign(x), x \neq 0$	$\varphi'(x) = x$	$\varphi'(x) = \dfrac{x(2 + |x|/\eta)}{(1 + |x|/\eta)^2}$	$\varphi'(x) = \dfrac{x}{\left(\dfrac{x^2}{2a^2} + 1\right)^{b+1}}$

It should be noted that Li et al. [11-12] dismissed $Huber(x) = \begin{cases} \frac{1}{2}x^2, & |x| < \eta \\ \eta x - \frac{1}{2}\eta^2, & |x| \geq \eta \end{cases}$ [35] as an objective function on the basis that second order derivative is not continuous. Although the bell-shaped objective function is shown to be superior, Pham et al. [17] did not provide any method for the selection of its shape-defining parameters to obtain the best performance, and they applied default values. These parameters are quite important when switching between L_1 and L_2 type influences on data. Consequently, these parameters affect the success of the result. In this paper, we propose a method to determine the parameters of the bell-shaped function for optimal restoration. According to Hwang and Haddad [18] and Ko and Lee [19], the outlier border is the point beyond which sample points are considered outliers. The goal is to reduce the contribution of samples outside the outlier border gradually. For that, the maximum of the influence function (the derivative of the objective function) should be on the outlier border. The region that contains insiders for the bell-shaped objective function is marked as gray in Figure 1. Taking the derivative of the influence function (second derivative of the objective function) and equating it to zero as

$$\frac{\partial f(x,a,b)}{\partial x} = \frac{1}{\left(\frac{x^2}{2a^2}+1\right)^{b+1}} - \frac{(b+1)x^2}{a^2\left(\frac{x^2}{2a^2}+1\right)^{b+2}} = 0 \tag{8}$$

we find

$$x = \pm \frac{\sqrt{2}a}{\sqrt{2b+1}} \tag{9}$$

as extremes and also outlier borders. The parameter b determines how steep the function is. Because it
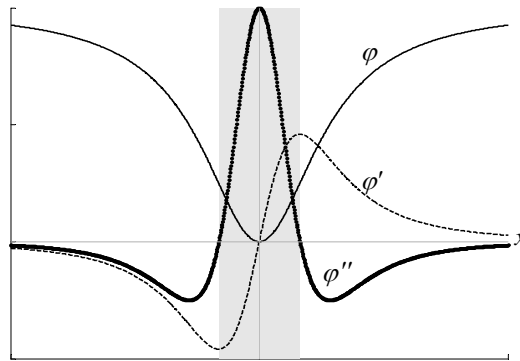

Figure 1. The bell-shaped error function and its first and second derivatives

is desired that the function behave like L_1 when the noise is high and like L_2 when the noise is low, the parameter b should assume values between 0.5 and 1.0, accordingly. Therefore, the empirical selector

$$b = \frac{\sigma_{image}^2 + \sigma_{noise}^2}{\sigma_{image}^2 + 2\sigma_{noise}^2} \tag{10}$$

can be used for parameter b. In most applications, noise statistics are not known, but some statistical parameters can be estimated. According to Donoho and Johnstone [20], the standard deviation of the noise can be calculated using $\sigma = median(D)/0.6745$ where D is the wavelet coefficient set in the HH band of the image. This is a good point to mention related multiscale and sparse representation research in image denoising. Wavelets can represent data in very sparse form and therefore can be used in denoising by thresholding [20-21]. Finally, replacing x in (9) with the border value estimated using the method described in [18] and [19] and inserting the value of b found using (10), we obtain

$$x = \pm \frac{\sqrt{2}a}{\sqrt{2b+1}} \Rightarrow a = \pm \frac{x\sqrt{2b+1}}{\sqrt{2}} \tag{11}$$

and calculate parameter a. The parameters a and b completely describe the function $\varphi'(x)$ to be used in minimizations. Among several approaches proposed for minimization given in (7), the steepest descent method is preferred over others for its simplicity. In that case, (7) becomes

$$X_{k+1} = X_k + \beta(-\nabla P_\lambda(Y,X)) = X_k - \beta\{\nabla(\varphi(Y,X) + \lambda\phi(X)\} \tag{12}$$

where β is a convergence parameter that is selected to reduce the error in every iteration. β can be found using

$$\beta_k = \arg\min_{\beta \in [0,1]} \varphi(Y, X_k - \beta \nabla P_\lambda(Y, X)) \tag{13}$$

minimization, so that it is bounded within (0,1) and decreases the data fidelity term adaptively [22]. The parameter λ,

$$\lambda = \|Y - HX_k\|^1 / \phi(X_k) \tag{14}$$

can be employed as described in [23].

