

Article

Non-Local Means Hole Repair Algorithm Based on Adaptive Block

Bohu Zhao ¹, Lebao Li ^{1,2} and Haipeng Pan ^{1,*} 

¹ School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310000, China; 202130605365@mails.zstu.edu.cn (B.Z.); llb@zstu.edu.cn (L.L.)

² College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China

* Correspondence: pan@zstu.edu.cn

Abstract: RGB-D cameras provide depth and color information and are widely used in 3D reconstruction and computer vision. In the majority of existing RGB-D cameras, a considerable portion of depth values is often lost due to severe occlusion or limited camera coverage, thereby adversely impacting the precise localization and three-dimensional reconstruction of objects. In this paper, to address the issue of poor-quality in-depth images captured by RGB-D cameras, a depth image hole repair algorithm based on non-local means is proposed first, leveraging the structural similarities between grayscale and depth images. Second, while considering the cumbersome parameter tuning associated with the non-local means hole repair method for determining the size of structural blocks for depth image hole repair, an intelligent block factor is introduced, which automatically determines the optimal search and repair block sizes for various hole sizes, resulting in the development of an adaptive block-based non-local means algorithm for repairing depth image holes. Furthermore, the proposed algorithm's performance are evaluated using both the Middlebury stereo matching dataset and a self-constructed RGB-D dataset, with performance assessment being carried out by comparing the algorithm against other methods using five metrics: RMSE, SSIM, PSNR, DE, and ALME. Finally, experimental results unequivocally demonstrate the innovative resolution of the parameter tuning complexity inherent in-depth image hole repair, effectively filling the holes, suppressing noise within depth images, enhancing image quality, and achieving elevated precision and accuracy, as affirmed by the attained results.



Citation: Zhao, B.; Li, L.; Pan, H. Non-Local Means Hole Repair Algorithm Based on Adaptive Block. *Appl. Sci.* **2024**, *14*, 159. <https://doi.org/10.3390/app14010159>

Academic Editors: Silvia Liberata Ullo and Li Zhang

Received: 19 November 2023

Revised: 21 December 2023

Accepted: 22 December 2023

Published: 24 December 2023



Copyright: © 2023 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/>).

Keywords: depth image inpainting; non-local means; adaptive block; RGB-D camera

1. Introduction

In recent years, depth cameras have been widely utilized in various fields, such as autonomous driving [1–3], robot grasping [4–6], and simultaneous localization and mapping (SLAM) [7–9]. However, due to object material, specular reflection, or occlusion, the captured depth data often exhibits holes at the edges where objects come into contact with the background. Additionally, limitations in-depth sensor accuracy and range, as well as the influence of optical noise, sometimes result in the reception of ineffective reflection information, thereby degrading the quality of depth images.

Considerable research efforts have been devoted by researchers both domestically and internationally to effectively fill the holes in-depth images and obtain high-quality depth images. CHO et al. [10] estimated the boundaries using color images and converged from the edges to the target boundary, effectively addressing holes between edges and multiple edge pixels to achieve seamless inpainting of object boundaries near the camera through an iterative process of alternating filling directions. Xiang et al. [11] extended Criminisi's inpainting approach [12] by incorporating a block matching algorithm for identifying irregularly shaped voids and a hybrid matching algorithm for selecting optimal repair patches with minimal differences in hole regions, leading to successful restoration,

particularly for larger irregular holes. Zhang et al. [13] introduced an object-oriented segmentation method to segment different objects in color images, locate source regions for repair patches, and utilize a search algorithm to match missing depth information regions with suitable objects, ultimately filling depth information gaps through optimal repair patch selection within the segmented objects. Zhang et al. [14] introduced the statistical nearest neighbors distance measurement (SNNNDM) into the NLM filter (NLM-SNN) to measure the distance between two patches, thereby enhancing perceived image quality. Lin et al. [15] introduced a fixing algorithm for the Kinect depth image based on non-local means (NLM), utilizing weights calculated on the corresponding grayscale image through a distance factor and a value-consistency factor to fill holes in the depth image. These methodologies collectively leverage diverse strategies in hole repair techniques to effectively address challenges in in-depth image repair. However, the methods proposed in [10,11] are somewhat complex. The ultimate repair outcome in [13] is influenced by the effectiveness of image segmentation. Additionally, the method in [14] combines the statistical nearest neighbor distance measurement so that it is more complex. In contrast, the approach proposed by [15], inspired by [14], integrates distance factors to enhance restoration efficiency for Kinect cameras but encounters challenges in intricate parameter adjustments. Drawing inspiration from [15], this paper exploits the structural similarity between grayscale and depth images. It utilizes joint grayscale images to estimate depth values within hole regions and applies the non-local means algorithm [14–21] for effective hole repair. Successful restoration is achieved through the application of structural blocks and NLM-based hole filling.

However, in the deep image hole repair algorithm based on the NLM, a requisite task involves the a priori determination of dimensions for both the repair and search blocks. If the dimensions are excessively large, they will significantly increase the computational runtime of the algorithm. Conversely, if the sizes are overly small, they may potentially compromise the quality of the inpainting outcomes. Consequently, a natural question arises: is there a method that can automatically determine the optimal sizes for the search and inpainting blocks of varying dimensions in deep missing regions while ensuring exceptional performance in inpainting quality? It is the motivation for this paper.

The main contributions of this paper can be given as follows:

- (1) A depth image inpainting algorithm based on NLM is proposed to address the issue of depth image hole repair. The utilization of the NLM algorithm for mitigating noise and speckle artifacts through filtering operations is discussed in [18–21]. However, the investigations highlighted above primarily concentrate on the realm of filtering procedures and have not encompassed the integration of the NLM algorithm within the context of depth image for hole restoration.
- (2) By introducing intelligent block factors, we enable the automatic determination of optimal search and repair patch sizes for voids of different dimensions, resulting in a significant reduction of the intricate debugging efforts in engineering applications. The strategy of improving the performance of the algorithm by manipulating the pixel weights is discussed in [5,14–17]. Nonetheless, it is worth noting that the acquisition of these weights relies predominantly on manual calibration, a process that demands a substantial investment of time.

