

CROSS-SCALE COLOR IMAGE RESTORATION UNDER HIGH DENSITY SALT-AND-PEPPER NOISE

Zecheng He[#], Ketan Tang^{##}, Lu Fang^{*}

[#]Princeton University ^{##}DJI Innovation ^{*}Tsinghua University

ABSTRACT

High-fidelity color image restoration is always of high demanding for high-density noise corrupted images. Such problem becomes more challenging if the degraded image and the expected restored image are of different resolutions, as conventional ‘cascaded: denoising followed by sampling’ and ‘operation on RGB channel independently’ methods induce error amplification and color artifacts. On the contrary, we propose a cross-scale and cross-RGB-channel Salt-and-Pepper Noise (SPN) removal scheme, i.e., the denoising and interpolation are conducted simultaneously within a unified model, which further takes account of RGB cross-channel correlations. Experiments show that the proposed algorithm outperforms conventional state-of-the-art methods in terms of both objective PSNR evaluation and subjective image quality, especially for extremely high-density noise corrupted color images.

Index Terms— Salt-and-Pepper Noise, Multiscale, AutoRegressive Model

1. INTRODUCTION

High-fidelity color image restoration has gathered much attention, especially with the massive production of digital images and movies, which are often grabbed in poor conditions. One typical degradation that affects digital image fidelity may be Salt-and-Pepper Noise (SPN) [1], which is usually caused by under exposure or excessive exposure.

A variety of research work has been proposed to deal with the annoying SPN, such as Standard Median Filter (SMF), Adaptive Median Filter (AMF) [2], Decision Based Algorithm (DBA) [3] etc. Chen *et al.* proposed an Edge Preserving Algorithm (EPA) based on observing the sample correlations of six different directions [4]. Madhu proposed advanced filters based on weighted median [5], direction [6] and pixel membership estimation [7], respectively. Esakkirajan *et al.* proposed Modified Design Based Unsymmetrical Trimmed Median Filter (MDBUTMF) [8], which is based on whether the neighbours of a given pixel are noisy or noise-free. However, these methods have the following problems:

- Conventional noise removal methods can not handle cross-scale cases. Although cascaded: denoising followed by interpolation’ scheme is a possible solution, the blurring and color artifacts induced in the denoising step tend to be amplified in the following interpolation.
- When dealing with color images, the conventional solution is usually performed in RGB channels independently or in Y (luminance) channel, which leads to color artifacts around strong edges. Some work[9], [10] exploit correlations between color channels, but they only discuss gaussian noise without scaling.

Considering the aforementioned problems, we propose a cross-scale denoising model to conduct denoising and interpolation simultaneously. Specifically, the intra-color spatial correlation within a single color channel is represented via a Piecewise AutoRegressive (PAR) model, which takes arbitrary cross-scale situations into account. While for inter-color correlation, a linear prediction model is leveraged to reflect the dependency of RGB channels. The unified model is further solved via quadratic programming to achieve a closed-form solution. Our contributions are (1) We propose a new model in cross-scale color image restoration, i.e., the denoising and interpolation are conducted simultaneously, which significantly eliminates the blurring effect induced in conventional methods. Our approach can be easily extended to more general cases in addition to SPN denoising. (2) Our model fully utilizes both intra and inter RGB color channel correlation to jointly remove the annoying color distortions, instead of separately denoising in each channel.

We elaborate the proposed cross-scale SPN removal model in Section 2. The experimental results are discussed in Section 3. We conclude our work in Section 4.

2. PROPOSED CROSS SCALE AND CHANNEL DENOISING

2.1. Spatial Correlation Model

To deal with the cross-scale situation between the degraded image and the restored image, we start with a Piecewise AutoRegressive (PAR) model for modeling the spatial correla-

