ICIC Express Letters
Part B: Applications
Volume **7**, Number **2**, February **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 345–350

MIXED GAUSSIAN-RAIN VIDEO NOISE REMOVAL BASED ON INEXACT RPCA AND MCA IN WIRELESS MULTIMEDIA SENSOR NETWORKS

Hui Luo[1], Yanli Zhu[1], Gong Li[1] and Manlin Luo[2]

[1]School of Information Engineering
East China Jiaotong University
No. 808, Shuanggang East Street, Nanchang 330013, P. R. China
{ lh_jxnc; zyl_jxnc }@163.com

[2]School of Physics and Micro Electronic
Hunan University
No. 2, Lushan South Road, Changsha 410000, P. R. China
lml_jxnc@163.com

Received August 2015; accepted November 2015

Abstract. *Rapid developments of multimedia applications have led to the necessity of stringent quality videos in Wireless Multimedia Sensor Networks (WMSN). However, inclement weather conditions and illumination variance often lead that WMSN videos contain many noises, especially, the mixed Gaussian-rain noise. Thus, WMSN video denoising has been recently studied extensively. In this paper, the mixed Gaussian-rain noise removal method based on inexact robust principle component analysis (Inexact RPCA) and morphological component analysis (MCA) in WMSN is proposed. Firstly, WMSN video is decomposed into three parts: low-rank part, Gaussian noise part, and sparse part via Inexact RPCA. Secondly, sparse part is decomposed into rain component and nonrain component by performing dictionary learning and sparse coding based on MCA algorithm. Lastly, the noise-removed version of WMSN video can be gained by integrating the nonrain component of sparse part with the low-rank part of WMSN video. Experimental results show that the performance of the proposed approach is competitive, qualitative, and has greater ability to retain video feature information.*
Keywords: Mixed Gaussian-rain noise removal, Inexact robust principle component analysis, Morphological component analysis, Dictionary learning

1. **Introduction.** Wireless Multimedia Sensor Networks (WMSN) can undertake monitoring task independently, widely used in various kinds of fields [1]. However, the visual quality of WMSN video is seriously affected by inclement weather and illumination variance, which will add multi-noises on the video [2]. In practice, the mixed Gaussian-rain noise is one of the most common noises, which impairs the visibility or interpretability of the video. Therefore, it is imperative to study new video denoising method to remove the mixed Gaussian-rain noise in WMSN video.

Recently, many classical algorithms for video denoising have been proposed, such as video block matching and 3D filtering (VBM3D) algorithm [3], non-local means (NLM) algorithm and the denoising method based on sparse coding [4]. These video denoising algorithms showed impressive results on suppressing Gaussian noise. However, in the presence of sparse noise such as rain streaks, impulse noise, the performance of these denoising algorithms noticeably decreases.

Inexact RPCA [5-7], an extended version of RPCA, can effectively eliminate the mixed Gaussian-sparse noise in WMSN video. However, RPCA [8] and Inexact RPCA take most feature information of moving objects and sparse noise in video as sparse part without the ability of further distinguishing them. In [9] the robust temporal-spatial decomposition

(RTSD) model is proposed to effectively remove impulse noise while retaining most feature information of moving objects via combining RPCA with TV-l_1 model. However, when the video is corrupted by mixed Gaussian-rain noise, the denoising effects of RTSD model is unsatisfactory as RPCA fails to separate Gaussian noise from video and TV-l_1 model cannot distinguish rain streaks from feature information of moving objects. Meantime, the MCA-based decomposition method, a decomposition method based on sparse representation, has excellent effect in the separation of rain streaks and feature information by utilizing the morphological diversity of different components contained in the video [10]. Therefore, inspired by this, we combine Inexact RPCA with MCA to suppress mixed Gaussian-rain noise in WMSN video.

The major contribution of this paper is twofold: 1) our method is among the first to combine Inexact RPCA with MCA to remove mixed Gaussian-rain noise while preserving feature information in video; 2) we do the dictionary learning every n frames and provide an extended dictionary and update it every n frames to enrich the dictionary. The first stage and the second stage of our proposed method are introduced in Section 2 and Section 3, respectively. Section 4 gives a performance comparison of other methods and the proposed method. Section 5 summarizes the paper.

2. **The First Stage: WMSN Video Decomposition via Inexact RPCA.**

In this paper, we proposed a video denoising method based on Inexact RPCA and MCA. Firstly, we employ Inexact RPCA to decompose the video into three parts. The specific process is as follows.

