

Article

Double-Constraint Inpainting Model of a Single-Depth Image

Wu Jin ¹, Li Zun ^{1,2,*} and Liu Yong ¹

¹ School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China; wujin@wust.edu.cn (W.J.); 141001@xxmu.edu.cn (L.Y.)

² School of Physics and Electronic Engineering, Xinxiang College, Xinxiang 453000, China

* Correspondence: lizun@xxu.edu.cn

Received: 14 February 2020; Accepted: 20 March 2020; Published: 24 March 2020



Abstract: In real applications, obtained depth images are incomplete; therefore, depth image inpainting is studied here. A novel model that is characterised by both a low-rank structure and nonlocal self-similarity is proposed. As a double constraint, the low-rank structure and nonlocal self-similarity can fully exploit the features of single-depth images to complete the inpainting task. First, according to the characteristics of pixel values, we divide the image into blocks, and similar block groups and three-dimensional arrangements are then formed. Then, the variable splitting technique is applied to effectively divide the inpainting problem into the sub-problems of the low-rank constraint and nonlocal self-similarity constraint. Finally, different strategies are used to solve different sub-problems, resulting in greater reliability. Experiments show that the proposed algorithm attains state-of-the-art performance.

Keywords: depth image inpainting; variable splitting technique; low-rank constraint; nonlocal self-similarity constraint

1. Introduction

With the rapid development of RGB-D (red green blue-depth) sensors [1–6], such as the Kinect sensor, colour images and depth images can be obtained simultaneously. Depth images are widely used in 3D reconstruction, 3D videos and medical intelligence and are therefore a research area focus for image processing and computer vision. Initially, the development of depth images was limited by the cost-effectiveness of devices used to acquire depth images [7–10]. In 2010, Microsoft launched the Kinect sensor to acquire depth images, and it attracted wide attention and expanded the associated applications.

In practical applications, depth images are of low quality and have black holes. Black holes represent missing depth information, and the black-hole filling problem is solved via depth image inpainting. At present, depth image inpainting methods can be divided into two categories according to whether the corresponding colour images are guided.

The first method relies on the use of corresponding colour images as a guide. Liu et al. [11] proposed a robust optimisation framework for colour image-guided depth image restoration. This method performs well in suppressing texture artefacts. Lee et al. [12] proposed an adaptive edge-oriented smoothing process based on the characteristics of holes with or without vertical lines in the colour image. The proposed method represents a good trade-off between time savings, hole reduction, and virtual view quality. Lei et al. [13] proposed a credibility-based multi-view depth image fusion strategy to refine images. This method considers the view synthesis quality and inter-view correlation in an improved repair approach.

The second method does not use corresponding colour images as a guide. Shen et al. [14] proposed the inpainting method using a weighted joint bilateral filter and fast marching. This method has obtained the best performance in improving depth images by producing smooth and edge regions. Buysens et al. [15] proposed a suitable method for recovering the lost structures of objects to in-paint depth images in a geometrically plausible manner. Lu et al. [16] proposed a method of inpainting depth images through similar patches in a matrix and enforced low-rank subspace constraints, thereby attaining good performance. Xue et al. [17] proposed the low-gradient regularisation method, an effective approach that reduces the penalty for gradient 1 while penalizing non-zero gradients to allow for gradual depth changes.

To reduce the complexity of the inpainting problem, we solve depth image inpainting without corresponding colour images and fully exploit the features of depth images to complete the inpainting task of single-depth image inpainting.

Based on previous work, the use of a single-image property as one constraint is not sufficient to obtain satisfying inpainting results. Consequently, we use more than one constraint to perform depth image inpainting.

Depth images can be regarded as textureless natural images that consist of many similar flat areas and few edge areas. Depth images therefore have the characteristics of a low rank and nonlocal self-similarity. Due to its textureless property, the effect of the low-rank constraint on inpainting will be too great and false details will be created. Therefore, we introduce the nonlocal self-similarity constraint. First, we regard the depth image as a matrix, and the corresponding low-rank reconstruction model is built based on the low-rank structure of the matrix in the image. We then introduce the nonlocal self-similarity constraint to improve the depth image results. The contributions of this paper are summarised as follows.

1. Rather than the traditional single-constraint method, we adopt a double-constraint method. According to the characteristics of the depth image, we combine the low-rank constraint and nonlocal self-similarity constraint.
2. We adopt the split Bregman algorithm, which is a variable splitting technique, to divide depth image inpainting into sub-problems, thus reducing the complexity of the solution.
3. We use different strategies to solve depth image inpainting: weighted Schatten p -norm minimisation as the low-rank constraint and nonlocal statistical modelling as the nonlocal self-similarity constraint. The proposed method achieves better performance.

The remainder of our paper is organised as follows. In Section 2, we present the related work. In Section 3, we describe the details of the depth image inpainting method based on the double-constraint. In Section 4, we present the experimental results. In Section 5, we summarise the paper.

2. Related Work

2.1. Depth Images

Depth images are greyscale images with pixel values of 0–255, and the greyscale value of the pixel represents the distance between the spatial scene and the camera. In general, the closer the area is to the camera, the greater the depth, while the farther away the area is to the camera, the smaller the depth. The depth image consists of the most similar flat regions and a few edge regions which contain a large number of regions with the same grey value. A continuous distribution with the same depth values exists inside the object, and the gradient is 0. Depth value mutations and gradients are observed at the edges. The Aloe depth image [18–20] is taken as an example, and the grey value of the depth image is as shown in Figure 1.

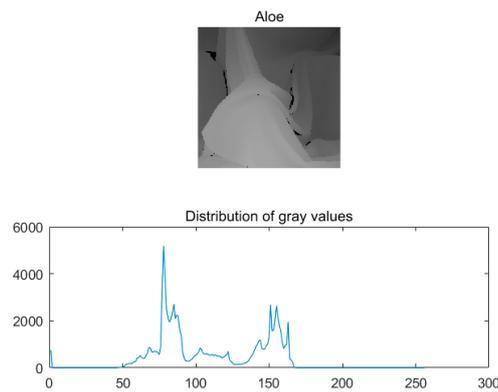


Figure 1. Distribution of grey values.

