



Article A Robust and Fast Method for Sidescan Sonar Image Segmentation Based on Region Growing

Xuyang Wang ¹, Luyu Wang ¹, Guolin Li ² and Xiang Xie ^{1,*}

- ¹ School of Integrated Circuits, Tsinghua University, Beijing 100084, China;
- xuyang-w17@mails.tsinghua.edu.cn (X.W.); luyu-wan19@mails.tsinghua.edu.cn (L.W.)
- ² Department of Electronic Engineering, Tsinghua University, Beijing 100084, China; guolinli@tsinghua.edu.cn
- Correspondence: xiexiang@tsinghua.edu.cn

Abstract: For high-resolution side scan sonar images, accurate and fast segmentation of sonar images is crucial for underwater target detection and recognition. However, due to the characteristics of low signal-to-noise ratio (*SNR*) and complex environmental noise of sonar, the existing methods with high accuracy and good robustness are mostly iterative methods with high complexity and poor real-time performance. For this purpose, a region growing based segmentation using the likelihood ratio testing method (RGLT) is proposed. This method obtains the seed points in the highlight and the shadow regions by likelihood ratio testing based on the statistical probability distribution and then grows them according to the similarity criterion. The growth avoids the processing of the seabed reverberation regions, which account for the largest proportion of sonar images, thus greatly reducing segmentation time and improving segmentation accuracy. In addition, a pre-processing filtering method called standard deviation filtering (*STDF*) is proposed to improve the *SNR* and remove the speckle noise. Experiments were conducted on three sonar databases, which showed that RGLT has significantly improved quantitative metrics such as accuracy, speed, and segmentation visual effects. The average accuracy and running times of the proposed segmentation method for 100 \times 400 images are separately 95.90% and 0.44 s.

Keywords: segmentation; sonar images; fast and accurate; region growing

1. Introduction

Sidescan sonar (SSS), which can provide high-resolution images of the seabed, is one of the most common sensors for various underwater applications, such as topography measurement [1], search for sunken vessels and submerged settlements [2], underwater mine detection [3], fish stocks detection, cable or pipeline detection [4-6], and offshore oil prospecting [7]. Accurate and efficient segmentation of SSS images is essential for underwater objects detection. Because segmenting sonar images into highlight areas with objects, regions of shadow, and seafloor reverberation is an effective method to obtain the region of interest (ROI), this is usually an important step before object classification. However, the existing issues on sonar images such as low SNR, intensity inhomogeneity, and complex marine environment noise severely affect the performance of segmentation algorithms especially for small targets. In addition, the characteristics of sonar images are unstable, for example, shadows are not always present. In addition, shadows and highlight areas are affected by sonar position and the properties of the environment, which adds challenges for this task and makes it difficult to obtain a dataset that is robust for all environments to be encountered. Moreover, the algorithms should be rapid and efficient to enable real-time detection, but it is difficult to do this due to the high resolution of sonar images. This work concentrates on fast and accurate segmentation of sonar images into highlighted areas with objects, shadowed areas, and reverberant areas of the seafloor.

Various methods have been put forward to carry out sidescan sonar imagery segmentation. Widely used algorithms in the field of sonar image segmentation include threshold



Citation: Wang, X.; Wang, L.; Li, G.; Xie, X. A Robust and Fast Method for Sidescan Sonar Image Segmentation Based on Region Growing. *Sensors* **2021**, *21*, 6960. https://doi.org/ 10.3390/s21216960

Academic Editor: Andrzej Stateczny

Received: 30 September 2021 Accepted: 18 October 2021 Published: 20 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). methods [8], clustering methods [9–11], Markov random field model (MRF) [12–15], curve evolution methods [16–18], convolutional neural networks (CNN) [19], etc. The threshold method usually faces the problem of varying illumination in sonar images, so it is difficult to find a suitable adaptive threshold. Methods based on clustering techniques (e.g., fuzzy C-means [11]) do not introduce modeling of the statistical distribution characteristics of luminance within the different types of regions, so these approaches lack robustness [15]. Active contour method [16,18] and level set method [17] based on curve evolution have been proposed and have been further investigated with good performance. MRF models are applied to sonar image segmentation and satisfactory results are achieved. However, these algorithms are quite complicated and require a large number of computational resources and running time [11], making them difficult to be applied for Autonomous Unmanned Vehicles (AUVs). Recently, CNN based methods have been researched on sonar segmentation [19]. However, the availability of a sufficient amount of training data to learn the classification model is still challenging, and the model results depend on the similarity between the training data and the test data. When the data are mismatched, the observed results may be inferior.

Filtering is commonly used to remove speckle noise to enhance the segmentation effect. Traditional image denoising algorithms can be generally divided into spatial domain and transform domain filtering methods. Lee [20], Frost [21], and SRAD [22] are typical spatial domain filtering methods with excellent filtering performance. However, the processed image appears too smooth and has blurred edges [23]. The transform domain filtering method contains common Discrete Cosine Transform [24], principal component analysis [25], and wavelet denoising algorithms [26]. Karthikeyan and Chandrasekar proposed a method combining the SRAD filter with the wavelet-based technique [27]. The method achieved a PSNR metric value of over 70 dB, as measured on a test dataset. Researchers used the stationary wavelet transform for sonar image denoising in [28], and the method has outstanding performance in terms of PSNR and SSIM. 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 [23].