3. EXPERIMENTAL WORKS

For a simple experiment, we generated a set of points representing noisy samples from a line in $\mathfrak{R}^2$ using $x_i = 10i/m, i = 1,...,m$ and $y_i = 3x_i + 2 + \varepsilon_i$, $N(0, \sigma^2)$ [24]. Denoising algorithm with proposed parameter calculation is applied upon the samples and compared against L_1, L_2 and the parameters/method used by Li et al [12]. MSE (mean squared errors) and the number of iterations performed for converging are given in Table 2. All MSE and iteration values are the averages of 100 experiments with different values of noise components. The results in table clearly show that the proposed parameter calculation is superior. Numbers of iterations do not extremely change between methods (somewhat decreases in the proposed method), showing that complexity of the methods are approximately same. This forms a basis for the image denoising algorithm with the proposed parameter calculation method.

Table 2. Test results on noisy samples of a line in $\mathbf{R}^2$

m	σ^2	Proposed		Li [12]		L_1		L_2	
		MSE	#iter.	MSE	#iter	MSE	#iter.	MSE	#iter.
10	1	0.09	41	0.1	68	0.16	96	0.09	93
	3	1.62	50	2.45	77	2.80	108	3.33	78
	5	3.64	80	3.95	85	4.62	120	5.59	87
100	1	1.07	66	1.17	76	1.31	82	1.28	79
	3	1.53	70	1.56	81	2.07	96	1.96	78
	5	1.94	81	2.01	86	3.00	103	3.27	84
500	1	1.04	78	1.25	79	1.36	91	1.26	81
	3	1.11	84	1.36	89	1.39	138	1.48	90
	5	1.20	86	1.42	96	1.44	142	1.77	98

The bell-shaped error measurement function with the proposed parameters is tested against L_1, L_2 and the images with varying levels of Gaussian and impulse noise. The test images are obtained using $Y_{ijk} = HX + \eta_i^\sigma + \eta_j^{rv} + \eta_k^{sp}$, where H represents a disk filter of 7 pixels in diameter for imposing blur onto clean images. Blurring is followed by adding various types and amounts of noise. Here, η_i^σ, η_k^{sp} and η_j^{rv} are noise images representing Gaussian noise, salt&pepper (sp) noise (or impulse noise) and random positioned/valued (rv) noise (rv is described as Gaussian noise added onto randomly selected image pixels) respectively, where indices i, j and k identify different amounts/characteristics such that test images with varying Peak-Signal-to-Noise-Ratios (PSNRs) are obtained.

For Gaussian noise, the standard deviations are selected to be 0 (no Gaussian noise), 3 and 5.

Salt&pepper noise is obtained by randomly assigning 0 or 255 (equal probability) to a randomly selected $\rho\%$ (where $\rho = \{0, 5, 10, \quad, 50\}$) of all pixels of images whose pixel values range from 0 to 255. Random-valued noise images, on the other hand, are generated in a manner similar to sp-type noise, but instead of assuming values of just 0 or 255, random values with uniform distribution between 0 and 255 are used as interference. Noise combinations of these types are applied on test images and four denoising methods are tested on these. It should be noted that for all methods that are compared, the same pre-generated noise image is added onto test images. Iterations are terminated with the stopping criteria $\|X_{k+1} - X_k\|_2 / \|X_k\|_2 < 10^{-8}$ which is determined empirically. After the restoration process, results are compared for all images whose original (not distorted) images are available. Selected results, images and graphs that best represent the rest are given in the following discussion, with a note that the results omitted here are characteristically similar to those presented. PSNR results for permutations of $\sigma = \{0, 3, 5\}$ (for Gaussian noise) and η_j^{rv} for $\{0, 25, 50\}$ percent (out of the tested $\{0, 5, 10, \quad, 50\}$ percent) are given in Table 3.

