Image Denoising based on Overlapped and Adaptive Gaussian Smoothing and Convolutional Refinement Networks

Yan-Tsung Peng*, Ming-Hao Lin†, Chun-Lin Tang*, and Chin-Hsien Wu†
*Pervasive Artificial Intelligence Research (PAIR) Labs,
Department of Computer Science, National Chengchi University, Taipei, Taiwan
Email: {ytpeng, 104703052}@cs.nccu.edu.tw
†Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan
Email: {d10002105, chwu}@mail.ntust.edu.tw

***Abstract*—We propose to use overlapped and adaptive Gaussian smoothing (OAGS) and convolutional refinement networks (CRN) to recover images corrupted by salt-and-pepper noise. First, the OAGS method identifies noise pixels and recover them. Then, CRN further improve and restore the recovered results with sharper and clearer edges. Experimental results demonstrate the proposed OAGS+CRN method significantly outperforms state-of-the-art denoising methods.**

I. Introduction

With the popularity of smartphones and digital cameras where the camera technology has advanced tremendously in recent years, everyone can take photos anywhere anytime. However, the quality of images captured by these devices may not be assured because it is inevitable that noise may be generated by camera sensors during image acquisition under different conditions. For example, images are prone to be contaminated by noise when taken in low light conditions. Therefore, it is crucial to develop an effective image denoising approach to remove unwanted noise.

In this work, we are focusing on dealing with one of the most common types of impulsive noise, which is "salt-and-pepper noise." Conventionally, the Median Filter [1] can be applied to removing such type of noise. Although median filtering is simple and fast, it blurs the entire image because it does not differentiate noise pixels from good ones. To further filter out noise, there were denoising methods proposed by switching median filters with different thresholds [2]–[4], where the decision is made based on pre-defined thresholds. Since it is in general hard to select appropriate thresholds, these methods often work for limited cases. Esakkirajan et al. designed the modified decision-based unsymmetric trimmed median filter (MDBUTMF) [5] to replace a noise pixel by the median or mean value in a fixed-size window (3×3) centered at the noise pixel based on predefined conditions. Because this method only considers a rather small window size for filtering, it does not do well when an image has higher density noise. To deal with high-density noise, filters with adaptive window sizes have been developed to deal with this problem. However, adopting a larger window may include many unwanted non-noise pixels not related to the corrupted pixel, which may instead introduce artifacts to the denoised image. Erkana et al. developed the different applied median filter [6], which is a two-pass median filtering with three different window sizes used (3×3, 5×5, and 7×7). In the first pass, it chooses the least window size with at least one non-noise pixel for median filtering. In the second pass, if the first-pass result still has noise pixels, 3×3 median filtering will then be applied to further remove the rest noise pixels. Although the two-pass filtering [6] provides decent results to combat high-density noise, it often introduces fake edges to its denoised results. Fareed et al. [7] proposed to use selective mean filtering with adaptive window sizes to include enough non-noise pixels, and mean filtering can ameliorate the problem of fake edges generated by smoothing them out. However, using simple mean filtering will take into account pixels far away from the corrupt one if the window grows too large, making the output image more blurry.

In this paper, we propose a method based on Overlapped and Adaptive Gaussian Smoothing (OAGS) and Convolutional Refinement Networks (CRN) for salt-and-pepper noise removal. The method first utilizes overlapped Gaussian smoothing with adaptive window sizes to filter out all the noise pixels and recover them. After that, we refine the filtered result with CRN to restore the sharpness of image edges and textures. The contributions of this paper are as follows. 1) We propose a conventional denoising method that outperforms state-of-the-art methods. 2) We adopt convolutional neural networks to refine the denoised output and further analyzing the denoising/refining ability of the CRN. 3) We present comprehensive experimental results using two full-reference metrics (PSNR and SSIM) on DIV2K dataset [8] and other often used images for this topic.

