Speckle Noise Reduction in Sonar Image Based on Adaptive Redundant Dictionary

Yifan Huang, Weixiang Li and Fei Yuan *

Key Laboratory of Underwater Acoustic Communication and Marine Information Technology, Ministry of Education, Xiamen University, Xiamen 361005, China; yfhuang@stu.xmu.edu.cn (Y.H.); 23320131153222@stu.xmu.edu.cn (W.L.)
* Correspondence: yuanfei@xmu.edu.cn; Tel.: +86-1330-602-9530

Received: 27 August 2020; Accepted: 26 September 2020; Published: 29 September 2020

Abstract: As acoustic waves are affected by the channel characteristics, such as scattering and reverberation when propagating in water, sonar images often exhibit speckle noise which will cause visual quality of the image to decrease. Therefore, denoising is a crucial preprocessing technique in sonar image applications. However, speckle noise is mainly caused by the sediment echo signals which are related to the background of seafloor sediment and can be obtained by prior modeling. Although deep learning-based denoising algorithms represent a research hotspot now, they are not suitable for such applications due to the high calculation amount and the large requirement of original images considering that sonar is carried by Autonomous Underwater Vehicles (AUVs) for collecting sonar images and performing calculation. In contrast, dictionary learning-based denoising method is more suitable and easier to be modeled. Compared with deep learning, it can greatly reduce the calculation amount and is more easily integrated into AUV systems. In addition, dictionary learning method based on image sparse representation can effectively achieve image denoising similarly. In order to solve the above problems, we propose a new adaptive dictionary learning method based on multi-resolution characteristics, which combines K-SVD dictionary learning with wavelet transform. Our method has the characteristics of dictionary learning and inherits the features of wavelet analysis as well. Compared with several classical methods, the proposed method is better at speckle noise reduction and edge detail preservation. At the same time, the calculation time is greatly reduced and the efficiency is significantly improved.

Keywords: speckle noise; sonar image; dictionary learning; sparse representation; K-SVD; wavelet analysis

1. Introduction

Acoustic waves have much smaller absorption coefficients underwater than light and electromagnetic waves. Therefore, sonar equipment with acoustic waves acting as the carrier plays a very important role in marine monitoring, maritime military operations, and underwater search and rescue. Due to the particularity and complexity of the underwater environment, the echo signals received by sonar are inevitably affected by factors such as channel propagation loss, ocean noise, multipath effects, and reverberation, resulting in the existence of features in sonar images such as low resolution, blurred target edge, and significant speckle noise [1–3]. In order to improve the visual effect of sonar images, denoising preprocessing technology has been widely used in feature extraction, target recognition, and image segmentation [4]. Moreover, the application characteristics of sonar images mean that speckle noise is mainly caused by the sediment echo signals which are related to the background of seafloor sediment and can be obtained by prior modeling. According to different processing domains, traditional image denoising algorithms can be generally divided into spatial domain filtering [5] and transform domain filtering methods [6].

J. Mar. Sci. Eng. 2020, 8, 761; doi:10.3390/jmse8100761 www.mdpi.com/journal/jmse
The spatial domain filtering method is one of the earlier classical image denoising techniques and there is much current research. An adaptive median filtering algorithm based on quadratic detection of noise and neighborhood pixel restoration was used to improve the shortcomings of classical median filtering in [7]. Lee, Kuan, Frost, and SRAD filtering are also categories of spatial domain filtering [8–10]. The authors of [11] introduced a new method based on the combination of three spatial domain filters. Although the spatial domain filtering method has achieved a certain denoising effect, there are certain drawbacks: the processed image is usually too smooth, and it can easily cause blurred edge and detail loss. However, the another transform domain filtering method contains common Discrete Cosine Transform (DCT) [12], principal component analysis (PCA) [13], and Wavelet denoising algorithms [14]. Based on wavelet transform, a direction parameter was added and the Curvelet transform threshold filtering was used for noise removal [15]. Researchers used the Curvelet method for sonar image denoising as well in [16,17]. The transform domain filtering method can achieve a certain denoising effect, but it will remove the high frequency components of the signal itself at the same time, which results in detail loss.

In addition, tensor-based methods have been also widely used in image denoising. A method based on the Tucker decomposition with automatically determined ranks of factoring tensors was proposed for speckle noise reduction in the side-scan sonar images [18]. In [19], a nonlocal low-rank regularized CANDECOMP/PARAFAC (CP) tensor decomposition (NLR-CPTD) was proposed to fully utilize these two intrinsic priors, which can greatly promote the denoising performance in various quality assessments. Researchers proposed a spatial-spectral TV regularized LR tensor factorization (SSTV-LRTF) method to remove mixed noise in HSIs [20]. However, the research on tensor decomposition methods is mainly focused on hyperspectral images, and there have been still few researches on sonar image denoising.

In order to obtain a better denoising effect, deep learning-based denoising algorithms have become a research hotspot in the field of signal processing in recent years. DnCNNs were combined with learning algorithms and regularization methods for image denoising in [21]. The authors of [22] used the automatic encoder algorithm based on convolutional neural network to perform sonar image denoising. In [23], a convolutional neural network based on a denoiser that could utilize the multi-scale redundancy of natural image was proposed. It is well known that deep learning can achieve better results than traditional methods. However, deep neural networks have the disadvantages of huge computation amount, slow training, time-consuming, and inconvenient storage. The required dataset is also very large. In particular, the purpose of denoising cannot be achieved when the training data is insufficient. Most importantly, it is not suitable for such applications due to the high calculation amount and the large requirement of original images considering that sonar is carried by AUV for capturing sonar images and performing calculation.

The dictionary learning-based denoising method is more suitable for sonar application scenarios and it is easier to model. Compared with deep learning methods, it can greatly reduce the calculation amount and more easily integrated into AUV systems. In addition, the dictionary learning method based on image sparse representation can effectively achieve denoising as well [24]. In [25], a novel supervised dictionary learning model with smooth shrinkage was proposed for image denoising. The uncertainty of estimation was considered and an adaptive Bayesian method was used to generate the sparse representation in [26]. Moreover, a polarization image sensor denoising algorithm based on K-singular value decomposition (K-SVD) was proposed in [27]. In summary, dictionary learning denoising method can increase the flexibility and adaptability for data processing. However, most of these methods are complicated and the performance should be improved.

