CENTER PIXEL WEIGHT ESTIMATION FOR NON-LOCAL MEANS FILTERING USING LOCAL JAMES-STEIN ESTIMATOR WITH BOUNDED SELF-WEIGHTS

Minh Phuong Nguyen and Se Young Chun

Ulsan National Institute of Science and Technology (UNIST), Republic of Korea

ABSTRACT

Assigning appropriate center pixel weights (CPW) in non-local means (NLM) filter is an important issue to affect the quality of filtered images. Using local James-Stein (LJS) type CPW yielded superior peak signal-to-noise ratio (PSNR) over using other existing methods of determining the contribution of center pixels in NLM. However, the original LJS CPW method assumed no upper bound for self-weights implicitly, it may yield excessively large self-weights. This issue may be addressed by setting the upper bound for the JS estimator, but there is no theoretical justification that this bounded JS estimator is still dominating locally. We propose a novel method, called bounded self-weights LJS (BLJS), to incorporate bounded self-weights into LJS such that this new estimator is dominating locally. Our proposed method was evaluated using a patient MR image with 3 levels of additive Gaussian noise. The proposed BLJS yielded lower variances than the original LJS for almost all bias levels. BLJS also achieved PSNR comparable to or better than LJS. Visual image quality assessment showed that BLJS produced less visual artifacts than LJS.

Index Terms— James-Stein estimator, non-local means, center pixel weight, image filtering

1. INTRODUCTION

Many image filtering methods are based on weighted average of noisy pixels to reduce noise. For example, conventional local Gaussian filtering assumes spatially smooth image and performs weighted averaging of pixels in local neighborhood such that more weights are assigned to closer pixels. However, for images with sharp edges and fine details, this assumption is violated and filtering noisy image results in blurred image. Non-local means (NLM) filter [1] uses the similarity between two local patches of noisy image to determine weights. For pixels with similar neighborhood patches, higher weights are assigned so that edges and details can be better-preserved after weighted averaging. This method also allows weighted averaging to use more pixels in non-local areas so that more powerful denoising can be possible.

In NLM filtering, assigning appropriate self-weights or center pixel weights (CPW, which is the weight between the same patches) affects filtered image quality significantly. The original NLM filter by Buades *et al.* set self-weights as one or as the maximum weight value in the neighborhood so that at least one or two weights are the same in weighted averaging, respectively [1]. Brox *et al.* proposed a method to have at least n number of same weights in weighted averaging [2]. Zimmer *et al.* considered self-weight as a free parameter to determine [3]. Salmon *et al.* developed a Stein's unbiased risk estimation (SURE) based method to address rare patch effect issue [4].

Recently, Wu *et al.* proposed a method to determine CPW using the James-Stein (JS) type estimator [5]. When their method applied to images locally (called local JS estimator or LJS), it outperformed other previous methods in terms of the peak signal-to-noise ratio (PSNR). However, this state-of-the-art LJS method has a few drawbacks. First of all, as smaller local areas are used for applying LJS, PSNR significantly decreases. Secondly, LJS can potentially yield self-weight much larger than one, which may lead to severe rare patch effect due to no assumption for the upper-bound on self-weights. Lastly, even though [5] suggested to use a bounded version of JS estimator to determine CPW, there is no theoretical justification for the issue that a modified bounded JS estimator is still dominating locally.

In this paper, we show that using smaller local areas for LJS is beneficial for better bias-variance trade-off. Then, we propose a novel method, called bounded self-weights LJS (BLJS), to incorporate self-weights with upper and lower bounds into LJS implicitly such that our new estimator is still dominating locally based on the positive part of JS estimator [6]. Our proposed method is evaluated using a real patient MRI image with various levels of additive Gaussian noise. Image quality was evaluated with PSNR, bias-variance from multiple realizations, and visual assessment and is compared with LJS CPW.

Section 2 reviews conventional NLM filtering and original LJS CPW method. Section 3 proposes a new method on LJS CPW method with bounded self-weights for NLM, which is called BLJS method. Section 4 describes filtering perfor-

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1007928).

mance results using MRI image.

