

Article

# Research on the Enhancement of Laser Radar Range Image Recognition Using a Super-Resolution Algorithm

Yu Zhai, Jieyu Lei, Wenze Xia, Shaokun Han \*, Fei Liu and Wenhao Li

Beijing Key Lab for Precision Optoelectronic Measurement Instrument and Technology, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; zhaiy\_0108@163.com (Y.Z.); jieyu\_lei@163.com (J.L.); xiawenze@126.com (W.X.); 00001liufei@163.com (F.L.); liwenhao\_bitgd@163.com (W.L.)

\* Correspondence: skhan@bit.edu.cn

Received: 3 August 2020; Accepted: 8 September 2020; Published: 11 September 2020



**Abstract:** This work introduces a super-resolution (SR) algorithm for range images on the basis of self-guided joint filtering (SGJF), adding the range information of the range image as a coefficient of the filter to reduce the influence of the intensity image texture on the super-resolved image. A range image SR recognition system is constructed to study the effect of four SR algorithms including the SGJF algorithm on the recognition of the laser radar (ladar) range image. The effects of different model library sizes, SR algorithms, SR factors and noise conditions on the recognition are tested via experiments. Results demonstrate that all tested SR algorithms can improve the recognition rate of low-resolution (low-res) range images to varying degrees and the proposed SGJF algorithm has a very good comprehensive recognition performance. Finally, suggestions for the use of SR algorithms in actual scene recognition are proposed on the basis of the experimental results.

**Keywords:** range image SR algorithm; target recognition; combined moment invariant; neural network

## 1. Introduction

With the advancement and development of technology, obtaining three-dimensional (3D) range images using laser radar (ladar) is easy and popular in recent years [1,2]. The recognition technology of the target object in the 3D range image is also extensively used in various fields such as high-altitude detection, precision guidance, automatic driving and smart homes and is a remarkable possibility for expansion remains in the future [3,4]. Owing to the limitation of hardware conditions and cost, the spatial resolution of the obtained 3D range images is typically low. Therefore, determining how to improve the spatial resolution of 3D range images through algorithms has become an important research direction of 3D imaging. This interest raises the question of whether we can improve the recognition rate of low-resolution (low-res) range images by increasing their spatial resolution.

Super-resolution (SR) of images can also be called up-sampling of images, which is a signal processing technique for estimating a high-resolution (high-res) image from a low-res image [5]. SR algorithms for range images are developed on the bases of SR algorithms for intensity images and can be classified into two categories according to the difference in input data sources. One uses only single- or multi-frame low-res range images such as bicubic interpolation (Bicubic) proposed by Hou [6], convex projections proposed by Stark [7], neighbor embedding proposed by Chang [8] and self-guided residual interpolation proposed by Konno [9]. Another uses the information of high-res intensity images of the same scene to guide the SR of range images [10] such as joint bilateral filter (JBF) proposed by Kopf [11], weighted mode filtering proposed by Min [12], Markov random field (MRF) proposed by Diebel [13] and total generalized variation (TGV) proposed by Ferstl [14]. Considering that these

algorithms must strictly align high-res intensity images and low-res range images, initially registering the two images through the registration algorithm is necessary. In addition, SR methods based on deep learning have also made significant progress in recent years [15,16]. However, the current research on the role of the SR algorithm in recognition is focused on specific fields such as 2D face recognition [17,18]. To the best of our knowledge, no relevant study covers the role of the SR algorithm in range image recognition.

This work aims to propose an SR algorithm for range images based on self-guided joint filtering (SGJF). The algorithm adds the interpolation result of the range image as a guide image to the filter kernel to reduce the influence of the intensity image texture on the super-resolved image. In addition, an SR recognition system for range images is constructed and a range image acquisition platform is designed to test the effect of SR algorithms in recognizing lidar range images. Four SR algorithms are tested, including Bicubic, JBF, MRF and our proposed SGJF. The SR recognition system uses combined moment invariants composed of seven Hu and three affine moment invariants to describe the global features of the image. Moreover, it uses a back propagation neural network (BPNN) to classify and recognize the features. The recognition rates of the SR recognition system under different SR algorithms, SR factors, model library sizes and noise conditions are tested via experiments. The results show that the SGJF algorithm we proposed can largely avoid the problem of ‘texture transfer’ in the JBF algorithm and is also robust to noise. All of the SR algorithms we tested can improve the recognition rate of low-res range images and a high SR factor increases the improvement. Regardless of the resolution of the model library, high resolution of the super-resolved scene raises the recognition rate and using the model library with the same resolution as the super-resolved scene can obtain the highest recognition rate. Finally, following the experimental results, the application suggestions of SR algorithms in actual scene recognition are proposed.