In order to solve the above problems, for sonar speckle noise reduction, we propose a new adaptive dictionary learning method based on multi-resolution characteristics, which combine K-SVD dictionary learning with wavelet transform. The proposed method not only has the characteristics of dictionary learning, but also inherits the multi-resolution and local characteristics of wavelet analysis. Its advantage is the dictionary contains atoms at different resolution scales, which can represent
the original image more effectively. Experiment results show that the proposed method has better ability for speckle noise reduction and edge detail preservation. At the same time, the calculation time is greatly reduced and the efficiency is significantly improved.

This paper is organized as follows. In Section 2, we introduce the statistical model of sonar imaging speckle noise. The proposed method and the principle of dictionary learning-based denoising are described in Section 3. The experiments and results are discussed in Section 4. Section 5 gives the conclusion.

2. Statistical Model of Sonar Imaging

2.1. Sonar Imaging Principle

In most sonar imaging systems, the sonar device actively emits acoustic waves that will be reflected when touching target or seafloor, and the echo is received by transducer. As shown in Figure 1, the pixel coordinates correspond to the arrival time of echo, and the gray pixel values correspond to the echo intensity. Specifically, the strong part of the received echo corresponds to the target, which is indicated in the white area of image; the black area corresponds to the weaker echo region. The remaining area consists of background noise, mostly from the background reflections of seafloor sediment. It is characterized by alternating very small bright and dark areas [28]. Finally, each received echo data is sequentially arranged, and a 2-D sonar image can be obtained.

![Figure 1. Echo intensity of side scan sonar.](image)

2.2. Speckle Noise Model of Sonar Image

Researches show that reverberation is the primary background interference of active sonar systems, which is the sum of received scattered waves reflecting on the seawater boundary and scattering on seawater volume. The echoes among different scatterers cancel each other out or reinforce each other, resulting in the occurrence of fluctuations in the amplitudes of echoes. This phenomenon appears as a grainy black and white texture on the sonar image, which is known as speckle noise.
According to the submarine reverberation model mentioned in [29], it is assumed that the signal emitted by acoustic source is $s(t)$ and the reverberation at time $t$ is expressed as

$$X(t) = \sum_{n=1}^{N} h(r_n) f(r_n) |\alpha_n| |s_0(t - t_n)| e^{j(\omega_0(t-t_n) + \Phi(t-t_n) + \Phi_n)}$$  \hspace{1cm} (1)

where $N$ represents the sum of all scattering elements contributing to time $t$, $h(r_n)$ is the number of scatterers (the value is 0 or 1) at $r_n$ in each scattering element $\Delta V_n$, $f(r_n)$ is the attenuation factor of scattered waves in the two-way propagation, $\alpha_n$ is the scattering coefficient, $t_n$ is the arrival time of echo, $\omega_0$ is the center frequency of transmitted signal, and $\Phi_n$ represents the phase of scattering coefficient $\alpha_n$. Besides, $h(r_n)$, $|\alpha_n|$, and $\Phi_n$ are independent of each other.

Suppose that $V_n(t)$ and $\varphi_n(t)$ represent the instantaneous amplitude and phase value of the $n$th scatterer, respectively, therefore

$$V_n(t) = h(r_n) f(r_n) |\alpha_n| |s_0(t - t_n)|$$  \hspace{1cm} (2)

$$\varphi_n(t) = \omega_0(t-t_n) + \Phi(t-t_n) + \Phi_n$$  \hspace{1cm} (3)

The real and imaginary parts of $X(t)$ can be obtained:

$$Re(t) = \sum_{n=1}^{N} V_n(t) \cos \varphi_n(t)$$  \hspace{1cm} (4)

$$Im(t) = \sum_{n=1}^{N} V_n(t) \sin \varphi_n(t)$$  \hspace{1cm} (5)

From the central limit theorem, we know that when $N$ is large enough, $Re(t)$ and $Im(t)$ are both random variables which are subject to Gaussian random distribution. For random scatterers, $V_n(t)$ and $\varphi_n(t)$, which are in a uniform distribution from 0 to $2\pi$, are independent random variables. It can be seen that $Re(t)$ and $Im(t)$ are obedient to Gaussian random distribution; they are independent of each other, the mean value is zero, and the variance is same. Furthermore, the reverberation amplitude $|X(t)| = \sqrt{(Re^2 + Im^2)}$ is a random variable following Rayleigh distribution. Therefore, it can be concluded that the envelope amplitude of reverberation follows Rayleigh distribution and its phase obeys uniform distribution.

The probability density function of amplitude is as follows. $\alpha$ represents the attenuation parameter of Rayleigh distribution.

$$p_X(x) = \begin{cases} 
0, & x < 0 \\
\frac{x}{\alpha^2} e^{-x^2/2\alpha^2}, & x \geq 0 
\end{cases}$$ \hspace{1cm} (6)

2.3. Speckle Noise Simulation for Sonar Images

From the above analysis, it can be seen that a large number of scatterers make the value of each pixel fluctuate randomly, forming a granular speckle noise. Moreover, the probability density function of its amplitude obeys Rayleigh distribution. Therefore, the multiplicative noise model of speckle noise can be obtained:

$$y = x \cdot n$$  \hspace{1cm} (7)

where $x$ represents the expected true amplitude. $n$ and $y$ are random variables with the same distribution, which both obey translational Rayleigh distribution. The mean value of $n$ is 1.
If considering complicated underwater environment, the impact of environmental noise should be added. However, compared with multiplicative speckle noise, the effect of additive Gaussian noise is much smaller. In order to simplify the model, we ignore the influence of additive noise in this paper. Therefore, the speckle noise in sonar image can be roughly regarded as a multiplication process [30].

Considering that the imaging mechanisms of sonar and ultrasound are similar, the Pizurica method for adding speckle noise to ultrasound images [31] can also be used to simulate the speckle noise for sonar images. The method can simulate characteristics of speckle noise well. The specific steps are as follows.