* The correspondence author is Lu Fang (luvision.net@gmail.com)

tions of neighbor pixels [11, 12],

$$y(i, j) = \sum_{p,q \in \{i,j\}^8} \alpha_{p,q} y(i+p, j+q) + n(i, j), \quad (1)$$

where $y(i, j)$ is a pixel of the ideal restored image at position (i, j) , $n(i, j)$ is a random noise independent of the image signal, $\{i, j\}^8$ is the set of 8-connected neighbors of position (i, j) , $\alpha_{p,q}$ are the corresponding weighting parameters. Note that $\alpha_{p,q}$ are assumed as constants within a local window T , and the validity of PAR model lies in the fact that in a local window T of a semantically meaningful image, the spatial structure is usually simple and stationary. For conciseness of presentation, the pixels within the block are represented in raster scanning order, and Eqn. (1) becomes

$$y(i) = \sum_{k=1}^8 \alpha_k y_k(i) + n(i), \quad i \in T, \quad (2)$$

where $y_k(i)$ is the k -th neighbor of $y(i)$. We further split the 8 neighbors into ‘diagonal’ $y_t^d(i)$, $t = 1, \dots, 4$ and ‘horizontal-and-vertical (HV)’ neighbors y_t^{hv} , $t = 1, \dots, 4$, due to their different distances from the center pixel and so as different influences. The coefficients α_k are split accordingly as $\alpha = (a, b)$ where a and b correspond to diagonal and HV neighbors respectively. As shown in Fig. 1, the black node represents the target pixel y_i , while blue and red ones denote its diagonal and HV neighbours, respectively. Consequently, the general energy function for optimizing the spatial correlation of pixels in T is

$$\begin{aligned} F(y, a, b) &= \sum_{i \in T} \left(\left\| y(i) - \sum_{t=1}^4 a_t y_t^d(i) \right\|^2 \right. \\ &\quad \left. + \lambda^2 \left\| y(i) - \sum_{t=1}^4 b_t y_t^{hv}(i) \right\|^2 \right), \end{aligned} \quad (3)$$

where λ^2 adjusts the importance of diagonal and HV neighbours.

2.2. Salt-and-Pepper Noise Corruption Model

Recall that the pixels in SPN corrupted image are either completely corrupted or completely clean. To discriminate noise-free and noisy pixels, we define a binary and diagonal weighting matrix W to detect the noise-free pixels, i.e.,

$$W(i) = \begin{cases} 0, & \text{if } x(i) = 0 \text{ or } 255, \\ 1, & \text{otherwise.} \end{cases} \quad (4)$$

W is only an approximate mask for separating the noisy and noise-free pixels. Advanced detection techniques[13],[14] can achieve slightly higher performance. Pixels are supposed to be consistent between the input image x and the restored

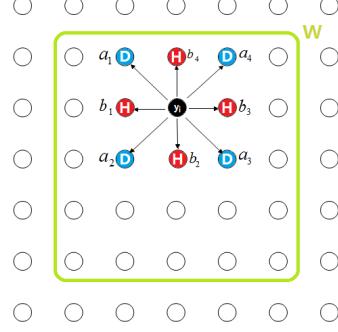


Fig. 1: Illustration of spatial correlations within Diagonal and HV neighbors. The black node is the central pixel, while blue and red ones are ‘diagonal neighbours’ and ‘horizontal-and-vertical(HV) neighbours’ respectively.

image y on noise-free positions. In particular, if x and y are of different resolutions, we have

$$WSy = Wx, \quad (5)$$

where S is subsampling matrix compensating for the cross-scale between x and y . Apparently, Eqn. (5) indicates that for noisy pixels, such consistent constraint does not apply.

Combine Eqn. (3) and Eqn. (5), the cross-scale SPN removal problem is formulated as a constrained quadratic programming problem,

$$\begin{aligned} \min_{\{y, a, b\}} \quad & F(y, a, b) = \\ & \sum_{i \in T} \left(\left\| y(i) - \sum_{t=1}^4 a_t y_t^d(i) \right\|^2 + \lambda^2 \left\| y(i) - \sum_{t=1}^4 b_t y_t^{hv}(i) \right\|^2 \right) \\ \text{s.t.} \quad & WSy = Wx. \end{aligned} \quad (6)$$

Note that the neighboring pixels of $y(i)$ may exceed the local window T for boundary pixels. Let v be the vector of pixels outside T but one pixel near y , \tilde{y} be the vector of all pixels including both y and v , then we have $\tilde{y} = \begin{bmatrix} y \\ v \end{bmatrix}$. To solve Eqn. (6), we introduce the assumption that the outer pixel vector v is known.