The rest of the paper is organized as follows. In Section II, we describe the proposed method. Qualitative and quantitative

978-1-7281-5606-4/19/$31.00 ©2019 IEEE
DOI 10.1109/ISM46123.2019.00032

experimental results are demonstrated in Section III. Finally, Section IV summarizes the conclusions.

II. Proposed Method

For image denoising, we propose to combine a conventional image denoising method with convolutional neural networks for refinement. First, we employ overlapped and adaptive Gaussian smoothing (OAGS) to preliminary remove all the noise pixels and recover them. Then, the convolutional refinement networks (CRN) use a rather simple network architecture for non-linearly refining the recovered results to further restore and to sharpen the edges and textures, instead of building large and complex networks that take up a lot of memory to handle the entire denoising task.

A. Overlapped and Adaptive Gaussian Smoothing

For an image with salt-and-pepper noise, we assume a noise pixel has either the highest or lowest intensity, which are normalized to be 1 and 0. Thus, first of all, we will generate a non-noise map $B \in Z^N$ for the input image $I \in R^N$ as

$$B(x) = \begin{cases} 0, I(x) = 0 \text{ or } 1; \\ 1, \text{otherwise}, \end{cases} \tag{1}$$

where N represents the total number of pixels, and x is the position of a pixel. Next, each pixel $I(p)$ will be visited, and the proper size of the square filtering window centered at $I(p)$ will be determined, which is that of the one with at least one non-noise pixel in it, denoted $\Omega_I^{p,r}$ with its radius r, where $1 \leq r \leq 6$. It means that r equals to the minimum value between 1 and 6 to contain at least one non-noise pixel in the filtering window. That is, the window size is $(2r+1)^2$, and in the paper. For each pixel in $\Omega_I^{p,r}$, if a noise one, it will be recovered by Gaussian smoothing and saved in the temporary result map T as

$$T(x) \leftarrow T(x) + \frac{\sum_{\forall y} G^r(\|x-y\|^2) \times \Omega_I^{p,r}(y) \times \Omega_B^{p,r}(y)}{\sum_{\forall z} G^r(\|x-z\|^2) \times \Omega_B^{p,r}(z)}, \tag{2}$$

where G^r is a Gaussian kernel using the zero mean and standard deviation r, and $\|x-y\|$ stands for the Euclidean distance between x and y. Note that $\forall y$ and $\forall z$ mean all the valid pixel indices in $\Omega_I^{p,r}$. Since using overlapped averaging to make the output more smooth, we need to save our temporary result in an accumulating way as Eq. (2). Additionally, we have a counter map $C \in Z^N$ that stores the number of times a recovered value is calculated for a noise pixel. In the end, the denoised result is attained by averaging all the recovered values calculated, which equals to T divided by C. The detailed algorithm is described in Algo. 1, where $.\times$ and $./$ are the element-wise multiplication and division.

B. Convolutional Refinement Networks

Following OAGS, CRN is devised to refine the recovered result. The architecture of the CRN is rather simple, which consists of five convolutional layers with 3×3 kernels used. The number of the input and output channels for each layer is 64 except for the number of the first input and the last output channel(s), which must be based on the input and output image. The activation function used is the Rectified Linear Unit, ReLU, and the loss function is the mean square error as

$$L_{mse} = \frac{1}{n} \sum_{i=1}^{n} \|F(I_{d_i}; \Theta) - Y_i\|^2, \tag{3}$$

where n is the training samples, the CRN can be represented in a form of the non-linear function F with I_d and Θ as its input and well-trained parameters, Y is the noise-free version of I, and I_d is the recovered result by OAGS. The flowchart of the proposed method is shown in Fig. 1. The training details will be stated in the Sec. III.

