Mixed noise removal based on a novel non-parametric Bayesian sparse outlier model

Peixian Zhuang, Yue Huang, Delu Zeng, Xinghao Ding*

Fujian Key Laboratory of Sensing and Computing for Smart City, School of Information Science and Engineering, Xiamen University, China

ARTICLE INFO

Article history:
Received 19 September 2014
Received in revised form 28 September 2015
Accepted 28 September 2015
Communicated by: Shiliang Sun
Available online 21 October 2015

Keywords:
Mixed noise removal
Non-parametric Bayesian model
Spike-slab
Automatic parameter estimation

ABSTRACT

We develop a novel non-parametric Bayesian sparse outlier model for the problem of mixed noise removal. Based on the assumptions of sparse data and isolated outliers, the proposed model is considered for decomposing the observed data into three components of ideal data, Gaussian noise and outlier noise. Then the spike-slab prior is employed for outlier noise and sparse coefficients of ideal data. The proposed method can automatically infer noise statistics (e.g., Gaussian noise variance) from the training data without changing model hyper-parameter settings. It is also robust to initialization without using adaptive median filter as in other denoising methods. Experimental results demonstrate proposed model can achieve better objective and subjective performances on mixed noise removal than other state-of-the-art methods.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

There has been significant interest in the problem of mixed noise removal based on the fact that image acquisition or transmission can be easily contaminated with different types of noise [1]. Fig. 1 shows the probability density function of mixed noise with Gaussian and salt-and-pepper noise, and this distribution has a longer heavy-tail than the Gaussian distribution. The larger value of salt-and-pepper noise ratio is, the longer of heavy-tail is in the corresponding distribution. And these results demonstrate that the commonly-encountered outlier noise will introduce more serious unreliable effects on image restoration. Therefore, mixed noise removal has received much research attention in these fields of image processing and computer vision [2,3].

For Gaussian noise removal, total variation (TV) method becomes one of most popular alternatives. However, the results of image detail structures (e.g. texture) can be over-smoothed or removed by TV [4]. To better preserve detail texture structures, other methods are proposed to improve the denoising performance by using the non-local prior of similar image patches [5,6]. A new algorithm called KSVD is an iterative method for alternative optimization, which alternates between sparse coding of the sample data based on the current dictionary and a process of updating the dictionary for fitting the training data [7]. But this method cannot achieve better performance if the noise statistics is unknown in advance. Beta process of factor analysis (BPFA) [8–10] can automatically infer the noise variance of noisy images by using non-parametric Bayesian technique of Bernoulli-Beta process. And block matching and 3D filtering (BM3D) is proposed to enhance the sparsity and thus achieves better performance than other leading methods on Gaussian noise removal [11]. However, all above methods consider the noise assumption as Gaussian distribution, therefore, they do not work well for mixed noise. For impulse noise removal, median filter is the classical method for its denoising ability and effective computation. Recently, various modified versions of median filter are proposed, e.g. adaptive median filter (AMF) [12], center weighted median filter [13] and multi-state median filter [14]. The general idea of these filters is that the location of noisy pixels can be directly detected and replaced by other pixel values within the neighboring windows, while the remained clean pixels keep unchanged. However, these methods cannot preserve local texture structures since the noisy pixels in the vicinity of edges are directly replaced by their median values according to the neighboring pixels [15]. Total variation based adaptively weighted L1 method (TVAWL1) [16] is formulated by weighted data-fidelity term and mixed exponential model, and it achieves better restoration performance under high density of impulse noise. However, TV method may over-smooth detail texture or edge structures in the recovered image. Note that these methods of impulse noise removal may not work for mixed noise with Gaussian noise incorporated.