Let us consider a video with N frames, denoted as $\tilde{Q}(1), \cdots, \tilde{Q}(N)$, where the i-th frame $\tilde{Q}(i) \in R^{K_1 \times K_2}$. For the sake of simplicity, each frame is reshaped as a column vector, for example, $\tilde{Q}(i)$ is reshaped as $q\,(i) \in R^{K_1 K_2 \times 1}$. These vectors are combined into a matrix $Q \in R^{K_1 K_2 \times N}$, with $q\,(i)$ being the i-th column. We will decompose Q into three parts: the low-rank part, Gaussian noise part and the sparse part via Inexact RPCA, as follows:

$$\min_{L,S} \|L\|_* + \lambda' \|S\|_1\,, \quad \text{s.t.}\ \|Q - L - S\|_F \leq \delta \tag{1}$$

Here, L is the low-rank part which represents the temporal-spatially correlated part, S is the sparse part, λ' is a suitable regularization parameter, in our approach, $\lambda' = 1/\sqrt{\max\,(K_1, K_2)}$, where K_1, K_2 are the number of rows and columns of the matrix, $\delta > 0$ is the Gaussian noise level, Q represents the matrix of input video, and $\|\cdot\|_*$, $\|\cdot\|_F$ are nuclear norm and Frobenius norm.

In our approach, we use the accelerated proximal gradient method (APG) to solve Equation (1). The APG method has been demonstrated to be efficient in solving various regularized convex optimization problems in compressed sensing, machine learning, and control [11].

3. **The Second Stage: MCA-based Sparse Part of Video Frames Decomposition.**

After the first stage, the Gaussian noise is removed, the low-rank part L and sparse part S are decomposed from the video. In this stage, we perform the MCA-based decomposition method to further decompose S into rain component and nonrain component. The detailed method shall be elaborated below.

3.1. **Preprocessing and problem formulation.**

For the sparse part of video frame, $S_{(i)}$, $i = 1, 2, \cdots, N$, in the preprocessing step, we apply a bilateral filter to roughly decomposing $S_{(i)}$ into low frequency part $S_{(i)LF}$ and high frequency part $S_{(i)HF}$. Then, we learn dictionary $D_{(i)HF}$ based on the training exemplar patches extracted from $S_{(i)HF}$ to further decompose $S_{(i)HF}$, where $D_{(i)HF}$ can be further divided into two subdictionaries $D_{(i)HF-N}$ and $D_{(i)HF-R}$ ($D_{(i)HF} = [D_{(i)HF-N}|D_{(i)HF-R}]$), for representing the nonrain component and rain component of $S_{(i)HF}$, respectively. As a result, we formulate the problem of rain streaks removal for $S_{(i)}$ as sparse coding-based image decomposition problem as follows:

$$\min_{\theta^k_{(i)HF}} \left\|f^k_{(i)HF} - D_{(i)HF}\alpha^k_{(i)HF}\right\|_2^2 \quad \text{s.t.} \left\|\alpha^k_{(i)HF}\right\|_0 \le L' \tag{2}$$

where $f^k_{(i)HF}$ represents the k-th patch extracted from $S_{(i)HF}$, $k = 1, 2, \ldots, P$. $\alpha^k_{(i)HF}$ are the sparse coefficients of $f^k_{(i)HF}$ with respect to $D_{(i)HF}$, and L' denotes the sparsity or maximum number of nonzero coefficients of $\alpha^k_{(i)HF}$. To solve Equation (2), we apply the efficient OMP provided in [12].

3.2. **Dictionary learning and partition.** We do a dictionary learning every n frames in our method. In this step, we extract from $S_{(i)HF}$ a set of overlapping patches $w^k_{(i)}$ as the training exemplars for learning dictionary $D_{(i)HF}$. We formulate the dictionary learning problem as

$$\min_{D_{(i)HF},\alpha^k_{(i)}} \sum_{k=1}^{P} \left(\frac{1}{2}\left\|w^k_{(i)} - D_{(i)HF}\alpha^k_{(i)}\right\|_2^2 + \lambda\left\|\alpha^k_{(i)}\right\|_1\right) \tag{3}$$

where $\alpha^k_{(i)}$ denotes the sparse coefficients of $w^k_{(i)}$ with respect to $D_{(i)HF}$, and λ is a regularization parameter. In this paper, we solve Equation (3) to obtain $D_{(i)HF}$ with an efficient online dictionary learning algorithm proposed in [12], which scales up gracefully to large data sets with millions of training samples, and extends naturally to various matrix factorization formulations.