Algorithm 1 Overlapped and Adaptive Gaussian Smoothing

Input: Noise image $I \in R^N$
Output: Denoised image $I_d \in R^N$
Lookup Table: Gaussian kernel G^σ
Internal Vars: Kernel radius r, Window centered at p with radius r in I: $\Omega_I^{p,r} \in R^{(2r+1)^2}$, Non-noise map $B \in Z^N$, Temporary result map $T \in R^N$, Smooth counter $C \in Z^N$.

function GAUSS($\Omega_I^{p,r}$, $\Omega_B^{p,r}$, x, r)
  return $\frac{\sum_{\forall y} G^r(\|x-y\|^2) \times \Omega_I^{p,r}(y) \times \Omega_B^{p,r}(y)}{\sum_{\forall z} G^r(\|x-z\|^2) \times \Omega_B^{p,r}(z)}$
end function

function OAGS(I)
  $B \leftarrow$ Get_NonNoise_Map(I)
  $T \leftarrow$ B.$\times$ I
  $C \leftarrow$ B
  for each pixel index p in I **do**
    $r \leftarrow$ Find_Kernel_Radius(I, p)
    for each pixel index x in $\Omega_I^{p,r}$ **do**
      if $\Omega_I^{p,r}(x)$ is a noise pixel **then**
        $T(x) \leftarrow T(x) +$ GAUSS($\Omega_I^{p,r}, \Omega_B^{p,r}, x, r$)
        $C(x) \leftarrow C(x) + 1$
      end if
    end for
  end for
  return $I_d \leftarrow T./C$
end function

III. Experimental Results

To evaluate the performance of denoising methods, we use 80 often used images for image processing tests plus the DIV2K dataset [8], which consists of 800 training and 100 testing images. Therefore, the number of test images is $100+80$ in total. Note that all the images are transformed into grayscale images and are resized to 512×512. To simulate images with different levels of noise, we choose to randomly add salt-and-pepper noise to images with 50%, 60%, 70%, 80%, and 90% noise levels because most denoising methods generally work well for lower noise levels. To train the CRN, we adopt the Adam optimizer with a fixed learning rate $1e-3$ to train our networks on the 4000 images (800×5 noise levels)

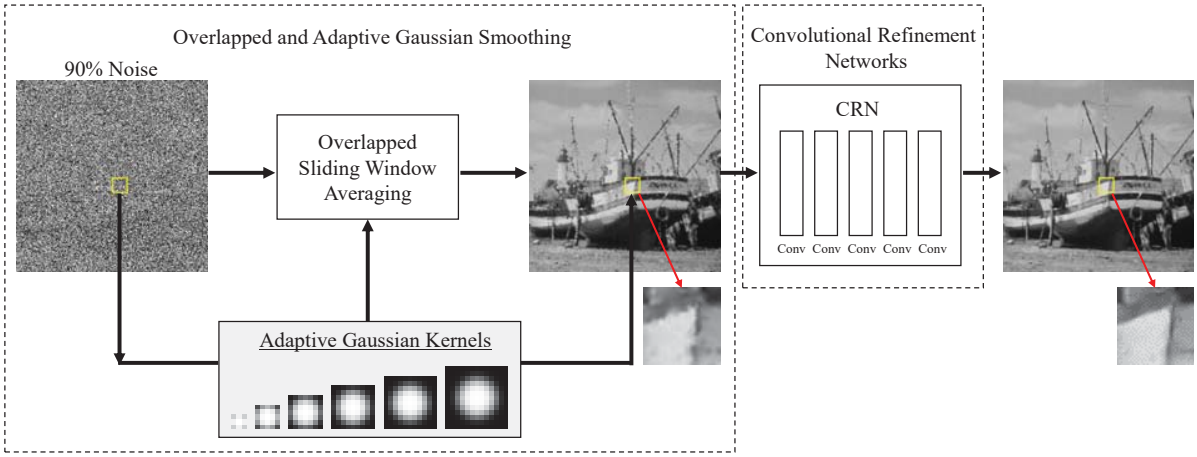

Fig. 1: A flowchart of the proposed method.

and run 500 epochs with batch size 16 without dropout or regularization terms used. The entire CRN has merely about 0.11 million parameters.