2.3. Adaptive RGB Inter-Color Correlation Model

The removal of SPN for color images is more challenging, since conventional methods usually conduct noise removal within RGB channels independently, leading to fake colors around sharp edges. As investigated in previous work for color image/video coding [15, 16, 17, 18], the high frequency of RGB color channels is highly resemble to each other. In other words, RGB inter-color correlation effect is non-negligible when dealing with color images, which can be exploited by inter-color prediction. For example, inter-color

prediction between R and G components is represented as

$$\hat{y}^r(i) = y^g(i) + o^{rg}, \quad i \in T, \quad (7)$$

where $y^r(i)$ and $y^g(i)$ denote the pixel intensities in R and G channels at position i respectively, and $\hat{y}^r(i)$ represents the prediction for $y^r(i)$. o^{rg} is the additive offset, which can be optimized via

$$\begin{aligned} \min_{o^{rg}} \quad & E[(\hat{y}^r - y^r)^2] \\ \text{s.t.} \quad & \hat{y}^r(i) = y^g(i) + o^{rg}. \end{aligned} \quad (8)$$

We have the closed-form solution of Eqn. (8) as

$$o^{rg} = \bar{y}^r - \bar{y}^g, \quad (9)$$

where \bar{y}^r and \bar{y}^g denote the mean of y^r and y^g within local window T respectively. The inter-color correlation between R and B, G and B are similar.

Taking inter-color correlation model into account, the energy function in Eqn. (6) is represented as

$$\begin{aligned} F(y, a, b) &= \|C\tilde{y}\|^2 + \beta^2 \sum_{\substack{c_1, c_2 \in \{r, g, b\} \\ c_1 \neq c_2}} \|y^{c_1} - y^{c_2} - O^{c_1 c_2}\|^2 \\ &= \left\| \begin{bmatrix} C^r & & \\ & C^g & \\ \beta I_{in} & -\beta I_{in} & C^b \\ -\beta I_{in} & \beta I_{in} & -\beta I_{in} \\ & & \beta I_{in} \end{bmatrix} \tilde{y} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \beta O^{rg} \\ \beta O^{gb} \\ \beta O^{br} \end{bmatrix} \right\|^2 \\ &\triangleq \|\tilde{C}\tilde{y} - O\|^2, \end{aligned} \quad (10)$$

where $c = \{r, g, b\}$, $C^c = \begin{bmatrix} C_D^c \\ C_H^c \end{bmatrix}$, C_D^c and C_H^c represent the coefficient matrix for ‘diagonal neighbors’ and ‘horizontal and vertical neighbors’ respectively, where

$$C_D^c(i, j) = \begin{cases} 1 & \text{if } y_j^c \text{ is the } i\text{-th pixel in } c \text{ channel,} \\ -a_{ij}^c & \text{if } y_j^c \text{ is the } t\text{-th diagonal neighbor of } y_i^c, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

and C_H^c has the same form. I_{in} extracts pixels inside T from \tilde{y} , i.e., $I_{in}\tilde{y} = y$. $O^{rg} = o^{rg}\mathbf{1}$, $O^{gb} = o^{gb}\mathbf{1}$, $O^{br} = o^{br}\mathbf{1}$, where $\mathbf{1}$ is a vector of all ones. β^2 controls the importance ratio of intra-color correlation and inter-color correlation¹.

2.4. Optimization with Unified Model

With the both intra and inter color correlation model represented in Eqn. (10), the cross-scale SPN removal problem in Eqn. (6) is reformulated as

$$\begin{aligned} \min_{\{y, a, b\}} \quad & F(y, a, b) = \|\tilde{C}\tilde{y} - O\|^2 \\ \text{s.t.} \quad & WSy = Wx, \end{aligned} \quad (12)$$

¹In our experiments, different β leads to tiny variation in SNR, thus $\beta = 1$ is used in our simulations.

which can be solved via Gauss-Seidel iterations [12] as follows,

$$\begin{aligned} \{a^{(n+1)}, b^{(n+1)}\} &= \arg \min_{a, b} F(y^{(n)}, a, b), \\ y^{(n+1)} &= \arg \min_y F(y, a^{(n+1)}, b^{(n+1)}), \end{aligned} \quad (13)$$

where $\{a^{(n)}, b^{(n)}\}$ are the coefficients at n -th iteration, $y^{(n)}$ is the estimated vector at n -th iteration².

