SCIENCE CHINA
Information Sciences

• HIGHLIGHT •

January 2018, Vol. 61 018103:1–018103:3
doi: 10.1007/s11432-017-9200-6

Robust video denoising with sparse and dense noise modelings

Guiping SHEN[1,2], Zhi HAN[1*], Xi'ai CHEN[1,2] & Yandong TANG[1]

[1]*State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang* 110016*, China;*
[2]*University of Chinese Academy of Sciences, Beijing* 100049*, China*

Received 22 March 2017/Accepted 21 June 2017/Published online 15 November 2017

Citation Shen G P, Han Z, Chen X A, et al. Robust video denoising with sparse and dense noise modelings. Sci China Inf Sci, 2018, 61(1): 018103, doi: 10.1007/s11432-017-9200-6

Videos can be contaminated by noise even when captured by high-quality cameras. Because video data has both spatial and temporal redundancies, low rank factorization has been developed. Originally, most denoising methods relied on a single statistical distribution to model noise, such as Gaussian distribution [1, 2]. Ji et al. [3] proposed a low rank matrix completion (LRMC) relying on a minimal assumption. Meng and Cao et al. [4, 5] proposed a low rank matrix factorization problem with the Mixture of Gaussian (MoG) noise model. These algorithms are optimal for noises with continuous distributions. More recently, another type of noise has received growing attention. This type of noise follows discrete distributions such as outliers. Wright et al. [6] used a robust Principal Component Analysis (PCA) method to recover a latent low rank matrix. Okutomi et al. [7] adopted the robust l_1 norm as the measurement to video denoising. Therefore, the noise appears as a combination of continuous and sparse forms.

To make the denoising algorithm more robust to varied situations, the main approach of this study is to use both the MoG and l_1 norm to model mixed noises, where MoG is adopted for modeling dense noises and the l_1 norm is adopted to model sparse noises. The proposed algorithm can therefore remove noise without the need for any assumptions on the statistical properties of noise. It is an extension to the research presented in our previous conference paper [8]. The main differences are the inclusion of the additional l_1 norm, and a thorough experimental evaluation.

Preprocessing. Movements often happen in the foreground or the background, so the simple alignment of video frames does not have a good low rank structure. We adopt the patch-based method for preprocessing to address this problem.

In the first step, we simultaneously search for the most similar patch block of pixels in the current and the consecutive frames. Considering the current ith patch block $p_{i,j,k}$ in the kth frame, which is centered on pixel j, we take MAD as the similarity criterion, which is described as

$$\mathrm{MAD}_{x_0,y_0}(x,y) = \frac{1}{WL}\sum_{i=0}^{W-1}\sum_{j=0}^{L-1}|p_k(x_0+i, y_0+j) - p_{k+1}(x_0+x+i, y_0+y+j)|, \quad (1)$$

(x_0, y_0) denotes the current block and (x_0+x, y_0+y) is the block to be compared with. The patch block size is $W \times L$ (The influence of the variations of W and L on performance is discussed in Appendix B. For all the experiments reported in this article, they were both chosen to be 8.). The sample interval was set to 4×4 pixels and the maximum displacement was 6 pixels. In the second step, we placed each similar patch into a column

* Corresponding author (email: hanzhi@sia.com)
The authors declare that they have no conflict of interest.

of a matrix $P_{j,k}$. The video denoising problem is written in a matrix form as

$$P_{j,k} = Q_{j,k} + N_{j,k}, \tag{2}$$

where $Q_{j,k}$ denotes the underlying clear patch matrix, and $N_{j,k}$ is the noise of the corresponding block.

Noise modeling. Many existing denoising algorithms perform well on either dense ones or sparse ones. We propose a mixed noise model to simultaneously deal with these two types of noise distributions. By using UV-decomposition $Q_{j,k} = U(V)^{\mathrm{T}}$ in (2), each element $p_{r,d}$ in $P_{j,k}$ can be written as

$$p_{r,d} = u^r(v^d)^{\mathrm{T}} + \varepsilon_{r,d}, \tag{3}$$

where u^r denotes the rth row of U, v^d is the dth row of V, and $\varepsilon_{r,d}$ is the video noise at $p_{r,d}$. The unknown video noise distribution $\varepsilon_{r,d}$ is divided into two parts: dense distribution $c_{r,d}$ and sparse distribution $s_{r,d}$, i.e.,

$$\varepsilon_{r,d} = c_{r,d} + s_{r,d}. \tag{4}$$

We use the MoG distribution to fit each dense component $c_{r,d}$ in (4). The probability of $p_{r,d}$ can be expressed as

$$p\left(p_{r,d}\left|u^r, v^d\right., \Pi, \Delta\right) = \sum_{n=1}^{N} \pi_n p(p_{r,d}|n), \tag{5}$$

where $p(p_{r,d}|n) = N(p_{r,d}|u^r(v^d)^{\mathrm{T}} + s_{r,d}, \sigma_n^2)$ is the Gaussian distribution, and the mixing proportion is π_n with $\pi_n \geqslant 0$ and $\sum_{n=1}^{N} \pi_n = 1$, $\Pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$, $\Delta = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$.