Step 1 First, generate a complex Gaussian random field. The imaginary and real parts are both independent distribution of Gaussian random variables with the same mean value being 1 and same standard deviation. It is noted that the variance determines different level of speckle noise.

Step 2 Considering the neighborhood correlation of speckle noise, a low-pass average filtering of $3 \times 3$ window is performed on the random field. Take out the amplitude of filtered output. Finally, multiply it with the reference image to obtain the simulated sonar image with speckle noise.

It is proved that there is an optimal structure for fitting the grayscale distribution of sonar image by using Gamma distribution [32]. Figure 2 shows fitting results of the background grayscale probability distribution of real sonar image and Gamma distribution, where the fitting area is the red box in the image. It can be seen that the background grayscale distribution of the original sonar image can be fitted accurately by using Gamma statistical distribution. Similarly, Figure 3 shows the fitting results of the simulated sonar image. Results show that using Gamma statistical distribution can accurately fit the gray distribution of the background area for the simulated noisy image. Therefore, the above speckle noise addition method can simulate the speckle noise of sonar image at an acceptable level.

In this paper, we use the above method to add speckle noise to sonar image, and then use the generated noisy images for subsequent dictionary training and experimental verification.

It should be noted that the general denoising method is mainly used to deal with the additive noise model. In order to be able to use these existing algorithms, the above multiplicative noise model must to be transformed to a simple form. We take the logarithm of both sides of Equation (7) to obtain a sonar image containing approximately Gaussian additive noise. After logarithmic transformation, the probability density function of random variable $\hat{y} = \ln(y)$ becomes

$$p_{\hat{y}} = \frac{\pi e^{2\hat{y}}}{2} e^{-\pi e^{2\hat{y}}/4}$$ (8)

The density function defined by the above formula is called a double exponential or Fisher–Tippet density function. The mean and variance can be calculated by

$$E(\hat{y}) = \frac{1}{2} \ln \left( \frac{4}{\pi} \right) + \frac{1}{2} \psi(1)$$ (9)

$$Var(\hat{y}) = \frac{1}{4} \psi(1, 1)$$ (10)

where $\psi(x)$ is Digamma function, defined as

$$\psi(x) = \frac{d}{dx} \ln \Gamma(x)$$ (11)

$\psi(1, 1)$ is the first-order Polygamma function. The m-order Polygamma function is defined as

$$\psi(m, x) = \frac{d^m}{dx^m} \psi(x) = \frac{d^{m+1}}{dx^{m+1}} \ln \Gamma(x)$$ (12)

The work in [33] proved that Fisher–Tippet distribution of Rayleigh multiplicative noise after logarithmic transformation is approximately in line with additive Gaussian random noise, except for a
small amount of large values. Therefore, denoising methods for additive Gaussian noise can be used for sonar images. Approximate additive Gaussian noise is obtained after logarithmic transformation. We then use similar Gaussian denoising methods to remove the speckle noise from sonar images, and finally perform logarithmic inverse transform to obtain the denoised image.

![Figure 2. Statistical distribution and fitting for the real sonar image background.](image1)

![Figure 3. Statistical distribution and fitting for the simulated sonar image background.](image2)

### 3. Proposed Denoising Algorithm

#### 3.1. Sparse Representation for Sonar Image

Most of the information and energy of an image are concentrated on the coefficients with larger values. Therefore, we consider representing the image by a linear combination of the atoms in a dictionary. Sonar images also have inherent sparse structural characteristics [34], and the distinction between useful information and noise in an image is determined by whether it is a sparse component of the image.

For an image, we consider a small image block $x$ size of $\sqrt{n} \times \sqrt{n}$, and arrange it as a $n$-dimension column vector $x \in \mathbb{R}^n$. Assume a redundant dictionary $D \in \mathbb{R}^{n \times k}$, where $k > n$. The following model can be obtained.

$$\hat{\alpha} = \arg\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad D\alpha \approx x \tag{13}$$

$\|\alpha\|_0$ is the number of nonzero elements in the coefficient $\alpha$, where $\|\hat{\alpha}\|_0 \ll n$. The solution of the above equation is sparse, and the basic idea is that each signal can be represented by a linear combination of some atoms in the redundant dictionary $D$.

In order to facilitate calculation, we use error constraint $\|D\hat{\alpha} - x\|_2^2 \leq \varepsilon$ to replace $D\alpha \approx x$. At the same time, define the sparse degree $L$, which meets $\|\hat{\alpha}\|_0 \leq L \leq n$. Actually, the number of
dictionary atoms in the form of a linear combination to represent image block \( x \) does not exceed \( L \). We use \( (\epsilon, L, D) \), which represents the above model.

In this paper, the speckle noise in sonar images can be roughly regarded as a multiplication process. To use the existing algorithms based on the additive noise model, the multiplicative noise model must be transformed to a simple form. We take the logarithm of both sides of the multiplication model to obtain a sonar image containing approximately additive Gaussian noise. Therefore, denoising methods for additive Gaussian noise can be used for sonar images. Approximate additive Gaussian noise is obtained after logarithmic transformation, then we use similar additive Gaussian denoising methods to remove speckle noise from sonar images, and finally perform logarithmic inverse transform to obtain the denoised image. Suppose \( x \) is a signal that satisfies model \( (\epsilon, L, D) \) and \( n \) is a Gaussian white noise signal with the mean value 0 and standard deviation \( \sigma \). There is a noisy signal \( y \) based on approximate additive noise model after logarithmic transformation:

\[
y = x + n
\]  

(14)

According to the Maximum Posteriori estimation (MAP), we can create the following model,

\[
\hat{\alpha} = \arg\min_\alpha \|\|D\alpha - y\|_2^2 \leq T
\]

(15)

where \( T \) depends on the values of \( \epsilon \) and \( \sigma \).

The denoised image is calculated by \( \hat{x} = D\hat{\alpha} \). Because there is a sparse representation of noise-free image under appropriate dictionary and the noise destroys sparsity of image information. The above model restores the original image by estimating sparse coefficients of the noise-free image so that noise can be removed.