As shown in Figure 1, the distribution of grey values in the depth image is very concentrated. Depth images have the characteristics of a low rank and nonlocal self-similarity.

2.2. Low-Rank Constraint and Nonlocal Self-Similarity Constraint

The single-depth image inpainting problem is transformed into the following mathematical expression:

$$\mathbf{x} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \cdot \psi(\mathbf{x}). \quad (1)$$

where \mathbf{x} is the intact depth image; \mathbf{y} is the degraded depth image; $\|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2$ is the data-fidelity term; $\psi(\mathbf{x})$ is the regularisation term; λ is the weight parameter; and \mathbf{H} is a binary template. We attempt to obtain a potential depth image \mathbf{x} from the degraded depth image \mathbf{y} . According to the characteristics of the depth image, we combine the low-rank constraint and nonlocal self-similarity constraint. Equation (1) can be converted into Equation (2):

$$\mathbf{x} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 \cdot \Psi_{\text{LR}}(\mathbf{x}) + \lambda_2 \cdot \Psi_{\text{NSS}}(\mathbf{x}). \quad (2)$$

where $\Psi_{\text{LR}}(\mathbf{x})$ represents the low-rank regularisation term, $\Psi_{\text{NSS}}(\mathbf{x})$ represents the nonlocal self-similarity regularisation term, and λ_1 and λ_2 are the weight parameters.

At present, the solution methods for low-rank matrices can be divided into two categories: low-rank matrix decomposition and rank minimisation. The commonly used low-rank matrix decomposition method mostly adopts the singular decomposition technique, which uses the f -norm fidelity loss to trim the singular value matrix to obtain the optimal rank approximation solution. The rank minimisation method mainly uses the relaxation method to minimise the rank and estimate the lowest rank for reconstruction. The latter method has a better recovery performance. Therefore, we use the nuclear norm minimisation-based (NNM-based) method for depth image inpainting.

Scholars have conducted some research into the solution of the NNM problem. In [21], under certain conditions, the NNM method is used to achieve reconstruction with limited information. In [22], the soft threshold operation is applied to the NNM method for matrix filling purposes in a very small storage space. In [23], low-level visual problems are solved by minimizing the sum of partial singular values. In [24], weighted nuclear model minimisation is proposed, and the method adaptively weights singular values differently, which improves the applicability and flexibility of low-quality images.

Compared with NNM, weighted Schatten p -norm minimisation (WSNM) [25] can better approximate the original low-order hypothesis and consider the importance of different components. WSNM can be effectively applied to obtain the global optimal solution. Therefore, we use WSNM as the low-rank constraint.

In addition to the low-rank constraint, nonlocal self-similarity is another important feature of depth images. This feature can describe the structure repetition characteristic of the nonlocal area of the depth image and preserve the edge and detail effectively.

The repeatability of the nonlocal self-similarity description advanced mode has enabled remarkable achievements in the field of image reconstruction. Buades et al. [26] proposed an effective denoising model called nonlocal means (NLMs) via the degree of similarity among surrounding pixels for denoising tasks. Jung et al. [27] proposed a class of restoration algorithms for colour images based upon the Mumford–Shah model and nonlocal image information. These algorithms are defined to work in a small local neighbourhood and are sufficient to denoise smooth regions with sharp boundaries. In [28], a nonlocal self-similarity constraint is introduced into the overall cost functional to improve the robustness of the model. The proposed method outperforms many existing image reconstruction methods. The nonlocal self-similarity constraint produces superior results with sharper image edges.

However, traditional nonlocal self-similarity constraints fail to recover accurate structures in depth images. We use the nonlocal self-similarity (NLSM) [29] of the three-dimensional transformation domain as the constraint term. Compared with traditional methods, the NLSM of the three-dimensional transformation domain represents self-similarity more effectively and has adaptive performance.

3. Double-Constraint Model

3.1. Similar Block Group and NLSM Model

The similar block group and NLSM are based on similar blocks. The construction is shown in Figure 2.

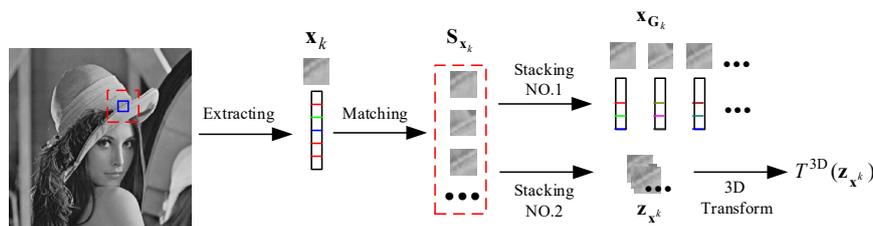


Figure 2. Construction of the similar block group and the non-local self-similar statistical model.

As shown in Figure 2, we first divide the image x into pixel blocks of size $\sqrt{B_s} \times \sqrt{B_s}$, and each pixel block is expressed in vector form x_k , where $k = 1, 2, 3, \dots, n$. Then, for each patch x_k denoted by a blue mark, in the red window, we determine its c similar patches, which compose set S_{x_k} .

In the first stacking, all patches in set S_{x_k} are arranged in a matrix to obtain similar groups, with $x_{G_k} \in \mathbb{R}^{B_s \times c}$. Due to the characteristics of the greyscale values in the depth image, the principle of similarity matching is the SSD principle.

In the second stacking, all patches in set S_{x_k} are stacked in a three-dimensional z_{x_k} . By orthogonal three-dimensional transformation T^{3D} , the coefficients of three-dimensional arrangement $T^{3D}(z_{x_k})$ are obtained.

3.2. Solution of Depth Image inpainting

By introducing variables \mathbf{u} and \mathbf{v} , we can transform Equation (2) into an equivalent constrained form as follows:

$$\mathbf{x} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 \cdot \Psi_{\text{LR}}(\mathbf{u}) + \lambda_2 \cdot \Psi_{\text{NSS}}(\mathbf{v}) \quad \text{s.t.} \quad \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \mathbf{G}\mathbf{x} \quad (3)$$

where $\mathbf{G} = [\mathbf{I}, \mathbf{I}]^T$.