The remaining structure of this paper unfolds in subsequent sections. The depth image hole repair algorithm is introduced in Section 2. The experimental results are presented along with their corresponding analysis in Section 3. And conclusions are drawn in Section 4.

2. Depth Image Inpainting Algorithm

Section 2 introduces the depth image repair algorithm, which aims to address holes in-depth images caused by various factors. This algorithm involves a multi-step process, starting with image preprocessing and progressing to the implementation of the NLM hole-repairing algorithm. This section also discusses an algorithm improvement, termed

Adaptive Block-based NLM (AB-NLM), which enhances the robustness of the original NLM algorithm.

2.1. Image Preprocessing

Image preprocessing is a critical stage in in-depth image repair, serving to enhance image quality and prepare the data for subsequent repair algorithms. In this section, we delve into the significance of image preprocessing, its objectives, and the specific operations employed to align grayscale and depth images.

Depth images captured using cameras often encounter voids caused by surface reflections, occlusions, and other factors, particularly in the interior and boundary regions of objects. Image preprocessing plays a crucial role in improving image quality, enhancing visualization effects, reducing noise interference, extracting meaningful features, and adapting to specific processing requirements, which provides a solid foundation for subsequent image processing and analysis tasks.

In the image preprocessing stage, this paper employs the following operations: First, a Gaussian filter is applied to the color image to reduce noise. Then, the color image is converted to a grayscale image. Subsequently, a perspective transformation matrix is used to align the grayscale image with the depth image. Finally, aligned depth and grayscale images are obtained, and the depth image is overlaid onto the grayscale one, as shown in Figure 1c. It can be observed that the aligned depth image and grayscale image exhibit similar structures [22]. To enhance the accuracy and quality of depth image hole repair, the similarity in structure between the grayscale and depth images can be leveraged to predict and repair the holes present in the depth image.

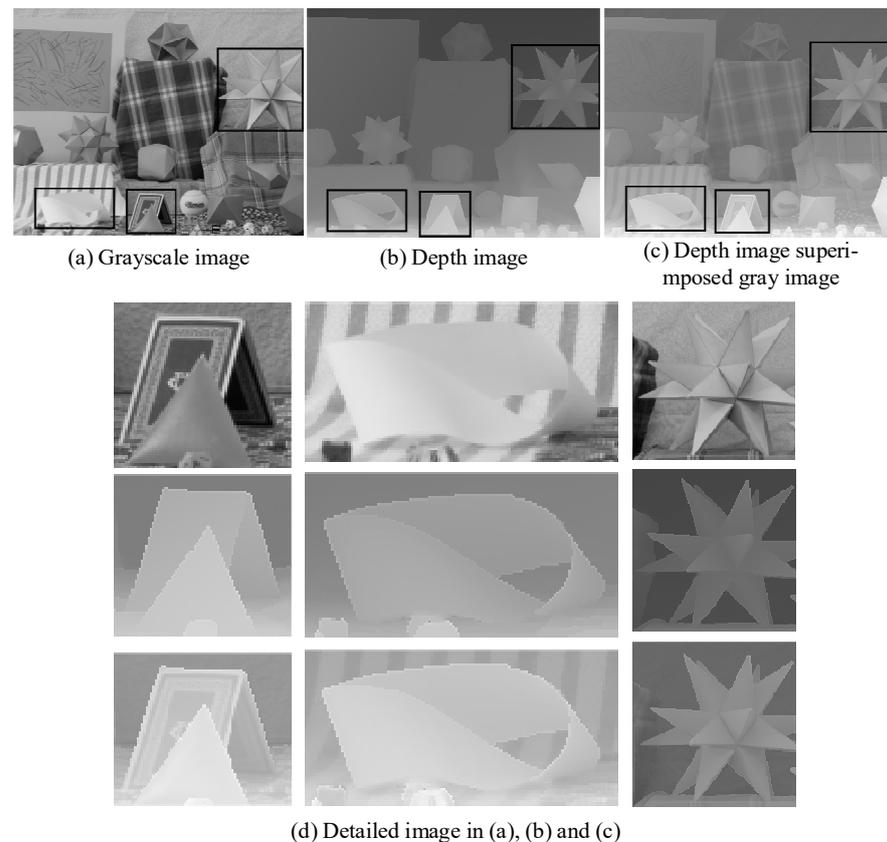


Figure 1. A schematic representation of the structural similarity between grayscale images and depth images.

2.2. Hole Filling Algorithm Based on Non-Local Means

Section 2.2 introduces the NLM hole-repairing algorithm, a block-based method selected for its unique approach to depth image repair. Unlike traditional pixel-based methods, NLM leverages vector-centered repairs and maximizes the use of redundant information in the image. This section outlines the fundamental principles of the NLM algorithm and its application to depth image hole filling.

The NLM Algorithm is a block-based method for depth image hole repair that considers the hole points as the centers and utilizes surrounding search blocks as processing units. By calculating the similarity between different repair blocks and the repair block centered at the hole point, similar weights are assigned to each repaired pixel. Then, a weighted average is employed to estimate the repaired depth value of the hole points. The NLM algorithm, which is different from traditional restoration methods based on individual pixels, uses the vector centered at the hole point for repair and makes full use of the redundant information in the image. The algorithm maximally preserves image details and enables efficient and accurate repair of in-depth images. The relationship between the search blocks and repair blocks of the hole point x and its neighboring non-hole point y is illustrated in Figure 2.