The remainder of this work is organized as follows. Section 2 presents the principles of the SR and recognition algorithm we used. Section 3 presents the experimental process and analyzes the effects of different SR algorithms, SR factors, model library sizes and noise conditions on the recognition rate. Section 4 provides the conclusions.

## 2. Theory

### 2.1. SR Algorithms

#### 2.1.1. Bicubic

The interpolation algorithm estimates unknown pixels on the basis of known pixel values in the neighborhood and image resolution conversion is one of its specific applications. Strictly, the resolution conversion must first reconstruct (or interpolate) discrete data into a continuous curve and then sample at different sampling rates. However, in a real interpolation algorithm, these two steps can be performed in one operation [6]. Bicubic is one of the most common interpolation algorithms. In this algorithm, the value of the function  $f$  at the sampling points  $(x, y)$  can be obtained by the weighted average of the nearest sixteen sampling points in the grid, which requires two polynomials to cubic interpolation in  $X$  and  $Y$  directions as shown in Equation (1):

$$f(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (1)$$

where  $a_{ij}$  is a weighting coefficient, which can be solved by constructing a bicubic basis function. Considering that the Bicubic can perform both up- and down-sampling, all down-sampling processes of images in the experiment in Section 3 use this algorithm. Compared with other SR algorithms, the Bicubic has a smaller calculation amount and has a smoothing effect on the image. However, the actual image is piecewise continuous and the global smoothness assumption is not satisfied in the discontinuous area (e.g., the edge of the image). Thus, when interpolating pixels near the edge

of the image, the resulting visual effect will become blurred. Furthermore, the interpolation result is susceptible to noise and pixels with large errors.

### 2.1.2. MRF

MRF is based on the Markov model and Bayesian theory. The Markov model describes the local statistical dependency relationship between image pixels and considers that a certain pixel value of an image is only related to some pixel values in its neighborhood. Bayesian theory is used to solve the uncertainty problem described by the MRF model, converts the prior knowledge of the image into the prior distribution model and then the maximum posterior probability is used to establish the objective function of the problem to transform the image processing problem into an optimization problem.

J. Diebel and S. Thru [13] used the MRF model for the first time to solve the problem of SR of range images and J. Park et al. [19] made improvements on this basis. The input is defined as the high-res intensity image  $I$  and the low-res range image  $D$  and the output as the estimated high-res range image  $\tilde{D}$ . The objective function  $E(\tilde{D})$  can be constructed as

$$E(\tilde{D}) = E_d(\tilde{D}) + \lambda_S E_S(\tilde{D}) + \lambda_N E_{NLS}(\tilde{D}), \quad (2)$$

where  $E_d(\tilde{D})$  is the data term,  $E_S(\tilde{D})$  is the neighborhood smoothing term and  $E_{NLS}(\tilde{D})$  is the NLS term.  $\lambda_S$  and  $\lambda_N$  are the relative weights to balance the three terms.

Specifically,  $E_d(\tilde{D})$  describes the quadratic distance between  $\tilde{D}$  and  $D$ :

$$E_d(\tilde{D}) = \sum_{p \in M} [\tilde{D}(p) - D(p)]^2, \quad (3)$$

where  $p$  is the index of the pixel position and  $M$  is the set of pixels of  $D$ .

$E_S(\tilde{D})$  represents the weighted quadratic distance between pixels in a neighborhood system:

$$\begin{aligned} E_S(\tilde{D}) &= \sum_p \sum_{q \in N(p)} \frac{w_{pq}}{W_p} [\tilde{D}(p) - \tilde{D}(q)]^2, \\ W_p &= \sum_q w_{pq}, \\ w_{pq} &= w_c w_s w_e w_d \end{aligned} \quad (4)$$

where  $N(p)$  is the first-order neighborhood of  $p$  and  $W_p$  is the normalization factor.  $w_{pq}$  is the confidence weight of pixels  $p$  and  $q$ , which defines the similarity of neighboring pixels and is the key to introducing the input  $I$  into the SR problem. The greater  $w_{pq}$  increases the probability that two neighboring pixels have the same range value.  $w_{pq}$  can be decomposed into four terms on the bases of color similarities  $w_c$ , segmentation  $w_s$ , edge saliency  $w_e$  and guided bicubic interpolated depth image  $w_d$ .

(1)  $w_c$  is defined in the YUV color space as

$$w_c = \exp \left\{ - \sum_{I \in YUV} \frac{[I(p) - I(q)]^2}{2\sigma_I^2} \right\}, \quad (5)$$

where  $\sigma_I$  controls the relative sensitivity of different color channels.

(2) The image segmentation algorithm in the VLFeat image processing library provided by [20] is used to segment the intensity image into super pixels. For the neighborhood pixels not within the same super pixel, a penalty term is defined as

$$w_s = \begin{cases} 1 & S_{co}(p) = S_{co}(q) \\ t_{se} & otherwise \end{cases}, \quad (6)$$

where  $S_{co}(\cdot)$  is the segmentation label and  $t_{se}$  is the penalty factor ( $0 < t_{se} < 1$ ).