Table 3. Resulting PSNR values for blur plus $\eta_i^{\sigma} + \eta_j^{rv}$ noise cases. In every cell, the PSNR values for the input image, L_1, L_2, Li et al. [11-12] and proposed method results are listed in that order.

η_i^{σ}		$\sigma = 0$					$\sigma = 3$					$\sigma = 5$				
η_j^{rv}		gold hill	house	camer aman	boat	bar bara	gold hill	house	camer aman	boat	bar bara	gold hill	house	camer aman	boat	bar bara
0%	input	25.3	35.3	22.9	25.2	23.4	25.1	33.7	22.8	25.0	23.3	24.7	31.7	22.6	24.7	23.1
	L_1	27.6	44.9	24.8	27.3	24.3	26.4	38.5	24.4	26.9	24.1	25.7	36.5	23.9	25.9	23.9
	L_2	32.7	45.4	34.1	33.3	30.3	27.3	40.5	27.0	28.3	25.0	26.4	37.9	25.4	27.0	**24.4**
	Li	32.4	45.1	30.8	34.0	29.6	27.0	40.1	25.3	28.1	24.4	26.5	38.4	25.3	27.2	24.2
	prop.	**32.9**	**45.7**	**34.5**	**35.5**	**31.4**	**27.4**	**40.7**	**27.1**	**28.5**	**25.3**	**26.6**	**38.9**	**25.6**	**27.3**	24.3
25%	input	14.8	14.6	14.0	15.1	14.4	14.8	14.6	13.9	15.1	14.4	14.8	14.7	13.9	15.1	14.4
	L_1	27.1	43.3	24.1	27.1	24.3	26.1	37.2	23.8	26.4	23.9	25.3	35.0	23.2	25.4	23.7
	L_2	29.6	40.0	27.7	30.7	27.8	26.7	38.8	25.3	27.2	24.5	26.2	36.7	24.5	26.5	24.1
	Li	**31.8**	43.1	28.5	32.1	28.8	26.8	39.0	25.2	27.6	24.2	26.3	37.6	24.7	26.8	24.1
	prop.	**31.8**	**43.9**	**29.5**	**33.7**	**29.1**	**27.0**	**39.7**	**26.3**	**27.8**	**24.6**	**26.4**	**38.2**	**25.3**	**26.8**	**24.2**
50%	input	12.0	11.7	11.3	12.3	11.7	11.9	11.6	11.3	12.3	11.7	11.9	11.7	11.3	12.3	11.7
	L_1	26.3	32.6	24.0	26.7	24.0	25.0	33.9	22.6	24.9	23.6	24.6	32.0	22.3	24.2	23.3
	L_2	24.9	31.3	22.8	25.1	23.4	24.9	31.5	22.9	25.3	23.5	24.8	30.1	22.8	24.7	23.3
	Li	27.8	35.3	25.6	29.3	25.8	26.2	36.8	23.9	26.7	23.9	25.6	34.8	23.6	25.7	23.7
	prop.	**28.7**	**37.7**	**27.3**	**30.1**	**26.7**	**26.5**	**37.8**	**24.8**	**27.0**	**24.1**	**25.9**	**36.7**	**24.2**	**26.4**	**23.9**

Similarly, results for permutations of $\sigma = \{0, 3, 5\}$ and η_k^{sp} for $\{0, 40, 80\}$ (out of the tested $\{0, 5, 10, \quad, 80\}$ percent) are listed in Table 4. In Table 3 and 4, rows show input PSNR, L_1 output, L_2 output, Li et al. [11-12] method output and the proposed method output in that order. Proposed method clearly outperforms the others in $\eta_i^{\sigma} + \eta_j^{rv}$ cases as shown in Table 3. In $\eta_i^{\sigma} + \eta_k^{sp}$ cases, however, the proposed method generated the best or next to best results in PSNR measure.