Four state-of-the-art denoising methods are chosen for comparisons in this paper, including [4]–[7]. In the following results, we will show our OAGS method has already outperformed the state-of-the-art methods. Moreover, the OAGS results can be further refined and improved by the proposed CRN. An ablation study on CRN, OAGS, and OAGS+CRN is also included.

As can be seen in Fig. 2, although the images corrupted with 90% noise pixels are very hard to be recovered, OAGS works better than the methods [4]–[7] with more smooth recovered results. Furthermore, OAGS+CRN refines the OAGS recovered results with sharper edges restored. In Table I, we can clearly see that OAGS+CRN outperforms all the other methods over the 180 test images in terms of average PSNR and SSIM values. In line with the subjective assessment, OAGS does better than all the compared state-of-the-art methods.

To further analyze whether basic convolutional networks can already outdo conventional image denoising methods, we train standalone CRN with noise images and their noise-free versions as the input and output in the same way as OAGS+CRN is trained. Interestingly and even surprisingly, it does better in almost every category in Table I than the methods [4]–[7] with simple convolutional neural networks, and only OAGS and OAGS+CRN can beat it. Fig. 3 shows subjective visual comparisons of denoised results of different noise levels by CRN, OAGS, and OAGS+CRN. We can observe noise still exists in the recovered results by the standalone CRN while one can barely find noise in the results using OAGS. However, CRN will restore sharper edges than OAGS. Thus, without a doubt, OAGS+CRN has the best denoised results.

TABLE I: Average PSNR/SSIM values of denoised results for 180 test images with $50\% - 90\%$ noise pixels from all the compared methods.

PSNR/SSIM	50%	60%	70%	80%	90%
[4]	27.02/.84	26.03/.80	25.00/.75	23.78/.69	21.10/.56
[5]	28.15/.87	25.49/.78	22.12/.60	18.32/.37	14.45/.17
[6]	28.91/.89	27.58/.85	26.25/.81	24.79/.75	22.71/.65
[7]	29.50/.90	28.16/.86	26.78/.82	25.24/.75	23.28/.65
CRN	30.07/.89	28.73/.86	27.25/.82	25.56/.77	23.30/.66
OAGS	30.36/**.91**	29.01/.88	27.60/.84	26.03/.78	24.09/.69
OAGS+CRN	**31.13**/**.91**	**29.97**/**.89**	**28.57**/**.85**	**26.88**/**.80**	**24.63**/**.71**

IV. Conclusion and Acknowledgment

We propose a denoising method using OAGS and CRN for salt-and-pepper noise. OAGS first recover pixels corrupted by noise, and CRN refine the recovered results. Based on experimental results, the proposed method remarkably outperforms the state-of-the-art methods subjectively and objectively. This work was supported in part by the Ministry of Science and Technology of Taiwan (MOST) AI Biomedical Research Center under Grant MOST 108-2634-F-019-001, by the MOST under Grant MOST 108-2634-F-004-001 through Pervasive Artificial Intelligence Research (PAIR) Labs, Taiwan, and by Qualcomm through a Taiwan University Research Collaboration Project.

References

[1] T. A. Nodes and N. C. Gallagher, "Median filters: Some modifications and their properties," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 5, pp. 739746, Oct. 1982.

[2] S. Zhang and M. A. Karim, "A new impulse detector for switching median filters," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 360363, Nov. 2002.

[3] P. E. Ng and K. K. Ma, "A switching median filter with boundary discriminative noise detection for extremely corrupted images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 15061516, Jun. 2006.

[4] K. K. V. Toh and N. A.M. Isa, "Noise adaptive fuzzy switching median filter for salt-and-pepper noise reduction," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 281284, Mar. 2010.

[5] S. Esakkirajan, T. Veerakumar, Adabala N. Subramanyam, and C. H. PremChand, "Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter," IEEE Signal Process. Lett., vol. 18, no. 5, pp. 287290, May 2011.