The l_0 norm is introduced for sparse expression, but it is an NP-hard problem. Here, the sparse noise $s_{r,d}$ is measured with the l_1 norm, which is the optimal convex approximation of the l_0 norm.

Denoising model. The video denoising model is a low rank matrix decomposition. It can be expressed as the minimization optimization problem:

$$\min_Q \|Q\|_* \quad \text{s.t } \|P - Q\|_{L_p}^2 \leqslant \varepsilon, \tag{6}$$

where $\|\cdot\|_*$ denotes the nuclear norm, $\|\cdot\|_{L_p}$ denotes the L_p norm, and ε is the default of allowable error. The aim of video denoising is to get the latent matrix Q by optimizing (6). Continuous noise C is modeled with mixed Gaussians and sparse noise S is with l_1 norm modeling. Hence, the video denoising problem can be formulated as

$$\min ||Q||_* + ||S||_1 \quad \text{s.t. } \|P - Q - S - C\| \leqslant \sigma. \tag{7}$$

After adding the regularization items, it is equivalent to minimizing the following equation:

$$\begin{aligned} L(Q,S,C,\mu) &= ||Q||_* + ||S||_1 \\ &+ \sum_{n=1}^{N} \pi_n p\left(C|n\right) + \mu \left\|P - Q - S - C\right\|_F^2. \end{aligned} \tag{8}$$

Algorithm. The proposed algorithm iterates through three steps until reaching convergence. The parameters of sparse part are updated in the first step and the parameters of the mixed Gaussians are updated simultaneously by the following two steps of the EM algorithm.

The parameters U, V, Π, Δ are alternatively updated in a closed form as

$$\pi_n = \frac{S_n^q}{D_n^q}, \tag{9}$$

$$\sigma_n^2 = \frac{1}{S_n^q} \sum_{r,d\in\Omega} \left[S_n^q \frac{\left(p_{r,d} - u^r\left(v^d\right)^{\mathrm{T}} - s_{r,d}\right)^2}{2\pi\sigma_n^2} \right]. \tag{10}$$

The dimension of $\{q_{r,d,n}\}$ is D_n^q and $S_n^q = \sum_{r,d\in\Omega} q_{r,d,n}$. The posterior responsibility of mixed Gaussians $q_{r,d,n}$ is calculated as

$$\begin{aligned} q_{r,d,n} &= p(z_{r,d,n}\left|p_{r,d}; U, V, \Pi, \Delta\right.) \\ &= \frac{\pi_n N\left(p_{r,d}\left|u^r\left(v^d\right)^{\mathrm{T}} + s_{r,d}, \sigma_n^2\right.\right)}{\sum_{n=1}^{N} \pi_n N\left(p_{r,d}\left|u^r(v^d)^{\mathrm{T}} + s_{r,d}, \sigma_n^2\right.\right)}. \end{aligned} \tag{11}$$

The number of Gaussian components (denoted by N) needs to be adjusted according to the variance of the Gaussian distribution, which can refer to our previous paper [8]. The low rank matrix components U and V are formulated using

$$\begin{aligned} &\sum_{r,d\in\Omega} \sum_{n=1}^{N} q_{r,d,n} \left(-\frac{\left(p_{r,d} - u^r\left(v^d\right)^{\mathrm{T}} - s_{r,d}\right)^2}{2\pi\sigma_n^2} \right) \\ &= -\sum_{r,d\in\Omega} \left(\sum_{n=1}^{N} \frac{q_{r,d,n}}{2\pi\sigma_n^2} \right) \left(p_{r,d} - u^r\left(v^d\right)^{\mathrm{T}} - s_{r,d}\right)^2 \\ &= -\left\|W \odot \left(P_{j,k} - UV^{\mathrm{T}} - S\right)\right\|_{L_2}^2, \end{aligned} \tag{12}$$

where W is a weighting indicator matrix. Several existing algorithms can be adopted to optimize (12), such as the Augmented Lagrange algorithm.

The matching blocks obtained in the preprocessing step exhibit some overlap. In order to avoid an artificial factor at each boundary between patches, the value of each pixel is calculated by the average of overlapping pixels at the same location.

More details of the algorithm can be found in Appendix A.