Figure 2 shows PSNR_in-vs-PSNR_out results for the 'cameraman' test image set. Figures 2a and 2c are the cases with rv- and sp-type noise, respectively, whereas Figures 2b and 2d illustrate the restoration abilities of four methods on images with Gaussian noise of $\sigma = 5$ in addition to rv and sp noise with varying levels. Analysis on the results reveals that the proposed function is clearly superior to others, except with some low input-PSNR cases. In high input-PSNR cases, all methods exhibit some performance drops. L_1 is always outperformed by the other methods.

Figures 3 to 7 visually compare four methods and demonstrate the behavior of the proposed parameters.

Figures 3 and 4 exemplify the clear superiority of the bell-shaped objective function with calculated parameters on over L_1, L_2 and the method of Li et al. [11-12] on the reduction of the artifacts and spurious pixels. These are the cases with output images with similar PSNR. Although the output image PSNRs are close, L_1 , L_2 and method of Li clearly shows some artifacts, while the proposed method output generated minimum artifact. With $\eta^{\sigma} + \eta_k^{sp}$ (11.1dB), all methods tend to lose texture.

Table 4. Resulting PSNR values for blur plus $\eta_i^{\sigma} + \eta_k^{sp}$ cases. In every cell, the PSNR values for the input image, L_1, L_2, Li et al. [11-12] and proposed method results are listed in that order.

η_i^{σ}		σ =0					σ =3					σ =5				
η_k^{sp}		gold hill	house	camer aman	boat	bar bara	gold hill	house	camer aman	boat	bar bara	gold hill	house	camer aman	boat	bar bara
0%	input	25.3	35.3	22.9	25.2	23.4	25.1	33.7	22.8	25.0	23.3	24.7	31.7	22.61	24.66	23.08
	L_1	27.2	44.9	24.8	27.3	24.3	26.4	38.5	24.4	26.9	24.1	25.7	36.5	23.93	25.87	23.85
	L_2	32.7	45.4	34.1	33.3	30.3	27.3	40.5	27.0	28.3	25.0	26.4	37.9	25.38	27.02	**24.39**
	Li	32.4	45.1	30.8	34.0	29.6	27.0	40.1	25.6	28.1	24.4	26.5	38.4	25.30	27.15	24.20
	prop.	**32.9**	**45.7**	**34.5**	**35.5**	**31.4**	**27.4**	**40.7**	**27.1**	**28.5**	**25.3**	**26.6**	**38.9**	**25.59**	**27.33**	24.26
40%	input	11.2	11.2	10.8	11.4	11.8	11.2	11.2	10.8	11.4	11.1	11.2	11.2	10.86	11.40	11.06
	L_1	27.1	**43.7**	24.6	27.4	24.4	26.1	36.6	24.4	26.4	24.0	25.3	35.0	23.28	25.31	23.70
	L_2	31.9	42.9	33.1	34.4	29.8	27.0	**39.9**	26.4	28.1	24.6	26.4	**37.9**	25.52	26.98	**24.25**
	Li	31.8	42.5	32.9	34.2	29.2	26.8	39.3	25.7	27.6	24.3	26.4	37.6	25.09	26.97	24.13
	prop.	**32.0**	43.0	**33.2**	34.4	**29.9**	**27.1**	39.8	26.2	**28.1**	**24.7**	**26.4**	37.9	25.47	**27.16**	24.23
80%	input	08.3	08.2	08.0	08.5	08.2	08.3	08.2	08.0	08.5	08.2	08.3	08.2	08.0	08.5	08.2
	L_1	26.7	**41.8**	24.7	27.1	24.3	25.7	35.6	23.1	25.7	23.7	25.0	33.7	22.8	24.9	23.6
	L_2	**30.6**	40.1	**30.9**	33.1	28.2	26.7	38.9	25.7	27.6	24.3	26.2	36.9	25.0	**26.8**	24.1
	Li	30.5	38.83	30.7	33.0	28.0	26.5	37.7	25.0	27.1	24.1	**26.2**	**37.2**	24.6	26.6	24.0
	prop.	30.5	39.9	30.5	32.9	27.7	26.6	38.1	25.2	27.2	24.3	26.1	37.0	24.6	26.6	24.0
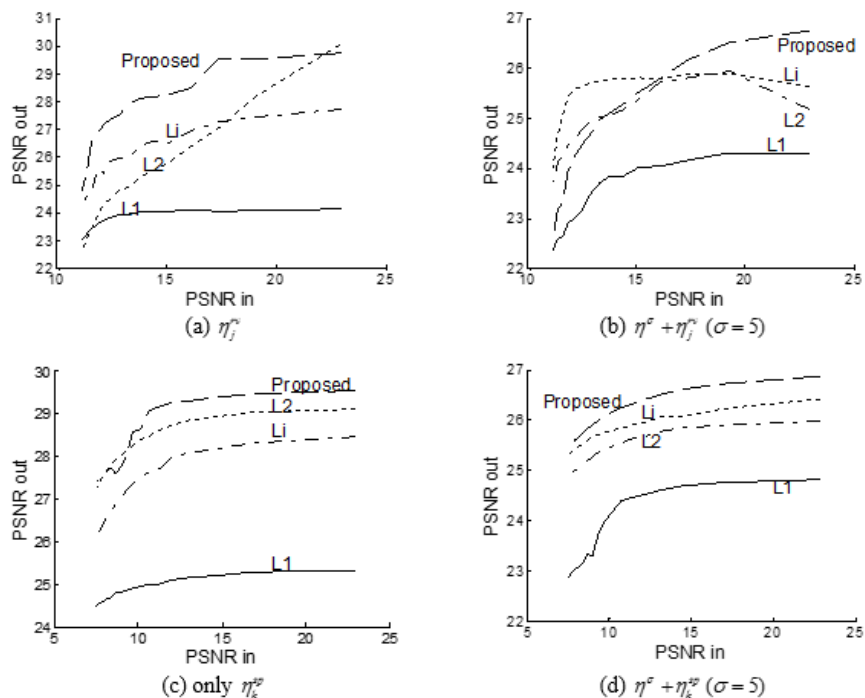
Figure 2. PSNR_in-vs-PSNR_out results for the 'cameraman' test set with the given noise types