[6] U. Erkana, L. Gökerm, and S. Enginoğluc, "Different applied median filter in salt and pepper noise," *Elsevier on Computers and Electrical Engineering*, pp. 789-798, 2018.

[7] S. B. S. Fareed and S. S. Khader, "Fast adaptive and selective mean filter for the removal of high-density salt and pepper noise," *IET Image Process.*, Vol. 12, Iss. 8, pp. 1378-1387, 2018.

[8] E. Agustsson and R. Timofte, "NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study," *Proc. Conf. Computer Vision and Pattern Recognition Workshops*, Jul. 2017.

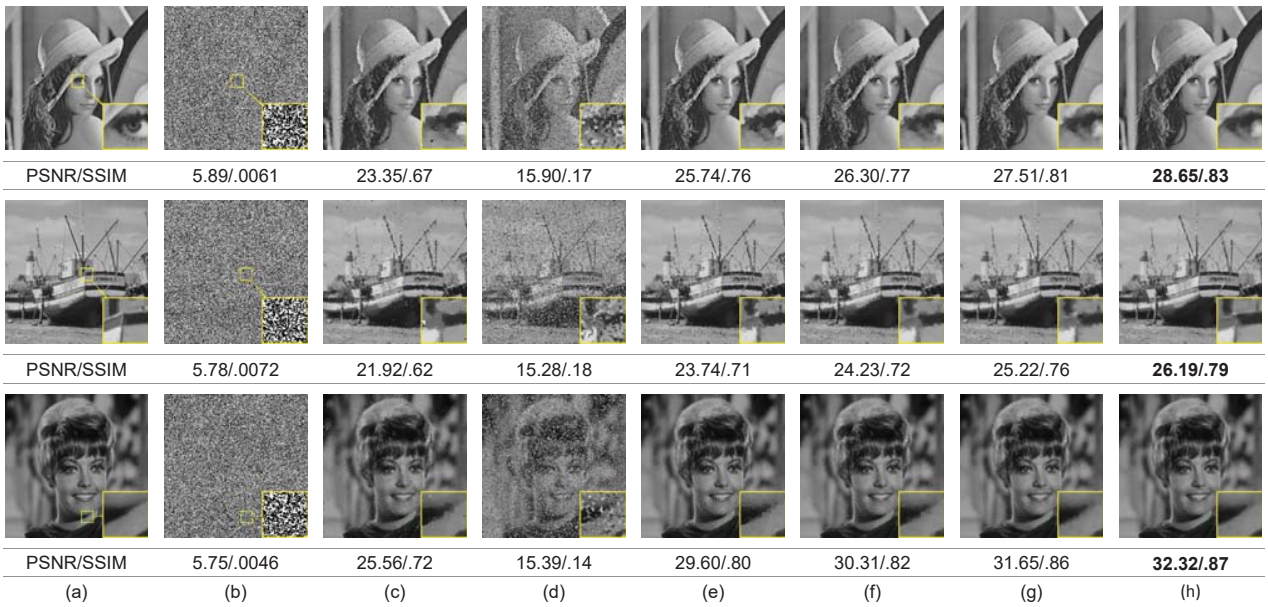

Fig. 2: Subjective comparisons with PSNR/SSIM shown. (a) Original images. (b) Images with 90% noise pixels added. The denoised results obtained using: (c) [4], (d) [5], (e) [6], (f) [7], (g) OAGS, and (h) OAGS+CRN.

Fig. 3: Subjective comparisons of CRN, OAGS, and OAGS+CRN with PSNR/SSIM shown. (a) Original image. The denoised results with (b) 50%, (c) 60%, (d) 70%, (e) 80%, and (f) 90% noise pixels added. From (b) to (f), the results in the first, second, and third row are obtained by CRN, OAGS, and OAGS+CRN, respectively.