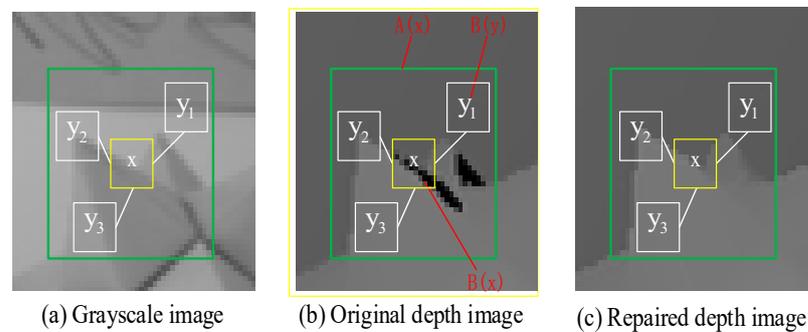


Figure 2. Illustration of the search block and the repair block.

According to Figure 2, it can be observed that the depth values within the holes are strongly correlated with non-hole points distributed in the neighborhood of the holes. Additionally, since the holes in-depth images often appear in the edge regions of objects, utilizing a single-point calculation for repair weights cannot effectively distinguish between points situated at the object edges and those within the smooth interior regions of objects. This leads to a certain degree of computational error. The proposed algorithm addresses this issue by computing similarity weights using vectors centered around the points within the holes rather than relying on individual pixels. This approach enhances the robustness of the algorithm in handling hole repair, providing a more accurate and reliable solution.

The weights in the NLM algorithm are composed of three components: grayscale similarity, distance similarity, and repairing weight. Grayscale similarity weight refers to the degree of likeness in grayscale values between the target hole point x within the grayscale image and the non-hole point y within the search block $A(x)$. The distance similarity weight concerns the extent of proximity between the remaining points within $A(x)$ and its central point. The computation of the repairing weight for each y within $A(x)$ is determined through the collaborative influence of grayscale weight and distance weight, thereby quantifying the similarity of depth values between x and y .

The grayscale similarity weight can be described as follows:

$$g(x, y) = e^{-\frac{\|v(N_x) - v(N_y)\|_{2,a}^2}{h^2}} \quad (1)$$

where h is the decay rate of the grayscale weight $v(N_x)$ and $v(N_y)$ represents the vectors of x and y , respectively. These two vectors are composed of pixels from the repair blocks centered at x and y , with a size of $q \times q$, denoted as $B(x)$ and $B(y)$ respectively. And the val-

ues are listed from left to right and from top to bottom. The Gaussian-weighted Euclidean distance is denoted by $\| \cdot \|_{2,a}^2$, where a is the standard deviation of the Gaussian kernel.

The distance similarity weight can be described as follows:

$$d(x, y) = e^{-\frac{\|x-y\|^2}{\sigma^2}} \tag{2}$$

where σ controls the decay rate of the exponential function and $\| \cdot \|^2$ represents the Euclidean distance between points x and y .

The repair weight can be described as follows:

$$r(x, y) = \frac{g(x, y)d(x, y)}{Z_1(x)} \tag{3}$$

where to ensure that $\sum_{y \in A(x)} r(x, y) = 1$, the normalization constant $Z_1(x)$ is introduced.

Before calculating the weights, the sizes of the search block $A(x)$ and the repair block $B(x)$ of the hole points x need to be manually set. The selection of the search and repair block sizes directly affects the algorithm’s runtime and repair effectiveness. Therefore, it is essential to strike a balance between the repair time and repair effectiveness of the algorithm by enabling the algorithm itself to compute appropriate sizes for the repair and search blocks. This encapsulates the challenges confronted by the NLM hole repair algorithm.

2.3. Algorithm Improvement

In response to the challenges present in the NLM hole repair algorithm, this section undertakes its enhancement and introduces a novel approach termed the adaptive block-based NLM hole repair algorithm. The detailed methodology is delineated as follows: The principle of the NLM hole repair algorithm based on adaptive blocks is shown in Figure 3.

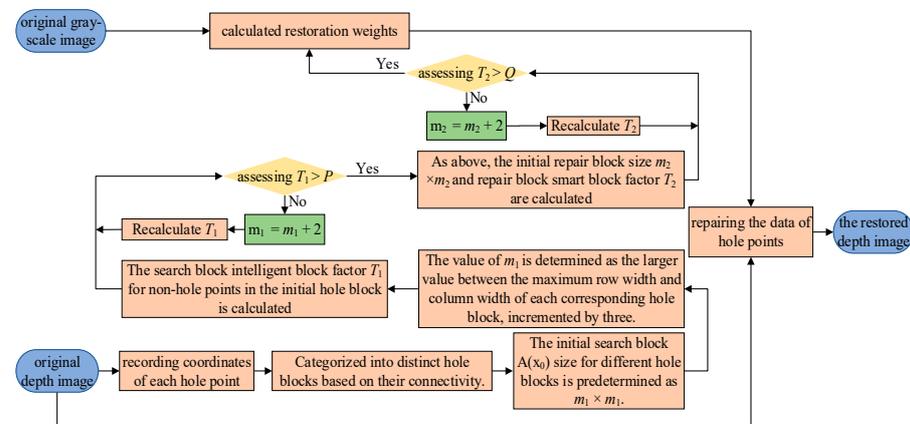


Figure 3. Framework of an Adaptive Block-based Non-Local Means Hole Repair Algorithm.

Step 1: Aligned depth images and grayscale images captured from the same viewpoint using an RGB-D camera facilitate the direct extraction of texture information from grayscale images, enabling algorithms to perform depth image restoration.

Step 2: The coordinates of each hole point are recorded, and the computation of connected components within the hole regions is performed. The initial value for the search block $A(x)$ is predetermined as $m_1 \times m_1$, where m_1 is the larger value between the maximum row width and the maximum column width of the respective hole block, incremented by three. The value of m_1 is defined as below:

$$m_1 = \max\{(X_r - X_l), (Y_t - Y_d)\} + 3 \tag{4}$$

where X_r and X_l represent the maximum and minimum values of the column indices of the hole points within the hole block, respectively. Y_t and Y_d represent the maximum and minimum values of the row indices of the hole points within the hole block, respectively.

