An improved two stage directional denoising using current and localized frames for video processing

Sanjay M.Malode[1] & Dr.V.M.Thakare[2]
[1]PhD Scholar, Professor & Head, Singhania University, Jhunjhunu, Rajasthan.
[2]P.G. Department of CSE, SGB Amravati University, Amravati (M.S.)

Received: July 16, 2018 Accepted: August 31, 2018

ABSTRACT *Removing noise in images and videos is a challenging task in the field of image and video processing today. Pixel dependent Poisson noise is one of the major problems in images and videos. Many literatures show various techniques to remove the unwanted noise for further applications. The main drawbacks of most of the work are low signal to noise ratio and loss of features. This work is focused on both issues for video processing. The Linear minimum mean square error estimate using directional denoising (DN) is used by localizing previous frame from the successive frame in a video. The directional denoising reduces the computational complexity since it is calculated by using mean, variance and covariance. The estimate of a noisy pixel is broken in three estimates along horizontal/vertical, diagonal and center direction and then fused to find the final estimate. The frame to be localized using the successive or the next immediate frame follows a simple block matching technique without thresholding. The block matching is performed using the minimum block size (3x3) and assumed that the block in successive frame is deflected within a 5x5 window due to camera motion. Further an average of both the current and the reproduced frame (localized frame) is applied with directional denoising which gave significant performance in terms of peak signal to noise ratio and the video frame quality preserving edges.*

Keywords: *Poisson noise, signal to noise ratio, linear minimum mean square error, directional denoising, block matching, thresholding, localized frame.*

Introduction

Image and video acquisitions are prone to signal dependent errors since they are captured by digital cameras [1-3]. The output of the camera should be modeled as a signal dependent Poisson-distributed random variable that is degraded by signal-independent additive Gaussian noise. Removing such signal dependent noise can be solved in broadly two ways, either by exploiting the Poisson noise properties directly [2-4] or to apply a nonlinear variance-stabilizing transformation (VST) to render the noise variance constant. In the second approach the noise is converted as a Gaussian noise with unity variance and then removed using various standard filters. Finally inverse variance stabilization gives the denoised image. Many researches [1, 5-6] have been oriented for the second approach since vast research has been done in removing Gaussian noise. The commonly used VST for Poisson–Gaussian noise is generalized Anscombe transformation (GAT) [7], which extends the classical Anscombe transformation [8]. The classical Anscombe transformation is aimed at pure Poisson noise. Researchers have focused on many video denoising algorithms for Gaussian noise, such as VBM3D [9] and SURE-LET [10], ST-GSM [11], DNLM [12]. These algorithms can effectively remove noise but are difficult to implement on hardware because they often consist of time- and resource-consumed computations, e.g., local similar patch search, iteration optimization, or domain transformation.

Large number of work was concentrated on Gaussian noise with a drawback that it requires proper tuning of parameters. The non-local mean filter [13, 14],the Gaussian smoothing model [15], Gaussian kernel density [16], Anisotropic Filtering method [17], Gaussian scale mixture modeling in [18], the principal component analysis approach in [19], the fast iterative shrinkage and thresholding method in [20, 21], and a model based on the Euler-Lagrange equations to reduce the noise [22].

There are many models that use sparse and redundant representations over trained dictionaries. Elad et al. [23] dictionary to describe image content, a sparse dictionary for the noisy image [24], the Principal Neighborhood Dictionary [25] and Mairal et al. [26] combined non-local means approach and dictionary learning.

Wavelet-based filters are very popular for image denoising too. Parrilli et al. [27] use non-local filtering and wavelet domain shrinkage. Yu et al. [28] incorporate wavelet-based trivariate shrinkage with a spatial-based filter. Yaroslavsky and Eden [29] and Yaroslavsky [30] utilize the neighborhood filtering method to attenuate the noise. Stein's unbiased risk estimate approach [31] is an orthonormal wavelet thresholding

approach. Eslami and Radha [32] implement the contourlet transform. The bilateral filter [33] is a nonlinear filter performing spatial averaging. Yan et al. [34] explore the sparsity of wavelets and employ hierarchical dictionary learning in each level of the wavelets.

In the current work, we use a video denoising algorithm by Lei Zhang, Xin Li, and David Zhang [35] with two stage denoising based for current and reproduced frame and then denoising the average frame from first stage denoising.