By introducing externalised constraints, we arrive at the following:

$$\mathbf{x} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 \cdot \|\mathbf{u}\|_{\mathbf{w},sp}^p + \lambda_2 \cdot \|\Theta_{\mathbf{v}}\|_1. \quad (4)$$

where $\|\mathbf{u}\|_{\mathbf{w},sp}^p = \sum_{i=1}^{\min\{n,m\}} \omega_i \sigma_i^p = \operatorname{tr}(\mathbf{W}\Delta^p)$, $\|\Theta_{\mathbf{v}}\|_1 = \sum_{i=1}^n \|T^{3D}(\mathbf{z}_{x^k})\|_1$

Then, we use the split Bregman algorithm [30] to transform complex problems into sub-problems that are easy to solve:

$$\mathbf{x}^{t+1} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2^2 + \frac{\mu}{2} \|\mathbf{x} - \mathbf{u}^t - \mathbf{b}^t\|_2^2 + \frac{\mu}{2} \|\mathbf{x} - \mathbf{v}^t - \mathbf{c}^t\|_2^2 \quad (5)$$

$$\mathbf{u}^{t+1} = \operatorname{argmin}_{\mathbf{u}} \lambda_1 \cdot \|\mathbf{u}\|_{\mathbf{w},sp}^p + \frac{\mu}{2} \|\mathbf{x}^{t+1} - \mathbf{u}^t - \mathbf{b}^t\|_2^2 \quad (6)$$

$$\mathbf{v}^{t+1} = \operatorname{argmin}_{\mathbf{v}} \lambda_2 \cdot \|\Theta_{\mathbf{v}}\|_1 + \frac{\mu}{2} \|\mathbf{x}^{t+1} - \mathbf{v}^t - \mathbf{c}^t\|_2^2 \quad (7)$$

where $\mathbf{b}^{t+1} = \mathbf{b}^t - (\mathbf{x}^{t+1} - \mathbf{u}^{t+1})$ and $\mathbf{c}^{t+1} = \mathbf{c}^t - (\mathbf{x}^{t+1} - \mathbf{v}^{t+1})$.

Other variable splitting techniques, such as the half quadratic splitting method, can also be used to transform complex problems into sub-problems.

For conciseness and to avoid ambiguity, the iterations are omitted in the following discussion of the sub-problems.

3.2.1. Sub-Problem \mathbf{x}

The split Bregman algorithm converts Equation (4) into three sub-problems. Equation (5) represents the minimisation problem of sub-problem \mathbf{x} transformed into a strict convex quadratic function. The closed solution of Equation (5) can be obtained as follows:

$$\mathbf{x} = (\mathbf{H}^T \mathbf{H} + 2\mu \mathbf{I})^{-1} [\mathbf{H}^T \mathbf{y} + \mu(\mathbf{u} + \mathbf{v} + \mathbf{b} + \mathbf{c})] \quad (8)$$

3.2.2. Sub-Problem \mathbf{u}

According to the solution of \mathbf{u} , Equation (6) can be converted into the following equation:

$$\mathbf{u} = \operatorname{argmin}_{\mathbf{u}} \|\mathbf{r}_{\mathbf{u}} - \mathbf{u}\|_2^2 + \frac{2\lambda_1}{\mu} \|\mathbf{u}\|_{\mathbf{w},sp}^p \quad (9)$$

where $\mathbf{r}_{\mathbf{u}} = \mathbf{x} - \mathbf{b}$. Let $\mathbf{e}_{\mathbf{u}} = \mathbf{r}_{\mathbf{u}} - \mathbf{u}$, which represents the residual. Taking the Aloe depth image as an example, we select deblurring replacement inpainting to carry out the simulation experiments. The reasons why we chose image deblurring as an example for verification are twofold: 1. we have an accurate original depth image for objective data comparison; and 2. image deblurring and inpainting both satisfy Equation (1).

We chose to approximate the Aloe depth image as \mathbf{u} . Then, the residual distribution of the k th iteration can be obtained.

We first performed 3×3 uniform blur kernel operations on the Aloe depth image and then added Gaussian noise with a standard deviation of 1 to obtain a blurred depth image. Figure 3 shows the distribution of the residuals in the experiments with three, five and seven iterations.

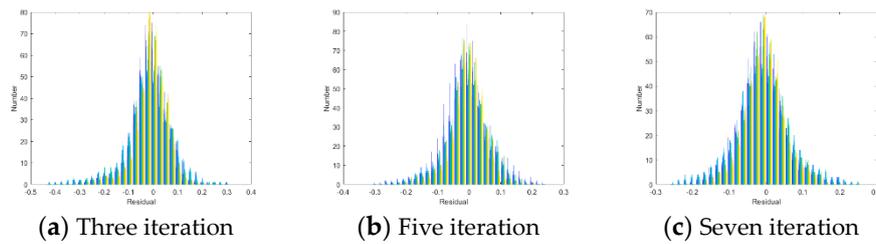


Figure 3. Distribution of \mathbf{e}_u for different iterations.

As shown in Figure 3, in each iteration, the distribution of \mathbf{e}_u is suitably characterised by a generalised Gaussian distribution with a zero mean.