For the estimation of coefficients $\{a, b\}$ in Eqn. (13), with the decomposition of the inter-color correlation penalty term and the sampling matrix, we have

$$\begin{cases} a &= \arg \min_a \|y - Y_a a\|^2 = (Y_a^T Y_a)^{-1} (Y_a^T y), \\ b &= \arg \min_b \|y - Y_b b\|^2 = (Y_b^T Y_b)^{-1} (Y_b^T y), \end{cases} \quad (14)$$

where Y_a and Y_b are the matrices of neighboring pixels of y . The i -th row of matrix Y_a consists of the diagonal neighbors $y_t^d(i)$ of $y(i)$, and the i -th row of matrix Y_b consists of the HV neighbors $y_t^{hv}(i)$ of $y(i)$, $t = 1, \dots, 4$, i.e., $Y_a(i, t) = y_t^d(i)$, $Y_b(i, t) = y_t^{hv}(i)$, for $t = 1, \dots, 4$.

For the estimation of y in Eqn. (13), by splitting coefficient matrix \tilde{C} into two sub-matrices corresponding to inner pixel vector y , and outer pixel vector v , i.e., $\tilde{C} = [D_y, D_v]$, we have

$$\begin{aligned} y &= \arg \min_y \|\tilde{C}\tilde{y} - O\|^2 \\ &= \arg \min_y \|D_y y + D_v v - O\|^2 \\ \text{s.t.} \quad & WSy = Wx, \end{aligned} \quad (15)$$

where D_y and D_v are coefficient matrices corresponding to y and v respectively. With the usage of intermediate variables, Eqn. (15) is represented as a standard equality constrained quadratic problem, with the closed form solution be

$$\begin{bmatrix} y \\ \lambda \end{bmatrix} = \begin{bmatrix} D_y^T D_y & S_{in}^T \\ S_{in} & 0 \end{bmatrix}^{-1} \begin{bmatrix} -D_y^T (D_v v - O) \\ x_{in} \end{bmatrix}. \quad (16)$$

3. EXPERIMENT RESULTS AND DISCUSSIONS

We generate test images by resizing original images with ratio 1/2 (equivalently, the cross-scale ratio between input and output is $\times 2$)³, followed by adding SPN whose density goes from 0 to 0.9 with increment of 0.1. We compare our method with conventional ‘cascaded: denoising followed by interpolation’ schemes. Specifically, the SPN removal algorithms contain

²Without causing any ambiguity, the super script index (n) or $(n+1)$ will be omitted in the following discussions. In our experiments, the optimization converges within 5 iterations

³Note that our proposed model covers arbitrary cross-scale cases between input and output images. For simply of presentation, we take cross-scale ratio $\times 2$ as an example.

EPA [4], DBA [3] and MDBUTMF [8]. The interpolation algorithms cover bicubic (BI), Sparse Representation based interpolation (SR) [19][20], and Nonlocally Centralized Sparse Representation (NCSR) [21].

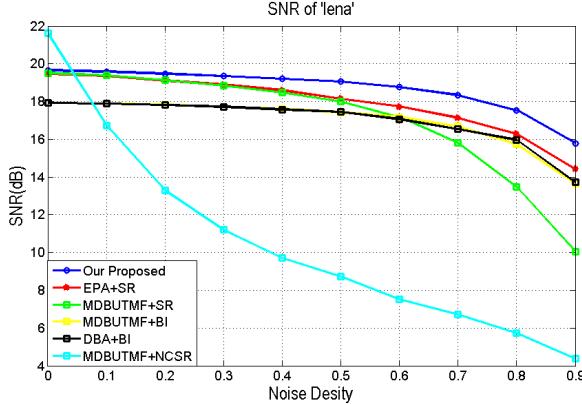


Fig. 2: Comparison of SNR results for representative image ‘Lena’.