In order to perform unsupervised segmentation effectively and efficiently, an algorithm based on region growing and a priori statistical distribution characteristics is proposed, called RGLT. The method first selects seed regions containing highlight and shadow areas and then grows them according to similarity criteria. Because the seabed reverberant areas take up the largest proportion of the sonar images while the highlight and shadow take up a small proportion, avoiding processing seabed areas when growing can save a considerable amount of time. The statistical characteristics are taken into account by a highlight-shadow likelihood ratio test at the stage of the seed regions selection, allowing the algorithm to efficiently and adaptively obtain preliminary segmentation. To solve the problem of low contrast, low SNR, and varying illumination, pre-process enhancement methods are presented to improve the contrast and SNR of a side scan sonar image, remove the speckle noise, and balance the illumination throughout the image. The results show significant gains in terms of quantitative metrics (speed and accuracy) and visual effects using RGLT. Compared with Fuzzy C-means [11] and active contour method [18], RGLT has obvious advantages in processing images with large differences in the ratio of background, highlight areas and shadows, and has better adaptability to noisy backgrounds and can clearly segment small targets.

The existing general issues on sonar images include low *SNR*, intensity inhomogeneity, and complex marine environment noise, which severely affect the performance of segmentation algorithms especially for small targets. In addition, most of the current methods with high accuracy and good robustness are iterative methods with high complexity and poor real-time performance.

The contributions of this article are summarized as follows:

1. This paper describes a pre-processing method for correcting time-varying gain (TVG) effects in sidescan sonar data, to balance the intensity throughout range.

- 2. To remove the speckle noise and enhance the *SNR* as well as the contrast of the sonar image, a filter *STDF* is proposed in this paper. The enhancement is beneficial to the subsequent segmentation.
- 3. An unsupervised segmentation algorithm is presented based on the assumption that the region of the seafloor obeys the Weibull distribution. Experiments showed the method has significantly improved quantitative metrics such as accuracy, speed, and segmentation visual effects. It is an effective and efficient algorithm based on region growing that can be used in the real-time sonar application.

2. Materials and Methods

2.1. Overview

As illustrated in Figure 1, in this paper, the side-scan sonar images are processed in four steps. Firstly, the time gain of sonars is re-compensated in order to make the echo level independent of range. The step can alleviate the problem of varying illumination in sonar images and make the seabed intensity distribution in different regions consistent. It contributes to improving the robustness of adaptively seeds selection and region growing during subsequent segmentation. Sonar images suffer from low SNR and speckle noise, which adds difficulties to fine segmentation. Therefore, a filter referring as STDF is used to enhance sonar images and remove the speckle noise. Secondly, the ratio of the standard deviation to the mean (RDM) is a constant under the assumption that the sonar images obey the Rayleigh distribution. Exploiting the feature that the highlight regions have higher RDM than the seabed reverberation regions, *STDF* improves the *SNR* while removing speckle noise. To speed up the segmentation process for high-resolution sonar images, an algorithm based on regional consistency is proposed instead of the global iteration. To avoid losing the highlight and the shadow regions and to reduce the probability of the seabed regions being misclassified into the foreground, seed points are selected based on the statistical distribution characteristics and likelihood ratio test. The seed points are selected in the highlight and the shadow areas adaptively and efficiently to obtain the preliminary segmentation. Then, the subsequent growing is based on the preliminary segmentation and is conducted on the proposed similarity criterion. The seed selection enables the growth process to save lots of time without processing the seafloor reverberation region with large areas. The segmentation obtains fine details and high accuracy by the growth step.



Figure 1. Block diagram of the proposed method process.

2.2. Time Gain Re-Compensation

Time gain compensation (TGC) is indispensable in sonars in order to compensate for transmission loss and inconsistency of reverberation so as to make the echo level independent of range. The usual TGC will be embedded in the hardware according to the formula and then fine-tuned by manual adjustment. However, because the estimation of the parameters in the TGC formula is coarse (the parameters are related to the bottom substrate, geometric spreading type, water temperature, sonar frequency, etc.), the dependence of the intensity on the range cannot be completely eliminated, and the uniformity of the intensity is essential for the subsequent segmentation task, so a time gain re-compensation technique is proposed to reduce the dependence further.

When we consider the submarine interface reverberation, the received echo level, *E*, for an active sound system can be expressed as [29]:

$$E = SL - \left(2TL + S_{s,v} + 10\log\frac{c\tau}{2}\phi R\right) \tag{1}$$

where *SL* is the source level, *TL* is the one-way transmission loss, $S_{s,v}$ is the surface/volume backscattering intensity. c, τ , ϕ and *R* are speed of sound, width of transmitted pulse, horizontal opening angle of the transducer, and the range, respectively. The transmission loss is mainly caused by the geometric spreading and absorption of the acoustic. The commonly used function has the form $2TL(R) = 40 \log R + 2\alpha R$, where $40 \log R$ represents the transmission loss caused by the geometric spreading, and this function takes the assumption of spherical spreading in the free field. However, the sonar signal transmission does not always follow this ideal assumption. The function will be $20 \log R$ or $60 \log R$, respectively, when taking the assumption of cylindrical spreading between parallel planes or hyperspherical spreading in free field with time expansion field. In addition, $2\alpha R$ represents the transmission loss caused by the absorption of the acoustic, and α represents the logarithmic absorption coefficient. α is estimated on the experience usually, but the estimation is complicated and inaccurate, for that α is related to various factors such as sonar frequency, water temperature, salinity, and pressure.