According to the experiments, we formulate the following hypothesis [24,31,32]: in each iteration, the residuals satisfy a generalised Gaussian distribution with a zero-mean. In each iteration, the following equation is satisfied:

$$\frac{1}{N} \|\mathbf{r}_u - \mathbf{u}\|_2^2 = \frac{1}{K} \sum_{k=1}^n \|\mathbf{r}_{G_k} - \mathbf{u}_{G_k}\|_F^2 \quad (10)$$

By substituting Equation (10) into Equation (9), we can obtain the following:

$$\mathbf{u} = \operatorname{argmin}_{\mathbf{u}} \sum_{k=1}^n \frac{\mu N}{2\lambda_1 K} \|\mathbf{r}_{G_k} - \mathbf{u}_{G_k}\|_F^2 + \|\mathbf{u}_{G_k}\|_{w, S_p}^p \quad (11)$$

For each similar group, we assume that the singular decomposition of \mathbf{r}_{G_k} is $\mathbf{r}_{G_k} = \mathbf{U}\mathbf{\Sigma}V^T$, with $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$, which is non-ascending. According to Von Neumann's trace inequality theorem [33], the solution of Equation (9) transforms the solution of $\mathbf{\Delta} = \operatorname{diag}(\delta_1, \dots, \delta_r)$ into $\mathbf{u}_{G_k} = \mathbf{Q}\mathbf{\Delta}R^T$. The solution equation is as follows.

$$\begin{cases} \min_{\delta_1, \dots, \delta_r} \sum_{i=1}^r [(\delta_i - \sigma_i)^2 + \omega_i \delta_i^p], i = 1, \dots, r \\ \text{s.t. } \delta_i \geq 0 \quad \text{and} \quad \delta_i \geq \delta_j, \quad \text{for } i \leq j \end{cases} \quad (12)$$

The solution of Equation (11) can be converted into the following equation:

$$\min_{\delta_i \geq 0} f_i(\delta) = (\delta_i - \sigma_i)^2 + \omega_i \delta_i^p, i = 1, \dots, r \quad (13)$$

Equation (13) can be solved by using the generalised soft threshold (GST) algorithm [34].

If p and ω_i are determined, according to the GST algorithm, there is a special threshold $\sigma_i \geq \tau_p^{GST}(\omega_i)$ that satisfies the following equation:

$$\tau_p^{GST}(\omega_i) = (2\omega_i(1-p))^{\frac{1}{2-p}} + \omega_i p (2\omega_i(1-p))^{\frac{p-1}{2-p}} \quad (14)$$

If $\sigma_i < \tau_p^{GST}(\omega_i)$, then the following holds:

$$f_i(\delta) = \sigma_i^2 + \omega_i \delta_i^p, i = 1, \dots, r \quad (15)$$

That is, $\delta_i = 0$, which is the global minimum.

If $\sigma_i \geq \tau_p^{GST}(\omega_i)$, then $f_i(\delta)$ has the minimum value of $S_p^{GST}(\sigma_i; \omega_i)$, which can be obtained by solving the following equation:

$$S_p^{GST}(\sigma_i; \omega_i) - \sigma_i + \omega_i p (S_p^{GST}(\sigma_i; \omega_i))^{p-1} = 0 \quad (16)$$

Then, sub-problem \mathbf{u} is solved.

3.2.3. Sub-Problem \mathbf{v}

According to the solution of \mathbf{v} , Equation (7) can be converted into the following equation:

$$\mathbf{v} = \operatorname{argmin}_{\mathbf{v}} \frac{1}{2} \|\mathbf{r}_{\mathbf{v}} - \mathbf{v}\|_2^2 + \frac{2\lambda_2}{\mu} \|\Theta_{\mathbf{v}}\|_1 \quad (17)$$

where $\mathbf{r}_{\mathbf{v}} = \mathbf{x} - \mathbf{c}$. Let $\mathbf{e}_{\mathbf{v}} = \mathbf{r}_{\mathbf{v}} - \mathbf{v}$, which represents the residual. $\mathbf{e}_{\mathbf{v}}$ and $\mathbf{e}_{\mathbf{u}}$ have the same property. As a result, Equation (17) can be converted into Equation (18):

$$\mathbf{v} = \operatorname{argmin}_{\mathbf{v}} \frac{1}{2} \|\Theta_{\mathbf{r}_{\mathbf{v}}} - \Theta_{\mathbf{v}}\|_2^2 + \frac{2K\lambda_2}{N\mu} \|\Theta_{\mathbf{v}}\|_1 \quad (18)$$

In the above equation, any elements of $\Theta_{\mathbf{v}}$ can be solved separately; therefore, we use the soft threshold [35]. to solve Equation (18).

$$\Theta_{\mathbf{v}} = \operatorname{soft}(\Theta_{\mathbf{r}_{\mathbf{v}}}, \frac{2K\lambda_2}{N\mu}) \quad (19)$$

Namely,

$$\Theta_{\mathbf{v}}(j) = \operatorname{sgn}((\Theta_{\mathbf{r}_{\mathbf{v}}}(j)) \max\{|\Theta_{\mathbf{r}_{\mathbf{v}}}(j) - \frac{2K\lambda_2}{N\mu}|, 0\}) \begin{cases} \Theta_{\mathbf{r}_{\mathbf{v}}}(j) - \frac{2K\lambda_2}{N\mu}, & \Theta_{\mathbf{r}_{\mathbf{v}}}(j) \in (\frac{2K\lambda_2}{N\mu}, +\infty) \\ 0, & \Theta_{\mathbf{r}_{\mathbf{v}}}(j) \in [-\frac{2K\lambda_2}{N\mu}, \frac{2K\lambda_2}{N\mu}] \\ \Theta_{\mathbf{r}_{\mathbf{v}}}(j) + \frac{2K\lambda_2}{N\mu}, & \Theta_{\mathbf{r}_{\mathbf{v}}}(j) \in (-\infty, -\frac{2K\lambda_2}{N\mu}) \end{cases} \quad (20)$$

In summary, all the sub-problems in our proposed algorithm are solved. The flow chart of the algorithm is shown in Table 1.

Table 1. Complete description of the proposed method.

| |
|--|
| Input: The observed depth image \mathbf{y} , the degraded operator \mathbf{H} |
| Output: The restored depth image \mathbf{x} |
| Repeat |
| Step 1: Update \mathbf{x} by Equation (8) |
| Step 2: For each group u_{G_k} |
| (1) The singular value decomposition of \mathbf{r}_{G_k} |
| (2) Update u_{G_k} by Equation (12) |
| Aggregate u_{G_k} to form \mathbf{u} |
| Step 3: Update \mathbf{v} by Equation (20) |
| Until maximum iteration number is reached |

4. Experiments

4.1. Depth Image Inpainting

In this paper, the hardware simulation platform was supported by a Lenovo R720 computer (Lenovo, Beijing, China) and the software simulation platform was MATLAB R2017a (The MathWorks, Inc., Massachusetts, USA).