In order to tackle the denoising problem of mixed noise with Gaussian and impulse noise, a few methods have been developed to make up the above problem. One of the well-known methods is

* Corresponding author.
E-mail address: dxh@xmu.edu.cn (X. Ding).

http://dx.doi.org/10.1016/j.neucom.2015.09.095
0925-2312/© 2015 Elsevier B.V. All rights reserved.
two-phase denoising algorithm, which firstly identifies the location of outlier noise by using median-type filters and then removes Gaussian and residual noise by TV or modified versions for image restoration [17,18]. However, the drawbacks of AMF and TV lead to over-smoothing or removing image detail structures. A novel robust Bayesian RVM algorithm in [19] has been proposed to achieve better performance of Gaussian and outlier noise removal. However, this cannot be well fit to various kinds of data with the fixed kernel regression technique. For the modification of KSVD, a new weighted model of combining KSVD with Gaussian mixture model (WKSVD) achieves better results of mixed noise removal. However, this method, with heavily depending on using median-type filter, is sensitive to the initialization, and may easily oversmooth image detail structures for texture images [20]. A new minimization method, using median-type filtering in combination with modified KSVD algorithm, is proposed to achieve state-of-the-art denoising performance, but the computational complexity is high since lots of iterative operations are needed for sparse coding and dictionary learning [2]. Henceforth, it is desirable to develop an efficient model for mixed noise removal.

For handling mixed noise removal, we develop a novel non-parametric Bayesian model with three contributions as follows: 1) The spike-slab prior [21–23] will be employed for the sparse coefficients and outlier noise. 2) The statistics of Gaussian noise variance can be automatically learned from the training data without tuning model hyper-parameter settings. 3) It is robust to the initialization without the requirement of median-type filter compared with other denoising methods. Experimental results finally validate that the proposed algorithm can outperform other state-of-the-art methods on mixed noise removal.

The remainder of the paper is organized as follows. In Section 2, we firstly present novel model construction. Next the inference process of proposed model will be detailed in Section 3. Then we validate the performance of proposed algorithm on numerous examples in Section 4. Conclusion and discussion are given in the final section.

2. Novel model construction

We use the following notation: $I_p$ is the identity matrix with the size of $P \times P$, and $N$ is denoted as the number of image pixels. Let $x \in \mathbb{R}^P$ be a $\sqrt{N} \times \sqrt{N}$ desired recovery image in a vectorized form, and $y$ be the noisy image. Let $n$ and $s$ be Gaussian noise and outlier noise, respectively. And all of them are the same size as $x$. For dictionary learning, $D = (d_1, d_2, ..., d_K) \in \mathbb{R}^{P \times K}$ is an overcompleted dictionary with the dimension of $P$ and column number of $N$, where $d_m \in \mathbb{R}^P$ with $m = 1, ..., K$. Let $R_i$ be the $i$-th patch extraction matrix, which is a $P \times N$ matrix of all zeros except for a one in each row that extracts a vectorized $\sqrt{P} \times \sqrt{P}$ patch from the image. Let $x_i = R_i x \in \mathbb{R}^P$ be the $i$-th patch with the size of $\sqrt{P} \times \sqrt{P}$, and $y_i = R_i y \in \mathbb{R}^P$ be the $i$-th patch extracted from $y$. The matrix $\alpha$ is sparse coefficients for all patches extracted from the image, which is the size of $K \times N$. And $\alpha_i \in \mathbb{R}^{P \times 1}$ as a column vector of sparse coefficients matrix is for the $i$-th patch. For noise notations, let $\sigma$ be the standard deviation of Gaussian noise, and $r$ be the ratio of salt-and-pepper noise. $n_i \in \mathbb{R}^{P \times 1}$ is in the vectorized form for the $i$-th patch of Gaussian noise, and $s_i \in \mathbb{R}^{P \times 1}$ is the similar vector for the $i$-th patch of outlier noise.