Fig. 2 depicts the comparisons of concerned algorithms for representative image ‘Lena’, where horizontal-axis represents the noise density from 0 (noise-free) to 0.9 (heavily noisy), and vertical-axis represents the SNR value of resultant image. The dark blue curve represents the performance of proposed method, while others are conventional cascaded methods. Examining Fig. 2, we have following observations:

- For relatively low SPN density, most algorithms have stable performance. However, the SNR of conventional cascaded methods tends to drop vastly when the noise increases. In particular, the degradation of NCSR (light blue curve in Fig.2) appears more severe, since the assumption of nonlocally self similarities in NCSR fails abruptly once the noise density increases.
- On the contrary, our method outperforms others in achieving stable results, and the superiority becomes more obvious when noise density increases, which originates from conducting denoising and interpolation simultaneously by incorporating both inter and intra color correlations.

As further illustrated in Figure.3, our method is capable in recovering strong edges with much less color artifacts than conventional methods, under very high (0.9) SPN density.

Considering that denoising problem is a special case of our algorithm, with the cross-scale ratio being $\times 1$, we compare our algorithm with conventional SPN removal algorithms in Table 1 (SPN density 0.6), where the proposed algorithm produces the highest PSNR results.

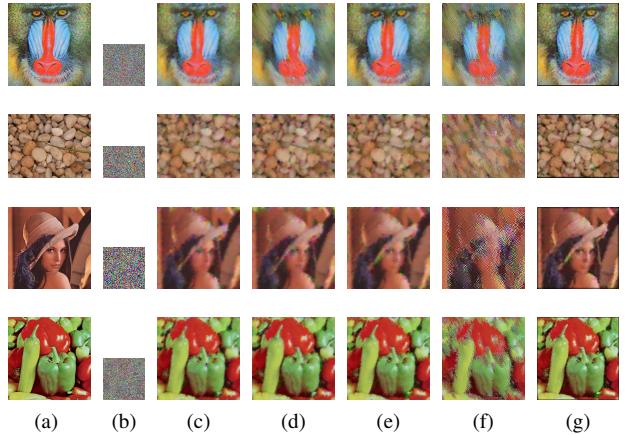


Fig. 3: The performance of concerned methods. a: Original. b: Noised and downsampled. c: EPA+BI. d: MDBUTMF+BI. e: EPA+SR. f: MDBUTMF+NCSR. g: Proposed. SPN density=0.9, cross-scale ratio = $\times 2$.

Image	AMF	DBA	MDBUTMF	EPA	Proposed	Gain
Baboon	21.77	21.67	22.70	24.92	28.76	3.84
Bike	21.10	21.02	22.18	24.85	31.13	6.28
Flower	19.35	19.31	20.25	21.74	22.66	0.92
Lena	28.28	27.78	29.83	33.22	35.55	2.33
Necklace	18.02	17.95	19.06	21.03	27.21	6.18
Parrot	27.78	27.43	29.28	31.67	33.42	1.73
Building	21.17	21.04	22.00	24.17	31.76	7.60
Tree	23.03	22.93	24.34	27.41	31.22	3.82
Average	22.57	22.39	23.71	26.82	30.09	3.26

Table 1: Comparisons between our algorithm (with cross-scale ratio= $\times 1$) and conventional SPN removal methods.

4. CONCLUSIONS

In this paper, we propose a cross-scale SPN removal algorithm that can deal with arbitrary scale ratios between SPN-corrupted image and resultant denoised image, taking account of both intra and inter correlations among RGB color channels. Experiments show the effectiveness of our algorithm especially for high density SPN corrupted images, when compared to conventional ‘cascaded: denoising followed by interpolation’ methods, which have defects in introducing blurring and color distortions. While setting the cross-scale ratio being $\times 1$, our algorithm becomes a sole SPN removal algorithm, which outperforms the previous SPN removal methods in producing visually pleasant results with much less color artifacts around strong edges.

Acknowledgment

This work is supported in part by the National key foundation for exploring scientific instrument No.2013YQ140517, in part by the NSFC fund (61571259, 61531014).