Our Contributions

Lei Zhang, Xin Li, and David Zhang [35] in their work had suggested an efficient directional denoising technique to remove poisson noise and preserved edges during interpolation. The process of directional denoising and interpolation were simultaneously applied to reduce computational complexity. The optimal estimation was modeled to estimate noiseless and missing samples. For each noisy sample, they computed multiple estimates of it along different directions and then fuse those directional estimates for a more accurate output.

Many frame stitching techniques are used by researchers in the literature. We propose a novel frame stitching technique using a 3x3 window from a current frame to locate in 7x7 neighborhood in the next frame, assuming that there is little deflection of the object features in the next frame with respect to the current frame. The main idea behind stitching two consecutive frames is to eliminate Poisson noise of the current frame by reproducing the current frame from the next frame. Results showed high peak signal to noise (psnr) ratio when directional estimate using LMMSE is made on the reproduced frame. Therefore our contribution is reproducing current from the successive or next frame and applying directional denoising.

Frame reproduction

To achieve greater block matching accuracy, the block size of the source (current frame) frame was taken to be 3x3. Considering camera movement to be little in the destination frame (successive or next frame) the source block is searched in a neighborhood of 7x7 window around the central block with same spatial coordinates as that of the source block. The successive frame is padded with border elements so as to create a neighborhood of 7x7 around the border elements. Figure 1 below shows both the source and the destination blocks from current and the next frame respectively.

	(m,n)	

	1	2	3	4	5	
	6	7	8	9	10	
	11	12	(m,n)	14	15	
	16	17	18	19	20	
	21	22	23	24	25	

Figure 1- The source 3x3 and the destination 7x7 blocks

The source block is compared with all the 25 destination blocks situated around the 25 elements. The distance error is calculated as

E=avg(abs(Bs-Ds))

where Bs represent all the nine elements of source block and Ds represents all the nine elements of the destination block. Individual elements are subtracted and the mean value is calculated. Thus we have 25 values corresponding to 25 blocks in the destination. Further the block index corresponding to the minimum distance is found. Here no thresholding is used, simply the block from destination whose distance is at least out of 25 blocks is considered. It may happen that more than one block have same distance value and that to be the minimum. The following conditions are applied for the best matching block.

1. If only one block is found to be having the minimum value, its index is stored.
2. If more than one block have the same minimum distance and block 13 (central block) is one of them, then block 13 is taken into account and rest are neglected.
3. If block 13 is not present, then the immediate neighbors 7, 8, 9, 12, 14, 17, 18 1nd 19 are searched in sequence. For example, if block 7 is found to have minimum distance, it is taken into account. Any one is considered and that to in sequence.
4. If none of the immediate block has minimum distance then remaining blocks 1, 2, 3, 4, 5, 6, 10, 11, 15, 16, 20, 21, 22, 23, 24 and 25 are searched again in sequence. This will result in motion error so the index is accompanied by an infinity value is stored. This will be corrected later when the camera alignment is calculated. For example if 11, 22 and 25 are the blocks with same minimum distance, then 11 is stored with infinity, whereas 22 and 25 are stored for future correction.

Now the total count for each index is calculated except for the block 13. This is done because block 13 has the same spatial coordinates as that of the source block. The block with the maximum count will give us the camera motion. For example if block 17 has the maximum count then the camera is deflected 2150 with respected to the current frame. Now the step 4 above index is corrected. At step 4, index 11 was stored, but as per camera movement, now it is corrected and replaced by index 22, since it is in the vicinity of 17.

The above frame reproduction is done after converting the frame from RGB color space to gray scale. Once the blocks are localized, the new frame is reproduced form the destination frame pixel value and finally all the components (R, G and B) are acquired from the same spatial coordinated of the destination frame. Figure 2 shows the current frame and the next successive frame and figure 3 shows the current frame with the reproduced frame.

Figure 2- The current frame and the successive frame

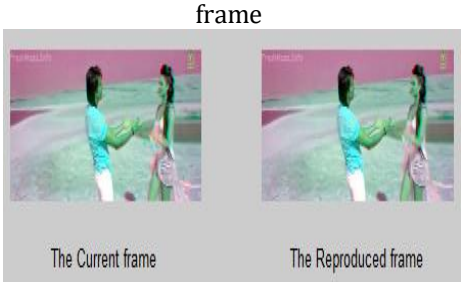

Figure 3- The current frame and the frame reproduced by block matching technique

Directional Denoising

The noisy frame In containing Poisson noise can be represented as

In=I+n;

Where, I is the noiseless frame, n is the Poisson noise.