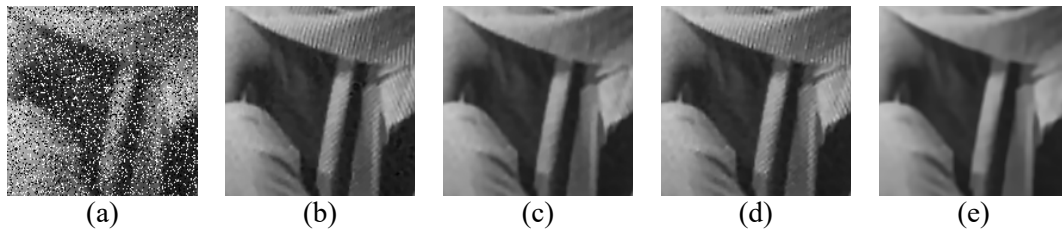| (a) | (b) | (c) | (d) | (e) |

Figure 3. A section of the 'barbara' test image with a) $\eta^\sigma + \eta_k^{sp}$ (11.10dB), b) restored with L_1 (24.00dB), c) L_2 (24.62dB), d) Li et al. [11-12] (24.28dB) and e) bell-shaped φ with optimized parameters (26.65 dB).
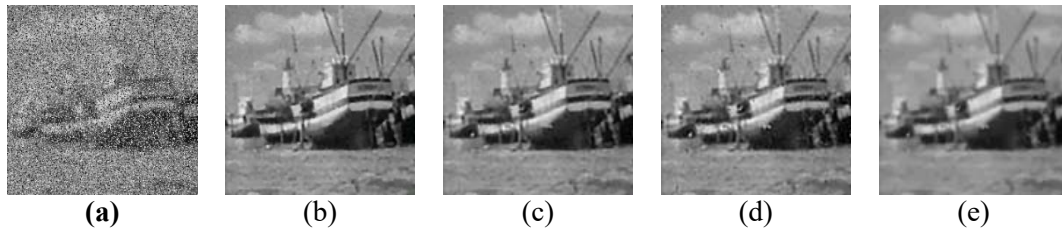

