

The Summary of Video Denoising Method

Zhijie Lin^{1,2, a}, Xiaohua Li^{3, b}, Zhijun Sun^{4, c}

¹ *Zhejiang University of Science and Technology*

² *The Department of Computer Science and Technology, Zhejiang University*

³ *The Information Center of Zhejiang University*

⁴ *The Center of Digital Cultural Heritage of Dunhuang Research Institute*

^a*bytelin@qq.com*, ^b*xiaohua@zju.edu.cn*, ^c*sunfoto@qq.com*

Abstract

The video remains noisy for low cost acquirement equipment at high sensitivities. The issue of removing video noise is still of acute and in fact growing very importance. In this paper, three approaches are introduced. Each method has its merit and advantage, the comparison to each other is done, and The effectiveness of algorithm is also validated in various experiments and each method compared favorably against other state-of-art algorithms.

Keywords: video noise; summary; low rank matrix completion; Multi-View denoising; 3D transform

Introduction

With current advances in camera sensor, the image/video is relatively free noise for high-fidelity digital cameras at low sensitivities, but it remains noisy for low cost cameras at high sensitivities, e.g., low light condition, high speed rate and high ISO setting. The issue of removing video image noise is still of acute and in fact growing very importance with the prevalence of webcams and mobile phone cameras. General speaking, video data tend to be more noisy than single image due to high speed capturing rate of video acquirement equipment. Video images denoising focus on efficiently removing noise from all image frames of a video by using information in both spatial and temporal domains. Such an hybrid approach is more optimal than independently applying a single-image denoising method on each image frame of the video, since algorithm could utilize relational data among

several image frame, as there exist high temporal redundancies in a video compared to a single image. Most existing video denoising algorithms rely on a statistical model of image noise, such as i.i.d. Gaussian noise, which is often violated in practice. For example, five major sources of image noise with different statistical distributions have been identified in [1]: fixed pattern noise, amplifier noise, photon shot noise, dark current noise and quantization noise. As a result, the performance of most existing denoising techniques will severely degrade when applied on those real noisy images with noises from multiple sources. In this paper, three approaches are introduced. Each method has its merit and advantage, the comparison to each other is done, and the effectiveness of algorithm is also validated in various experiments and each method compared favorably against other state-of-art algorithms.

Overview of Denoising Method

1. Multiple View Image Denoising[17]

This approach acquires pinhole images using many pinhole cameras[17]. The cameras can be distributed spatially to monitor a common scene, or compactly assembled as a camera array[17]. Each camera uses a small aperture and short exposure to ensure minimal optical defocus and motion blur[17]. Under such camera settings, the incoming light is very weak and the images are extremely noisy[17]. The algorithm casts pinhole imaging as a denoising problem and seek to restore all the pinhole images by jointly removing noise in different viewpoints[17]. Using multi-view images for noise reduction has a following advantage: pixel correspondence from one image to all other images is determined by its single depth map[17]. This advantage contrasts with video denoising, where motion between frames in general has many more degrees of freedom. The algorithm uses it for finding similar image patches in multi-view denoising[17]. Specifically, algorithm is built upon the recent development in image denoising literature, where similar image patches are grouped together and “collaboratively” filtered to reduce noise. When considering whether a pair of patches in one image is similar or not, the similarity between corresponding patches in all other views

using depth estimation is simultaneously considered[17]. This depth-guided patch matching improves patch grouping accuracy and substantially boosts denoising performance. The main contributions of multi-view denoising include[17]:

Depth-guided denoising: Using depth estimation as a constraint, multi-view method is able to group similar image patches in the presence of large noise and exploit data redundancy across views for noise removal[17].

Removing signal-dependent noise: In low-light conditions, photon noise is manifest whose variance depends on its mean, which uses the principal component analysis and tensor analysis to remove such noise[17].

Adaptive noise reduction: For both PCA and tensor analysis, an effective scheme is proposed to automatically choose dimensionalities in a way that is adaptive to the complexity of image structures in the patches[17].

Tolerance to depth imperfection: multi-view method is based on a probabilistic formulation that marginalizes depth maps as hidden variables and therefore does not require perfect depth estimation. From an application perspective, approach does not require any change in camera optics or image detectors[17]. All it uses is a set of pinhole cameras, such as those equipped on cell phones[17]. Such flexibilities make this method applicable to places that can only take miniaturized cameras with simple optics, such as low-power video surveillance networks, portable camera arrays, and multi-camera laparoscopy[17].