Motivated by the superposition model of [24], we decompose the noisy image $y$ corrupted by mixed noise into three components: the ideal term $x$, Gaussian noise term $n$ and outlier noise term $s$, thereby the formulation of the proposed model can be expressed as follows:

$$y = x + n + s$$ (1)

The final objective $x$ needs to be recovered from the observed image $y$, which is an ill-posed inverse problem. And this can be solved by employing the regularization scheme of image priors. Recent studies have shown that image reconstruction of patch-level is an effective means for sparse representation, in addition to reducing computing time and saving memory. Therefore, we take advantage of the scheme of overlapped image patches to reconstruct the corresponding image. Based on compressed sensing and sparse representation, $X = [x_1, x_2, ..., x_N] \in \mathbb{R}^{P \times N}$ can be represented by the dictionary $D$ and sparse coefficients $\alpha$, namely $X = Da$. And each patch $x_i = Da_i$, is with $i = 1, ..., N$. Then the model of each patch $y_i$ can be formulated as

$$y_i = Da_i + n_i + s_i$$ (2)

We will impose the corresponding prior distribution on every parameter in the proposed model. Then the model is detailed as follows: we assume the dictionary matrix $D$ of independent and identically distributed random variables, which will be adaptively learned to fit the training data. The dictionary element $d_{ki}$ as the $k$-th column of $D$, is drawn from a zero-mean Gaussian distribution. Then initial prior values can be expressed as

$$d_{ki} \sim N(0, P^{-1}I_p)$$ (3)
Based on the sparse property of image patch and outlier noise, we impose the spike-slab prior [25–27] on the sparse coefficients \( \alpha_i \) and outlier noise \( s_i \). And they can be formulated below:

\[
\alpha_i \sim \prod_{k=1}^{K} ((1-\pi_k)\delta + \pi_k N(0, \gamma_{\alpha}^{-1}))
\]

(4)

\[
\pi_k \sim \text{Beta}(a_0, b_0)
\]

(5)

\[
\gamma_{\alpha} \sim \Gamma(c_0, d_0)
\]

(6)

\[
s_i \sim \prod_{p=1}^{P} ((1-q)\delta + qN(0, \gamma_s^{-1}))
\]

(7)

\[
q \sim \text{Beta}(m_0, n_0)
\]

(8)

\[
\gamma_s \sim \Gamma(e_0, f_0)
\]

(9)

where \( \delta \) is the delta function to enhance the sparsity of the spike part with a high probability. \( \pi_k \) is drawn from a beta distribution with two hyper-parameters of \( a_0 \) and \( b_0 \), and it is associated with the probability of \( d_k \) being used or not to represent sample data. And all patches share the same \( \pi_k \). For the independent and sparse property of outlier noise, \( q \) is also derived from the beta distribution with parameters of \( m_0 \) and \( n_0 \), and it corresponds to the probability of current pixel as outlier noise or not. Non-informative gamma hyper-priors are typically imposed on \( \gamma_{\alpha} \) and \( \gamma_s \), which are used to infer the precision or inverse variance. And \( c_0, d_0, e_0 \) and \( f_0 \) are the corresponding gamma parameters. In term of the noise \( n_i \), we assume a Gaussian distribution with zero mean and variance \( \gamma_{\nu}^{-1} \), and this can be expressed as:

\[
n_i \sim N(0, \gamma_{\nu}^{-1})
\]

(10)

\[
\gamma_{\nu} \sim \Gamma(g_0, h_0)
\]

(11)

where \( \gamma_{\nu} \) is drawn from a Gamma distribution with hyper-parameters of \( g_0 \) and \( h_0 \), and it can be used to infer the Gaussian variance.

3. Model inference

In the proposed model, every parameter will be imposed on a prior distribution. And then the posterior distribution can be inferred from the training data and Bayesian inference. In this section, we derive a Markov Chain Monte Carlo (MCMC) sampling algorithm [28–30] to compute the approximation inference of the proposed model. And the data samples are drawn from the following conditional distributions:

**Step 1. Sampling \( D \):** The posterior distribution of \( d_k \) obeys the Gaussian distribution, and its corresponding mean and variance can be expressed as follows:

\[
p(d_k | \cdot) \sim N(\mu_{d_k}, \Sigma_{d_k})
\]

(12)

\[
\Sigma_{d_k} = (\mathbf{I} + \gamma_{\nu} \sum_{i=1}^{N} \alpha_i^2)^{-1}
\]

(13)

\[
\mu_{d_k} = \gamma_{\nu} \Sigma_{d_k} \sum_{i=1}^{N} \alpha_i \tilde{y}_i^{-k}
\]

(14)

where \( \tilde{y}_i^{-k} = y_i - D\alpha - s_i + d_k \).