- (3)  $w_e$  indicates the weight that depends on the edge saliency response. The edge saliency responses are detected by a set of Gabor filters with different sizes and orientations, which can enhance the smoothness at the edge of the image. The weight is calculated as

$$w_e = \frac{1}{\sqrt{S_x(p)^2 + S_x(q)^2 + 1}}, \quad (7)$$

where  $S_x(\cdot)$  is the value of the  $x$ -axis edge saliency image. If  $p$  and  $q$  are  $x$ -axis neighborhoods,  $S_x(\cdot)$  is calculated. If they are  $y$ -axis neighborhoods,  $S_y(\cdot)$  is calculated.

- (4) Allowing the range values to propagate freely with only very sparse data constraints can lead to notable range ambiguity. Therefore, the guide range image  $D_g$  obtained from the low-res range image through Bicubic is introduced to solve this problem. Similar to the bilateral filter, the weight of the guide range image is defined as

$$w_d = \exp\left\{-\frac{[D_g(p) - D_g(q)]^2}{2\sigma_g^2}\right\}. \quad (8)$$

$E_{NLS}(\tilde{D})$  is defined by the anisotropic structural-aware filter [21]:

$$E_{NLS}(\tilde{D}) = \frac{\sum_p \sum_{r \in H(p)} \frac{k_{pr}}{K_p} [\tilde{D}(p) - \tilde{D}(r)]^2}{K_p = \sum_r k_{pr}}, \quad (9)$$

where  $H(p)$  is the local window (e.g.,  $7 \times 7$ ) of  $p$  and  $K_p$  is the normalization constant.  $k_{pr}$  is the weight of the filter and is defined as

$$k_{pr} = \frac{1}{2} \left\{ \exp\left[-(p-r)^T \Sigma_p^{-1} (p-r)\right] + \exp\left[-(p-r)^T \Sigma_r^{-1} (p-r)\right] \right\},$$

$$\Sigma_p = \frac{1}{|H|} \sum_{p' \in H(p)} \nabla I(p') \nabla I(p')^T, \quad (10)$$

$$\nabla I(p) = [\nabla_x I(p), \nabla_y I(p)]^T,$$

where  $\nabla I(p)$  is the  $x$ - and  $y$ - image gradient vector at pixel  $p$ . The structural-aware filter defines the structural similarity of  $p$  and  $r$  in  $I$ .

Finally, the conjugate gradient method is used to solve  $\tilde{D}$ . The first derivative w.r.t.  $\tilde{D}$  on Equation (2) is taken and the derivative equal to zero is set:

$$\frac{\Delta E(\tilde{D})}{\Delta \tilde{D}} = 0. \quad (11)$$

After finishing, the system of linear equations can be obtained:

$$A\tilde{d} = d, \quad (12)$$

where  $A$  is a  $n \times n$  matrix with weight terms and  $n$  is the number of pixels in  $I$ .  $\tilde{d}$  is the desired range values and  $d$  is an observed range that is conditionally filled with  $A$ . For  $\tilde{D}(p)$ , the elements of  $A$  and  $d$  are filled as

$$A_{pp} = \begin{cases} 1 + \lambda_S + \lambda_N & \text{if } p \in M \\ \lambda_S + \lambda_N & \text{otherwise} \end{cases}, \quad (13)$$

$$A_{pq} = -\frac{\lambda_S w_{pq}}{W_p}, \quad (14)$$

$$A_{pr} = -\frac{\lambda_N k_{pr}}{K_p}, \quad (15)$$

$$d_p = \begin{cases} D(p) & \text{if } p \in M \\ 0 & \text{otherwise} \end{cases}, \quad (16)$$

where  $A_{pp}$  is the element on the diagonal of  $A$ ,  $A_{pq}$  is  $A$ 's element at the  $p$ -th row and  $q$ -th column and  $d_p$  is the  $p$ -th element of  $d$ .  $q \in N(p)$ ,  $r \in H(p)$  indicate neighborhoods for  $E_S(\tilde{D})$  and  $E_{NLS}(\tilde{D})$ , respectively. Finally,  $\tilde{d}$  can be obtained by solving Equation (12) and then  $\tilde{D}$  can be reconstructed from  $\tilde{d}$ .

The specific implementation steps of the MRF algorithm are summarized as follows:

1. Input low-res range image  $D$  and high-res intensity image  $I$ .
2. Set parameters  $\lambda_S$ ,  $\lambda_N$ ,  $\sigma_I$ ,  $\sigma_g$  and  $t_{se}$ .
3. Calculate parameters  $w_{pq}$ ,  $W_p$ ,  $k_{pr}$  and  $K_p$  for all pixels according to Equations (4) to (10).
4. Calculate the matrix  $A$  and vector  $d$  according to Equations (13) to (16).
5. Solve Equation (12) to obtain vector  $\tilde{d}$ .
6. Reconstruct  $\tilde{d}$  to obtain the high-res range image  $\tilde{D}$ .