From the above analysis, the gain compensation related to the range is abbreviated as follows:

$$20\log\left(\frac{I(R_{ref})}{I(R)}\right) = Gain(R,\lambda,\alpha,\gamma) = 10\beta\log R + 2\alpha R + \gamma$$
(2)

where $I(R_{ref})$ and I(R) are the echo intensity at reference range R_{ref} and range R. β and γ are constants. It is often the case that the image intensity still depends on the range due to imprecise parameters estimation in fact. Therefore, it is necessary to re-compensate the gain that has been compensated by hardware. In this paper, the parameters estimation problem is described as the nonlinear least squares problem as follows:

$$\min \sum_{R} (20 \log \frac{\widetilde{I}(R_{ref})}{\widetilde{I}(R)} - Gain(R, \gamma, \alpha, C))^2$$
(3)

Then, this problem is solved by the Gauss–Newton method to estimate parameters β , α , and γ . In the above equation, the mean of intensity at *R* and R_{ref} (i.e., $\tilde{I}(R)$ and $\tilde{I}(R_{ref})$) obtained from data are used as the estimation of I(R) and $I(R_{ref})$. The mean of the intensity at range R for origin data and re-compensated data are shown in Figure 2c. The re-compensation can balance the intensity inhomogeneity between the range. The intensity of sonar image without re-compensation is dependent on range obviously in Figure 2a. The sonar image with the presented re-compensation has a uniform intensity distribution over different ranges in contrast as shown in Figure 2b and can contribute to subsequent segmentation.

2.3. Filter for Enhancement

Since SSS systems generate images by coherent processing of the scattered signals, they are severely suffered from speckle noise [30]. Due to high levels of speckle noise, the underwater images often are generally poor quality with low *SNR* and contrast. Therefore, the pre-processing should be done to suppress noises and enhance the images before segmentation, feature extraction, and object detection. A pre-processing filter based on the statistical properties of the sonar scattering signal is proposed to remove the speckle noise. We called the proposed filter *STDF*.



Figure 2. Comparison of a sonar image with and without time gain re-compensation. (**a**) A sonar image without time gain re-compensation. (**b**) A sonar image with time gain re-compensation. (**c**) The mean of the intensity at range R for origin data and re-compensated data.

2.3.1. Speckle Noise Analysis

Speckle is a granular interference that inherently exists in and degrades the quality of the active radar, medical ultrasound, and sonar images [31]. In a sidescan sonar imaging system, the echo signal can be regarded as the coherent accumulation of backscatter signals generated by a large number of scattering points.

$$Z = u e^{j\phi} = \sum_{i=0}^{N-1} z_i e^{j\theta_i}$$
(4)

where $Z = ue^{j\phi}$ is the overall impulse response of the SSS imaging system, and z_i , θ_i as well as N are the signal intensity, phase of the *i*th scattering point, and the number of scattering points within a resolution cell, respectively. If θ_i is a uniform distributed random variable taking values between 0 and 2π , and z_i meets independent and identical distribution, the probability density function of u can be written as follows by using the central limit theorem [19]. where a is a constant:

$$p(u) = \frac{u}{a^2} exp\left(-\frac{u^2}{2a^2}\right) \tag{5}$$

The ratio of the deviation σ and the mean μ (RDM) of the above distribution can be easily calculated as

$$\frac{\sigma}{\mu} = \sqrt{\frac{2 - \pi/2}{\pi/2}} \approx 0.523 \tag{6}$$

Speckle is caused by the coherent accumulation of backscattered signals and can be modeled as multiplicative noises [20]. When the number of scatters which have random phases θ with uniform distribution tends to infinity, the above result will be obtained. The seabed reverberation areas are more consistent with the assumption than the highlight areas with targets. Experiments show that the RDM in the seabed reverberation areas is closer to the theoretical value of 0.523, and the RDM of the highlight areas is higher than that of seedbed reverberation areas. The RDM of highlight areas (μ_h/σ_h) and seabed reverberation areas (μ_b/σ_b) in 10 sonar images are calculated and shown in Table 1.

Sonar Images	σ_h/μ_h	σ_b/μ_b
Image 1	0.8926	0.5590
Image 2	0.9266	0.5622
Image 3	0.9957	0.5665
Image 4	0.8484	0.5582
Image 5	0.8347	0.5537
Image 6	0.8882	0.5648
Image 7	0.9822	0.5639
Image 8	0.9187	0.5762
Image 9	1.0921	0.5768
Image 10	0.9791	0.5766

Table 1. Ratios of mean and deviation of highlight areas and seafloor reverberation areas in 10 sonar images.