Consider a noisy pixel In(m,n), the goal is to estimate the noiseless value I'(m,n) of it using it neighbors Iv(m,n). We had utilized the technique mentioned in [32] to estimate I'(m,n) nearer to original pixel I(m,n). Optimal estimation technique such as the LMMSE [33] is used to find the estimate I'(m,n) of the original pixel I(m,n). If a 3×3 window around I'(m,n) is used, Sv will be a 9×1 variable vector and its variance matrix var(Sv) is a 9×9 matrix. The inverse of the 9×9 matrix var(Sv) will cost much computation. So, we divide the estimation of into several sub-problems, each of which yields a directional estimate of I'(m,n), and then fuse those directional estimates into a more robust one.

Refer to Figure 4, we partition the nine noisy samples within the 3×3 window centered on (m, n) into three groups along different directions: horizontal/vertical, diagonal and the noisy sample Iv(m,n). The red circle represents the noiseless pixel to be estimated and the blue circles represent the available noisy measurements. Each of the first two groups has four elements and the last group has one member only. With the three groups, we are able to calculate three directional estimates of I'(m,n). The three estimates can then be adaptively fused to obtain a more robust and accurate estimation of I'(m,n).

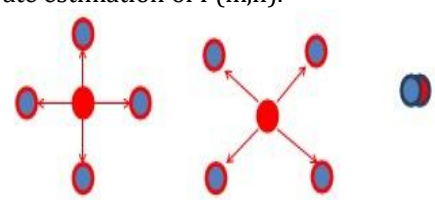

Figure 4 - Partition of the nine measurements into 3 groups to estimate the unknown noiseless sample.

Considering four neighbors of I'(m,n) above, then the LMMSE can be calculated by finding mean, variance and covariance for all the three groups separately. The above mean, variance and covariance are calculated by considering a training window around the 3x3 matrix under consideration as suggested in [32]. The training window centered on the 3x3 window considered is of 5x5. The directional estimates are then fused to find the actual estimate. The brief estimation can be found in [32]. Here we had not taken into account the weight vectors they had calculated for interpolation.

Methodology & Results

1. Read the video.
2. Store the frames in memory.
3. Select any two consecutive frames.
4. LMMSE using directional denoising to current frame – IM1
5. Frame stitching
6. LMMSE using directional Denoising to reproduced frame – IM2
7. LMMSE using directional Denoising to (IM1+IM2)/2

e ISSN 2348 –1269, Print ISSN 2349-5138
Cosmos Impact Factor 4.236

Figure 5 – Results of Directional denoising to R, G and B components of current frame. First column represent original R, G and B frames. Second column represent frames with poisson noise and the third column is the result after denoising.

Figure 6 – First frame is the RGB frame with poisson noise and the second frame is the result of directional denoising. Second frame is the combined view of all components.

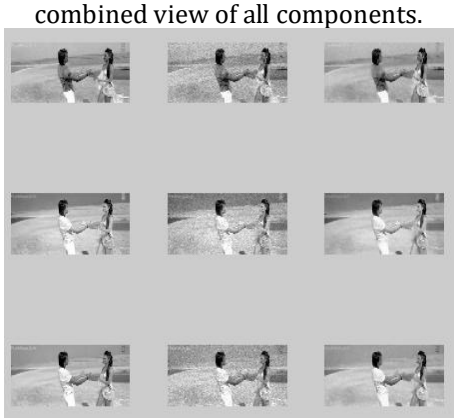

Figure 7 – Results of Directional denoising to R, G and B components of reproduced frame. First column represent original R, G and B frames. Second column represent frames with Poisson noise and the third column is the result after denoising.

Figure 8 – First frame is the original RGB frame and the second frame is the result of directional denoising. Second frame is the combined view of all reproduced R, G and B components.

Figure 9 – Results of Directional denoising to R, G and B components of average of denoised current and denoised reproduced frame. First column represent original R, G and B frames. Second column represent frames with Poisson noise and the third column is the result after denoising.