Experiments and conclusion. The competitive methods include cVBM3D [1], MoG [8], LRMC [3] and WNNM [9]. We synthesized mixed noises with

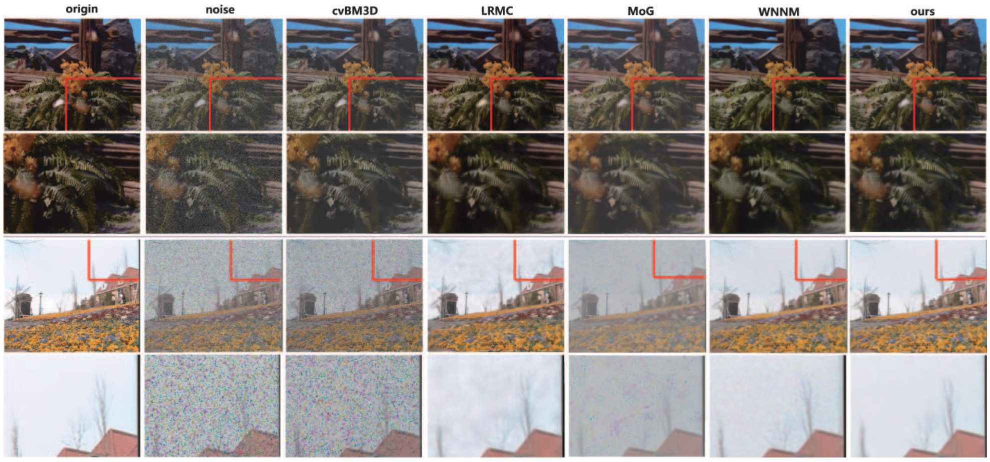

Figure 1 (Color online) Experiments on videos with mixed noise: Gaussian noise variance ($\delta_1 = 20$), Poisson noise parameter ($p = 10$), salt and pepper(10%). In each group, the upper line and the lower line correspond to global and detailed (zoomed-in) results, respectively.

two categories: continuous and sparse. The continuous noise was generated from Gaussian distributions and Poisson distributions, and the sparse noise from salt & pepper noise. Figure 1 shows the qualitative results in which the comparisons of some detailed sections are provided in the zoomed-in image in the red box. As shown, our method gives the best visual effect on both the global view and the detailed view. More experiments and quantitative results are provided in Appendix B. For better visual effects, we show the results of applying our proposed algorithm to videos[1)].

In this article, we proposed a new method for video denoising. By using the spatial and temporal redundancies of videos, the new model is based on a low-rank matrix decomposition technique. To make the model more robust to various noises, we introduced MoG for modeling dense noises and the l_1 norm for sparse noises. The algorithm therefore shows the advantage of dealing with a wide range of noise types without requiring prior statistical knowledge of the video noise. Compared with other competitive video denoising algorithms, the proposed algorithm performs more effectively and robustly, and excels in preserving the local structure of video frames.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61303168). The authors also thank the support by Youth Innovation Promotion Association CAS.

Supporting information Appendixes A and B. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Boracchi G, Foi A, Egiazarian K. Video denoising using separable 4D nonlocal spatiotemporal transforms. Int Soc Opt Eng, 2011, 7870: 1–12
2 Sorel M, Bartos M. Fast Bayesian JPEG decompression and denoising with tight frame priors. IEEE Trans Image Process, 2016, 26: 490–501
3 Ji H, Liu C Q, Shen Z W, et al. Robust video denoising using low rank matrix completion. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, 2010. 1791–1798
4 Meng D Y, Fernando D L T. Robust matrix factorization with unknown noise. In: Proceedings of IEEE International Conference on Computer Vision, Sydney, 2013. 1337–1344
5 Cao X Y, Chen Y, Zhao Q, et al. Low-rank matrix factorization under general mixture noise distributions. In: Proceedings of IEEE International Conference on Computer Vision, Santiago, 2015. 1493–1501
6 Wright J, Peng Y G, Ma Y, et al. Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization. In: Proceedings of the 22nd International Conference on Neural Information Processing Systems, Vancouver, 2009. 2080–2088
7 Okutomi M, Yan S, Sugimoto S, et al. Practical low-rank matrix approximation under robust L1-norm. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, 2012. 1410–1417
8 Shen G P, Han Z, Tang Y D. Robust video denoising by low-rank decomposition and modeling noises with mixture of Gaussian. In: Proceedings of IEEE International Conference on Robotics and Biomimetics (ROBIO), Bali, 2014. 2226–2231
9 Gu S H, Zhang L, Zuo W M, et al. Weighted Nuclear Norm Minimization with Application to Image Denoising. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014. 2862–2869

1) http://vision.sia.cn/our%20team/Hanzhi-homepage/vision-ZhiHan(English).html.