2.3.2. The Proposed Filter STDF

The mean filtering can suppress the speckle noises and improve image quality by calculating the mean of the neighborhood to estimate the value of the centroid pixels [32]. As illustrated above, the mean of the neighborhood of a pixel can be estimated by estimating the standard deviation of the neighborhood, because the RDM is a constant. In our method, the deviation of each pixel is chosen instead of the mean of the pixel to enhance the targets and suppress speckle noise. The above description shows that the statistical properties of the highlight areas with targets are different from those of the seabed reverberation areas, i.e., the RDM of the highlight areas is higher than that of seabed reverberation areas. Therefore, using the standard deviation as an estimate for each pixel point can widen the margin between classes, amplify the signal of object areas, suppress the signal of seafloor background, and then improve the SNR and contrast of sonar images. The filter is called STDF and is defined as calculating the standard deviation of its 8-neighborhood as the estimated illumination of the pixel as $Y(x,y) = \sqrt{\frac{1}{n-1}\sum_{s \in S} (J(x_s,y_s) - \overline{J}(x,y))^2}$, where $J(x_s, y_s)$ is the intensity at pixel (x_s, y_s) in the original sonar images. $\overline{J}(x, y)$ and *n* are the mean of pixels' intensity and the number of pixels, in the 8-neigborhood S centering at (x, y). Y(x, y) is the STDF result at pixel (x, y). The filter can be calculated in matrix form and be expressed by Equation (7):

$$I_{STDF} \odot I_{STDF} = \frac{1}{n-1} \{ (I \odot I) \otimes h \} + \frac{1}{n(n-1)} \{ (I \otimes h) \odot (I \otimes h) \}$$
(7)

where *I* is the original sonar images, I_{STDF} is the *STDF* result. *h* is a filter operator of size 3×3 whose elements are all 1. \otimes indicates the convolution operation. \bigcirc indicates the dot multiplication.

2.4. Proposed Segmentation

Because the sidescan sonar typically has a centimeter range resolution and a large range of hundreds of meters, there are numerous pixels in a single scan line. Most current segmentation (e.g., active contour [16,18], Fuzzy C-means [11], MRF [12–15]) are algorithms with high numbers of iterations and struggle to meet real-time requirements. Each iteration step of these algorithms involves the processing of reverberant regions of the seabed. The highlight regions with targets occupy a small proportion of the target, while the seabed reverberant regions occupy a large area. Aiming at concentrating on the highlight areas and shadow areas, the method based on seeded region growing (SRG) [33] is considered which focuses on the growth of the boundary pixels rather than on global processing. The region growing algorithm can correctly separate the regions that have similar properties as defined and is stable to noise. Yu et al. [34] proposes an image segmentation method that uses the region growing technique and edge strength information to improve the traditional MRF-based approaches. Jiao et al. [35] proposes an SRG algorithm with superior performance based on the Gaussian pyramid, which automatically selects seed points and optimizes

the growth path. Wu et al. [36] proposes an effective segmentation method combined with MRF and region growing for sonar images. However, the speed of current region growing based methods combined with MRF does not meet the real-time requirement, and the region growing methods not employing the statistical features of sonar lack robustness.

The proposed segmentation method is composed of the following main steps: first, seed pixels or regions as started points are selected, similar to [33], and then, pixels with the same or similar properties in the vicinity of the seed pixel are merged into the seed pixel domain according to the growth criteria. In this paper, the seed selection principle and the similarity criterion for the region growing step are redesigned for segmenting sonar images fast and accuracy. The seeds are selected from the foreground (i.e., the highlight and shadow areas) to avoid the long-time iterative growth of the seabed reverberation area. The seed points are grouped into two sets A_i , i.e., the highlight areas A_1 and shadow areas A_2 . The selection should be adaptive with no human interaction. Aiming at reducing the appearance of false targets without losing the foreground areas, every highlight region which needs to be extracted should have one seed point at least, and the probability that the seed points are in the seabed background area should be reduced as much as possible. Therefore, the method of likelihood ratio testing is adopted on the statistical distribution of the sonar images. Take the distinction between the seabed and the highlight points as an example. $y_s \in \{l_0, l_1\}$ stands for the class of image points, and l_0 stands for the "highlight" label while l_1 corresponds to the "sea-bottom" class. Type I error is the rejection of a true null hypothesis, and type II error is failing to reject a false null hypothesis. The criterion for the selection of the rejection domain is to choose the rejection domain with the smallest possible probability of making the type II error while ensuring that the probability of making the type I error does not exceed a certain level. The criterion is consistent with the requirements of seed points selection.

According to the Section 2.3.1, the luminance within the seafloor reverberation area follows the Rayleigh distribution as Equation (5). In this section, we assume that the luminance within the seafloor reverberation area follows the Weibull distribution including the Rayleigh distribution, and the highlight and shadow areas follow the Gaussian law [15]. The conditional density function for the sea-bottom is modeled by a Weibull law and for highlight is modeled by a Gauss law as

$$P(u|\lambda,k,l_1) = \begin{cases} \frac{k}{\lambda} \left(\frac{u}{\lambda}\right)^{k-1} e^{-\left(\frac{u}{\lambda}\right)^k} & u \ge 0\\ 0 & u < 0 \end{cases}$$
(8)

$$P(u|\sigma_g, \mu_g, l_0) = \frac{1}{\sqrt{2\pi\sigma_g^2}} exp(-\frac{(u-\mu_g)^2}{2\sigma^2})$$
(9)