2. PREVIOUS WORKS

2.1. Classical Non-Local Means Filtering

Let's assume that the image is contaminated by noise with the model:

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{n}$ is $i.i.d.$ zero-mean additive Gaussian noise. The NLM filtering at the pixel x_i is the weighted average of all pixels y_j within search region Ω_i:

$$\hat{x}_i = \frac{\sum_{j\in\Omega_i} w_{i,j} y_j}{\sum_{j\in\Omega_i} w_{i,j}}. \tag{2}$$

where $\hat{x}_i$ is the ith element of $\hat{\mathbf{x}}$. Ω_i can be constrained to a small area or the entire image. The similarity weight is defined as:

$$w_{i,j} = \exp\left(\frac{-\|\mathbf{P}_i\mathbf{y} - \mathbf{P}_j\mathbf{y}\|^2}{2\,|\mathbf{P}|\,h^2}\right) \tag{3}$$

where $\mathbf{P}_i$ is an operator to extract a square patch centered at the ith pixel, $\|\cdot\|$ is Euclidean norm, $|\mathbf{P}|$ is the number of pixels within the patch, and h is a global smoothing parameter. This definition implies that the center pixel weights $w_{i,i}$ are always equal to 1 since the two patches are identical. However, several CPW methods [1, 2, 3, 5] exist to assign special values to $w_{i,i}$ to improve image quality further.

2.2. Local James-Stein Center Pixel Weight Method

Wu *et al.* [5] proposed LJS CPW method using the JS type estimator to determine $w_{i,i}$. LJS CPW method first decomposes the NLM filter (2) into two terms:

$$\hat{x}_i = \frac{W_i}{W_i + w_{i,i}}\hat{z}_i + \frac{w_{i,i}}{W_i + w_{i,i}}y_i \tag{4}$$

where

$$W_i = \sum\nolimits_{j\in\Omega_i\setminus\{i\}} w_{i,j}$$

and

$$\hat{z}_i = \frac{\sum_{j\in\Omega_i\setminus\{i\}} w_{i,j} y_j}{W_i}. \tag{5}$$

Note that $\hat{z}_i$ does not contain $w_{i,i}$. Instead of determining $w_{i,i}$, the authors of [5] re-parametrized (4) using

$$p_i = \frac{w_{i,i}}{W_i + w_{i,i}} \tag{6}$$

so that (4) becomes

$$\hat{x}_i = (1 - p_i)\,\hat{z}_i + p_i y_i. \tag{7}$$

In [5], the parameter p_i was estimated using the JS estimator that shrinks the noisy observation y_i toward the pre-filtered value $\hat{z}_i$ [7] over a small neighborhood of the ith pixel

$$p_i^{\mathrm{LJS}} = 1 - \frac{(|\mathbf{B}| - 2)\,\sigma^2}{\|\mathbf{B}_i\mathbf{y} - \mathbf{B}_i\hat{\mathbf{z}}\|^2} \tag{8}$$

where $\mathbf{B}_i$ is an operator to extract a square neighborhood centered at the ith pixel and $|\mathbf{B}|$ is the number of pixels within the neighborhood. Based on the results of James and Stein [7], if $|\mathbf{B}| \geq 3$, then for a small neighborhood extracted by $\mathbf{B}_i$,

$$\mathbf{B}_i\hat{\mathbf{x}} = \left(1 - p_i^{\mathrm{LJS}}\right)\mathbf{B}_i\hat{\mathbf{z}} + p_i^{\mathrm{LJS}}\mathbf{B}_i\mathbf{y} \tag{9}$$

is a dominating estimator "locally". [5] proposed to use this result for denoising the center pixel of the neighborhood only:

$$\hat{x}_i^{\mathrm{LJS}} = \left(1 - p_i^{\mathrm{LJS}}\right)\hat{z}_i + p_i^{\mathrm{LJS}} y_i. \tag{10}$$

[5] used a small neighborhood with the size of 15×15 pixels. (8) suggests that $p_i^{\mathrm{LJS}} \in (-\infty, 1]$. However, since (6) implies that $p_i^{\mathrm{LJS}} \geq 0$, the previous work [5] suggested to use the positive part of p_i^{LJS} so that

$$\hat{x}_i^{\mathrm{LJS}_+} = \left(1 - p_i^{\mathrm{LJS}_+}\right)\hat{z}_i + p_i^{\mathrm{LJS}_+} y_i \tag{11}$$

where $p_i^{\mathrm{LJS}_+} = \max(p_i^{\mathrm{LJS}}, 0)$ without any justification. Actually, this estimator is still a (locally) dominating estimator according to the work of Baranchik [6] and this estimator even dominates the original JS estimator (10) locally. Wu *et al.* [5] also suggested to use user-defined upper bound for $p_i^{\mathrm{LJS}_+}$, but this variant was not studied. To the author's knowledge, there is no theoretical support that such a bounded estimator is dominating.