**Step 2. Sampling \( \alpha \):** The posterior distribution of \( \alpha_i \) can be written as:

\[
p(\alpha_i | \cdot) \sim \prod_{k=1}^{K} ((1-\pi_k)\delta + \pi_k N(\mu_{\alpha_i}, \Sigma_{\alpha_i}))
\]

(15)

\[
\tilde{\Sigma}_{\alpha_i} = (\gamma_{\alpha} + \gamma_{\nu} \alpha_i d_k^{-1})^{-1}
\]

\[
\tilde{\mu}_{\alpha_i} = \gamma_{\nu} \tilde{\Sigma}_{\alpha_i} (\alpha_i \tilde{y}_i^{-k})
\]

(16)

\[
\omega_{\alpha_i} = \frac{\pi_k}{\Gamma(\alpha_i + c_0, d_0)}
\]

(17)

**Step 3. Sampling \( \pi_k \):**

\[
p(\pi_k | \cdot) \sim \text{Beta}\left(a_0 + \sum_{i=1}^{N} 1(\alpha_i = 0), b_0 + \sum_{i=1}^{N} 1(\alpha_i = 1)\right)
\]

(18)

**Step 4. Sampling \( \gamma_s \):**

\[
p(\gamma_s | \cdot) \sim \Gamma\left(c_0 + \frac{1}{2} |N|, d_0 + \frac{1}{2} \sum_{i=1}^{N} |\alpha_i| \right)
\]

(19)

 Then we exclude the ideal data before the calculation of outlier noise parameters, and thus let

\[
\tilde{\gamma}_s = y_i - D\alpha - s_i
\]

(20)

**Step 5. Sampling \( s \):** The posterior distribution of \( s_p \) is described below:

\[
p(s_p | \cdot) \sim \text{Beta}\left(n_0 + \sum_{p=1}^{M} 1(s_p \neq 0), n_0 + \sum_{p=1}^{M} 1(s_p = 0)\right)
\]

(21)

**Step 6. Sampling \( q \):**

\[
p(q | \cdot) \sim \text{Beta}\left(n_0 + \sum_{p=1}^{M} 1(s_p \neq 0), n_0 + \sum_{p=1}^{M} 1(s_p = 0)\right)
\]

(22)

**Step 7. Sampling \( \gamma_s \):**

\[
p(p | \cdot) \sim \Gamma\left(e_0 + \frac{1}{2} |N|, f_0 + \frac{1}{2} \sum_{i=1}^{N} |\gamma_s|\right)
\]

(23)

**Step 8. Sampling \( \gamma_s \):**

\[
p(p | \cdot) \sim \Gamma\left(g_0 + \frac{1}{2} |N|, h_0 + \frac{1}{2} \sum_{i=1}^{N} |\gamma_s|\right)
\]

(24)

\[\sum_{p=1}^{M} 1(s_p = 0)\]

**4. Experimental validation**

In this section, experimental results are reported to validate the performance of the proposed algorithm. Experimental parameters will be set as follows: image resolution is 256 \times 256, the image patches are overlapped with a shift of one pixel, and each patch size is 8 \times 8. The dimension of the dictionary atom is 64, while the number of dictionary atoms is set to 256. Hyper-parameters of \( c_0, d_0, e_0, f_0, g_0 \) and \( h_0 \) are all set to \( 10^{-5} \), along the lines suggested in [31]. \( a_0, b_0, m_0 \) and \( n_0 \) are set to 1, 0.8, 0.1 and 0.9, respectively. Peak Signal to Noise Ratio (PSNR) is measured as the objective
evaluation, and image visual quality is utilized as the subjective evaluation. We firstly add different noise levels into images of Fig. 2(a) and (b), and compare proposed method with KSVD [7], BPFA [8], BM3D [11], AMF [12], TVAWL1 [16] and WKSVD [20] on mixed noise removal respectively. Then the real brain image in [20] is used to validate the performance of the noise removal in real-world application by proposed method in comparison with other algorithms. The state-of-the-art Weighted Encoding with Sparse Nonlocal Regularization (WESNR) [1] and WKSVD [20] are employed for comparison to validate the proposed performance on texture images (Fig. 2(c) and (d)).