The inpainting effect as analysed by subjective visual requirements and objective parameters. The objective metrics could be assessed on the basis of two objective metrics: the peak signal-to-noise ratio (PSNR) [36] and the feature similarity (FSIM) [37]. PSNR uses the ratio of the maximum semaphore to

noise intensity to measure image quality, which is easy to calculate and understand and can reflect the image quality. FSIM is a novel low-level feature similarity parameter. Phase congruency is a dimensionless measure of the significance of a local structure, and it is used as the primary feature in FSIM.

We used a dataset that included the Middlebury datasets [20–22] and NYU v2 dataset [38]. For the comparison algorithm, we proposed similar algorithms, namely NNM [8] and WSNM [18].

In experiment 1, the depth images were obtained directly from the Middlebury datasets. The area requiring repair was the actual situation. as shown in Figure 4.

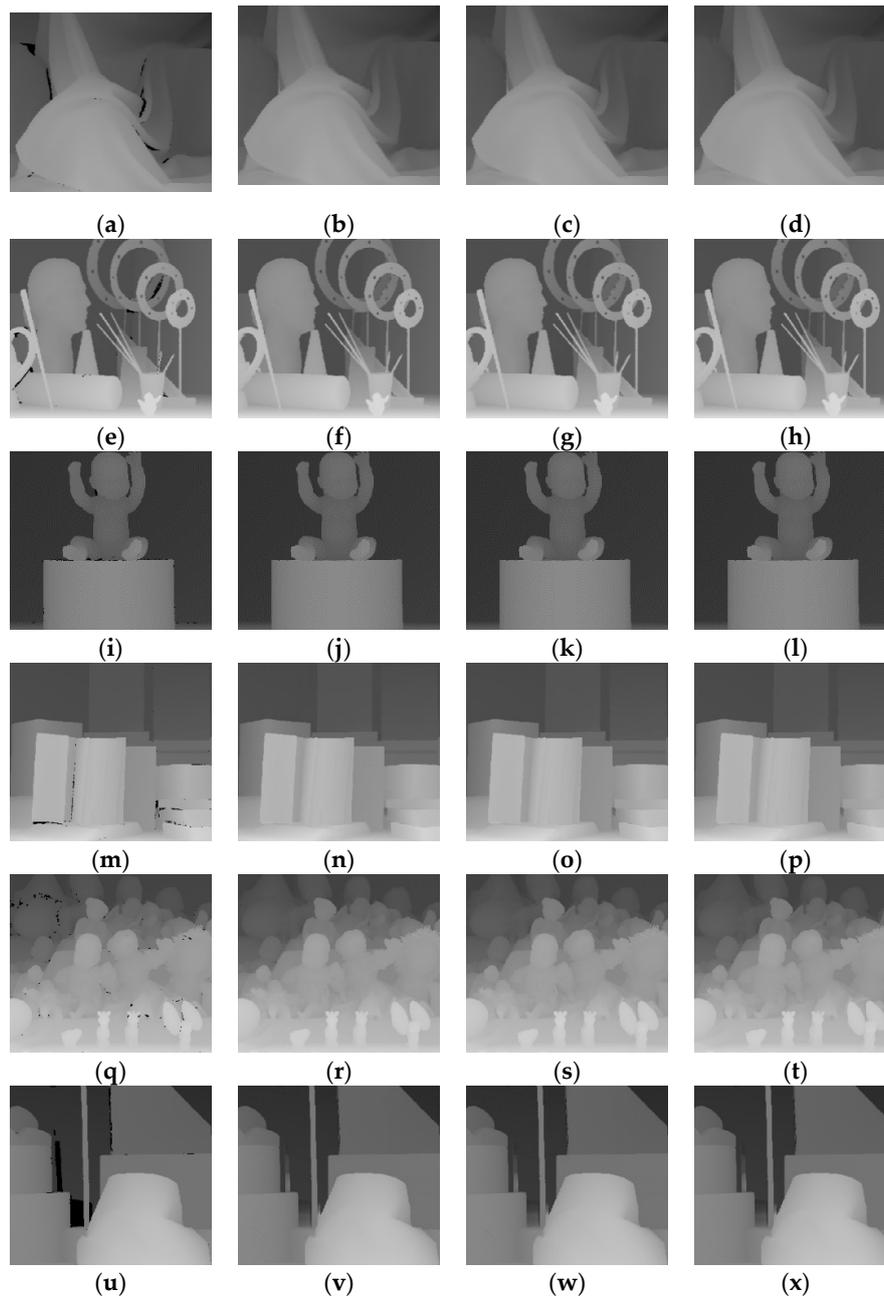


Figure 4. Visual quality comparison of the inpainting result (1): (a) Aloe depth image; (b–d) inpainting effects of the nuclear norm minimisation (NNM) algorithm, weighted Schatten p -norm minimisation (WSNM) algorithm and proposed algorithm for the Aloe depth image; (e) Art depth image; (f–h) inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm for the Art depth

image; (i) Baby depth image; (j–l) inpainting effects of the (nuclear norm minimisation) NNM algorithm, WSNM algorithm and the proposed algorithm for the Baby depth image; (m) Books depth image; (n–p) inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm for the Books depth image; (q) Dolls depth image; (r–t) inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm for the Dolls depth image; (u) Lam depth image; and (v–x) inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm for the Lam depth image.

As shown in Figure 4, all three algorithms met the visual requirements and no obvious repair marks occurred. However, the details varied from image to image. The NNM algorithm and WSNM algorithm both resulted in the smoothing of boundaries. Our proposed algorithm could reduce this situation, as shown in the enlarged portion of the figure in the red box.

As summarized in Table 2, the proposed algorithm was superior to the other two algorithms and the objective data were improved.

