Mixed Noise Removal Method Based on Sparse Representation and Dictionary learning: WESNR

Ekta Desai
EC-M.Tech UTU university
Bardoli-394601,
Gujarat India
ektadesai123@gmail.com

Asst.Prof.Manish Nakrani
EC-Department UTU university
Bardoli-394601,
Gujarat,India
manish.nakrani@utu.ac.in

ABSTRACT
Noise removal is the fundamental problem in image processing. Knowledge of Noise Distribution is important in image denoising. Removing mixed noise from an image is since a difficult task as the characteristics of different types of noises are different. The commonly experienced mixed noise is impulse Noise(IN) together mixed with additive White Gaussian noise(AWGN). Various mixed noise removal methods are available but they are detection based methods. These methods first of all find the location of IN pixels and then remove the other noises. But these methods produce many artifacts when the density of mixed noise is high. Here, we introduce a very effective method named as weighted encoding with sparse nonlocal regularization (WESNR). In this method there is not impulse pixel detection and AWGN removal performs separately but it performs this both task in unified framework. In WESNR impulse pixel detection by weighted encoding is done to deal with IN and AWGN simultaneously. Nonlocal self-similarity and sparsity is combined in to a regularization term. WESNR achieves best mixed noise removal performance than any other methods.

Keywords
sparse representation, weighted encoding, regularization, nonlocal means.

1. INTRODUCTION
Digital images may get affected by noise at time of image acquisition and transmission. Noise removal is the most important task in image processing. Most of the noises added to images are represented mainly by two types of noises: additive white Gaussian noise (AWGN) and impulse noise (IN). Most of the operations like image segmentation, image compression and image analysis are done on image so restoration of original image from its noise corrupted observation is very necessary. Impulse noise is represented by replacing a portion of an image pixel with noise values and keeping the remainder unchanged. The pixel of an image which is corrupted by AWGN, pixel value is independently sampled from a zero-mean Gaussian distribution and it is added to the pixel gray level. In past so many papers have been published to remove either IN[1]-[6] or AWGN[7]-[13]. But as mixture of IN and AWGN are the commonly experienced mixed noise so removal of that is necessary. Many mixed noise removal methods [14]-[20] are also available but they don’t give efficient results when the mixed noise is strong. Impulse noise is mostly occurred due to malfunctioning pixels in camera, faulty memory locations in devices or bit errors in transmission[14]. There are mostly two types of impulse noises are encountered in practice that are salt and pepper Impulse noise(SPIN) and random valued impulse noise(RVIN). The most popular Nonlinear filters the median filter[1] with good denoising power, very simple implementation, and high efficiency which exploits the rank order information of pixel intensities within a filtering window and replaces the center pixel with the median. Even though there is one disadvantage of the median filter is that sometimes it removes the desirable details and it makes the denoised image looks unnatural. Various improvements of median filter are proposed such as the weighted median filter[2], the multisate median (MSM) filter, the center-weighted median filter, and the stack filter. However these filters do not distinguish whether the current pixel is noise pixel or not. This would inevitably alter the intensities and remove the image details contributed from uncorrupted pixels so the image quality will degrade. To deal with this problem, one solution is there to introduce a noise-detection mechanism to identify the corrupted pixels and to leave the uncorrupted pixels unaltered. Incorporating such noise-detection mechanism in to the median filter that so-called the switching median filter is proposed. After that better noise removal methods with different kinds of noise detectors have been introduced such as MSM[3] filter, tristate median (TRI) filter, adaptive center-weighted median (ACWM) filter[4], the pixel-wise median absolute deviation (PWMAD) filter, the adaptive switching median (ASWM) filter, a directional weighted median (DWM) filter for IN removal. For AWGN, traditional linear filters such as Gaussian filters can smooth noise efficiently but they will over-smooth the image edges at the same time. To overcome this problem, nonlinear filtering methods have been introduced. Bilateral filter (BF)[7] is good at edge preservation which estimates each pixel as the weighted average of the neighboring pixels and the weights are determined by both the intensity similarity and spatial similarity. In the nonlocal means (NLM) filtering[8] method similar pixels in an image can be spatially far from each other and each pixel is estimated as the weighted average of all its similar pixels in the image and the weights are determined by the similarity between them. Grouping the similar patches into a matrix and to apply the principal component analysis (PCA) is used to remove AWGN which is called LPG-PCA algorithm achieves...
very good edge preservation. Nowadays the sparse representation and dictionary learning based methods have given the efficient results in image restoration. The K-SVD[10] initiates the study of learning a dictionary from natural images for AWGN removal. Sparse representation and nonlocal self-similarity regularization together has lead to state-of-the-art AWGN removal performance[12]. For removing mixed noise the median based rank ordered mean filter can remove impulse noise effectively but when it applied to images with the mixed noise it produces inefficient result. The trilateral filter[14] incorporated with rank-order absolute difference (ROAD) mechanism for impulse noise and Gaussian noise removal. But ROAD provide false value when half of the pixels in the image are corrupted. The sorted quadrant median vector switching bilateral filter (SQMV-SBF)[15] which is the modification of bilateral filter incorporated with the SQMV for impulse noise detection. It removes the mixed noise by switching the bilateral filter between the Gaussian noise and the impulse noise. The HDR filter removes the mixed noise by the kernel regression framework. There are so many mixed noise removal methods are existing but they are detection based methods and they involve two sequential steps. First they detect the IN pixels and then remove the noise. This two-phase strategy will become less effective when the AWGN or IN is strong. Here we propose a simple and though effective method weighted encoding with sparse nonlocal regularization (WESNR)[21]. In that there is no explicit impulse pixel detection and we encode each noise-corrupted patch over a pre-learned dictionary to remove the IN and AWGN simultaneously in a soft impulse pixel detection manner. The conventional $l_2$-norm data fidelity term which is used to characterize the Gaussian distributed data fitting residual is not suitable to suppress the mixed noise with complex non-Gaussian distribution. Here in WESNR the mixed noise is removed by weighting the encoding residual so that the final encoding residual will tend to follow Gaussian distribution. Weighted encoding and sparse nonlocal regularization are unified into a variational framework.