Step 3: An intelligent block factor T is introduced. The factor T dynamically adjusts the initial value of the search block $A(x)$ based on the proportion of non-hole points within the hole block. Through T , a large-size search and repair block is used to repair a large area of hole block. Conversely, for small-area hole blocks, small-sized search and repair blocks are utilized. This dynamic adjustment enhances the generality of our algorithm, enabling it to adapt to various hole sizes and image features. The intelligent block factor T is defined as

$$T = \frac{\sum_{k \in H(p)} I_k}{|H(p)|} \quad (5)$$

where $|H(p)|$ represents the sum of the total number of pixels in the image and I_k denotes the non-hole points within the target block.

Step 4: The size of the repair block $A(x)$ is determined as follows: We calculate the proportion T_1 of non-zero pixels within the $m_1 \times m_1$ neighborhood of the hole point. If T_1 is greater than the given threshold P for the proportion of non-hole points in $A(x)$, then the final search block size is set as $p = m_1$. Otherwise, increase m_1 by 2 ($m_1 = m_1 + 2$), and reevaluate the relationship between T_1 and P .

Step 5: The size of the repair block $B(x)$ is determined as follows: The initial value of the repair block is set as $m_2 \times m_2$, where $m_2 = m_1 - 3$. The proportion T_2 of non-zero pixel points within the $m_2 \times m_2$ neighborhood of the hole points is calculated. Should T_2 surpass Q (where Q represents the threshold for the proportion of non-vacant hole points within $B(x)$), the ultimate determination for the size of the rectification block is $q = m_2/2$. Otherwise, m_2 is incremented by 2 ($m_2 = m_2 + 2$), and the relationship between T_2 and Q is reevaluated.

Step 6: For each hole point x , calculate the repair weight $r(x, y)$ for each non-hole point y in the neighborhood $A(x)$ of x . Utilize the repair weights to restore the hole point. By iterating through all hole points in the depth image and performing the same operation, the final restoration of the depth image is achieved.

3. Experimental Results and Analysis

The algorithm presented in this study is implemented on a system with the following specifications: Windows 10 operating system, 2.80 GHz Intel(R) Core(TM) i7-7700HQ CPU, and 16 GB of memory. The experiments are conducted using Visual Studio 2019 and OpenCV 4.6.0. To evaluate the performance of the proposed depth image optimization algorithm, employ the NLM algorithm, the Multiple Edge Converge Inpainting Algorithm (MECI) [10], and the AB-NLM algorithm as comparative benchmarks.

The experimental phase involves the utilization of fifty sets of depth and grayscale images sourced from the Middlebury dataset [23], along with one hundred and twenty-six self-captured sets of depth and grayscale images acquired through Orbbec's imaging system. The chosen scenes for evaluation consist of five intricate scenarios from the Middlebury dataset, namely "Art", "Teddy", "Moebius", "Recycle", and "Jadeplant". Additionally, five self-captured scenes from the Orbbec dataset, identified as "Bag", "Case", "Corner1", "Corner2", and "Miscellany", constitute the designated test set employed for the assessment of visual performance. The depth image dataset are now partitioned into a training set comprising 70% (123 images), a validation set consisting of 15% (26 images), and a test set comprising the remaining 15% (27 images). The partitioning process are randomized, ensuring the maintenance of data balance between the training and test sets.

In certain studies, metrics are computed as part of an objective analysis to assess the correspondence between reconstructed depth images and the unprocessed raw depth images. Nevertheless, owing to the existence of holes, the original depth images exhibit incompleteness. Metrics comparing the original and reconstructed depth images may introduce inaccuracies into the analysis and assessment of the efficacy of the restoration

process. To ensure precise effectiveness evaluation, all depth images within the training set undergo manual augmentation, wherein holes are filled with ground truth depth values.

The evaluation of the results is carried out both qualitatively and quantitatively to assess the effectiveness of the proposed method. The first set of experiments utilizes the Middlebury stereo matching dataset. The second set of experiments involves using grayscale and depth images captured by the Orbbec [24,25] Astra S sensor to perform depth restoration in indoor scenes.

Based on extensive experimentation, the threshold values for the proportion of non-hole points in $A(x)$ and Q in $B(x)$ are determined to be 0.7 and 0.5, respectively. The decay rate for the grayscale weight h is set to 2, and the decay rate for the distance weight σ is also set to 2. Additionally, the standard deviation of the Gaussian kernel, denoted as a , is set to 2.

3.1. Qualitative Evaluation

In this section, we present a qualitative evaluation of the proposed algorithm's performance on two distinct datasets: the Middlebury dataset and data captured using the Orbbec Astra S monocular structured light depth camera. The assessment entails a detailed analysis of the algorithm's effectiveness in hole repair, with a focus on specific scenes and comparisons with existing algorithms. Scrutinizing the algorithm's performance on the Middlebury dataset emphasizes subjective quality assessments, while the subsequent exploration of the Orbbec Depth Camera Data highlights the algorithm's adeptness in addressing hardware-specific challenges.

3.1.1. Result on Middlebury Database

The proposed algorithm is applied to the Middlebury dataset, which contains the depth images of 2014 and 2021 and their corresponding grayscale images for depth recovery. Five scenes from the Middlebury dataset, namely "Art", "Teddy", "Moebius", "Recycle", and "Jadeplant", are specifically selected to compare and evaluate the subjective quality of the restoration results.

From Figure 4, it can be observed that the three algorithms perform well in repairing small areas and internal voids. However, the MECI algorithm exhibits errors when dealing with edge holes, as evident in the repair details of "Teddy" and "Jadeplant" shown in Figure 4c. Additionally, the NLM algorithm shows issues with edge distortion and blurring when repairing large voids, as demonstrated in the repair details of "Recycle" and "Jadeplant" in Figure 4d.