3. PROPOSED WORKS

3.1. Bounded Self-Weights

According to Baranchik [6], $p_i^{\mathrm{LJS}_+} = \max(p_i^{\mathrm{LJS}}, 0)$ results in a dominating JS estimator and in this case, $p_i^{\mathrm{LJS}_+} \in [0, 1]$. However, (6) suggests that if $p_i = 1$ and $W_i > 0$, then $w_{i,i} \gg 1$. This implies that the filtered image may have severe rare patch artifacts.

Usual choice for the upper bound of $w_{i,i}$ is one [1] or some values that are less than one [1, 2, 4]. If we bound self-weights $w_{i,i}$ to be $0 \leq w_{i,i} \leq w_{i,i}^{\max}$, then the range of p_i is

$$0 \leq p_i \leq \frac{w_{i,i}^{\max}}{W_i + w_{i,i}^{\max}} = p_i^{\max}$$

where the upper bound $p_i^{\max}$ is less than one if $W_i > 0$. One heuristic method is to apply this finite range for p_i directly to (8) using $\min(\max(0, p_i), p_i^{\max})$, which may be sub-optimal. In the next section, we present a new method to incorporate bounded p_i into the the positive part of the JS estimator.

Preprint submitted to 2016 IEEE International Symposium on Biomedical Imaging.
Received October 21, 2015.

3.2. James-Stein Estimator with Bounded Self-Weights

Let us decompose (4) in a different way from [5]:

$$\begin{aligned}\hat{x}_i &= \hat{z}_i - \frac{w_{i,i}}{W_i + w_{i,i}}\hat{z}_i + \frac{w_{i,i}}{W_i + w_{i,i}}y_i \\ &= \hat{z}_i - \frac{w_{i,i}}{W_i + w_{i,i}}\frac{p_i^{\max}}{p_i^{\max}}\hat{z}_i + \frac{w_{i,i}}{W_i + w_{i,i}}\frac{p_i^{\max}}{p_i^{\max}}y_i \\ &= \hat{z}_i - p_i^{\mathrm{B}}\hat{z}_i^{\mathrm{B}} + p_i^{\mathrm{B}}y_i^{\mathrm{B}} \\ &= \hat{z}_i - \hat{z}_i^{\mathrm{B}} + (1 - p_i^{\mathrm{B}})\hat{z}_i^{\mathrm{B}} + p_i^{\mathrm{B}}y_i^{\mathrm{B}}\end{aligned} \quad (12)$$

where $\hat{z}_i^{\mathrm{B}} = p_i^{\max}\hat{z}_i$, $y_i^{\mathrm{B}} = p_i^{\max}y_i$, and

$$p_i^{\mathrm{B}} = \frac{w_{i,i}}{W_i + w_{i,i}}\frac{1}{p_i^{\max}}. \quad (13)$$

Note that $p_i^{\mathrm{B}} \in [0, 1]$ is an increasing function of $w_{i,i}$ for $0 \le w_{i,i} \le w_{i,i}^{\max}$ and $\hat{z}_i - \hat{z}_i^{\mathrm{B}}$ is a constant term for p_i^{B}. Therefore, if one applies the JS shrinkage to estimate

$$(1 - p_i^{\mathrm{B}})\hat{z}_i^{\mathrm{B}} + p_i^{\mathrm{B}}y_i^{\mathrm{B}},$$

then, the following local positive part JS estimator is obtained:

$$\hat{x}_i^{\mathrm{BLJS}_+} = \hat{z}_i + p_i^{\mathrm{BLJS}_+}\left(y_i^{\mathrm{B}} - \hat{z}_i^{\mathrm{B}}\right) \quad (14)$$

where

$$p_i^{\mathrm{BLJS}_+} = \max\left(1 - \frac{(|\mathbf{B}| - 2)\,\sigma^2}{\|\mathbf{B}_i\mathbf{y}^{\mathrm{B}} - \mathbf{B}_i\hat{\mathbf{z}}^{\mathrm{B}}\|^2}, 0\right). \quad (15)$$

(15) implies that $0 \le p_i^{\mathrm{B}} \le 1$ and as shown in (9), this estimator is dominating "locally" [6].