Figure 10 – First frame is the original RGB frame and the second frame is the result of directional denoising. Second frame is the combined view of

all reproduced R, G and B components.

The results can be better understood from signal to noise ratios.

Table 1- Comparison of SNR value when denoising is applied to the components of original frame.

Sr. No.	Frames	Initial SNR	SNR using DN for current Frame
1	R	26.1944	37.3713
2	G	25.8389	37.6302
3	B	25.8801	37.8484

Table 2- Comparison of SNR value when denoising is applied to the components of reproduced frame.

Sr. No.	Frames	Initial SNR	SNR using DN for reproduced Frame
1	R	23.8219	34.6812
2	G	25.3181	35.4727
3	B	24.5509	35.6084

Table 3- Comparison of SNR value when denoising is applied to the components of average of denoised frame.

Sr. No.	Frames	Initial SNR	SNR using DN for average of denoised current frame & denoised reproduced Frame
1	R	31.4118	36.2483
2	G	33.5079	36.9786
3	B	33.5231	37.3920

Table 4 – Structural Similarity (SSIM) index between two images

Sr. No.	Frame 1	Frame 2	SSIM
1	Current frame	Denoised frame	0.9316
2	Reproduced frame	Denoised frame	0.9146
3	Averaged frame	Denoised frame	0.9313

Conclusions

Figure 6, 8 and 10 clearly shows the difference in visual context. The reproduced frame when applied for directional denoising produces much better results than applying denoising to the frame itself. The visual quality is much improved in figure 10 when average of the denoised frames of current frame and the denoised reproduced frame are denoised. The SNR values for all the components independently as seen from table 1, 2 and 3 shows a further improvements. The edges as seen in figure 10 are preserved with small amount of loss. Table 4 indicates the structural similarity index [36]. The computational complexity of the denoising and the frame reproduction is low. The color perception has been degraded to some extent but as far as poisson noise is concerned it is worth. Also if video is concerned it may not be an issue when high frame rate. Further work will be focused on enhancing the image quality with better color perception and preserving edges. Also we have not used minimum threshold value in block matching, which may have represented an incorrect block from the successive frame to the source block. The block matching algorithm can be modified to improve the performance.

References

1. Makitalo, M., Foi, A.: Optimal inversion of the generalized Anscombe transformation for Poisson–Gaussian noise. IEEE Trans. Image Process. 22(1), 91–103 (2013)
2. Luisier, F., Blu, T., Unser, M.: Image denoising in mixed Poisson-Gaussian noise. IEEE Trans. Image Process. 20(3), 696–708 (2011)
3. Salmon, J., Harmony, Z., Deledalle, C.A., Willett, R.: Poisson noise reduction with non-local PCA, In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012 (2012)
4. Deledalle, C., Tupin, F., Denis, L.: Poisson NL means: unsupervised non local means for Poisson noise. In: Proceedings of IEEE International Conference on Image Processing, ICIP 2010 (2010)
5. Ma¨kitalo, M., Foi, A.: Optimal inversion of the Anscombe transformation in low-count Poisson image denoising. IEEE Trans. Image Process. 20(1), 99–109 (2011)
6. Zhang, B., Fadili, J.M., Starck, J.-L.: Wavelets, ridgelets, and curvelets for Poisson noise removal. IEEE Trans. Image Process. 17(7), 1093–1108 (2008)
7. Starck, J.L., Murtagh, F., Bijaoui, A.: Image Processing and Data Analysis. Cambridge University Press, Cambridge (1998)
8. Anscombe, F.J.: The transformation of Poisson, binomial and negative-binomial data. Biometrika 35(3/4), 246–254 (1948)
9. Dabov, K., Foi, A., Egiazarian, K.: Video denoising by sparse 3-D transform-domain collaborative filtering. In: Proceedings of IEEE transactions on Europen Signal Processing Conference, (EUSIPCO), Poznan, Poland, pp. 1257–1260 (2007)
10. Luiser, F., Blu, T., Unser, M.: SURE-LET for orthonormal wavelet-domain video denoising. IEEE Trans. Circuits Syst. Video Technol. 20(6), 913–919 (2010)
11. Varghese, G., Wang, Z.: Video denoising based on a spatiotemporal Gaussian scale mixture model. IEEE Trans. Circuits Syst. Video Technol. 20(7), 1032–1040 (2010)
12. Han, Y., Chen, R.: Efficient video denoising based on dynamic nonlocal means. Image Vis. Comput. 30(2), 78–85 (2012)
13. A Buades, B Coll, J-M Morel, Image denoising methods. A new nonlocal principle. SIAM Rev. 52(1), 113–147 (2010).
14. A Buades, B Coll, JM Morel, A review of image denoising algorithms, with a new one. Multiscale model. Simul. 4(2), 490–530 (2005).
15. M Lindenbaum, M Fischer, A Bruckstein, On gabor contribution to image enhancement. Pattern Recognit. 27: 1–8 (1994).
16. D Gabor, Information theory in electron microscopy. Lab. Investig. 14(6), 801–807 (1965).
17. P Perona, J Malik, Scale space and edge detection using anisotropic diffusion. IEEE Trans. Patt. Anal. Mach. Intell. 12:, 629–639 (1990).
18. M Miller, N Kingsbury, Image denoising using derotated complex wavelet coefficients. IEEE Trans. Image Process. 17(9), 1500–1511 (2008).