In the proposed method, we utilize the histogram of oriented gradients (HOG) descriptor to describe each atom in $D_{(i)HF}$. Then we apply the K-means algorithm to classify all of the atoms in $D_{(i)HF}$ into two clusters D_1 and D_2 based on their HOG feature descriptors. The following procedure is to identify which cluster consists of rain atoms and which cluster consists of nonrain atoms. First, we calculate the variance of gradient direction for each atom $d_{t\beta}$, $\beta = 1, 2, \ldots, N_t$, in cluster D_t as $F_{t\beta}$, where N_t denotes the number of atoms in D_t, $t = 1, 2$. Then, we calculate the mean of $F_{t\beta}$ for each cluster D_t as $MF_{t\beta}$. Based on the fact that the edge directions of rain streaks in an atom are usually consistent, i.e., the variance of gradient direction for a rain atom should be small, we identify the cluster with the smaller $MF_{t\beta}$ as rain subdictionary $D_{(i)HF-R}$ and the other one as nonrain subdictionary $D_{(i)HF-N}$.

On the other hand, an extended dictionary D_E is provided to further improve decomposition performance. We first learn extended dictionary D_E by collecting a set of exemplar patches from the HF parts of the sparse part in some training nonrain frames, and then update it every n frames by collecting training exemplar patches from the HF part of the sparse part in the adjacent rain-removed frames. Subsequently, we integrate D_E with $D_{(i)HF-N}$ to form the final nonrain subdictionary $D'_{(i)HF-N}$. $D_{(i)HF-R}$ obtained by dictionary partition is the final rain streaks subdictionary.

3.3. **Removal of rain streaks.** Based on the two dictionaries $D_{(i)HF-R}$ and $D'_{(i)HF-N}$, we perform sparse coding for each patch $f^k_{(i)HF}$ extracted from $S_{(i)HF}$ via minimization of Equation (2) to find its sparse coefficients $\tilde{\alpha}^k_{(i)HF}$. We perform sparse coding only once for each patch $f^k_{(i)HF}$. Then, each reconstructed patch $f^k_{(i)HF}$ can be used to recover either nonrain component $S^N_{(i)HF}$ or rain component $S^R_{(i)HF}$ based on the sparse coefficients as follows. We set the coefficients corresponding to $D'_{(i)HF-N}$ in $\tilde{\alpha}^k_{(i)HF}$ to zeros to obtain $\tilde{\alpha}^k_{(i)HF-R}$, whereas the coefficients corresponding to $D_{(i)HF-R}$ in $\tilde{\alpha}^k_{(i)HF}$ to zeros to obtain $\tilde{\alpha}^k_{(i)HF-N}$. Therefore, each patch $f^k_{(i)HF}$ can be re-expressed as:

$$\tilde{f}^k_{(i)HF-N} = D'_{(i)HF-N} \times \tilde{\alpha}^k_{(i)HF-N} \tag{4}$$

$$\tilde{f}^k_{(i)HF-R} = D_{(i)HF-R} \times \tilde{\alpha}^k_{(i)HF-R} \tag{5}$$

where $\widetilde{f}^k_{(i)HF-N}$ and $\widetilde{f}^k_{(i)HF-R}$ are used to recover $S^N_{(i)HF}$ and $S^R_{(i)HF}$, respectively, by averaging the pixel values in overlapping regions. Then, the nonrain component of sparse part in the i-th frame $S_{1(i)}$ and the noise-removed version of the i-th frame $V_{(i)}$ can be obtained via:

$$S_{1(i)} = S^N_{(i)HF} + S_{(i)LF} \tag{6}$$

$$V_{(i)} = S_{1(i)} + L_{(i)} \tag{7}$$

where $L_{(i)}$ is the low-rank part of the i-th frame. The noise-removed version of every video frame can be easily acquired via implementing the above operation for every frame, to achieve the final denoised video. In conclusion, the process of the proposed method can be shown in Figure 1.