If we translate the constraint of the above formula into a penalty item, we can get the model as follows,

\[
\hat{\alpha} = \arg\min_\alpha \|D\alpha - y\|_2^2 + \mu\|\alpha\|_0
\]

(16)

where \( \mu \) is a penalty factor. Given an appropriate value \( \mu \), the above two questions are consistent. The later discussion of this paper is based on model (16).

This problem is equivalent to minimizing the number of nonzero elements, that is, using as few atoms as possible to represent the image. \( \epsilon \) is the approximate error of this process. In general, the above problem is difficult to solve. We mainly use Orthogonal Matching Pursuit (OMP) [35] to solve the problem in this paper.

3.2. Dictionary Learning and Updating

In this paper, we use the K-SVD framework for dictionary learning and updating [27]. Noisy sonar images are used for training to construct an adaptive dictionary. Because K-SVD has the ability to suppress noise during the learning process, samples used for training can also be selected from noisy images. For a larger image \( X \) size of \( \sqrt{N} \times \sqrt{N} (N \gg n) \), we divide the image first and then rebuild the results. In order to prevent the occurrence of manually processing traces, the results of overlapping blocks can be averaged.

Assume that each block of the image belongs to sparse model \( (\epsilon, L, D) \) and \( D \) is an unknown parameter. This problem can be modeled as follows,

\[
\{D, \hat{\alpha}_{ij}, \hat{X}\} = \arg\min_{\alpha, X} \lambda\|X - Y\|_2^2 + \sum_{ij} \mu_{ij}\|\alpha_{ij}\|_0 + \sum_{ij} \|D\alpha_{ij} - R_{ij}X\|_2^2
\]

(17)

where the first term is a log-likelihood global constraint, representing the degree of approximation between image \( Y \) and its reconstructed image \( X \). The second and third terms are image priors, which guarantee that each block \( X_{ij} = R_{ij}X \) has a sparse representation of bounded error at each
position. $R_{ij}$ is a matrix block size of $n \times N$ extracted from the position $(i, j)$. Moreover, the coefficients $\mu_{ij}$ must be independent in order to satisfy the constraint form $\| D\mu_{ij} - x_{ij} \|_2^2 \leq T$.

First, dictionaries $D$ and $X$ are assumed to be fixed, and $\mu_{ij}$ can be calculated by OMP sparse decomposition method. Then, the dictionary is updated by K-SVD. After dictionary updating, we repeat the stage of sparse decomposition and update the dictionary continuously. These two steps are iterated until the algorithm converges or satisfies the predefined stop criterion. Finally, we can get sparse representation coefficients and a suitable dictionary. The output image can be calculated by using the updated dictionary and sparse coefficient matrix. The process of dictionary construction and updating is as Algorithm 1 [34].

| Algorithm 1 Dictionary learning-based denoising algorithm |
|----------------------------------------------------------|
| 1. Input: noisy images                                    |
| 2. Output: denoised images                                |
| 3. Begin                                                  |
| 4. Initialization: $j = 10$, $\lambda = 30/\sigma$, $C = 1.15$, $X = Y$, $D$ is a redundant DCT dictionary |
| 5. For $j$ times                                          |
| 6. Do                                                     |
| 7. Use OMP to approximate $\forall i,j \min_{a_{ij}} \| a_{ij} \|_0$ s.t. $\| R_{ij} X - D a_{ij} \|_2^2 \leq (C\sigma)^2$; |
| Calculate sparse representation vector $a_{ij}$ of each block $R_{ij}X$ |
| 8. For each column $(l = 1, 2, \ldots, k)$ in $D$, find the set of these atoms $\omega_l = \{(i, j)| a_{ij}(l) \neq 0\}$; |
| Index $(i, j)$ in $\omega_l$, calculate representation error: $e_{ij} = R_{ij}X_{ij} - \sum_{m \in \omega_l} d_m a_{ij}(m)$ |
| 9. Assume the column vector of matrix $E_l$ is $\{e_{ij}\}_{(i, j) \in \omega_l}$; |
| Use SVD to decompose it: $E_l = U\Delta V^T$ |
| 10. Select the first column of $U$ to update column $d_1$ of the dictionary; |
| Use $\Delta(1, 1) \times V$ to update coefficient value $\{a_{ij}(l)\}_{(i, j) \in \omega_l}$ |
| 11. For all the columns of dictionary, repeat the above process to get a new dictionary and sparse representation coefficients. |
| 12. End                                                   |
| 13. Give all of $\hat{a}_{ij}$, fix these coefficients to update $X$; |
| obtain: $\hat{X} = \text{argmin}_X \lambda \| X - Y \|_2^2 + \sum_{ij} \| D\hat{a}_{ij} - R_{ij}X \|_2^2$ |
| 14. Calculate the denoised image: $\hat{X} = \left(IA + \sum_{ij} R_{ij}^T R_{ij}\right)^{-1} \left(\lambda Y + \sum_{ij} R_{ij}^T D\hat{a}_{ij}\right)$ |
| Where $I$ is a unit matrix and $\lambda I + \sum_{ij} R_{ij}^T R_{ij}$ is a diagonal matrix. |

where $j$ is the number of iterations. $\lambda$ and $C$ are parameters. $\sigma$ is the standard deviation of Gaussian white noise signal $y$. $X$ is the noisy image, and $Y$ is the reconstructed image.

It should be noted that K-SVD can adaptively update the atoms in dictionary. K-SVD includes two steps: sparse decomposition and dictionary updating. The biggest difference between K-SVD and other algorithms is the dictionary updating process. K-SVD only updates one atom at a time; until all atoms in the dictionaries are updated, repeat the above process several times to get the representation coefficient of signal and a dictionary after training.

After sparse decomposition, the atom needs to be updated to get a better dictionary. In this stage, dictionary $D$ updates one column each time, and fixes other columns in the dictionary. $d_m$ is the $m$th column in the dictionary. $\omega_l$ is the index of signal atom $d_l$. Through decomposing $E_l$ by using Singular Value Decomposition (SVD), we can get $E_l = U\Delta V^T$. The $k$th column of the dictionary is updated with the first column of $U$, and the coefficient is updated by $\Delta(1, 1) \times V$. Repeat this process for all columns in the dictionary (hence named K-SVD). After dictionary updating, the sparse decomposition stage is repeated, and the dictionary is updated again. These two steps are iteratively performed.
until the algorithm converges or meets the predefined stopping criterion, and finally the sparse representation coefficient and a suitable dictionary are obtained.