2. WEIGHTED ENCODING WITH SPARSE NONLOCAL REGULARIZATION (WESNR)

2.1 Adding Of a Mixed Noise

Let take $x$ an image and $x_{i,j}$ its pixel at location $(i, j)$. Suppose $y$ is the noisy observation of $x$. There are mostly two types of impulse noises are the salt and pepper and Random valued impulse Noise Denote by $[d_{noise, white}]$ the dynamic range of $y$ and For additive white Gaussian noise (AWGN) each noisy pixel $y_{i,j}$ in $y$ is modeled as $y_{i,j} + v_{i,j}$. The corrupted images by mixture of IN and AWGN is shown in fig 2 & fig 3.

2.2 Noise Removal Model

Let take one image $x \in \mathbb{R}^n$ and $x_i = R_ix \in \mathbb{R}^n$ is the stretched vector of an image patch of size $\sqrt{n} \times \sqrt{n}$. $R_i$ is the matrix operator used to extract patch $x_i$ from $x$ at location $i$. We can find an over-complete dictionary based on the sparse representation[22] that is $\mathbf{Phi} = [\phi_1, \phi_2, \ldots, \phi_s]$. $x_i$ over dictionary $\mathbf{Phi}$ can be written as $x_i = \mathbf{Phi}a$ where $a$ is a sparse coding vector which is of few non-zero entities. Noisy observation $y$ is encoded over the dictionary $\mathbf{Phi}$ to obtain the desired $a$. Encoding model for AWGN can be written as follows:

$$\hat{a} = \arg\min_a \|y - \Phi a\|^2 + \lambda R(a) \quad \text{...(1)}$$

Where $R(a)$ is some regularization term and $\alpha$ and $\lambda$ is the regularization parameter. The distribution of corrupted images is generally far from Gaussian and thus the $l_2$-norm data fidelity term in above equation will not lead to a MAP solution. If we modify the data fidelity term so that the residual can be more Gaussian. Using robust estimation theory[24][25] to weight the data fitting residual so that its distribution can be more regular. Let

$$e = [e_1, e_2, \ldots, e_N] = y - \Phi a$$

where $e_i = (y - \Phi a_i)$. And let $e_1, e_2, \ldots, e_N$ are samples. Here we use the robust estimation technique[25] to minimize the following loss:

$$\min \sum_{i=1}^{N} f(e_i) \quad \text{...(2)}$$

Assigning each residual a proper weight the resulting weighted residual will as follows:

$$e_i^w = w_i^{1/2} e_i \quad \text{...(3)}$$

Residuals obtained at the pixels corrupted by AWGN will mostly follows Gaussian distribution so they should be assigned with weights close to 1. The residuals obtained at other pixels corrupted by IN they should be assigned with smaller weights to reduce the heavy tail.