Table 2. Peak signal-to-noise ratio (PSNR)/feature similarity (FSIM) in experiment (1).

| Image | Algorithm (PSNR/FSIM) | | |
|-------|-----------------------|----------------|-----------------------|
| | NNM | WSNM | Proposed |
| Aloe | 26.0395/0.9571 | 26.0767/0.9628 | 26.1296/0.9705 |
| Art | 26.8853/0.9366 | 27.1790/0.9825 | 27.1833/0.9835 |
| Baby | 30.0559/0.9413 | 30.2569/0.9902 | 30.3200/0.9932 |
| Books | 27.4590/0.9674 | 28.1774/0.9632 | 28.1806/0.9752 |
| Dolls | 29.2717/0.9758 | 29.0181/0.9739 | 29.1254/0.9745 |
| Lam | 23.5473/0.9756 | 24.4459/0.9761 | 24.4534/0.9761 |

In experiment 2, the depth images were obtained directly from the NYU v2 dataset. The area requiring repair was the 10% data loss area, as shown in Figure 5.

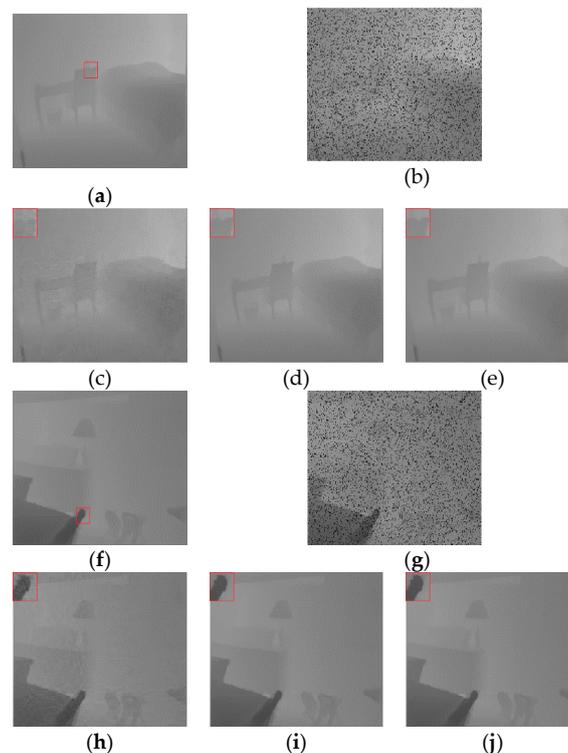


Figure 5. Cont.

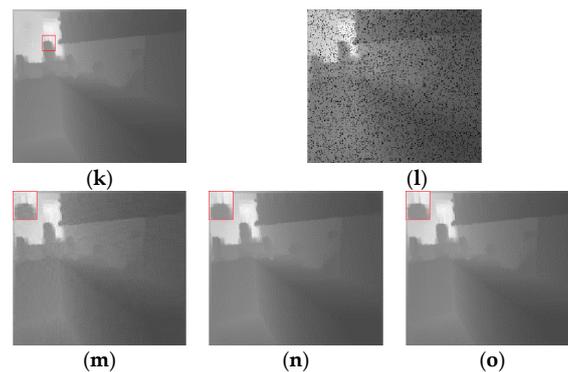


Figure 5. Visual quality comparison of the inpainting result (2): (a) Bedroom depth image; (b) Corrupted bedroom depth image with 10% pixels missing; (c–e) Inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm; (f) Lamp depth image; (g) Corrupted lamp depth image with 10% pixels missing; (h–j) Inpainting inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm; (k) Kitchen depth image; (l) Corrupted kitchen depth image with 10% pixels missing; (m–o) Inpainting effects of the NNM algorithm, WSNM algorithm and the proposed algorithm.

Subjectively, as shown in Figure 5, the NNM algorithms were not able to meet the visual requirements, and the image was blurred. The WSNM algorithm and our algorithm both met the visual requirements. However, the edge processing of algorithm WSNM was not good. Our proposed algorithm could reduce this situation, as shown in the enlarged portion of the figure in the red box.

Objectively, as summarized in Table 3, the proposed algorithm was superior to the other two algorithms. All the objective data were improved.

Table 3. PSNR/FSIM in experiment (2).

| Image | Algorithm (PSNR/FSIM) | | |
|---------|-----------------------|----------------|-----------------------|
| | NNM | WSNM | Proposed |
| Bedroom | 23.0866/0.9475 | 23.4500/0.9798 | 23.5326/0.9820 |
| Lamp | 23.9880/0.8252 | 24.1906/0.8594 | 24.2457/0.8823 |
| Kitchen | 24.7820/0.8763 | 26.3996/0.8822 | 26.5009/0.9029 |

In summary, the proposed algorithm has certain advantages and can be used in depth image inpainting applications.

4.2. Parameter Influence

To discuss the influence of different parameters on the proposed algorithm, we analysed the Aloe, Art and Books depth images.

4.2.1. Number of Best-Matched Patches

The manual damage in this section was 10% data loss and 20% data loss. As shown in Figure 6, when were number of pixel blocks of a similar group was in the range of 20–100, the experimental curves are relatively flat. That is, the proposed algorithm was insensitive to the number of pixel blocks. Therefore, the number of pixel blocks was 60.

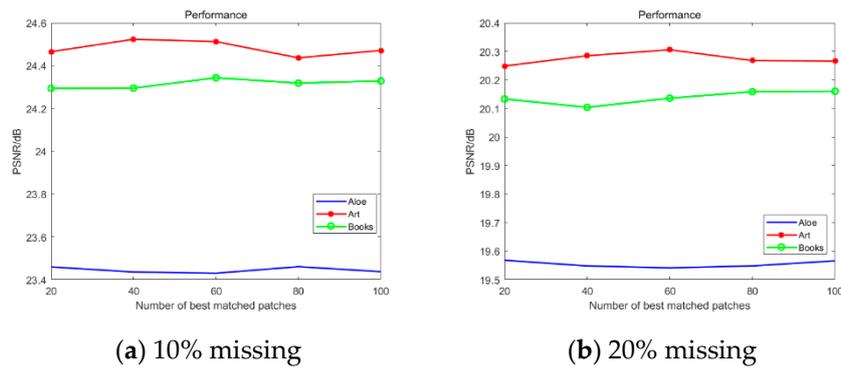


Figure 6. Performance comparison with best-matched patches.