### 3.3. Proposed Solution for Dictionary Training

Based on the adaptive dictionary constructed by K-SVD in the previous section, we propose a new method for dictionary training in this section. On the framework of K-SVD, we combine multi-resolution and local characteristics of wavelets to obtain a learning dictionary with structural information at different scales (multi-resolution adaptive dictionary). This constructed dictionary contains atoms at different resolution scales, where the large-scale atoms can effectively describe the overall structure of image and the small-scale atoms can capture the small features. Therefore, the newly constructed dictionary can effectively represent original image.

Multi-resolution analysis can be visually represented by using a set of nested multi-resolution subspaces. It is assumed that the frequency space of original signal is expressed as \( V_0 \), which is decomposed into two subspaces of low frequency \( V_1 \) and high frequency \( W_1 \), and then \( V_1 \) is subdivided into low-frequency components \( V_2 \) and high-frequency components \( W_2 \), and so on. The following characteristics of subspace can be obtained,

\[
V_0 = V_1 \oplus W_1 = V_2 \oplus W_2 \oplus W_1 = V_N \oplus W_N \oplus W_{N-1} \oplus \ldots W_2 \oplus W_1
\]  

(18)

where \( \oplus \) means the orthogonal sum of two subspaces; \( V_j \) represents the subspace corresponding to resolution \( 2^{-j} \); the space vector \( W_j \) formed by the expansion and translation of wavelet function is the orthogonal complement space of \( V_j \); \( W_j \) reflects the high frequency subspace of the spatial signal details of \( V_{j-1} \); \( V_j \) shows the low frequency subspace of the spatial signal profile of \( V_{j-1} \). Moreover, the model’s physical meaning is that a space \( V_0 \) with resolution \( 2^0 = 1 \) can be approximated by finite subspaces.

As the image is a two-dimensional signal which is similar to the multi-resolution analysis of a one-dimensional signal, we need to perform corresponding expansion and translate the space \( L^2(R) \) into \( L^2(R \times R) \), where the one-dimensional scale function \( \phi(x) \) becomes \( \phi(x, y) \).

Assuming \( \{V_j\}_{j \in Z} \) is the multi-resolution analysis of space \( L^2(R) \), there is a tensor space \( \{V_j^2\}_{j \in Z} \), where \( V_j^2 = V_j \oplus V_j \) constitutes the multi-resolution analysis of space \( L^2(R \times R) \) and the two-dimensional scale function of multi-resolution analysis \( \{V_j^2\}_{j \in Z} \) is

\[
\phi(x, y) = \phi(x)\phi(y)
\]

(19)

where \( \phi(x) \) is the one-dimensional scale function of \( \{V_j\}_{j \in Z} \). For each \( j \in Z \), the normative orthogonal basis of \( \{V_j^2\}_{j \in Z} \) consists of function systems \( \phi_{j,n,m}(x,y) = \phi_{j,n}(x)\phi_{j,m}(y) \). Moreover, \( \{V_j^2\}_{j \in Z} \) is called separable multi-resolution analysis of \( L^2(R \times R) \). Because \( \phi(x) \) and \( \phi(y) \) are both low-pass scale functions, \( \{V_j^2\}_{j \in Z} \) is a smooth low-pass space.

If \( \psi(x) \) represents the orthogonal wavelet basis of one-dimensional multi-resolution analysis \( \{V_j\}_{j \in Z} \), the integer translation sequences of three wavelet functions of two-dimensional multi-resolution analysis is expressed as

\[
\begin{align*}
\psi_{j,n,m}^{1}(x,y) &= \phi_{j,n}(x)\psi_{j,m}(y) \\
\psi_{j,n,m}^{2}(x,y) &= \psi_{j,n}(x)\phi_{j,m}(y) \\
\psi_{j,n,m}^{3}(x,y) &= \psi_{j,n}(x)\psi_{j,m}(y)
\end{align*}
\]

(20)
Noted that the superscripts in the equation are not exponents, just index terms, which form the normalized orthogonal basis of \( \{ W^2_j \} \), Moreover, these three orthogonal basis reflecting the detailed information of the two-dimensional signal are band-pass. Specifically, \( \varepsilon = 1, 2, 3 \) corresponds to the horizontal, vertical, and diagonal directions of the two-dimensional signal, respectively. For any arbitrary \( f(x, y) \in L^2(R \times R) \) at resolution of \( 2^{-j} \), there are:

\[
A_jf = \langle f(x, y), \phi_{j,n,m}(x,y) \rangle, (n,m) \in \mathbb{Z}^2
\]
\[
D^1_jf = \langle f(x, y), \psi^1_{j,n,m}(x,y) \rangle, (n,m) \in \mathbb{Z}^2
\]
\[
D^2_jf = \langle f(x, y), \psi^2_{j,n,m}(x,y) \rangle, (n,m) \in \mathbb{Z}^2
\]
\[
D^3_jf = \langle f(x, y), \psi^3_{j,n,m}(x,y) \rangle, (n,m) \in \mathbb{Z}^2
\]

The above equation indicates that the image is decomposed into sub-bands of these four directions: \( A_jf, D^1_jf, D^2_jf, \) and \( D^3_jf \) at resolution of \( 2^{-j} \), where \( A_jf \) represents the approximate components of the image (the low frequency portion of image), and \( D^\varepsilon_jf \) indicates the high frequency components or detail part at each direction. Specifically, \( D^1_jf \) corresponds to the high frequency component of vertical direction, \( D^2_jf \) belongs to the horizontal high frequency component and \( D^3_jf \) corresponds to the diagonal.