2. Sparse 3-D Transform-Domain denoising[18]

The sparse 3D transform domain denoising is based on an enhanced sparse representation in transform domain[18]. The enhancement of the sparsity is achieved by grouping similar 2-D image fragments (e.g., blocks) into 3-D data arrays which we call “groups.” Collaborative filtering is a special procedure developed to deal with these 3-D groups, which is realized using the three successive steps: 3-D transformation of a group, shrinkage of the transform spectrum, and inverse 3-D transformation[18]. The result is a 3-D estimate that consists of the jointly filtered grouped image blocks. By attenuating the noise, the collaborative filtering reveals even the finest details shared by grouped blocks and,

at the same time, it preserves the essential unique features of each individual block. The filtered blocks are then returned to their original positions[18]. Because these blocks are overlapping, for each pixel, many different estimates which need to be combined is obtained. Aggregation is a particular averaging procedure which is exploited to take advantage of this redundancy[18]. A significant improvement is obtained by a specially developed collaborative Wiener filtering. An algorithm based on this novel denoising strategy and its efficient implementation are presented in full detail[18]; In this algorithm, the grouping is realized by block-matching and the collaborative filtering is accomplished by shrinkage in a 3-D transform domain[18]. The used image fragments are square blocks of fixed size. The general procedure carried out in the algorithm is as follows. The input noisy image is processed by successively extracting reference blocks from it and for each such block[18]:

- find blocks that are similar to the reference one (blockmatching) and stack them together to form a 3-D array (group) [18];
- perform collaborative filtering of the group and return the obtained 2-D estimates of all grouped blocks to their original locations[18].

After processing all reference blocks, the obtained block estimates can overlap, and, thus, there are multiple estimates for each pixel. This general procedure is implemented in two different forms to compose a two-step algorithm. This algorithm includes basic estimate and final estimate[18]. Basic estimate consists of block-wise estimate and grouping. For each block in the noisy image, do the following. Find blocks that are similar to the currently processed one and then stack them together in a 3-D array (group) [18]. Apply a 3-D transform to the formed group, attenuate the noise by hard-thresholding of the transform coefficients, invert the 3-D transform to produce estimates of all grouped blocks, and return the estimates of the blocks to their original positions[18]. Compute the basic estimate of the true-image by weighted averaging all of the obtained block-wise estimates that are overlapping[18].

Final estimate mainly Use the basic estimate to perform improved grouping and collaborative Wiener one using the energy spectrum of the basic estimate as the

true (pilot) energy spectrum[18]. Produce estimates of all grouped blocks by applying the inverse 3-D transform on the filtered coefficients and return the estimates of the blocks to their original positions[18]. Compute a final estimate of the true-image by aggregating all of the obtained local estimates using a weighted average[18].

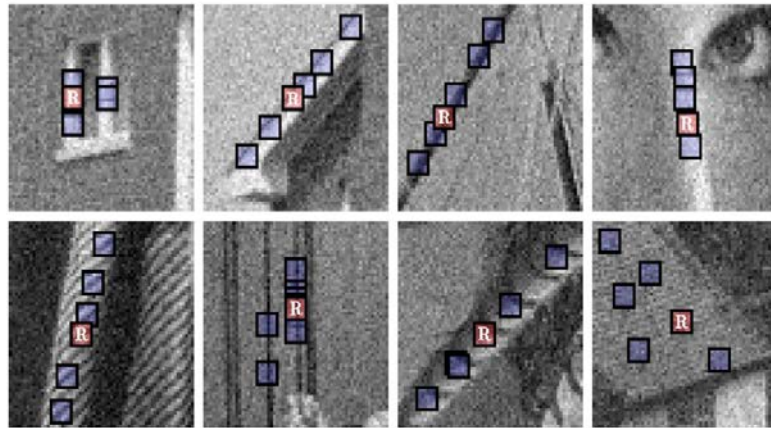


Figure 1. Illustration of grouping blocks from noisy natural images corrupted by white Gaussian noise with standard deviation 15 and zero mean. Each fragment shows a reference block marked with “R” and a few of the blocks matched to it[17].

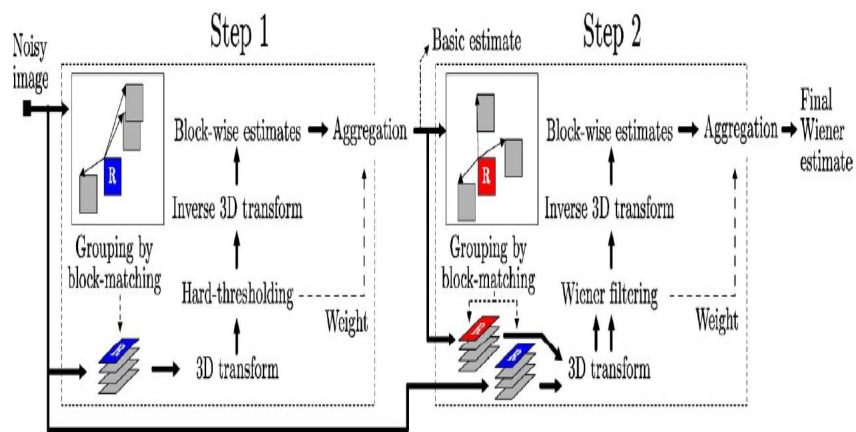


Figure 2. Flowchart of the proposed image denoising algorithm. The operations surrounded by dashed lines are repeated for each processed block (marked with “R”) [17].