4. RESULTS

4.1. Setup

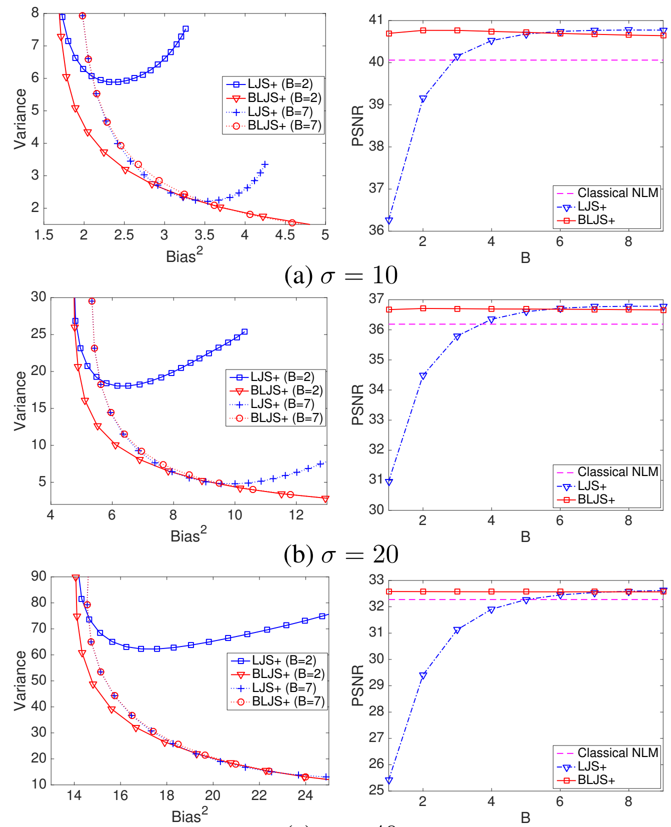
A patient MR image (512×512 pixels, 8 bits) was used as the true image, which was acquired and processed under institutional review board (IRB) approved protocols. A white Gaussian noise was added to this image with various standard deviations $\sigma \in \{10, 20, 40\}$. All algorithms were implemented using MATLAB R2015b (The Mathworks, Inc., Natick, MA, United States). Patch size and search window size were chosen to be 7×7 and 31×31, respectively. Both conventional LJS_+ and proposed BLJS_+ algorithms were tested for $\mathrm{B} = 1, \cdots, 9$ where $|\mathbf{B}| = (2\mathrm{B} + 1)^2 > 3$. For BLJS_+, we set $w_{i,i}^{\max} = 1$. A smoothing parameter h was chosen empirically to yield the best PSNR:

$$\mathrm{PSNR}\,(\hat{\mathbf{x}}) = 10\log_{10}\frac{255^2}{\|\hat{\mathbf{x}} - \mathbf{x}\|^2/N} \quad (16)$$

where N is an image size. In addition to PSNR, we also used mean bias vs. mean variance curves for different smoothing parameter values h from 30 noisy realizations as performance measure. Visual quality assessment was also performed.

4.2. Results

Fig. 1 [LEFT] shows mean bias vs. mean variance curves for varying h for different noise levels. These results show the benefit of using smaller neighborhood size B for both original LJS_+ and proposed BLJS_+. Original method LJS_+ yielded lower variance for higher bias when using large B as shown in square ($\mathrm{B} = 2$) and cross curves ($\mathrm{B} = 7$). However, for lower bias, LJS_+ achieved lower variance when using small B. It seems that using B is beneficial for LJS_+ to achieve lower bias with lower variance. Our proposed method BLJS_+ with small B yielded lower variance for all bias levels than BLJS_+ with large B and original LJS_+. The same tendency was observed at all 3 noise levels.



(a) $\sigma = 10$

(b) $\sigma = 20$

(c) $\sigma = 40$

Fig. 1. [LEFT] Bias-variance curves: each curve is obtained by varying h. [RIGHT] PSNR: for each block size, the maximum PSNR with respect to h is plotted on the vertical axis.

Fig. 1 [RIGHT] shows the PSNR performance of each method with varying B. For large B, LJS_+ and BLJS_+ achieve comparable PSNR. When B is small, LJS_+ yielded significantly lower PSNR, which is much lower than the classical NLM. However, our proposed BLJS_+ still achieved consistent PSNR for varying B. The above observed trends are similar for different noise levels even though relatively more distinction is noticed at less noise. Table 1 summarizes

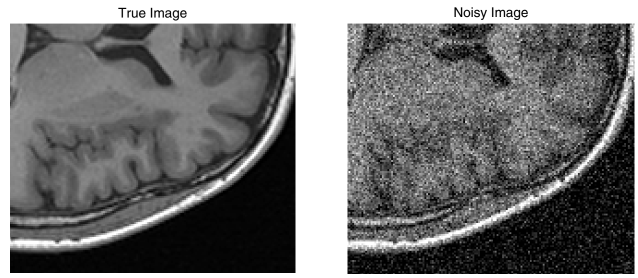
Preprint submitted to 2016 IEEE International Symposium on Biomedical Imaging.
Received October 21, 2015.

these trends for $\mathrm{B} = 2, 7$ quantitatively.