In this experiment, the original image of Barbara is corrupted by mixed noise of Gaussian and salt-and-pepper noise with $\sigma = 10$ and $r = 10\%$. As the denoising results shown in Fig. 3(b)–(d), BM3D,
KSVD and BPFA have worse performances, due to their assumptions that the noise distribution is Gaussian. AMF is able to remove the salt-and-pepper noise effectively, (Fig. 3(e)), however, it is still limited in dealing with Gaussian noise. TVAWL1 and WKSVD (Fig. 3 (f) and (g)) have better performances than previous methods since they consider Gaussian and salt-and-pepper noise simultaneously. Some noises are still existed in the results of TVAWL1, while WKSVD over-smooths the structures. As shown in Fig. 3(h), proposed method outperforms than other methods based on the strong ability of removing mixed noise and preserving the structures simultaneously. Next, we demonstrate the dictionaries learned from Barbara with mixed noise $\sigma = 20$ and $r = 20\%$. As the illustrations of Fig. 4(a) and (b), KSVD and BPFA have no structure atoms in the dictionaries, which leads to the worse denoising ability. However, the dictionary of the proposed approach (Fig. 4(d)) has much more structure atoms than that of WKSVD (Fig. 4(c)),
KSVD and BPFA. Moreover, the dictionary learned by proposed method is much clear and can describe image features more effectively, which results in better denoising performance.

Fig. 5 shows PSNR curves of the denoising results for different methods with different noise levels. Barbara and House are corrupted by salt-and-pepper noise ratio increasing from 10% to 30% under Gaussian noise \( \sigma \) fixed as 10, 20 and 30, respectively. PSNR plots of proposed method are always higher than that of other methods on images with various mixed noises, which demonstrate the better performances and robustness in comparison. In addition, the proposed model can automatically infer noise statistics. As in Table 1, estimations of various Gaussian noise variances from different algorithms on Barbara are proposed. Noise estimation by BPFA is relatively worse due to the assumption that the noise model is Gaussian distribution. Though KSVD is able to achieve better estimation performances than TAWAL1, it strongly depends on the initialization based on AMF in mixed noise removal. Compared with other methods, proposed method is able to estimate the noise variances without any initialization.

Denoising performances across different methods on the real brain MR noisy image are illustrated in Fig. 6 [20]. KSVD has the worst effect with some blurs and contrast distortions in the recovery image. BPFA, BM3D, AMF and TVAWL1 have relative poor visual quality with some residual noise remained in the denoised images. WKSVD over-smoothes image detail structures and produces some blurs. However, the proposed method can achieve the better visual quality and preserve image detail structures than other methods on the actual noise removal.

To validate the effectiveness of the proposed method and make comparison with WESNR and WKSVD, we use texture images of Fig. 2(c) and (d) corrupted by different noise levels for denoising test. From Fig. 7(c) and (g), it can be observed that WKSVD fails in recovering desired texture images when mixed noise is large (e.g., Gaussian noise standard deviation \( \sigma = 20 \) and salt-and-pepper noise ratio \( r = 20\% \)). As shown in Fig. 7(d), WESNR achieves the worst performance of visual quality when mixed noise is small (e.g., Gaussian noise standard deviation \( \sigma = 10 \) and salt-and-pepper noise ratio \( r = 10\% \)). However, the proposed method gains about 1.5 dB improvement over WESNR and 5.8 dB improvement over WKSVD when the added noise is mixed Gaussian noise \( \sigma = 10 \) with salt-and-pepper noise \( r = 10\% \). And proposed approach can also achieve about 0.37 dB improvement over WESNR and 4.8 dB improvement over WKSVD with mixed noise of Gaussian noise \( \sigma = 20 \) with salt-and-pepper noise \( r = 20\% \). As the results in Fig. 7, it can be concluded that proposed method achieves best PSNR values and preserves detail texture structures better than other methods across various mixed noises.