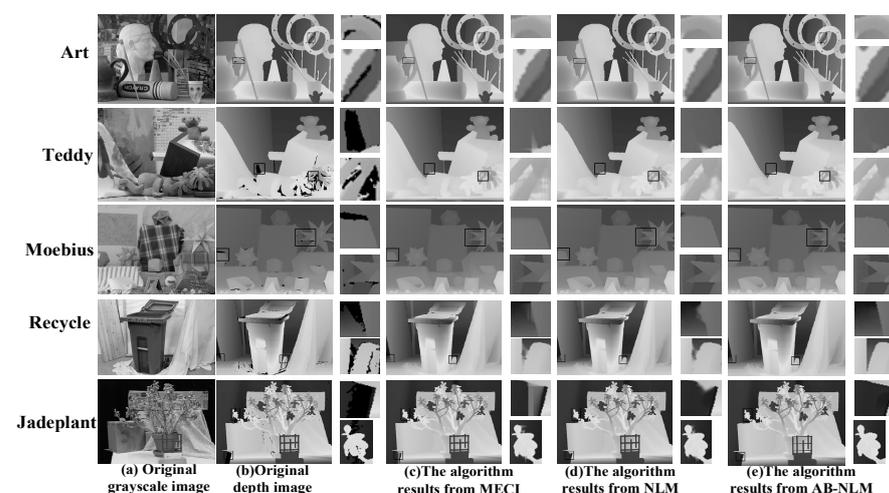


Figure 4. Illustrates a comparative evaluation of hole-inpainting outcomes from the MECI, NLM, and AB-NLM algorithms using the Middlebury dataset.

The AB-NLM algorithm proposed in this paper achieves adaptability to different sizes of hole blocks by automatically adjusting the search and repair block sizes. For larger hole blocks, larger repair blocks are selected, while smaller repair blocks are used for smaller hole blocks. This approach yields excellent results for hole blocks of various sizes. As shown in Figure 4e, the repaired images not only effectively restore the hole areas but also enhance the details of occluded objects, resulting in sharper edges.

3.1.2. Result on Orbbec Depth Camera Data

Orbbec is a domestic company specializing in 3D cameras and computer vision technology. It is dedicated to the research and development, as well as the production, of high-performance, high-precision, and cost-effective 3D cameras and sensors. In this study, we utilize the Orbbec Astra S monocular structured light depth camera, which is suitable for scenes with less stringent accuracy requirements and within a distance range of 40 cm to 200 cm.

Due to the hardware design of the Orbbec camera, there are data missing in several leftmost columns of the captured depth image, as shown in Figure 5b. From the details observed in Figure 5c,d, it can be deduced that the MECI algorithm successfully accomplishes the hole recovery, yet disparities from the ground truth depth map persist. The contours of objects exhibit a certain degree of blurring. In comparison, the algorithm presented in this paper achieves more accurate object contours in most cases and demonstrates superior performance in filling the holes within the depth map. However, the depth restoration of transparent objects is not optimal in our algorithm, as illustrated in the restoration details of the “Corner2” scene in Figure 5d. The repaired values for the missing region in the computer host area are close to the depth values of the curtain behind it, which is not appropriate for that region.

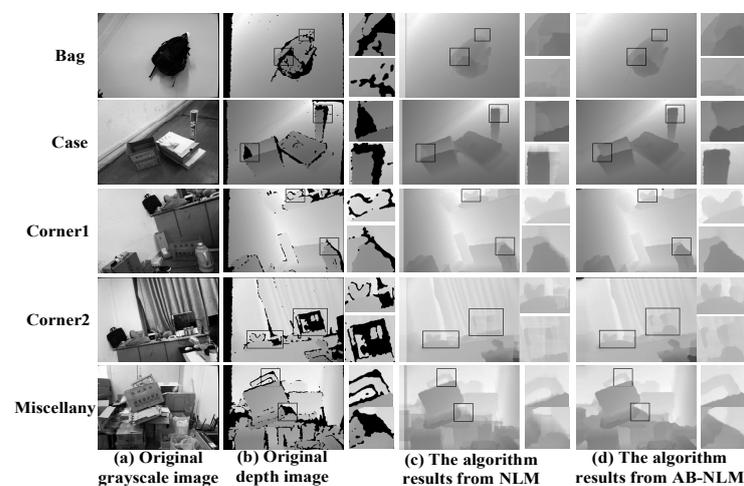


Figure 5. Presents a comparison of hole inpainting results between the NLM and AB-NLM algorithms utilizing a self-constructed dataset from the Orbbec camera.

3.2. Quantitative Evaluation

In this section, we evaluate the efficacy of the aforementioned methods in the Middlebury dataset using five metrics: Root Mean Square Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Depth Error (DE), and Average Local Mean Error (ALME). These metrics serve to measure the quality of the restoration outcomes.

In the restoration of in-depth images, RMSE is employed to quantify the disparity between the repaired image and the original image. A lower RMSE value indicates that the restored image closely approximates the original image. The RMSE is defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{(x,y) \in D} (d_g(x,y) - d_i(x,y))^2} \quad (6)$$

where n represents the number of pixels contained in D and (x, y) denotes the pixel coordinates. $d_g(x, y)$ and $d_i(x, y)$ refer to the depth values of the original image and the restored image at point (x, y) , respectively.

PSNR is commonly utilized to assess the quality of repaired images. A higher PSNR value signifies a reduced disparity between the repaired image and the original image, highlighting the commendable performance of the algorithm. The PSNR is defined as follows:

$$PSNR = 10 \log_{10} \left(\frac{\max(d_i)^2}{MSE} \right) \quad (7)$$

where $\max(d_i)$ represents the maximum depth value of the restored depth image, and MSE represents the mean square error between the original and restored images.

SSIM measures the structural similarity of image quality between two images. We strive for an SSIM value close to one to ensure a high degree of consistency between the repaired image and the original image in terms of structural features. The SSIM is defined as follows:

$$SSIM = \frac{(2\mu_{xy} + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)$$

where μ_x and μ_y represent the mean values of depth values $d_g(x, y)$ and $d_i(x, y)$, respectively. σ_x^2 and σ_y^2 represent the variances of $d_g(x, y)$ and $d_i(x, y)$, respectively. μ_{xy} and σ_{xy} represent the covariances of $d_g(x, y)$ and $d_i(x, y)$, respectively. Additionally, C_1 and C_2 are chosen as 0.0001 and 0.0009, respectively, to prevent division by zero and avoid potential system errors.