The distribution of residuals $e_i$ and the fitting Gaussian function based on the variance of $e_i$ and the distribution of weighted residuals $w_i^{1/2}e_i$ and the fitting Gaussian function based on the variance of $w_i^{1/2} e_i$ clearly shows that the distribution of weighted residuals is much closer to Gaussian distribution[21]. we get a new loss function $f(e_i) = (w_i^{1/2} e_i)^2$. so, we have new model for mixed noise removal:

$$\hat{a} = \arg\min_a \left\| w_i^{1/2}(y - \Phi a) \right\|^2 + \lambda R(a) \quad \text{...(5)}$$

where $W$ is a diagonal weight matrix. To make the above weighted encoding model more effective for mixed noise removal, some regularization terms $R(a)$ can be used. Two priors are widely used in image denoising that is local sparsity and nonlocal self-similarity (NSS). The local sparsity of encoding coefficients $a$ can be characterized by the $l_1$-norm of $a$ and the Nonlocal self similarity can be characterized by the prediction error of a patch by its similar patches. Therefore, we approximately assume that $y$ follows laplacian distribution, and hence the $l_1$-norm regularization on $a$ could lead to a MAP-like estimation. Finally, the proposed model becomes:

$$\hat{a} = \arg\min_a \left\| w_i^{1/2}(y - \Phi a) \right\|^2 + \lambda ||a - \mu||_1 \quad \text{...(6)}$$

In the above model, the data fidelity term weights the encoding residual, while the regularization term integrates sparsity and NSS priors. We call the proposed model weighted encoding with sparse nonlocal regularization (WESNR). In the WESNR model $W$ is a diagonal weight matrix. Clearly, the pixels corrupted by IN should have small weights to reduce their effect on the encoding of $y$ over $\Phi$, while the weights assigned to uncorrupted pixels should be...
close to 1. In our algorithm, the dictionary $\Phi$ is pre-learned from clean natural images and the pixels corrupted by IN will have big coding residuals. Therefore, the coding residual $e_i$ can be used to guide the setting of weight $W_{il}$, where $W_{il}$ should be inversely proportional to the strength of $e_i$. In order to make the weighted encoding stable and easy to control, we set $W_{il} \in [0, 1]$. One simple and appropriate choice of $W_{il}$ is:

$$W_{il} = \exp(-a*e_i^2)$$

where $a$ is a positive constant to control the decreasing rate of $W_{il}$ w.r.t. $e_i$. With Eq. (1), the pixels corrupted by IN will be adaptively assigned with lower weights to reduce their impact in the process of encoding. Once $W$ is given, the WESNR model above equation becomes an $l_1$-norm sparse coding problem and many existing $l_1$ norm minimization techniques [26][27] can be used to solve it. Here we solve it via the iteratively reweighted scheme for its simplicity. Let $V$ be a diagonal matrix. First we initialize it as an identity matrix, and then in the $(k + 1)^{th}$ iteration each element of $V$ is updated as:

$$V_{il}^{(k+1)} = \frac{\lambda}{((\alpha_i^{(k)} - \mu_i)^2 + \epsilon^2)^{1/2}}$$

where $\epsilon$ is a scalar and $\alpha_i(k)$ is the $i$th element of coding vector $a$ in the $k$th iteration. Then we update $a$ as $\hat{a}$

$$\hat{a}^{(k+1)} = (\varphi^TW\varphi + V^{(k+1)})^{-1}(\varphi^TWy - \varphi^T\varphi\hat{a}) + \mu$$

By iteratively updating $V$ and $a$, the desired $a$ can be efficiently obtained. The convergence of the iteratively reweighted scheme has been proved in [27].

2.3 Algorithm Of WESNR

Algorithm:

Input: Take Corrupted Image $y$; 
Initialize $e$ by Equation $e^{(0)} = y - x^{(0)}$; and then Initialize $W$ By Equation $W_{il} = \exp(-a*e_i^2)$;

Initialize Nonlocal coding Vector $\mu$ to zero and Generate Dictionary $\Phi$;

Loop: Iterate on $k = 1, 2, 3, \ldots K$;
1. First Compute $\alpha^{(k)}$ by $\hat{a}^{(k+1)} = (\varphi^TW\varphi + V^{(k+1)})^{-1}(\varphi^TWy - \varphi^T\varphi\hat{a}) + \mu$;
2. Compute $x^{(k)} = \varphi^T\hat{a}$;
3. Update the nonlocal coding vector $\mu$;
4. Evaluate the residual $e^{(k)} = y - x^{(k)}$;
5. Compute the weights $W$ by $e^{(k)}$ using equation $W_{il} = \exp(-a*e_i^2)$;

End

First of all Determine the Dictionary $\Phi$ for a given Patch so that the given WESNR[21] model can be solved by iteratively updating the $W$ and the coding vector. Updating of $W$ depends on the coding residual $e$. As said in the mixed AWGN and SPIN noise removal methods AMF[5] is widely used to detect SPIN. So in the case of AWGN+SPIN noise removal here we apply AMF to Noisy image $y$ to obtain an initialized image $x^{(0)}$. After that we initialize $e$ as:

$$e^{(0)} = y - x^{(0)}$$

As in the case of AWGN+RVIN+SPIN noise removal AMF cannot be applied to $y$ to initialize $x$ so we initialize $e$ as

$$e^{(0)} = y - \mu; 1$$

Where $\mu_i$ is the mean value of all pixels in Noisy image $y$ and $i$ is a column vector whose elements are all 1. Here we simply use the mean value of $y$ for the initialization of $x$. If once the coding residual $e^{(0)}$ is initialized then $W$ can be initialized. Mostly, WESNR algorithm will terminate in six to twelve iterations[21].