where k, λ , σ_g , and μ_g are parameters to be estimated. To estimate the parameters for the seabed area luminance model precisely, pre-segmentation will be conducted using the threshold segmentation method with *STDF* images. The highlight regions, the shadow regions, and the seabed regions obtained by the pre-segmentation refer to the highlight data \mathcal{H} , the shadow data \mathcal{S} , and the seabed data \mathcal{B} . The points numbers in three kinds of regions are $n_{\mathcal{H}}$, n_S , and n_B . The parameters of the Weibull model are estimated based on the raw data after time gain re-compensation but without *STDF*, as *STDF* will change the luminance distribution. The following expression of the maximum likelihood estimator can be obtained as follows, where $n_{\mathcal{F}}$ stands for $n_{\mathcal{H}}$ or n_S , and \mathcal{F} stands for \mathcal{H} or \mathcal{S} .

$$\hat{\lambda} = \left(\frac{1}{n_B} \sum_{i \in \mathcal{B}}^{n_B} u_i^{\ k}\right)^{\frac{1}{k}} \tag{10}$$

$$\frac{1}{\hat{k}} = \frac{\sum_{i \in \mathcal{B}} u_i^k \ln u_i}{\sum_{i \in \mathcal{B}} u_i^k} - \frac{1}{n_B} \sum_{i=1}^{n_B} \ln u_i$$
(11)

$$\hat{\mu}_g = \frac{1}{n_F} \sum_{i \in \mathcal{F}} u_i \tag{12}$$

$$\hat{\sigma_g}^2 = \frac{1}{n_F} \sum_{i \in \mathcal{F}} (u_i - \hat{\mu_g})^2 \tag{13}$$

The likelihood ratio is calculated as

$$\beta(u) = \frac{P(u|H_0)}{P(u|H_1)} = \frac{P(u|\sigma_g, \mu_g, l_0)}{P(u|\lambda, k, l_1)}$$
(14)

The assumption test is as Equation (15), where η is the threshold for seed points selection. η is chosen by restricting the probability of making type I error e_1 to a certain percentage. Next, a test is chosen which minimizes class e_2 . When $\beta(u) \leq \eta$, pixels are assigned to the seed point sets; else, pixels are assigned to unallocated pixels.

$$\beta(u) \stackrel{H_0}{\underset{H_1}{>}} \eta \tag{15}$$

The details of the process of region growing are shown in Appendix A. Firstly, the neighbors of the points in A_i are marked as boundary pixels. Next, it is judged whether the bounding pixels set is empty. If the set is not empty, it is judged whether the points in the set satisfy the similarity criteria. The similarity criteria can be represented by the formula: $I_n \ge \overline{\mu_B} + \sigma_B$, where I_n represents the intensity value of points in the boundary set. $\overline{I_B}$ and σ_B represent the average intensity and the standard deviation of the intensity of the seabed reverberation area. If the points meet the similarity criteria, they are assigned to A_i , otherwise they are labeled as unallocated points. The iteration continues until the boundary set is empty, and all unallocated points are marked as seabed reverberation points.

3. Results and Discussion

The testing environment of the algorithm proposed in this paper is as follows: Intel (R) Core (TM) i58250U CPU, 8G RAM (Intel, Santa Clara, CA, USA), Windows, Matlab.

3.1. Dataset

In order to verify the effectiveness and superiority of the proposed algorithm, this paper conducts experiments on real sonar datasets. The seabed materials include sandy bottom and muddy bottom. The data are collected by a 500 kHz multi-beam sonar. Each scan line on each side covers a range of 175 m with 7000 pixels. The datasets contain complex landforms, such as terraces, ridges, ravines, gravel areas, etc.

3.2. Results of Proposed STDF

Aiming at evaluating the performance of the *STDF*, the results are compared with mean filter, median filter, Gaussian filter, and Lee filter [20] in Table 2 in terms of *SNR* and contrast, defined as

$$SNR = (I_M - I_B) / \sigma_B \tag{16}$$

$$contrast = \left(\overline{I_H} - \overline{I_B}\right) / \overline{I_B} \tag{17}$$

where I_M is the maximum intensity value of the highlight areas; $\overline{I_B}$ and $\overline{I_H}$ represent the average intensity of the highlight areas and the seabed background areas, and σ_B represents the standard deviation of the intensity of the seabed background area. The results demonstrate that the proposed *STDF* has higher *SNR* and contrast than other methods. The RDM of the highlight areas is higher than that of the seabed reverberation areas, thus calculating the standard deviation of its 8-neighborhood as the estimated illumination of the pixel can significantly improve *SNR* and contrast. The visual results of different filters are shown in Figure 3. The subgraph (a) shows a raw SSS image without filtering, and subgraphs (b), (c), (d), and (e) respectively show the results for mean filtering, median filtering, Gaussian filtering, Lee filtering, and the *STDF* proposed in this paper. The results in (b), (c), (d) suffer from details missing, namely, the tail of the cable in white boxes is not clear. The results in (e) still have obvious intensity inhomogeneity in the seabed area indicated by the green boxes. In addition, as illustrated in blue boxes, the results of *STDF* shown in (f) have better contrast of highlight areas with seabed areas than other methods, i.e., the *STDF* can improve the *SNR* and contrast of the sonar images meanwhile preserving details and edges. To verify the sensitivity characteristics with respect to the resolution of the sonar images, sonar images are downsampled in the column with the sampling rate reduction factors of 2, 4, and 8 in the column's direction as shown in subgraphs (A), (B), and (C) in Figure 4, respectively. The subgraphs (a)–(f) in Figure 4 correspond to the same methods as Figure 3. As the downsampling rate increases, the resolution of the image decreases. It can be seen that the proposed *STDF* still has finer edges and higher contrast than other methods under low resolution conditions. It proves that the proposed method has low sensitivity.