5. REFERENCES

- [1] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Prentice Hall, 3 edition, 2007.
- [2] H. Hwang and R. A. Haddad, “Adaptive median filters: new algorithms and results,” *Image Processing, IEEE Transactions on*, vol. 4, no. 4, pp. 499–502, 1995.
- [3] K. S. Srinivasan and D. Ebenezer, “A new fast and efficient decision-based algorithm for removal of high-density impulse noises,” *Signal Processing Letters, IEEE*, vol. 14, no. 3, pp. 189–192, 2007.
- [4] Pei-Yin Chen and Chih-Yuan Lien, “An efficient edge-preserving algorithm for removal of salt-and-pepper noise,” *Signal Processing Letters, IEEE*, vol. 15, pp. 833–836, 2008.
- [5] Madhu S Nair and PM Ameera Mol, “An efficient adaptive weighted switching median filter for removing high density impulse noise,” *Journal of The Institution of Engineers (India): Series B*, vol. 95, no. 3, pp. 255–278, 2014.
- [6] Madhu S Nair and PM Ameera Mol, “Direction based adaptive weighted switching median filter for removing high density impulse noise,” *Computers & Electrical Engineering*, vol. 39, no. 2, pp. 663–689, 2013.
- [7] Madhu S Nair and G Raju, “A new fuzzy-based decision algorithm for high-density impulse noise removal,” *Signal, Image and Video Processing*, vol. 6, no. 4, pp. 579–595, 2012.
- [8] 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,” *Signal Processing Letters*, vol. 18, pp. 287–290, 2011.
- [9] Lazhar Khriji and Moncef Gabbouj, “Vector median-rational hybrid filters for multichannel image processing,” *IEEE signal processing letters*, vol. 6, no. 7, pp. 186–190, 1999.
- [10] Julien Mairal, Michael Elad, and Guillermo Sapiro, “Sparse representation for color image restoration,” *IEEE Transactions on image processing*, vol. 17, no. 1, pp. 53–69, 2008.
- [11] Xiangjun Zhang and Xiaolin Wu, “Image interpolation by adaptive 2-d autoregressive modeling and soft decision estimation,” *Image Processing, IEEE Transactions on*, vol. 17, no. 6, pp. 887–896, 2008.
- [12] Ketan Tang, Oscar C. Au, Lu Fang, Zhiding Yu, and Yuanfang Guo, “Image interpolation using autoregressive model and gauss-seidel optimization,” in *Image and Graphics (ICIG), International Conference on*, 2011, pp. 66–69.
- [13] Fei Duan and Yu-Jin Zhang, “A highly effective impulse noise detection algorithm for switching median filters,” *IEEE Signal Processing Letters*, vol. 17, no. 7, pp. 647–650, 2010.
- [14] Xuming Zhang and Youlun Xiong, “Impulse noise removal using directional difference based noise detector and adaptive weighted mean filter,” *IEEE Signal processing letters*, vol. 16, no. 4, pp. 295–298, 2009.
- [15] L. Goffman-Vinopal and M. Porat, “Color image compression using inter-color correlation,” in *Image Processing (ICIP), International Conference on*, 2002, vol. 2, pp. 353–356.
- [16] S. Benierbah and M. Khamadja, “Compression of colour images by inter-band compensated prediction,” *Vision, Image and Signal Processing, IEE Proceedings*, vol. 153, no. 2, pp. 237–243, 2006.
- [17] Kim Yong-Hwan, Choi Byeongho, and Paik Joonki, “High-fidelity rgb video coding using adaptive inter-plane weighted prediction,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 19, no. 7, pp. 1051–1056, 2009.
- [18] Bahadir K. Gunturk, Yucel Altunbasak, and Russell Mersereau, “Color plane interpolation using alternating projections,” in *Acoustics, Speech, and Signal Processing (ICASSP), IEEE International Conference on*, 2002, vol. 4, pp. IV–3333–IV–3336.
- [19] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma, “Image super-resolution as sparse representation of raw image patches,” in *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, 2008.
- [20] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, “Image super-resolution via sparse representation,” *IEEE transactions on image processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [21] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li, “Nonlocally centralized sparse representation for image restoration,” *Image Processing, IEEE Transactions on*, , no. 4, pp. 1620–1630, 2013.