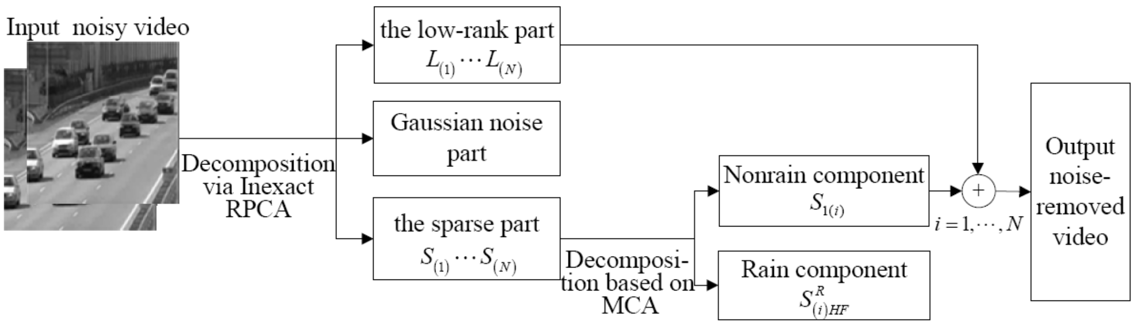

Figure 1. The process of the proposed method

4. **Experiment Simulation and Results Analysis.** For validating the performance of proposed method, we compare our method with the VBM3D method and the RTSD model. These methods are tested upon two WMSN videos: "Bus (352×288)" and "Road (352×288)", all of which are corrupted by mixed Gaussian-rain noise. The standard deviation of Gaussian noise δ varies from 5 to 20, and the percentage of pixels corrupted by rain streaks γ varies from 10% to 30%.

In our experiment, $N = 50$ video frames are used. The parameter λ' in (1) is set to $1/\sqrt{(352 \times 288)}$. The regularization parameter λ in (3) is set to 0.15. The patch size, number of training patches, dictionary size, and the number of training iterations are set to 16×16, 337×273, 1024 and 100, respectively. L' in (2) is set to 10. In the extended dictionary D_E learning step, the patch size, dictionary size, and number of training iterations are also set to 16×16, 1024 and 200, respectively. The dictionary is updated every n frames and n is set to 5. Firstly, Gaussian noise with $\sigma = 10$ and rain with 20% are added to the videos, and then under the same experimental condition, VBM3D, RTSD model and the proposed method are used to denoise for corrupted videos, respectively. The key frames of Bus and Road are shown in Figures 2(a) and 2(f), respectively. The different denoising results are shown in Figure 2. Table 1 tabulates average PSNR (Peak Signal to Noise Ratio) values of the results for video denoising in the presence of mixed Gaussian-rain noise with different levels.

Through analyzing Table 1 and Figure 2, We can find our method not only depresses noise effectively, but also preserves the feature information better. From Table 1, it can be observed that the proposed method achieves better performance in terms of average PSNR to the two comparisons. No matter from the visual effect or PSNR, it is noticeable that the proposed algorithm is superior to VBM3D and RTSD model. Therefore, the proposed method is suitable for the WMSN video denoising as compared to the other algorithms.

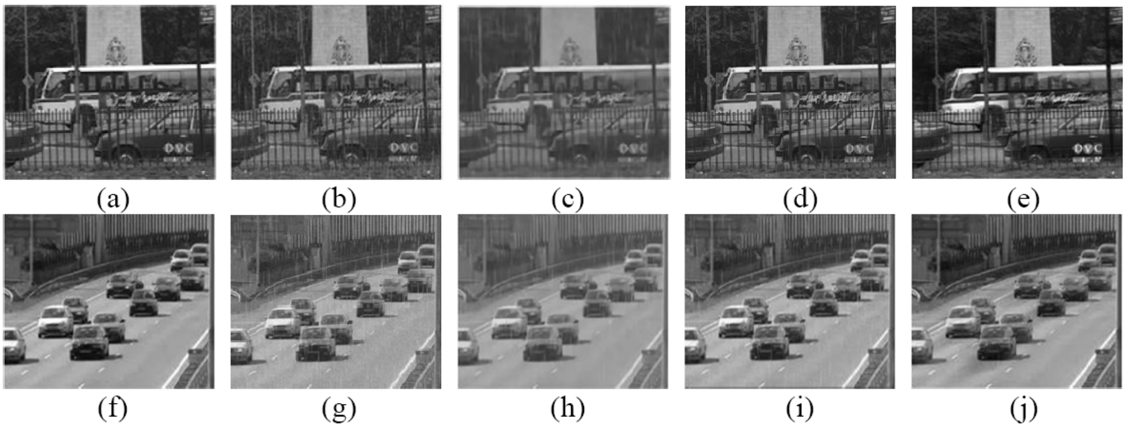

Figure 2. Visual comparison of denoising results for Bus and Road of different methods: (a) and (f) original data, (b) and (g) noisy input ($\sigma = 10$, $\gamma = 20\%$), (c) and (h) VBM3D method, (d) and (i) RTSD model, (e) and (j) proposed method