| (a) | (b) | (c) | (d) | (e) |

Figure 4. a) Boat image with $\eta^\sigma + \eta_j^{rv}$ noise where σ =5 and s=50 (50% of all pixels are randomized with uniform distribution). Restored by the b) L_1, c) L_2 and d) Li et al. [11-12] methods and e) the bell-shaped function with calculated parameters. Visible artifacts are scattered around the images in the results from the L_1, L_2 and Li et al methods.
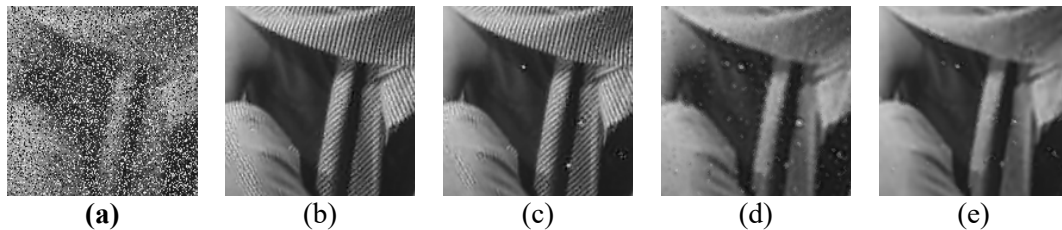

| (a) | (b) | (c) | (d) | (e) |

Figure 5. A section of the 'barbara' test image with a) $\eta^\sigma + \eta_j^{rv}$ (11.71dB), b) restored with L_1 (24.06dB), c) L_2 (24.43dB) d) Li et al. [11-12] (25.82dB) and e) bell-shaped φ with optimized parameters (26.68 dB). Although L_1 and L_2 were able to retain texture, artifacts introduced by L_1, L_2 and Li et al. [11-12] are clearly visible.
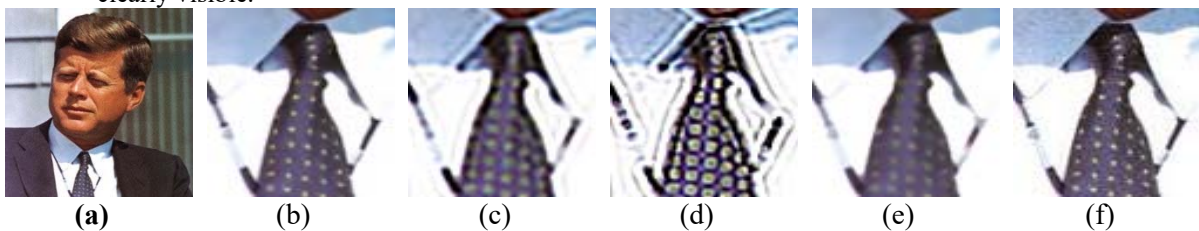

| (a) | (b) | (c) | (d) | (e) | (f) |

Figure 6. a) Original 'kennedy' image and b) tie close-up. Images restored by the c) L_1, d) L_2, e) Li et al. [11-12] methods and f) the bell-shaped φ with optimized parameters.