### 5. Conclusion and discussion

In this work, a novel non-parametric Bayesian model has been presented for mixed noise removal. With consideration of the sparsity for data and outliers, the observed data is composed of three components including ideal data, Gaussian noise and outlier noise. Then the spike-slab sparse prior has been used to fit sparse coefficients and outlier noise. The proposed non-parametric Bayesian model is able to automatically estimate Gaussian noise variance without changing model hyper-parameter settings during different noises. And the proposed method, compared with other denoising methods, is robust to initialization without using adaptive median filter. Finally simulation results demonstrate that

<table>
<thead>
<tr>
<th>Mixed noise</th>
<th>Ground-truth noise variance</th>
<th>Proposed</th>
<th>KSVD</th>
<th>KSVD without AMF</th>
<th>BPFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 + 10%</td>
<td>10</td>
<td>10.48</td>
<td>10.37</td>
<td>45.12</td>
<td>42.40</td>
</tr>
<tr>
<td>15 + 15%</td>
<td>15</td>
<td>15.21</td>
<td>14.63</td>
<td>55.11</td>
<td>50.72</td>
</tr>
<tr>
<td>20 + 20%</td>
<td>20</td>
<td>21.09</td>
<td>18.98</td>
<td>64.01</td>
<td>59.68</td>
</tr>
<tr>
<td>25 + 25%</td>
<td>25</td>
<td>26.53</td>
<td>22.97</td>
<td>71.68</td>
<td>69.82</td>
</tr>
<tr>
<td>30 + 30%</td>
<td>30</td>
<td>34.10</td>
<td>24.75</td>
<td>78.16</td>
<td>77.08</td>
</tr>
</tbody>
</table>

Fig. 6. Denoising results of different methods on the real brain MR image.
the proposed model can subjectively and objectively outperform other leading methods for mixed noise removal. For the limitation of proposed model, the proposed method runs slower than WKSVD and WESNR since the iterative operation and cascade calculation of MCMC algorithm are time-consuming. However, the parallel computation of parameters in MCMC can be adapted to reduce the computation time by GPU hardware. And this will be explored in the near future.

Acknowledgments

The project was supported in part by the National Natural Science Foundation of China under Grants 30900328, 61172179, 61013021, 81301278, 61571382 and 61571005, in part by the Fundamental Research Funds for the Central Universities under Grants 20720150169 and 20720150093, in part by the Natural Fundamental Research Funds for the Central Universities under Grants 61103121, 81301278, 61571382 and 61571005, in part by the Science Foundation of China under Grants 30900328, 61172179, 11971082 and 11871005, in part by the Science Foundation of Fujian Province under Grant 2012J05160, and in part by the Research Fund for the Doctoral Program of Higher Education under Grant 20120121120043.

References


Peixian Zhuang received the B.S. from School of Physics and Electrical Engineering, Taishan University, and the M.S. degree from College of Information Science and Engineering, Huqiao University, in 2010 and 2013, respectively. Now he is a Ph.D. candidate from School of Information Science and Technology in Xiamen University. His main research interests include image processing, sparse signal representation and machine learning.

Yue Huang received the B.S. from Department of Electrical Engineering, Xiamen University, and Ph.D. degrees from Department of Biomedical Engineering, Tsinghua University, Beijing, China, in 2005 and 2010, respectively. Since 2010, she has been an assistant professor with the Xiamen University, Xiamen, China. Her main research interests include sparse signal representation, and machine learning.

Delu Zeng received the B.S. and M.S. degrees in applied mathematics and the Ph.D. degree in electronic and information engineering from South China University of Technology, Guangzhou, China, in 2003, 2005, and 2009, respectively. He is now an associate professor in the School of Information Science and Technology in Xiamen University, China. His research interests include partial differential equations, machine learning, image and video processing.

Xinghao Ding received the B.S. and Ph.D. degrees from Department of Precision Instruments, Hefei University of Technology, Hefei, China, in 1998 and 2003, respectively. Since 2006, he has been a Professor with the Xiamen University, Xiamen, China. His main research interests include image processing, sparse signal representation, and machine learning.