Table 1. Average PSNR values of the results for video denoising

Video	(δ, γ)	Noisy frame	VBM3D	RTSD model	Proposed
	$(5, 10\%)$	22.5480	22.7512	23.2416	25.7403
Bus	$(10, 20\%)$	21.1902	21.7666	22.4632	23.4212
	$(20, 30\%)$	18.4298	19.3241	20.3975	21.5179
	$(5, 10\%)$	22.6734	22.8905	24.9327	28.4723
Road	$(10, 20\%)$	20.3140	20.7878	23.9401	25.9810
	$(20, 30\%)$	17.2835	17.9165	20.0609	21.7361

5. **Conclusions.** In this paper, a mixed Gaussian-rain noise removal method based on Inexact RPCA and MCA is proposed, which consists of two stages, namely WMSN video decomposition via Inexact RPCA and MCA-based sparse part of video frames decomposition. Compared with some state-of-the-art methods, our method shows greater ability to retain video feature information, and achieves better visual quality. Therefore, the proposed method is suitable for removing the mixed Gaussian-rain noise in WMSN video, especially in video with moving objects. For future work, the performance of our method may be further improved by enhancing the sparse coding, dictionary learning, and dictionary partitioning processes.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (No. 61261040).

REFERENCES

[1] I. F. Akyildiz, T. Melodia and K. R. Chowdhury, A survey on wireless multimedia sensor networks, *Comput. Networks*, vol.51, no.4, pp.921-960, 2007.
[2] P. Barnum, S. Narasimhan and T. Kanade, Analysis of rain and snow in frequency space, *Int. J. Comput. Vis.*, vol.86, no.23, pp.256-274, 2010.
[3] K Dabov, A. Foi and K. Egiazarian, Video denoising by sparse 3D transform-domain collaborative filtering, *Proc. of the 15th European Signal Processing Conference*, vol.1, no.2, p.7, 2007.
[4] J. Mairal, F. Bach, J. Ponce, G. Sapiro and A. Zisserman, Non-local sparse models for image restoration, *Proc. of IEEE International Conference on Computer Vision*, pp.2272-2279, 2009.
[5] Z. Zhou, X. Li, J. Wright, E. Candes and Y. Ma, Stable principal component pursuit, *Proc. of IEEE Int. Symp. Inf. Theory*, pp.1518-1522, 2010.
[6] H. Ji, S. Huang and Z. Shen, Robust video restoration by joint sparse and low rank matrix approximation, *SIAM Journal on Imaging Sciences*, vol.4, no.4, pp.1122-1142, 2011.

[7] Z. Wang, H. Li and Q. Ling, Mixed Gaussian-impulse video noise removal via temporal-spatial decomposition, *IEEE International Symposium on Circuits and Systems*, pp.1851-1854, 2012.

[8] E. J. Candès, X. Li and Y. Ma, Robust principal component analysis, *Journal of the ACM*, vol.58, no.3, p.11, 2011.

[9] Z. Wang, H. Li et al., Robust temporal-spatial decomposition and its applications in video processing, *IEEE Trans. Circuits and Systems for Video Technology*, vol.23, no.3, pp.387-400, 2013.

[10] L. W. Kang, C. W. Lin and Y. H. Fu, Automatic single-image-based rain streaks removal via image decomposition, *IEEE Trans. Image Processing*, vol.21, no.4, pp.1742-1755, 2013.

[11] Z. Shen, K. C. Toh et al., An accelerated proximal gradient algorithm for frame-based image restoration via the balanced approach, *SIAM Journal on Imaging Sciences*, vol.4, no.2, pp.573-596, 2011.

[12] J. Mairal, F. Bach and J. Ponce, Online learning for matrix factorization and sparse coding, *The Journal of Machine Learning Research*, no.11, pp.19-60, 2010.