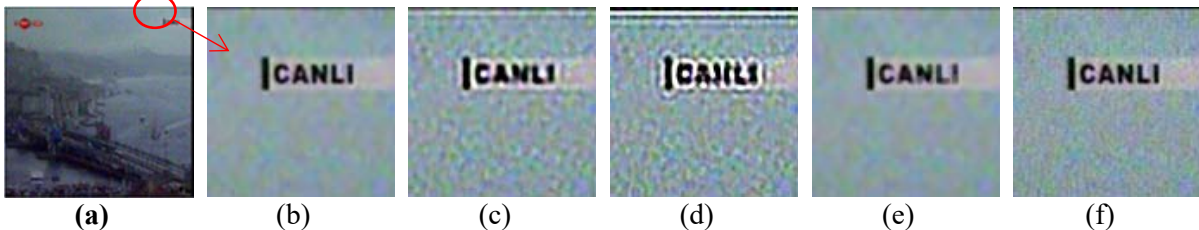| (a) | (b) | (c) | (d) | (e) | (f) |

Figure 7. a) Original 'goldenhorn' image and b) a close-up of the corner text. Images restored by the c) L_1, d) L_2 e) Li et al. [11-12] methods and f) the bell-shaped φ with optimized parameters.

Figures 6 and 7 present examples in which clean originals are not available so that a visual comparison

needs to be made. The bell-shaped objective function φ with the proposed parameters is superior to the other objective functions that were examined, in PSNR measure and reduced artifacts. This is due to its balancing nature between the L_1 and L_2 norms, consequently allowing decent deblurring/sharpening while doing quite a good job at removing noise of different types.

4. CONCLUSION

Both L_1 and L_2 error measures have their own virtues in image denoising algorithms. It is reasonable to employ L_2 when noise is low and to use L_1 otherwise. Selection of the error function and its parameters are crucial for a successful denoising implementation. We employed a bell-shaped weighted switching function whose parameters are calculated from the image statistics and obtained successful results. Although the bell-shaped function has been used by Pham and Schutte[16], we additionally provided a mechanism to calculate shape-defining parameters of the function where they used default values. The overall performance is also affected by the definition of local outliers which needs and aimed to be researched in further study.

REFERENCES

[1] Banham MR, Katsaggelos AK. Digital image restoration. IEEE Signal Proc Mag 1997; 14 (2): 24-41.

[2] Farsiu S, Robinson D, Elad M, Milanfar P. Advances and challenges in Super-Resolution. Int J Imag Syst Tech 2004; 14 (2): 47-57.

[3] Shi M, Han T, Liu S. Total variation image restoration using hyper-Laplacian prior with overlapping group sparsity. Signal Proc 2016; 126: 65-76.

[4] He H, Kondi LP. An image super-resolution algorithm for different error levels per frame. IEEE T Imag Process 2006; 15 (3): 592-603.

[5] Shen X, Wang K, Guo Q. Local thresholding with adaptive window shrinkage in the contourlet domain for image denoising. Sci China Inform Sci 2013; 56 (9): 1-9.

[6] Farsiu S, Elad M, Milanfar P. Multiframe demosaicing and super-resolution of color images. IEEE T Image Process 2006; 15 (1): 141-159.

[7] Wang S, Huang T.Z, Liu J, Guang X. An alternating iterative algorithm for image deblurring and denoising problems. Commun Nonlinear Sci Numer Simulat 2014; 19: 617–626.

[8] Zibetti MVW, Mayer J. A robust and computationally efficient simultaneous super-resolution scheme for image sequences. IEEE T Circ Syst Vid 2007; 17(10): 1288-1300.

[9] Rudin LI, Osher S, Fatemi E. Nonlinear Total Variation Based Noise Removal Algorithms. Physica D: Nonlin Phen. 1992; 60(1-4): 259-268.

[10] Li HJ, Zhao ZM, Zhu Y, Yu XL. Novel model for iteration step selection in image denoising using total variation technique. Imag Sci J 2010; 58(4): 222 – 230.

[11] Li YR, Dai DQ, Shen LX. Multiframe Super-Resolution Reconstruction Using Sparse Directional Regularization. IEEE T Circ Syst Vid 2010; 20(7): 945-956.

[12] Li YR, Shen LX, Dai DQ, Suter BW. Framelet Algorithms for De-Blurring Images Corrupted by Impulse Plus Gaussian Noise IEEE T Imag Process 2011; 20(7): 1822-1837.

[13] Levenberg K. A method for the solution of certain non-linear problems in least squares. Quat Appl Math 1944; 2(2): 164-168.