3. Low rank matrix completion denoising[19]

Low rank matrix completion denoising is a new patch-based video denoising algorithm capable of removing serious mixed noise from the video data. By grouping similar patches in both spatial and temporal domain, the problem of removing mixed noise as a low-rank matrix completion problem is formulated, which leads to a denoising scheme without strong assumptions on the statistical properties of noise. The resulting nuclear norm related minimization problem can be efficiently solved by many recently developed methods. The robustness and effectiveness of our proposed denoising algorithm on removing mixed noise. a robust video denoising algorithm capable of removing mixed noise from image sequences. The method is built upon the same methodology “grouping and collaboratively filtering” as many patch-based methods (e.g. [3, 12, 13, 14]) do. Different from existing methods, the algorithm is derived with minimal assumptions on the statistical properties of image noise. The basic idea is to convert the problem of removing noise from the stack of matched patches to a low rank matrix completion problem, which can be efficiently solved by minimizing the nuclear norm of the matrix with linear constraints. It is shown in the experiments that our low rank matrix completion based approach can efficiently remove complex noise mixed from multiple statistical distributions.

Conclusion

In this paper, three approaches are introduced. Each method has its merit and advantage, the comparison to each other is done, and The effectiveness of algorithm is also validated in various experiments and each method compared favorably against other state-of-art algorithms. This work is partially supported by Zhejiang Provincial projects (2014C33069), and the National Key Technology R&D Program projects (2013BAH27F01, 2013BAH27F02, 2012BAH43F06).

References

- [1]. G. Healey and R. Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. PAMI*, 16(3):267–276, 1994.
- [2]. S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *CVPR*, pages 860–867, 2005.
- [3]. A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Simulation*, 4, 2005.
- [4]. M. Elad and M. Aharon. Image denoising via learned dictionaries and sparse representation. In *CVPR*, 2006.
- [5]. S. Lyu and E. P. Simoncelli. Statistical modeling of images with fields of gaussian scale mixtures. In *NIPS*, 2006.
- [6]. K. Dabov, R. Foi, V. Katkovnik, K. Egiazarian, and S. Member. Image denoising by sparse 3d transform-domain collaborative filtering. *TIP*, 16:2007, 2007.
- [7]. M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning gaussian conditional random fields for low-level vision. In *CVPR*, pages 1–8, 2007.
- [8]. A. Foi, V. Katkovnik, and K. Egiazarian. Pointwise shapeadaptive dct for high-quality denoising and deblocking of grayscale and color images. *TIP*, 16(5):1395–1411, 2007.
- [9]. E. P. Bennett and L. McMillan. Video enhancement using perpixel virtual exposures. In *SIGGRAPH*, 2005.
- [10]. K. Dabov, A. Foi, and K. Egiazarian. Video denoising by sparse 3d transform-domain collaborative filtering. In *Proc. 15th European Signal Processing Conference*, 2007.
- [11]. J. Chen and C.-K. Tang. Spatio-temporal markov random field for video denoising. In *CVPR*, 2007.
- [12]. A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multisc. Model. Simulat.*, 4(2):490–530, 05.

- [13]. K. Dabov, A. Foi, and K. Egiazarian. Video denoising by sparse 3d transform-domain collaborative filtering. In Proc. 15th European Signal Processing Conference, 2007.
- [14]. K. Dabov, V. K. R. Foi, K. Egiazarian, and S. Member. Image denoising by sparse 3d transform-domain collaborative filtering. IEEE Trans. Image Processing, 16, 2007.
- [15]. L. Zhang, S. Vaddadi, H. Jin, and S. K. Nayar. Multiple view image denoising. In ICCV, 2009.
- [16]. M. Okutomi and T. Kanade. A multiple-baseline stereo. TPAMI, 15(4):353–363, 1993.
- [17]. Li Zhang, Sundeep Vaddadi, Hailin Jin, Shree K. Nayar. Multiple view image denoising. In CVPR, 2009. p1542-1549.
- [18]. Kostadin Dabov. Image denoising by sparse 3D transform domain collaborative filtering. IEEE Transactions on image processing, vol 16, no. 8, p2080-2095. august 2007.
- [19]. Hui Ji, Chaoqiang Liu, Zuowei Shen and Yuhong Xu. Robust video denoising using low rank matrix completion. In CVPR 2010, p1791-1798.