4.2.2. Algorithm Stability

The manual damage in this section was 10% data loss and 20% data loss. Since the objective function is non-convex, it was difficult to mathematically prove the global convergence of the proposed algorithm. We obtained experimental data and empirically verified the stability of the proposed algorithm. As shown in Figure 7, the number of iterations increased, the PSNR increased monotonically and eventually stabilised. Therefore, the stability of the proposed algorithm could be verified.

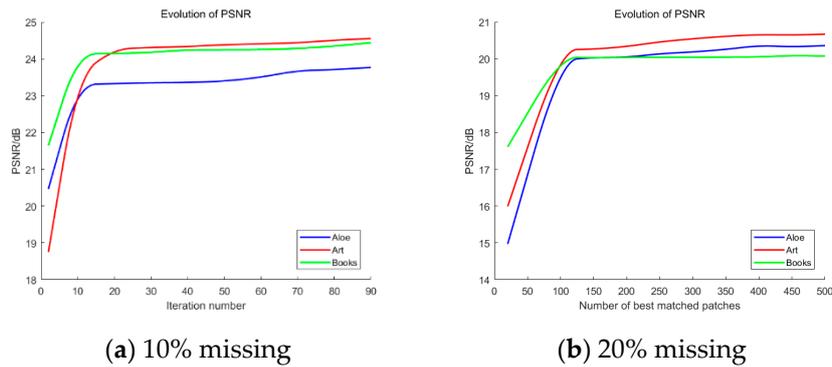


Figure 7. Stability of the proposed algorithm.

4.2.3. Influence of p

The manual damage in this section was 10% and 20% data loss. Figure 8 shows that the value of p was small and the objective data were improved. However, p was too small and excessive smoothing occurred. As shown in Figure 9, p was 0.05. Therefore, we chose $p = 0.2$, which agreed with [39].

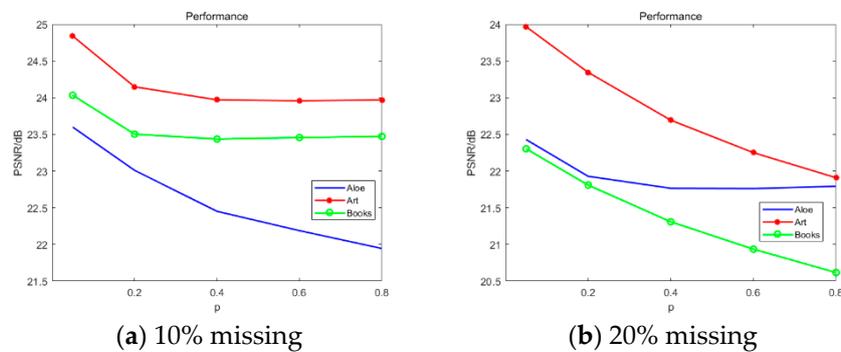


Figure 8. Performance comparison with p .



Figure 9. Visual quality.

5. Conclusions

The main research topic in our paper is depth image inpainting. The proposed method is based on the fact that depth in-painted images using a low-rank structure and nonlocal self-similarity can fully exploit the features of depth images to complete the inpainting task. The experiments prove that, regardless of the subjective visual effect and objective contrast data, the proposed algorithm can obtain a better repair effect and has a certain practical application value.

Author Contributions: W.J.: methodology, writing—review and editing; L.Z.: methodology, software, validation, analysis and writing; L.Y.: writing—review. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant Nos. U1704132 and 11747089 and by The Ninth Group of Key Disciplines in Henan Province under Grant No. 2018119.

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

References

- Deng, H.; Xu, T.; Zhou, Y.; Miao, T. Depth Density Achieves a Better Result for Semantic Segmentation with the Kinect System. *Sensors* **2020**, *20*, 812. [[CrossRef](#)]
- Dybedal, J.; Aalerud, A.; Hovland, G. Embedded Processing and Compression of 3D Sensor Data for Large Scale Industrial Environments. *Sensors* **2019**, *19*, 636. [[CrossRef](#)] [[PubMed](#)]
- Örücü, S.; Selek, M. Design and Validation of Rule-Based Expert System by Using Kinect V2 for Real-Time Athlete Support. *Appl. Sci.* **2020**, *10*, 611. [[CrossRef](#)]
- Zhang, C.; Huang, T.; Zhao, Q. A New Model of RGB-D Camera Calibration Based on 3D Control Field. *Sensors* **2019**, *19*, 5082. [[CrossRef](#)] [[PubMed](#)]
- Yazdi, M.Z. Depth-Based Lip Localisation and Identification of Open or Closed Mouth, Using Kinect 2. In Proceedings of the 15th International Workshop on Advanced Infrared Technology and Applications, Firenze, Italy, 17–19 September 2019; Volume 27, p. 22.
- Ophoff, T.; Van Beeck, K.; Goedemé, T. Exploring RGB-Depth Fusion for Real-Time Object Detection. *Sensors* **2019**, *19*, 866. [[CrossRef](#)] [[PubMed](#)]
- Dogan, S.; Haddad, N.; Ekmekcioglu, E.; Kondoz, A.M. No-Reference Depth Map Quality Evaluation Model Based on Depth Map Edge Confidence Measurement in Immersive Video Applications. *Future Internet* **2019**, *11*, 204. [[CrossRef](#)]
- Lie, W.-N.; Ho, C.-C. Multi-Focus Image Fusion and Depth Map Estimation Based on Iterative Region Splitting Techniques. *J. Imaging* **2019**, *5*, 73. [[CrossRef](#)]
- Dai, Y.; Fu, Y.; Li, B.; Zhang, X.; Yu, T.; Wang, W. A New Filtering System for Using a Consumer Depth Camera at Close Range. *Sensors* **2019**, *19*, 3460. [[CrossRef](#)]
- He, W.; Xie, Z.; Li, Y.; Wang, X.; Cai, W. Synthesizing Depth Hand Images with GANs and Style Transfer for Hand Pose Estimation. *Sensors* **2019**, *19*, 2919. [[CrossRef](#)]
- Liu, W.; Chen, X.; Yang, J. Robust Color Guide Depth Map Restoration. *IEEE Trans. Image Process.* **2017**, *26*, 315–327. [[CrossRef](#)]
- Lee, P.J. Nongeometric Distortion Smoothing Approach for Depth Map Preprocessing. *IEEE Trans. Multimed.* **2011**, *13*, 246–254. [[CrossRef](#)]
- Lei, J.; Li, L.; Yue, H.; Wu, F.; Ling, N.; Hou, C. Depth map super-resolution considering view synthesis quality. *IEEE Trans. Image Process.* **2017**, *26*, 1732–1745. [[CrossRef](#)] [[PubMed](#)]