DE conducts a holistic error measurement on the repaired depth images, utilizing absolute depth error to quantify the disparities between them. A smaller absolute depth error indicates a more favorable restoration outcome. The definition of DE is as follows:

$$DE = \frac{1}{N} \sum_{(x,y)} |d_g(x,y) - d_i(x,y)| \quad (9)$$

where N represents the total number of pixels in the image.

ALME involves a comparison of the local means between the repaired and original depth images, with a smaller value indicating greater consistency of the repaired depth image with the original one in terms of local means. The definition of the local mean error for the k th local region is as follows:

$$LME_k = \frac{1}{m} \sum_{(x,y) \in R} |d_g(x,y) - d_i(x,y)| \quad (10)$$

where R denotes a fixed-size local region (a 10×10 window) that sequentially slides 10 steps from left to right and top to bottom across the image. m represents the pixels within R , resulting in a value of 100. The definition of ALME is expressed as follows:

$$ALME = \frac{1}{M} \sum_{k=1}^M LME_k \quad (11)$$

where M represents the total number of local regions in the image.

To assess the robustness of this research findings, Monte Carlo simulation was employed to estimate the 95% confidence intervals for various evaluation metrics. For RMSE, experimental results based on the dataset yielded a 95% confidence interval of [0, 10]. The confidence intervals for SSIM, PSNR, DE, and ALME were determined as [0, 1], [0, 50] (with values above 30 considered good and above 40 considered excellent for PSNR), [0, 40], and [0, 30], respectively.

As shown in Figure 6, an in-depth analysis of the Root Mean Square Error (RMSE) for the MECI, NLM, and AB-NLM algorithms on various images reveals noteworthy observations. On the Art image, AB-NLM exhibits a significantly lower RMSE (3.297) compared to MECI (6.291) and NLM (5.294). Similarly, in the Teddy image, AB-NLM demonstrates superior performance, with an RMSE of 1.873, exhibiting a significant difference compared to NLM (2.358) and MECI (3.185). This trend persists across other images (Moebius, Recycle, and Jadeplant), showcasing the comparatively more accurate depth image restoration capabilities of AB-NLM.

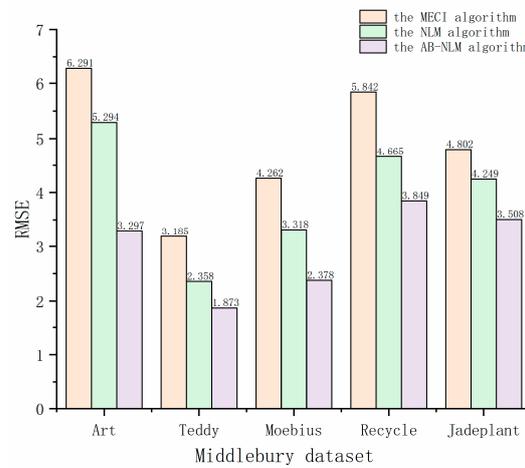


Figure 6. Displays a comparative analysis of RMSE results among the MECI, NLM, and AB-NLM algorithms based on the Middlebury dataset.

Examining the Peak Signal-to-Noise Ratio (PSNR) results in Figure 7 reveals the consistent superiority of the AB-NLM algorithm across diverse images. For instance, on the Art image, AB-NLM achieves a remarkable PSNR of 45.148, surpassing MECI (42.156) and NLM (43.153). This pattern extends to other images (Teddy, Moebius, Recycle, and Jadeplant), underscoring the consistently higher PSNR values and reaffirming AB-NLMs robust performance in image reconstruction tasks.

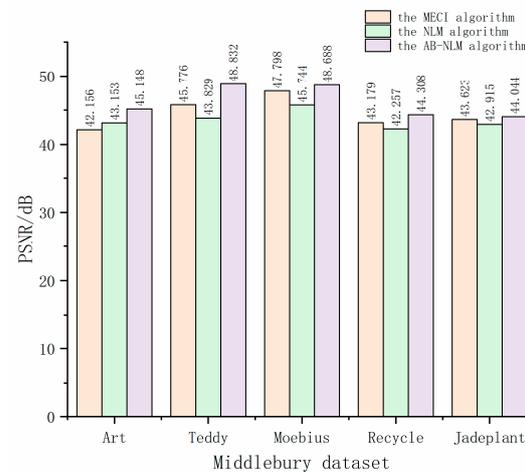


Figure 7. Displays a comparative analysis of PSNR results among the MECI, NLM, and AB-NLM algorithms based on the Middlebury dataset.

Figure 8 presents a scientific analysis of the Structural Similarity Index (SSIM) results for the MECI, NLM, and AB-NLM algorithms on various images. Notably, AB-NLM consistently demonstrates higher SSIM values, indicating its excellence in image quality assessment. Taking the Art image as an example, AB-NLM achieves an SSIM of 0.986, significantly surpassing MECI (0.982) and NLM (0.983). This consistent trend across different images (Teddy, Moebius, Recycle, and Jadeplant) emphasizes AB-NLMs ability to provide higher structural similarity, indicative of a more accurate reconstruction of the original images.