Method	SNR	Contrast
Raw image	110.97	1.09
mean filter	87.53	1.08
median Filter	83.84	0.99
Gaussian Filter	84.54	1.08
Lee Filter	113.17	1.07
STDF	126.42	1.36





Figure 3. Results of the different filters. The subgraph (**a**) shows a raw SSS image without filtering, and subgraph (**b**–**f**), respectively, show the results for mean filtering, median filtering, Gaussian filtering, Lee filtering, and the *STDF* proposed in this paper.



Figure 4. Results of different filters for downsampled sonar images. The column downsampling factors for subgraph (**A**), subgraph (**B**), and subgraph (**C**) are 2, 4, and 8, respectively. (**a**) Raw SSS image. (**b**) Results for mean filtering. (**c**) Results for median filtering. (**d**) Results for Gaussian filtering (**e**) Results for Lee filtering. (**f**) Results for proposed *STDF*.

3.3. Segmentation Results of Proposed Method

The method is compared with two robust unsupervised segmentation methods, namely, the fuzzy C-means clustering [11] and the active contour method [18]. The cor-

responding ground-truth segmentation maps of the real sonar images are obtained by manual segmentation. The segmentation accuracy ρ is used for the overall evaluation of the whole segmentation results, which is calculated as follows.

$$p = N_c / N_t \times 100\% \tag{18}$$

where N_c is the number of correctly segmented pixels, and N_t is the total number of pixels. The ratio of correct to incorrect is defined as $\zeta = N_c / (N_t - N_c)$.

The quantitative measures of the results are given in Table 3. Our proposed algorithm is named RGLT I (w/o STDF) or RGLT II (w/ STDF). Table 3 shows that the performance of both proposed methods RGLT I and RGLT II is better than that of the fuzzy C-means clustering and active Contour method in segmentation accuracy. In addition, the accuracy and the correct to incorrect ratio of RGLT II is higher than that of RGLT I, which indicates that STDF filtering not only contributes to the improvement of SNR, contrast, and visual effect but also improves the final segmentation effect significantly. Table 3 shows the average running times of the four segmentation methods used for 100×400 images segmentation. The proposed segmentation method is much faster than the fuzzy C-means and the active contour method. The data generation speed is 1 s/ping, and the speeds of RGLT I and RGLT II are 0.75 s/ping and 0.77 s/ping on average; thus, the proposed methods have an obvious advantage for real-time sonar detection. The results demonstrate that only growth in the highlight areas and shadow areas can effectively increase the speed of segmentation. Some segmentation results of the four methods are given in Figure 5 and are summarized in Table 4. As can be seen from Figure 5c, the fuzzy C-means method is sensitive to noise, and the problem of misclassifying the background into highlighted or shadowed areas occurs. As shown in Figure 5d, the active contour method appears to misclassify the background reverberant region into the highlight region when the area size of the three types of regions is very different, for example, when the bright region in the image is very small. In addition, for small targets, the segmentation results of the active contour method are not accurate enough, as shown in the red bounding box in Figure 5d. Figure 5e shows that our proposed RGLT I is robust to noise and has a low probability of background misclassification. This was a benefit from the low false alarm obtained by the maximum likelihood ratio test during the seed region selection phase. By comparing Figure 5e,f, it can be seen that the segmentation results are more accurate at the edges after using *STDF*. The details in the blue bounding box show that the growth of RGLT II is finer and more complete in the presence of noises than RGLT I. It shows that STDF can alleviate the interference of noise on the target edge.

Methods	Accuracy (%)	Correct to Incorrect Ratio	Running Time (s)
	(0.02	1.75	10.71
Fuzzy C-means	62.23	1.65	13.71
active contour	86.70	6.52	6.23
KGLI I	95.76	22.58	0.43
KGLI II	95.90	23.39	0.44

Table 3. Comparison of the overall accuracy and speed using several segmentation methods.

 Table 4. Summary of the visual effects of the segmentation methods.

Methods	Visual Effects
Fuzzy C-means	Is sensitive to noise and has the problem of misclassifying the background into highlighted or shadowed areas.
active contour	 Misclassify the background when the area size of the three types of regions is very different. For small targets, the segmentation results of the active contour method are not accurate enough.
RGLT I	Is robust to noise and has a low probability of background misclassification.
RGLT II	The growth of RGLT II is finer and more complete in the presence of noises than RGLT I.



Figure 5. Segmentation results of the four methods: (a) The real sonar images. (b) The corresponding manually labeled ground truth segmented images. The green areas represent the shadow areas, and the red areas represent the highlighted areas. (c) The segmentation results of fuzzy C-means. (d) The segmentation results of the active contour. (e) The segmentation results of RGLT I. (f) The segmentation results of RGLT II.