19. T Tasdizen, in ICIP 15th IEEE International Conference on Image Processing. Principal components for non-local means image denoising (IEEESan Diego, 2008), pp. 1728–1731.
20. A Beck, M Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Trans. Image Process. 18(11), 2419–2434 (2009).
21. H Liu, R Xiong, J Zhang, W Gao, in IEEE Conference on Computer Vision and Pattern Recognition. Image denoising via adaptive soft-thresholding based on non-local samples (IEEEBoston, 2015), pp. 484–492.
22. J Bai, X Feng, Fractional-order anisotropic diffusion for image denoising. IEEE Trans. Image Process. 16(10), 2492–2502 (2007).
23. M Elad, M Aharon, Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15(12), 3736–3745 (2006).
24. M Protter, M Elad, Image sequence denoising via sparse and redundant representations. IEEE Trans. Image Process. 18(1), 27–35 (2009).
25. T Tasdizen, Principal neighborhood dictionaries for nonlocal means image denoising. IEEE Trans. Image Process. 18(12), 2649–2660 (2009).
26. J Mairal, F Bach, J Ponce, G Sapiro, A Zisserman, in 2009 IEEE 12th International Conference on Computer Vision. Non-local sparse models for image restoration (IEEEKyoto, 2009), pp. 2272–2279.
27. S Parrilli, M Poderico, CV Angelino, L Verdoliva, A nonlocal SAR image denoising algorithm based on LLMMSE wavelet shrinkage. IEEE Trans. Geosci. Remote Sensing. 50(2), 606–616 (2012).
28. H Yu, L Zhao, H Wang, Image denoising using trivariate shrinkage filter in the wavelet domain and joint bilateral filter in the spatial domain. IEEE Trans. Image Process. 18(10), 2364–2369 (2009).
29. L Yaroslavsky, M Eden, Fundamentals of digital optics (Birkhauser Boston, Boston, 1996).
30. L Yaroslavsky, in Proceedings of wavelet applications in signal and image processing IV. Local adaptive image restoration and enhancement with the use of dft and dct in a running window (SPIE, the International Society for Optics and PhotonicsDenver, 1996), pp. 1–13.
31. F Luisier, T Blu, M Unser, A new sure approach to image denoising: interscale orthonormal wavelet thresholding. IEEE Trans. Image Process. 16(3), 593–606 (2007).
32. R Eslami, H Radha, Translation-invariant contourlet transform and its application to image denoising. IEEE Trans. Image Process. 15(11), 3362–3374 (2006).
33. M Zhang, BK Gunturk, Multiresolution bilateral filtering for image denoising. IEEE Trans.Image Process. 17(12), 2324–2333 (2008).
34. R Yan, L Shao, Y Liu, Nonlocal hierarchical dictionary learning using wavelets for image denoising. IEEE Trans. Image Process. 22(12), 4689–4698 (2013).
35. Lei Zhang, Xin Li, and David Zhang," Image Denoising and Zooming under the LMMSE Framework", IET_IP 2010.
36. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Imagequality assessment: From error visibility to structural similarity,"IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612,Apr. 2004.