We propose an adaptive dictionary learning method with multi-resolution characteristics by combining K-SVD with wavelet transform. The proposed method not only has the characteristics of dictionary learning, but also inherits the multi-resolution and local characteristics of wavelet analysis. The advantage is the dictionary contains atoms at different resolution scales. The large-scale atoms can effectively describe the overall structure of image and the small-scale atoms can capture the small features. Therefore, the newly constructed dictionary can effectively represent original image. Our proposed method has better ability for speckle noise reduction and edge detail preservation. At the same time, the calculation time is greatly reduced and the efficiency is significantly improved. Algorithm 2 shows the proposed method.
Algorithm 2 The proposed denoising algorithm

1: Input: noisy images  
2: Output: denoised images  
3: Begin  
4: Perform 3-layers wavelet transform on the noisy sonar images  
5: Extract the low-frequency coefficients of the first scale, and the high-frequency  
6: coefficients of each scale to train dictionaries respectively in steps: 6-15  
7: Initialization  
8: Fix dictionary $D$ and $X$ in sparse model $(\varepsilon, L, D)$  
9: For $j = 10$ times  
10: Use OMP to calculate the sparse representation vector  
11: Update the dictionary by K-SVD  
12: Repeat the above steps to update dictionary and coefficient value by column  
13: End  
14: Obtain the newly constructed multiple adaptive dictionary and coefficient of  
15: sparse representation  
16: Give all of the sparse coefficients, fix these coefficients to update $X$  
17: Cascade these dictionaries to get the final multi-resolution adaptive dictionary  
18: Perform logarithmic transform of the simulated noisy sonar image to transform  
19: multiplicative Rayleigh noise into approximate additive Gaussian noise  
20: Decompose and reconstitute the image by using the above dictionaries  
21: Calculate the denoised image  
22: Perform inverse logarithmic transform for the denoised image  
23: Obtain the final denoised image.  
24: End

4. Experiments and Discussions

The denoising results of sonar images are given in this section, including the simulated noisy  
sonar images and the real sonar images. All the simulation results are obtained in the configuration of  
Pentium dual-core CPU, dominant frequency at 3.2GHz, 8G memory, Windows XP 32-bit operating  
system and MATLAB R2010b.

4.1. Denoising Results of Simulated Noisy Image

Above all, select the noise-free image as original image, and then add speckle noise of different  
levels to the image. The simulation method of speckle noise is mentioned in Section 2. We use  
the proposed denoising method for simulated noisy sonar image, and compare it with several methods.  
The denoising comparison algorithms used in this paper include Lee filter [36], Kuan filter [37],  
Frost filter [38], SRAD filter [39], Wavelet-based denoising method [14], Curvelet-based denoising  
method [15], DCT-based denoising method [12], and K-SVD denoising method [34].

Figure 4 shows the original image and simulated noisy image. Figure 5 shows the denoising  
results of different algorithms for the simulated image. The standard deviation of speckle noise is 0.9.  
It can be seen that the edge details are lost and there is still residual noise of spatial domain filtering  
methods such as Lee, Kuan, Frost, and SRAD. The Wavelet denoising retains edges well, but a lot of  
noise remains. Curvelet has a good effect on edge retention, but scratches exist. The K-SVD algorithm  
achieves better noise removal and edge retention. However, the performance of the proposed method  
is still the best in all respects. In addition, the proposed algorithm is far superior to K-SVD in terms of  
calculation efficiency. When $\sigma = 0.3$, K-SVD took 1270.11 seconds, and the proposed took 550.10 s.  
When $\sigma = 0.6$, K-SVD took 267.68 seconds, and the proposed took 169.39 seconds. When $\sigma = 0.9$,  

K-SVD took 315.76 seconds, and the proposed took 146.81 seconds. Therefore, the time efficiency of the proposed algorithm is nearly double that of K-SVD.

![Figure 4. Original image and simulated noisy image.](image)

In addition, we use the full reference indicators for objective quality evaluation, including structural similarity index measurement (SSIM) and equivalent numbers of looks (ENL). SSIM evaluates image quality by comparing the structural characteristics of the original image and the distorted image, without error merging. The theoretical basis of this method is that natural images are based on texture characteristics. Pixels that are especially adjacent in space have strong correlations, and these correlations carry a lot of important image information. SSIM measures image similarity from three aspects: brightness, contrast, and structure. The larger the value, the better the algorithm. The goal of speckle noise suppression is to effectively suppress speckles while keeping the structural information in the image scene as much as possible. Therefore, ENL is introduced in this paper to measure the smoothness of uniform area, which can be used to measure the relative intensity of speckle noise in an image. It is also an index to measure the filtering performance of a filter. In the texture area, the smaller the ENL value, the better the texture information is maintained. In the uniform area, the larger the ENL value, the better the filtering effect.

Tables 1–3 show the values of the objective evaluation indexes of different algorithms when the standard deviation of noise is 0.3, 0.6, and 0.9, respectively. Each data in the table is a value obtained by averaging 10 simulation experiments. It can be seen that the performance of the proposed method is better than others by comparing the values of SSIM and ENL.
Table 1. The objective evaluation values of denoising results.

| Noise  | Method  | SSIM  | ENL   |
|--------|---------|-------|-------|
| σ = 0.3| Lee     | 0.8779| 741.96|
|        | Kuan    | 0.8879| 741.96|
|        | Frost   | 0.8371| 1333.30|
|        | SRAD    | 0.8868| 1304.20|
|        | Wavelet | 0.8849| 611.85 |
|        | Curvelet| 0.8907| 612.93 |
|        | DCT     | 0.9685| 3354.10|
|        | K-SVD   | 0.9623| 2899.80|
|        | Proposed| 0.9812| 6254.90|

Table 2. The objective evaluation values of denoising results.

| Noise  | Method  | SSIM  | ENL   |
|--------|---------|-------|-------|
| σ = 0.6| Lee     | 0.7794| 216.28|
|        | Kuan    | 0.7794| 216.28|
|        | Frost   | 0.7959| 497.52|
|        | SRAD    | 0.8306| 644.30 |
|        | Wavelet | 0.7782| 207.83 |
|        | Curvelet| 0.7738| 235.84 |
|        | DCT     | 0.9240| 1687.70|
|        | K-SVD   | 0.9177| 1346.10|
|        | Proposed| 0.9592| 2278.00|

Table 3. The objective evaluation values of denoising results.