4. Conclusions

In this paper, an unsupervised segmentation algorithm that using likelihood ratio testing is presented, which can rapidly and accurately segment noisy sonar images. Firstly, a time gain re-compensation method is used to balance the intensity throughout the range. Then a filter referring to as *STDF* is proposed in this paper to enhance the *SNR* and contrast of the sonar image. The results demonstrate that the proposed *STDF* has higher *SNR* (126.42) and contrast (1.36) than other methods. The enhancement can preserve details as well as edges and be beneficial to the subsequent segmentation. Lastly, the proposed segmentation method, RGLT, is applied with real sonar images. The results demonstrate that the proposed method achieved favorable performance in both visual quality and qualitative quality in comparison with other outstanding methods. The performance of RGLT is better than that of the fuzzy C-means clustering and active Contour method in segmentation accuracy (95.90) and the correct to incorrect ratio (23.39). As for the running time, the proposed segmentation method is much faster than other methods, and thus, the proposed method has an obvious advantage for real-time sonar detection. In the future, we will study the integration of RGLT and detection algorithms to achieve real-time and intelligent detection and identification on AUVs in search and rescue at sea. Future work also involves generalizing existing algorithms to other sonar images, such as forward-looking sonar and synthetic aperture sonar.

Author Contributions: Conceptualization, X.X., X.W. and G.L.; methodology, X.X., X.W., G.L. and L.W.; software, X.W., L.W.; validation, X.W., L.W.; formal analysis, X.X., X.W. and G.L.; investigation, X.X.; X.W.; G.L. and L.W.; writing—original draft preparation, X.W. and L.W.; writing—review and editing, X.X. and G.L.; visualization X.W., L.W.; project administration, X.X. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by National Key R&D Program of China (2019YFB2204800).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank all subject members for their hard experimental work and active participation in the project.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Algorithm A1 The pseudo-code for the proposed segmentation process.
Seed selection based on likelihood ratio test
Input: The original image and the pre-segmentation using the threshold method.
Output: foreground points set A_i and sedbed reverberation points set.
Begin:
1: Estimate the distribution parameters using (10)–(13)
2: Calculate the likelihood ratio $\beta(u)$ using (14)
3: if $\beta(u) \leq \eta$ then
4: Pixels are assigned to the seed points sets A_i ;
5: else
6: Pixels are assigned to unallocated pixels.
7: end if
Region growing
8: Label boundary pixels set B and their candidate labels $i(x)$
9: while (<i>B</i> is not empty) then
10: If the points <i>x</i> in set <i>B</i> satisfy the similarity criterion then
11: Append x to A_i ;
12: else
13: Append <i>x</i> to unallocated pixels set.
14: end if
15: Upate the Label boundary pixels set <i>B</i> and their candidate labels $i(x)$
16: end while
End

References

- 1. Shang, X.; Zhao, J.; Zhang, H. Obtaining high-resolution seabed topography and surface details by co-registration of side-scan sonar and multibeam echo sounder images. *Remote Sens.* **2019**, *11*, 1496. [CrossRef]
- 2. Singh, H.; Adams, J.; Mindell, D.; Foley, B. Imaging underwater for archaeology. J. Field Archaeol. 2000, 27, 319–328. [CrossRef]
- 3. Barngrover, C.; Althoff, A.; DeGuzman, P.; Kastner, R. A brain–computer interface (BCI) for the detection of mine-like objects in sidescan sonar imagery. *IEEE J. Ocean. Eng.* 2015, *41*, 123–138. [CrossRef]
- 4. Fatan, M.; Daliri, M.R.; Shahri, A.M. Underwater cable detection in the images using edge classification based on texture information. *Measurement* **2016**, *91*, 309–317. [CrossRef]
- 5. Williams, D.P. Fast target detection in synthetic aperture sonar imagery: A new algorithm and large-scale performance analysis. *IEEE J. Ocean. Eng.* **2014**, *40*, 71–92. [CrossRef]
- Petillot, Y.; Reed, S.; Bell, J.M. Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder. In Proceedings of the OCEANS'02 MTS/IEEE, Biloxi, MI, USA, 29–31 October 2002; pp. 217–222. [CrossRef]
- 7. Celik, T.; Tjahjadi, T. A novel method for sidescan sonar image segmentation. *IEEE J. Ocean. Eng.* 2011, 36, 186–194. [CrossRef]
- 8. Yen, J.C.; Chang, F.J.; Chang, S. A new criterion for automatic multilevel thresholding. *IEEE Trans. Image Process.* **1995**, *4*, 370–378. [CrossRef]
- 9. Gou, S.P.; Chen, P.J.; Yang, X.Y.; Jiao, L.C. Image segmentation based on fussing multi-feature and spatial spectral clustering. In Proceedings of the 2008 Congress on Image and Signal Processing, Sanya, China, 27 May 2008; pp. 667–671. [CrossRef]
- Guillaudeux, S.; Daniel, S.; Maillard, E. Optimization of a sonar image processing chain: A fuzzy rules based expert system approach. In Proceedings of the OCEANS 96 MTS/IEEE Conference Proceedings. The Coastal Ocean-Prospects for the 21st Century, Fort Lauderdale, FL, USA, 23–26 September 1996; pp. 1319–1323. [CrossRef]
- 11. Ye, X.; Wang, L.; Wang, T. Sonar image segmentation on Fuzzy C-Mean using local texture feature. In Proceedings of the 2011 IEEE/ICME International Conference on Complex Medical Engineering, Harbin, China, 22–25 May 2011; pp. 115–120. [CrossRef]
- 12. Mignotte, M.; Collet, C.; Pérez, P.; Bouthemy, P. Three-class Markovian segmentation of high-resolution sonar images. *Comput. Vis. Image Underst.* **1999**, *76*, 191–204. [CrossRef]
- 13. Jiang, M.; Stewart, W.; Marra, M. Segmentation of seafloor sidescan imagery using Markov random fields and neural networks. In Proceedings of the OCEANS'93, Victoria, BC, Canada, 18–21 October 1993; Volume 453, pp. III456–III461. [CrossRef]
- 14. Choi, H.; Baraniuk, R.G. Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Trans. Image Process.* **2001**, *10*, 1309–1321. [CrossRef]