2.4 Generating a Dictionary

The overcomplete dictionary $D$ that leads to sparse representations can either be chosen as a prespecified set of functions or designed by adapting its content to better fit a given set of signals[7]. To choose a prespecified transform matrix is appealing because it is simpler and it leads to sparse and fast algorithms for the evaluation of the sparse representation. Success of such dictionaries in applications depends on how suitable they are to sparsely describe the signals. For the sparse coding and reconstruction of a signal selection of a dictionary is an important issue. In noise removal model above described we assumed that the dictionary $\Phi$ is given. In image restoration the dictionaries learned from image patches have shown effective results. The K-SVD [7] learns an over-complete to process any input patch even though it is not adaptive to the content of the given patch but it is not efficient due to the large number of atoms in the over-complete dictionary. Here in WESNR we adopt the strategy to learn a set of local PCA dictionaries from natural images. We have use one high quality image to train the PCA dictionaries. A number of patches (size: $7 \times 7$) are extracted from the image and they are clustered into 200 clusters by using the K-means clustering algorithm. For every cluster a local PCA dictionary is learned. In the meantime the centroid of each cluster is calculated. For a given image patch the euclidian distance between it and the centroid of each cluster is computed and the PCA dictionary associated with this closest cluster is chosen to encoding the given patch[21]. The selected dictionary which is denoted by $\Phi_i$ is orthogonal, $\mu_i$ for the given patch $x_i$ can be simply computed as $\mu_i = \Phi_i^T x_i$.

3. RESULT

3.1 Setting of a Parameter

The parameter to control the termination of iteration $\tau_i$ to balance the number of iterations we set it to 0.003. The parameter $a$ which is a positive constant that controls the decreasing rate of weights w.r.t. $e$ and we set it to 0.0008. There are two parameters to compute the diagonal matrix $V$: $\lambda$ and $\epsilon$ in equation (9). In this method the sparse nonlocal regularization is used to mainly remove AWGN. In the first loop of the algorithm since the IN is severe so the
block-matching based nonlocal similar patch searching process is not accurate. So the nonlocal regularization is not very helpful and we assign \( \lambda \) a small value (0.0001) to weaken the role of nonlocal regularization. In the second loop IN is largely reduced and thus the nonlocal similar patch searching becomes more accurate. After that we assign \( \lambda \) a large value to remove AWGN. Whenever the standard deviation of AWGN is more than 10 at that time we set \( \lambda = 1 \) else we set \( \lambda = 0.5 \) to suppress AWGN. The parameter \( \varepsilon \) used to increase the numerical stability of computing equation(9) we set it to 

\[
\varepsilon^{(k+1)} = \min(\varepsilon^{(k)}, \text{median}(|\alpha^{(k)} - \mu|))
\]

3.2 Results
Here, we have implemented whole WESNR algorithm and the experimental results are shown in the figures. We take one noise free image then added mixed noise to it where for AWGN we take \( \sigma = 5 \) and for impulse noise \( s = 10\% \). Then image is recovered by adaptive median filter now we conduct the whole WESNR algorithm and then we get the total noise free image which is as shown in fig 5. The PSNR value is 15.42692945 and total time taken for WESNR is 495.566000 s. WESNR algorithm shows very efficient result than any other method.

Fig 1 : Original image
Fig 2 : Image with AWGN noise (\( \sigma = 5 \))
Fig 3 : Image with impulse noise and AWGN (Mixed noise)
Fig 4 : Image recovered by adaptive median filter
Fig 5 : Denoised image by WESNR
4. CONCLUSION

Here we implemented a mixed noise removal method namely weighted encoding with sparse nonlocal regularization (WESNR). Distribution of mixed noise for e.g. additive white Gaussian noise mixed with impulse noise is much more irregular than Gaussian noise alone and often has a heavy tail. So to solve this difficulty we proposed the weighted encoding technique to remove Gaussian noise and impulse noise jointly. We encoded the image patches over a set of PCA dictionaries and weighted the coding residuals to suppress the heavy tail of the distribution. Weights were adaptively updated to decide whether a pixel is heavily corrupted by impulse noise or not. Image sparsity prior and nonlocal self-similarity prior were integrated into a single nonlocal sparse regularization term to enhance the stability of weighted encoding. The results clearly shows that WESNR outperforms much better than other noise removal methods.

5. REFERENCES