| Noise  | Method  | SSIM  | ENL   |
|--------|---------|-------|-------|
| σ = 0.9| Lee     | 0.6840| 88.03 |
|        | Kuan    | 0.6840| 88.03 |
|        | Frost   | 0.7408| 184.78|
|        | SRAD    | 0.7400| 293.31|
|        | Wavelet | 0.6952| 95.12 |
|        | Curvelet| 0.6730| 110.86|
|        | DCT     | 0.8874| 581.47|
|        | K-SVD   | 0.8881| 595.71|
|        | Proposed| 0.9104| 693.74|

Figure 6a shows the values of objective evaluation index SSIM after filtering by various algorithms under different noise levels. Under the same noise pollution level, the proposed method has the highest SSIM value, which indicates that our method is better at removing speckle noise and preserves the image structure information well. Figure 6b shows the ENL values under different noise levels. As can be seen from the graph, the proposed method has higher equivalent visual value, which means its speckle suppression effect is better.
4.2. Denoising Results of Real Sonar Images

The experiment images are selected from the side scan sonar image library Fenn Enterprises on the website: http://www.fennent.com/sonar.html. First, we grayscale the images and adjust the size to 256 × 256. As there is no reference image, subjective observation method is selected to illustrate the denoising results of various algorithms. The three side scan sonar images are selected from the image library Fenn Enterprises of size 256 × 256.

It can be seen that the three original images contain speckle noise with different levels. Figure 7–10 correspondingly show the denoising results. The denoising effect of the different methods has consistent characteristics on these three images and the results are consistent with the simulated noisy image. As can be seen from the figures, the traditional methods such as Lee, Kuan, Frost, Wavelet, and Curvelet have more residual noise. Moreover, SRAD appears visual blurring. Predictably, the proposed method achieves better results in suppressing speckle noise and detail retention than several other methods.

Figure 6. Comparison of the objective evaluation values.

Figure 7. Denoising results of real sonar images.
Figure 8. Denoising results of real sonar images.

Figure 9. Denoising results of real sonar images.
5. Conclusions

Based on the study of the characteristics of sonar imaging and the statistical characteristics of speckle noise, we propose an adaptive dictionary learning method with multi-resolution characteristics and apply it to speckle noise removal in sonar images. The proposed method not only has the characteristics of dictionary learning, but also inherits the multi-resolution and local characteristics of wavelet so that the dictionary contains atoms at different resolution scales, which can represent the original image more effectively. Experiment results show that the proposed method can achieve a better denoising effect than the classical K-SVD algorithm and several traditional methods, and can preserve the detail information of image. At the same time, the calculation speed of the proposed algorithm is increased approximately twofold and the running time is reduced by half that of the classical K-SVD. Therefore, the proposed method effectively reduces the computation time and improves the efficiency as well.

Author Contributions: Conceptualization, Y.H. and W.L.; data curation, W.L.; formal analysis, Y.H. and F.Y.; funding acquisition, F.Y.; investigation, Y.H.; methodology, Y.H. and W.L.; project administration, F.Y.; resources, F.Y.; software, Y.H. and W.L.; supervision, F.Y.; validation, Y.H. and W.L.; visualization, Y.H. and W.L.; writing—original draft, Y.H. and W.L.; writing—review & editing, Y.H.. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to thank the National Natural Science Foundation of China (grant no 62071401, 61771412, and 61871336) and the Fundamental Research Funds for the Central Universities (grant no 20720180068).

Conflicts of Interest: The authors declare no conflicts of interest.
References

1. Marsal, A.B.; David, J.P.; Daniel, C. Synthetic aperture sonar speckle noise reduction performance evaluation. *J. Acoust. Soc. Am.* 2018, 143, 1856.

2. Chen, D.S.; Chu, X.M.; Ma, F.; Teng, X.X. A variational approach for adaptive underwater sonar image denoising. *Int. Conf. Transp. Inf. Saf. ICTIS Proc.* 2017, 1177–1181.

3. Chen, W.L.; Gu, K.; Lin, W.S.; Xia, Z.F.; Patrick, L.C.; Cheng, E. Reference-Free Quality Assessment of Sonar Images via Contour Degradation Measurement. *IEEE Trans. Image Process.* 2019, 28, 5336–5351. [CrossRef]

4. Huo, G.Y.; Yang, S.X.; Li, Q.W.; Zhou, Y. A Robust and Fast Method for Sidescan Sonar Image Segmentation Using Nonlocal Despeckling and Active Contour Model. *IEEE Trans. Cybern.* 2017, 47, 855–872. [CrossRef]

5. Su, J.; Yu, L.Z. Research of Image Spatial Domain Filtering Based on MATLAB. *Softw. Guide* 2017, 16, 132–134.

6. Xiao, J.; Zou, W.; Zhang, S.; Lei, J.; Wang, W.; Wang, Y.F. Video denoising algorithm based on improved dual-domain filtering and 3D block matching. *IET Image Process.* 2018, 12, 1751–9659. [CrossRef]

7. Li, T.T.; Tang, S.F.; Wang, F.; Tong, M.M.; Xu, C.L. Image enhancement study based on adaptive median filtering with secondary noise detection and neighborhood pixel recovery. In *Proceedings of the 2018 International Conference on Robots & Intelligent System (ICRIS)*, Changsha, China, 26–27 May 2018; pp. 134–136.

8. Guo, H.T.; Xu, L.; Zhao, H.Y.; Jiao, S.X.; Hou, Y.M. A morphological filter for despeckling of a sonar image. *Chin. J. Sci. Instrum.* 2015, 36, 654–659.

9. Tripathi, A.; Bhatia, V.; Sharma, A. Kuan modified anisotropic diffusion approach for speckle filtering. *Adv. Intell. Syst. Comput.* 2017, 458, 537–545.

10. Choi, H.; Jeong, J. Speckle noise reduction in ultrasound images using SRAD and guided filter. In *Proceedings of the 2018 International Workshop on Advanced Image Technology (IWAIT)*, Chiang Mai, Thailand, 7–9 January 2018; pp. 1–4.