- 15. Mignotte, M.; Collet, C.; Perez, P.; Bouthemy, P. Sonar image segmentation using an unsupervised hierarchical MRF model. *IEEE Trans. Image Process.* 2000, *9*, 1216–1231. [CrossRef]
- Köhntopp, D.; Lehmann, B.; Kraus, D.; Birk, A. Segmentation and classification using active contours based superellipse fitting on side scan sonar images for marine demining. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26 May 2015; pp. 3380–3387. [CrossRef]
- Ye, X.; Wang, L.; Wang, T.; Wang, G. A level-set model driven by Bidimensional EMD for sonar image segmentation. In Proceedings of the 2011 IEEE International Conference on Mechatronics and Automation, Beijing, China, 7–10 August 2011; pp. 703–708. [CrossRef]
- 18. Huo, G.; Yang, S.X.; Li, Q.; Zhou, Y. A robust and fast method for sidescan sonar image segmentation using nonlocal despeckling and active contour model. *IEEE Trans. Cybern.* **2016**, *47*, 855–872. [CrossRef]
- 19. Song, Y.; He, B.; Liu, P. Real-time object detection for AUVs using self-cascaded convolutional neural networks. *IEEE J. Ocean. Eng.* **2019**, *46*, 56–67. [CrossRef]
- 20. Lee, J.S. Speckle analysis and smoothing of synthetic aperture radar images. *Comput. Graph. Image Process.* **1981**, *17*, 24–32. [CrossRef]
- 21. Frost, V.; Stiles, J.; Shanmugam, K.S.; Holtzman, J.; Smith, S. An adaptive filter for smoothing noisy radar images. *Proc. IEEE* **1981**, 69, 133–135. [CrossRef]
- 22. 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. [CrossRef]
- Huang, Y.; Li, W.; Yuan, F. Speckle Noise Reduction in Sonar Image Based on Adaptive Redundant Dictionary. J. Mar. Sci. Eng. 2020, 8, 761. [CrossRef]
- 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. [CrossRef]
- 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] [PubMed]
- 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]
- 27. Grabek, J.; Cyganek, B. Speckle Noise Filtering in Side-Scan Sonar Images Based on the Tucker Tensor Decomposition. *Sensors* **2019**, *19*, 2903. [CrossRef] [PubMed]
- 28. Priyadharsini, R.; Sree Sharmila, T.; Rajendran, V. A wavelet transform based contrast enhancement method for underwater acoustic images. *Multidimens. Syst. Signal Process.* **2018**, *29*, 1845–1859. [CrossRef]
- Chu, D.; Hufnagle, L. Time varying gain (TVG) measurements of a multibeam echo sounder for applications to quantitative acoustics. In Proceedings of the OCEANS 2006, Boston, MA, USA, 18 September 2006; pp. 1–5. [CrossRef]
- Park, J.-M.; Song, W.-J.; Pearlman, W. Speckle filtering of SAR images based on adaptive windowing. *IEE Proc. Vis. Image Signal Process.* 1999, 146, 191–197. [CrossRef]
- Shruthi, B.; Renukalatha, S.; Siddappa, M. Detection of Kidney Abnormalities in Noisy Ultrasound Images. Int. J. Comput. Appl. 2015, 120. [CrossRef]
- Abu, A.; Diamant, R. Enhanced fuzzy-based local information algorithm for sonar image segmentation. *IEEE Trans. Image Process.* 2019, 29, 445–460. [CrossRef]
- 33. Adams, R.; Bischof, L. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 1994, 16, 641–647. [CrossRef]
- Yu, Q.; Clausi, D.A. IRGS: Image segmentation using edge penalties and region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* 2008, 30, 2126–2139. [CrossRef]
- 35. Jiao, J.; Wang, X.; Zhang, J.; Wang, Q. Salient region growing based on Gaussian pyramid. *IET Image Process.* **2021**, *15*, 3142–3152. [CrossRef]
- Wu, J.; Guo, H. A method for sonar image segmentation based on combination of MRF and region growing. In Proceedings of the 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, India, 4–6 April 2015; pp. 457–460. [CrossRef]