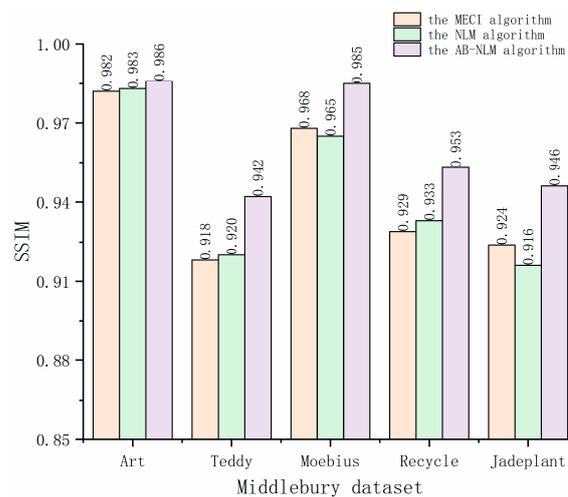


Figure 8. Displays a comparative analysis of SSIM results among the MECI, NLM, and AB-NLM algorithms based on the Middlebury dataset.

In Table 1, when $ndisp = 256$, it implies that the depth map’s disparity values do not exceed 256 during the restoration evaluation. Increasing $ndisp$ results in higher DE and ALME. For instance, in the Jadeplant image, both DE and ALME values are notably higher, yet this does not necessarily mean poorer restoration. Notably, within the five images in Table 1, all DE values surpass their respective ALME values, indicating most restoration outcomes closely approximate true depth values, with a small fraction exhibiting suboptimal results. Analysis of the MECI, NLM, and AB-NLM algorithms in Table 1 reveals that AB-NLM outperforms others in terms of DE and ALME values across various images, indicating superior performance in-depth image repair.

Table 1. Comparative analysis of DE and ALME for NLM, MECI, and AB-NLM algorithms based on the Middlebury dataset.

Image of Scenes	ndisp *	NLM		MECI		AB-NLM	
		DE	ALME	DE	ALME	DE	ALME
Art	256	2.75	2.29	2.31	1.98	2.08	1.67
Teddy	256	2.68	2.61	2.54	2.35	2.41	2.28
Moebius	256	2.88	2.59	2.69	2.55	2.21	2.11
Recycle	260	3.04	2.81	2.75	2.58	2.19	1.78
Jadeplant	640	10.70	10.31	9.11	8.43	8.31	7.92

* The ‘ndisp’ commonly refers to the number of potential disparity values in the disparity map, representing the levels of disparity corresponding to depth differences.

4. Conclusions

The effectiveness of the proposed adaptive block-based deep image hole repair algorithm is successfully demonstrated in this study. Initially, the algorithm based on NLM for

hole repair of void regions is expounded upon. Subsequently, an intelligent block factor is introduced upon the foundation of the NLM hole repair algorithm. This factor facilitates automatic adjustments of search and repair block sizes to accommodate void regions of various dimensions, thereby mitigating the complexities associated with parameter tuning. Moreover, optimal search and repair blocks are computed for each void region, aiding in their proficient restoration. Finally, a series of experiments is conducted utilizing the Middlebury stereo matching dataset and scenes captured by the Orbbec Astra S depth camera. The NLM, AB-NLM, and MECI algorithms proposed in this paper are subjected to both qualitative and quantitative assessments employing the RMSE, SSIM, PSNR, DE, and ALME metrics. The outcomes validate that the proposed algorithm consistently yields more authentic object contours in the majority of scenarios. Furthermore, its performance in void region restoration within in-depth images is notably superior.

Our future work will focus primarily on the following research areas:

- (1) Through an extensive series of experimental investigations, it has been ascertained that the algorithm proposed within the context of this study exhibits certain deficiencies in the realm of transparent object depth restoration. Consequently, a prospective avenue for further inquiry entails the formulation of algorithms capable of addressing the task of transparent object depth restoration.
- (2) As the proportion of missing regions in the image gradually increases, the restoration effectiveness of various methods, including the algorithm proposed in this paper, will decrease. This is manifested by varying degrees of distortion at the restoration locations. Regardless of whether based on traditional algorithms or deep learning-based approaches, the fundamental principle of image restoration involves filling in the missing regions using known information in a certain manner. Naturally, better restoration results are achieved when the unknown regions are minimized. However, when the proportion of masked areas becomes excessively large, the limited amount of known information cannot adequately support the predictive function of the restoration algorithm. Therefore, in the future, the incorporation of the concept of style transfer will be explored. Particularly for large-scale restoration, combining image-style prior information with the idea of style transfer will be considered. Two models will synchronize output features and mutually supervise each other, thereby enhancing the model's capability to restore large-scale missing regions.

Author Contributions: Conceptualization, L.L.; methodology, B.Z.; software, B.Z.; data curation, B.Z.; writing—original draft preparation, B.Z.; writing—review and editing, B.Z. and L.L.; supervision, H.P.; funding acquisition, H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the Zhejiang Provincial Natural Science Foundation of China (LQ23F030019).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated during the current study are available from the corresponding author on reasonable request. The data are not publicly available due to all the data used in the experiments are from a private database.

Acknowledgments: The authors would like to thank the anonymous reviewers and the Associate Editor for their valuable comments and suggestions to improve the quality of the manuscript. The authors would like to express their sincere gratitude to Donghai Su for his valuable contributions to this work, particularly in the areas of validation and data curation. His dedication greatly contributed to the success of this research. The authors would also like to express their gratitude to the anonymous reviewers and the Associate Editor for their invaluable comments and insightful suggestions, which greatly contributed to enhancing the overall quality of this manuscript.