11. Garg, A.; Khandelwal, V. Combination of Spatial Domain Filters for Speckle Noise Reduction in Ultrasound Medical Images. *Adv. Electr. Electron. Eng.* 2017, 15, 857–865. [CrossRef]

12. Yuki, K.; Yoshihiro, M.; Norishige, F. Parallelized and vectorized implementation of DCT denoising with FMA instructions. In *Proceedings of the 2018 International Workshop on Advanced Image Technology (IWAIT)*, Chiang Mai, Thailand, 7–9 January 2018.

13. Zhang, J.C.; Luo, H.B.; Liang, R.G.; Zhou, W.; Hui, B.; Chang, Z. PCA-based denoising method for division of focal plane polarimeters. *Opt. Express* 2017, 25, 2391–2400. [CrossRef]

14. Shen, Y.; Liu, Q.; Lou, S.Q.; Hou, Y.L. Wavelet-Based Total Variation and Nonlocal Similarity Model for Image Denoising. *IEEE Signal Process Lett.* 2017, 24, 877–881. [CrossRef]

15. Biswas, R.; Purkayastha, D.; Roy, S. Denoising of MRI Images Using Curvelet Transform. *PLect. Notes Electr. Eng.* 2018, 442, 575–583.

16. Yao, Z.A.; Sun, C.Y.; Shi, X.L.; Liu, M.T. A combined denoising method based on Curvelet transform and anisotropic diffusion filtering. *Acta Pet. Sin.* 2016, 37, 490–498.

17. Wen, N.; Yang, S.Z.; Cui, S.C. High resolution remote sensing image denoising based on Curvelet-Wavelet transform. *Eng. Sci.* 2015, 49, 79–86.

18. Grabek, J.; Cyganek, B. Speckle Noise Filtering in Side-Scan Sonar Images Based on the Tucker Tensor Decomposition. *Sensors 2019*, 19, 2903. [CrossRef] [PubMed]

19. Xue, J.; Zhao, Y.Q.; Liao, W.Z.; Cheung-Wai Chan, J. Nonlocal Low-Rank Regularized Tensor Decomposition for Hyperspectral Image Denoising. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 5174–5189. [CrossRef]

20. Fan, H.Y.; Li, C.; Guo, Y.L.; Kuang, G.Y.; Ma, G.Y. Spatial–Spectral Total Variation Regularized Low-Rank Tensor Decomposition for Hyperspectral Image Denoising. *IEEE Trans. Geosci. Remote Sens.* 2018, 56, 6196–6213. [CrossRef]

21. Zhang, K.; Zuo, W.M.; Chen, Y.J.; Meng, D.; Zhang, L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* 2017, 26, 3142–3155. [CrossRef]

22. Kim, J.; Song, S.; Yu, S. Denoising auto-encoder based image enhancement for high resolution sonar image. In *Proceedings of the 2017 IEEE Underwater Technology (UT)*, Busan, Korea, 21–24 February 2017.

23. Dong, W.S.; Wang, P.Y.; Yin, W.T.; Shi, G.M.; Wu, F.F.; Lu, X.T. Denoising Prior Driven Deep Neural Network for Image Restoration. *Comput. Sci.* 2018, 41, 2305–2318. [CrossRef]
24. Liu, J.; Osher, S. Block Matching Local SVD Operator Based Sparsity and TV Regularization for Image Denoising. *J. Sci. Comput.* 2018, 78, 607–624. [CrossRef]
25. Zhang, K.T.; Zhang, L.Q. Supervised Dictionary Learning with Smooth Shrinkage for Image Denoising. *Neural Process Lett.* 2018, 47, 535–548. [CrossRef]
26. Serra, J.G.; Testa, M.; Molina, R.; Katsaggelos, A.K. Bayesian K-SVD using fast variational inference. *IEEE Trans. Image Process.* 2017, 26, 3344–3359. [CrossRef]
27. Li, S.; Ye, W.; Liang, H.; Pan, X.; Lou, X.; Zhao, X. K-SVD Based Denoising Algorithm for DoFP Polarization Image Sensors. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018.
28. Li, Q.W.; Ma, G.C.; Huo, G.Y. New segmentation method of side-scan sonar image based on edge detection in NSCT domain. *Chin. J. Sci. Instrum.* 2013, 34, 1795–1801.
29. Middleton, D. A Statistical Theory of Reverberation. *J. Acoust. Soc. Am.* 1967, 42, 1200–1200. [CrossRef]
30. Jennifer-Ranjani, J.; Chithra, M.S. Bayesian denoising of ultrasound images using heavy-tailed Levy distribution. *IET. Image Process.* 2015, 9, 323–331. [CrossRef]
31. Pizurica, A.; Philips, W.; Lemahieu, I.; Acheroy, M. A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging. *IEEE Trans. Med. Imaging* 2003, 22, 323–331. [CrossRef]
32. Tian, X.D.; Liu, Z. A new image processing algorithm for sonar. *Acoust. Eng.* 2007, 2, 8–11.
33. Xie, H.; Pierce, L.E.; Ulaby, F.T. Statistical properties of Logarithamically transformed speckle. *IEEE Trans. Geosci. Remote Sens.* 2002, 40, 721–727. [CrossRef]
34. Ma, L.Y.; Ma, H.B.; Liu, L. Speckle Noise Reduction in SAR image based on K-SVD. In Proceedings of the 2015 International Symposium on Computers & Informatics, Beijing, China, 17–18 January 2015; pp. 1615–1622.
35. Pati, Y.C.; Rezaiifar, R.; Krishnaprasad, P.S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 1–3 November 1993; pp. 40–44.
36. Lee, J. Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Trans. Pattern Anal. Mach. Intell.* 1980, PAMI-2, 165–168. [CrossRef]
37. Kaun, D.T.; Sawchuk, A.A.; Strand, T.C.; Chavel, P. Adaptive Noise Smoothing for Images with Signal-Dependent Noise. *IEEE Trans. Pattern Anal. Mach. Intell.* 1985, PAMI-7, 165–177. [CrossRef]
38. Frost, V.S.; Stiles, J.A.; Shanmugan, K.S.; Holtzman, J.C. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Pattern Anal. Mach. Intell.* 1982, PAMI-4, 157–166. [CrossRef] [PubMed]
39. Yu, Y.; Acton, S.T. Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.* 2002, 11, 1260–1270. [PubMed]