14. Shen, Y.; Li, J.; Lu, C. Depth map enhancement method based on joint bilateral filter. In Proceedings of the 7th International Congress on Image and Signal Processing, Dalian, China, 14–16 October 2014; pp. 153–158.
15. Buyskens, P.; Le Meur, O.; Daisy, M. Depth-guided disocclusion inpainting of synthesized RGB-D images. *IEEE Trans. Image Process.* **2017**, *26*, 525–538. [[CrossRef](#)] [[PubMed](#)]
16. Lu, S.; Ren, X.; Liu, F. Depth enhancement via low-rank matrix completion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 3390–3397.
17. Xue, H.; Zhang, S.; Cai, D. Depth image inpainting: Improve Low Rank Matrix completion with Low Gradient Regularisation. *IEEE Trans. Image Process.* **2017**, *26*, 4311–4320. [[CrossRef](#)]
18. Scharstein, D.; Szeliski, R. High-accuracy stereo depth maps using structured light. In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA, 18–20 June 2003; pp. 195–202.
19. Scharstein, D.; Pal, C. Learning conditional random fields for stereo. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007.
20. Hirschmüller, H.; Scharstein, D. Evaluation of cost functions for stereo matching. In Proceedings of the IEEE Conference on CVPR, Minneapolis, MN, USA, 17–22 June 2007.
21. Candes, E.J.; Recht, B. Exact matrix completion via convex optimisation. *Found. Comput. Math.* **2009**, *9*, 717–772. [[CrossRef](#)]
22. Cai, J.F.; Candes, E.J.; Shen, Z.W. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **2010**, *20*, 1956–1982. [[CrossRef](#)]
23. Oh, T.H.; Tai, Y.W.; Bazin, J.; Kim, H. Partial sum minimisation of singular values in RPCA for low-level vision. In Proceedings of the IEEE CVPR, Columbus, OH, USA, 25–27 June 2013; pp. 744–758.
24. Xie, Y.; Gu, S.; Liu, Y.; Zuo, W.; Zhang, W.; Zhang, L. Weighted Schatten p-norm Minimisation for Image Denoising and Background Subtraction. *IEEE Trans. Image Process.* **2016**, *25*, 4842–4857. [[CrossRef](#)]
25. Gu, S.H.; Xie, Q.; Meng, D.Y. Weight Nuclear Norm Minimisation and Its Applications to Low Level Vision. *Int. J. Comput. Vis.* **2017**, *121*, 183–208. [[CrossRef](#)]
26. Buades, A.; Coll, B.; Morel, M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005; pp. 60–65.
27. Jung, M.; Bresson, X.; Chan, T. Nonlocal Mumford Shah regularizers for color image restoration. *IEEE Trans. Image Process.* **2011**, *20*, 1583–1598. [[CrossRef](#)]
28. Dong, W.; Zhang, L.; Shi, G.; Xu, W. Image deblurring and superresolution by adaptive sparse domain selection and adaptive regularisation. *IEEE Trans. Image Process.* **2011**, *20*, 1338–1857.
29. Zhang, J.; Zhao, D.; Xiong, R.; Ma, S.; Gao, W. Image Restoration Using Joint Statistical Modeling in a Space-Transform Domain. *IEEE Trans. Image Process.* **2014**, *24*, 915–928. [[CrossRef](#)]
30. Goldstein, T.; Osher, S. The split Bregman algorithm for L_1 regularized problem. *SIAM J. Imaging Sci. (SIIMS)* **2009**, *2*, 323–343. [[CrossRef](#)]
31. Zhang, J.; Zhao, D.; Gao, W. Group-based Sparse Representation for Image Restoration. *IEEE Trans. Image Process.* **2014**, *23*, 3336–3351. [[CrossRef](#)] [[PubMed](#)]
32. Candès, E.J.; Wakin, M.B.; Boyd, S. Enhancing sparsity by reweighted L_1 minimisation. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [[CrossRef](#)]
33. Mirshy, L. A trace inequality of John von Neumann. *Monatshefte Mathematik.* **1975**, *79*, 303–306. [[CrossRef](#)]
34. Zuo, W.; Meng, D.; Zhang, L.; Feng, X.; Zhang, D. A generalized iterated shrinkage algorithm for non-convex sparse coding. In Proceedings of the IEEE CVPR, Columbus, OH, USA, 25–27 June 2013; pp. 217–224.
35. Zhang, J.; Zhao, D.; Zhao, C.; Xiong, R.; Ma, S.; Gao, W. Image compressive sensing recovery via collaborative sparsity. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2012**, *2*, 380–391. [[CrossRef](#)]
36. Sheikh, H.R.; Sabir, M.F.; Bovik, A.K. A statistical evaluation of recent full reference image quality assessment algorithm. *IEEE Trans. Image Process.* **2006**, *15*, 3440–3451. [[CrossRef](#)]
37. Zhang, L.; Zhang, L.; Mou, X.Q.; Zhang, D. FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)]

38. Nathan, S.; Derek, H.; Pushmeet, K.; Rob, F. Indoor Segmentation and Support Inference from RGBD images. In Proceedings of the European Conference on Computer Vision, Florence, Italy, 7–13 October 2012.
39. Qiu, Y.F. Research on Image Completion Algorithm Based on Low Rank and Smooth Prior Information. Master's Thesis, Southwest University, Chongqing, China, 2018.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).