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

Abbreviations

NLM	Non-local means
AB-NLM	Adaptive Block-based Non-local means
MECI	Multiple Edge Converge Inpainting
RMSE	Root Mean Square Error
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity Index
DE	Depth Error
ALME	Average Local Mean Error
ndisp	number of disparities

References

- Roman-Rivera, L.-R.; Sotelo-Rodríguez, I.; Pedraza-Ortega, J.C.; Aceves-Fernandez, M.A.; Ramos-Arreguín, J.M.; Gorrostieta-Hurtado, E. Reduced Calibration Strategy Using a Basketball for RGB-D Cameras. *Mathematics* **2022**, *10*, 2085. [\[CrossRef\]](#)
- Wang, H.; Guo, E.; Chen, F.; Chen, P. Depth Completion in Autonomous Driving: Adaptive Spatial Feature Fusion and Semi-Quantitative Visualization. *Appl. Sci.* **2023**, *13*, 9804. [\[CrossRef\]](#)
- Ceccarelli, A.; Secci, F. RGB Cameras Failures and Their Effects in Autonomous Driving Applications. *IEEE Trans. Dependable Secur. Comput.* **2023**, *20*, 2731–2745. [\[CrossRef\]](#)
- Cong, Y.; Chen, R.; Ma, B.; Liu, H.; Hou, D.; Yang, C. A Comprehensive Study of 3-D Vision-Based Robot Manipulation. *IEEE Trans. Cybern.* **2023**, *53*, 1682–1698. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, X.; Zhang, X.; Zhou, X.; Chen, I.-M. UPG: 3D vision-based prediction framework for robotic grasping in multi-object scenes. *Knowl.-Based Syst.* **2023**, *270*, 110491. [\[CrossRef\]](#)
- Alhamazani, F.; Lai, Y.-K.; Paul, Rosin, L. 3DCascade-GAN: Shape completion from single-view depth images. *Comput. Graph.* **2023**, *115*, 412–422. [\[CrossRef\]](#)
- Kim, W.; Luong, T.; Ha, Y.; Doh, M.; Yax, J.F.M.; Moon, H. High-Fidelity Drone Simulation with Depth Camera Noise and Improved Air Drag Force Models. *Appl. Sci.* **2023**, *13*, 10631. [\[CrossRef\]](#)
- Herrera-Granda, E.P.; Torres-Cantero, J.C.; Rosales, A.; Peluffo-Ordóñez, D.H. A Comparison of Monocular Visual SLAM and Visual Odometry Methods Applied to 3D Reconstruction. *Appl. Sci.* **2023**, *13*, 8837. [\[CrossRef\]](#)
- Cheng, S.; Sun, C.; Zhang, S.; Zhang, D. SG-SLAM: A Real-Time RGB-D Visual SLAM Toward Dynamic Scenes with Semantic and Geometric Information. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 7501012. [\[CrossRef\]](#)
- Sun, C.; Wu, X.; Sun, J.; Sun, C.; Xu, M.; Ge, Q. Saliency-Induced Moving Object Detection for Robust RGB-D Vision Navigation Under Complex Dynamic Environments. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 10716–10734. [\[CrossRef\]](#)
- Xiang, S.; Deng, H.; Zhu, L. Exemplar-based depth inpainting with arbitrary-shape patches and cross-modal matching. *Signal Process. Image Commun.* **2019**, *71*, 56–65. [\[CrossRef\]](#)
- Criminisi, A.; Pérez, P.; Toyama, K. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Process.* **2004**, *9*, 1200–1212. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, L.; Xia, H.; Qiao, Y. Texture synthesis repair of RealSense D435i depth images with object-oriented RGB image segmentation. *Sensors* **2020**, *23*, 6725. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, X. A modified non-local means using bilateral thresholding for image denoising. *Multimed. Tools Appl.* **2023**. [\[CrossRef\]](#)
- Wang, L.; Liao, C.; Yao, R. Fixing algorithm of Kinect depth image based on non-local means. *Multimed. Tools Appl.* **2023**, 1–20. [\[CrossRef\]](#)
- Cheng, J.; Xie, Y.; Zhou, S.; Lu, A.; Peng, X.; Liu, W. Improved Weighted Non-Local Mean Filtering Algorithm for Laser Image Speckle Suppression. *Micromachines* **2023**, *1*, 98. [\[CrossRef\]](#)
- Deshe, G.; Abookasis, D. Imaging targets hidden in scattering and viscous liquid-based media by combining multiple projections and applying a non-local mean filtering algorithm. *Optik* **2021**, *247*, 167988. [\[CrossRef\]](#)
- Wu, J.; Gomez, L.; Frery, A.C. *A Non-Local Means Filters for Sar Speckle Reduction with Likelihood Ratio Test*; IEEE International Geoscience and Remote Sensing Symposium: Kuala Lumpur, Malaysia, 2022; pp. 2319–2322.
- Zhang, X.; Li, M.; Liu, X.; Zhou, Z.; Wang, X.; Xu, Z. An improved wavelet denoising algorithm for SAR interferogram using fast non-local means filtering. *Geocarto Int.* **2022**, *37*, 18600–18617. [\[CrossRef\]](#)
- Manjón, J.V.; Coupé, P.; Martí-Bonmatí, L.; Collins, D.L.; Robles, M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J. Magn. Reson. Imaging* **2010**, *31*, 1053–1807. [\[CrossRef\]](#)
- Gastal, E.S.L.; de Oliveira Neto, M.M. Adaptive manifolds for real-time high-dimensional filtering. *ACM Trans. Graph. (TOG)* **2012**, *31*, 1–13. [\[CrossRef\]](#)
- Wang, R.; Liang, D.; Cao, X.; Guo, Y. Semantic Correspondence with Geometric Structure Analysis. *ACM Trans. Multimed. Comput. Commun. Appl.* **2021**, *17*, 21. [\[CrossRef\]](#)
- Bao, W.; Wang, W.; Xu, Y.; Guo, Y.; Hong, S.; Zhang, X. InStereo2K: A large real dataset for stereo matching in indoor scenes. *Sci. China Inf. Sci.* **2023**, *63*, 212101. [\[CrossRef\]](#)

24. Vila, O.; Boada, I.; Raba, D.; Farres, E. A Method to Compensate for the Errors Caused by Temperature in Structured-Light 3D Cameras. *Sensors* **2021**, *21*, 2073. [[CrossRef](#)] [[PubMed](#)]
25. Köhler, N.A.; Nöh, C.; Geis, M.; Kerzel, S.; Frey, J.; Groß, V.; Sohrabi, K. Influence of Ambient Factors on the Acquisition of 3-D Respiratory Motion Measurements in Infants—A Preclinical Assessment. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3271724. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.