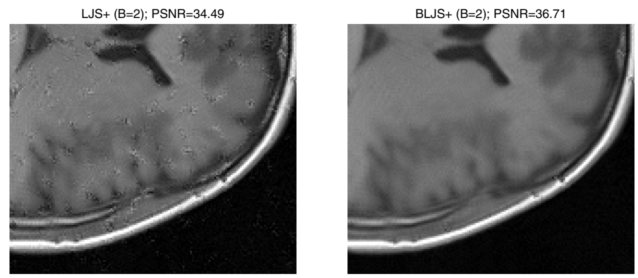
Table 1. PSNR summary table.

		B = 2		B = 7	
	Classical NLM	LJS+	BLJS+	LJS+	BLJS+
$\sigma = 10$	40.06	39.15	40.77	40.77	40.67
$\sigma = 20$	36.19	34.49	36.71	36.77	36.68
$\sigma = 40$	32.27	29.41	32.58	32.54	32.57

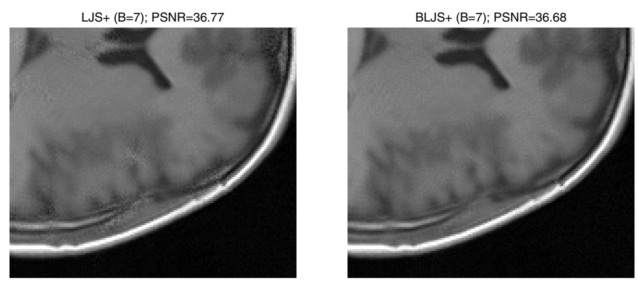
Fig. 2 (a) shows true and noisy MR images. When using small B for lower variance at small bias, the original LJS_+ yielded severe artifacts as shown in Fig. 2 (b), left. However, our proposed method BLJS_+ suppressed these artifacts significantly as in Fig. 2 (b), right. When using large B, both LJS_+ and proposed BLJS_+ achieved comparable PSNR as shown in Table 1. However, Fig. 2 (c) shows that LJS_+ still yielded slight artifacts near edges and detail areas. In contrast, proposed BLJS_+ yielded image with less artifacts.



(a) True and Noisy Images



(b) LJS_+ and BLJS_+ with $\mathrm{B} = 2$



(c) LJS_+ and BLJS_+ with $\mathrm{B} = 7$

Fig. 2. True, Noisy, and Filtered images using LJS_+ and BLJS_+ with the neighborhood sizes of $\mathrm{B} = 2, 7$ at $\sigma = 20$.

5. CONCLUSION

We proposed a new method, BLJS_+, to estimate CPW for NLM using local positive-part JS type estimator with bounded self-weights. Our proposed method with small neighborhood size yielded better variance for all levels of bias than original LJS_+, consistently good PSNR results for varying neighborhood size, and visually pleasing images with less artifacts than the original LJS CPW method.

6. REFERENCES

[1] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[2] Thomas Brox and Daniel Cremers, "Iterated nonlocal means for texture restoration," in *Scale Space and Variational Methods in Computer Vision*, pp. 13–24. Springer, 2007.

[3] Sebastian Zimmer, Stephan Didas, and Joachim Weickert, "A rotationally invariant block matching strategy improving image denoising with non-local means," in *Proc. 2008 International Workshop on Local and Non-Local Approximation in Image Processing*, 2008, pp. 135–142.

[4] Joseph Salmon, "On two parameters for denoising with non-local means," *Signal Processing Letters, IEEE*, vol. 17, no. 3, pp. 269–272, 2010.

[5] Yue Wu, Brian Tracey, Prem Natarajan, and Joseph P Noonan, "James–stein type center pixel weights for non-local means image denoising," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 411–414, 2013.

[6] Alvin J Baranchik et al., "A family of minimax estimators of the mean of a multivariate normal distribution," *The Annals of Mathematical Statistics*, vol. 41, no. 2, pp. 642–645, 1970.

[7] William James and Charles Stein, "Estimation with quadratic loss," in *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, 1961, vol. 1, pp. 361–379.

Preprint submitted to 2016 IEEE International Symposium on Biomedical Imaging.
Received October 21, 2015.