[14] Marquardt DW. An algorithm for least squares estimation of non-linear parameters. J Soc Ind Appl Math 1963; 11: 431-441.

[15] Tikhonov AN, Arsenin VY. Solutions of Ill-Posed Problems. Washington DC: Winston, 1977.

[16] Osher S, Rudin LI. Feature-Oriented Image-Enhancement Using Shock Filters. Siam J Numer Anal 1990; 27(4): 919-940.

[17] Pham TQ, Vliet LJv, Schutte K. Robust super-resolution by minimizing a Gaussian-weighted L2 error norm, J Phy: Conf Ser 2008; 124(1): 1-20.

[18] Hwang H, Haddad RA. Adaptive Median Filters - New Algorithms and Results. IEEE T Image Process 1995; 4(4): 499-502.

[19] Ko SJ, Lee YH. Adaptive center weighted median filter. IEEE Trans Circ Sys II. 1998; 38: 984-993.

[20] Donoho DL, Johnstone IM. Ideal Spatial Adaptation by Wavelet Shrinkage. Biometrika 1994; 81(3): 425-455.

[21] Khare A, Tiwary US, Pedrycz W, Moongu J. Multilevel adaptive thresholding and shrinkage technique for denoising using Daubechies complex wavelet transform. Imag Sci J 2010; 58(6): 340 – 358.

[22] Kim Y, Vese LA. Image Recovery using Functions of Bounded Variation and Sobolev Spaces of Negative Differentiability. Inv Probl Imag 2009; 3(1): 43-68.

[23] Starck JL, Elad M, Donoho DL. Image decomposition via the combination of sparse representations and a variational approach. IEEE T Image Process 2005; 14(10): 1570-1582.

[24] Sabo K, Scitovski R. The best least absolute deviations line – properties and two efficient methods for its derivation. ANZIAM J 2008; 50: 185–198.

[25] Fan F, Yong M., Chang L, Xiaoguang M, Jun H, Jiayi M, Hyperspectral image denoising with superpixel segmentation and low-rank representation, Information Sciences 397–398 (2017) 48–68.

[26] Tom F, Rowan KL, Paul AM. Denoising time-resolved microscopy image sequences with singular value thresholding, Ultramicroscopy 178 (2017) 112–124.

[27] Oh-Young L, Jae-Won L, Jong-Ok K. Combining Self-Learning Based Super-Resolution with Denoising for Noisy Images Journal of Visual Communication and Image Representation, in press, doi: doi.org/10.1016/j.jvcir.2017.05.010.

[28] Antonio GL, Pedro CM. Denoising portal images by minimizing the SURE estimator on a parameterized family of shrinkage functions Physica Medica 38 (2017) 59–65.

[29] Muthukumaran M, Gopalakrishnan S, Purna CR B, Soumitra K, Thangavel S, Arunmuthu K. Anisotropic diffusion based denoising on X-radiography images to detect weld defects Digital Signal Processing ,in press, doi.org/10.1016/j.dsp.2017.05.014.

[30] Jain P, Tyagi V. LAPB: Locally adaptive patch-based wavelet domain edgepreserving image denoising, Inf. Sci. 294 (February) (2015) 164–181.

[31] Rodriguez P, Wohlberg B. Efficient minimization method for a generalized total variation functional, IEEE Trans. Image Process. 18 (2) (2009) 322–332.

[32] Zhong H, Ma K, Zhou Y. Modified BM3D algorithm for image denoising using nonlocal centralization prior, Signal Process. 106 (January) (2015) 342–347.

[33] Dong W, Li X., Zhang D, Shi G. Sparsity-based image denoising via dictionary learning and structural clustering, in: IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, 2011, pp. 457–464.

[34] Mairal J, Bach F, Ponce J, Sapiro G, A. Zisserman. Non-local sparse models for image restoration, in: Proceedings of the IEEE International Conference on Computer Vision, Kyoto, Japan, 2009, pp. 2272–2279.

[35] Huber, PJ. Robust regression: Asymptotics, conjectures, and Monte Carlo, Ann. Statist.,1, 1973, 799-821.