### 2.1.3. JBF

The edge is one of the most important structural features of the image, which reflects the details of the image and belongs to the sensitive high-frequency information of the human eye. Generally, the smoothing effect of the filter will cause blurring in the image edge area, which will reduce people's ability to distinguish the details of objects. Therefore, in many computer vision applications, filters are required to have edge preservation capabilities.

The bilateral filter is an edge-preserving filter, which has edge-preserving capability whilst smoothing the image and has been widely used in image processing [22]. The kernel function of the bilateral filter can be divided into two parts: the spatial and range filter kernels. Among them, the spatial filter kernel is the same as the Gaussian filter and the range filter kernel is a Gaussian function that calculates the difference between the central pixel and its neighborhood. For point  $p$  in image  $G$ , the bilateral filtering result can be described as

$$J(p) = \frac{1}{k_p} \sum_{q \in \Omega} w_{pq} G(q),$$

$$w_{pq} = \exp\left(-\frac{|p-q|^2}{\sigma_d}\right) \exp\left(-\frac{|G(p)-G(q)|^2}{\sigma_c}\right), \quad (17)$$

where  $p$  and  $q$  are pixel coordinate.,  $\Omega$  is the filter window,  $k_p$  is the normalization factor,  $w_{pq}$  is the filter kernel and  $\sigma_d$  and  $\sigma_c$  are the corresponding weight parameters. Given that the bilateral filter includes a spatial filter and also combines a range filter that describes the intensity similarity, the filter coefficients calculated for the smooth and edge areas of the image are different. The filter coefficient is smaller in the edge area of the image, which can protect the edges better.

J. Kopf et al. [11] first proposed the idea of JBF and applied it to the image SR problem. The principle was to construct the spatial filter from a low-res range image  $D$  and the range filter from a high-res intensity image  $I$  as shown in Equation (18):

$$\tilde{D}(p) = \frac{1}{k_p} \sum_{q_1 \in \Omega} w_{pq} D(q_1),$$

$$w_{pq} = \exp\left(-\frac{|p-q_1|^2}{\sigma_d}\right) \exp\left(-\frac{|I(p)-I(q)|^2}{\sigma_c}\right), \quad (18)$$

where  $\tilde{D}$  is the high-res range image to be restored and  $p_1$  and  $q_1$  denote (possibly fractional) coordinates of pixels in  $D$  corresponding to  $p$  and  $q$  in  $\tilde{D}$ .

### 2.1.4. SGJF

The filter of the JBF algorithm only considers the spatial Gaussian kernel and the intensity Gaussian kernel; that is, it assumes that where the intensity image gradient is large, the range image gradient is also large. However, the range image actually often makes smooth changes in sections; that is, the range information belonging to the same area of the object is constant or gradual, whereas the corresponding color guide image may have a certain texture structure that will cause the latter to be introduced into the super-resolved images after filtering. This phenomenon is called ‘texture transfer’.

To reduce the effect of intensity image texture on the super-resolved images, we propose the SGJF algorithm. This algorithm adds the range information of the range image as a coefficient of the filter and reduces the influence of the intensity image texture by strengthening its weight. The specific method is to interpolate the low-res range image (e.g., using the Bicubic algorithm) to obtain a high-res range guided image  $R_g$  and then express it in a form of similar intensity as a Gaussian filter kernel as

$$\exp\left(-\frac{|R_g(p) - R_g(q)|^2}{\sigma_u}\right), \quad (19)$$

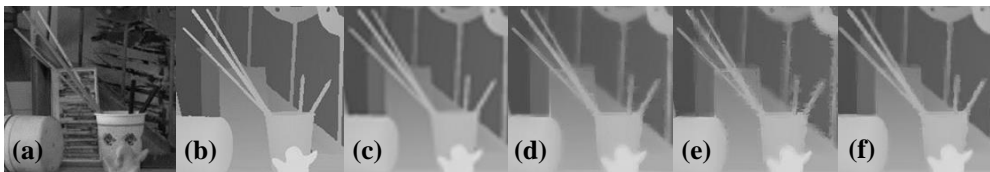
where  $\sigma_u$  is the weight parameter; a larger value increases the weight. Finally, Equation (19) is added as a coefficient to Equation (18) to obtain the final filter kernel as

$$w_{pq} = \exp\left(-\frac{|p_{\downarrow} - q_{\downarrow}|^2}{\sigma_d}\right) \exp\left(-\frac{|I(p) - I(q)|^2}{\sigma_c}\right) \exp\left(-\frac{|R_g(p) - R_g(q)|^2}{\sigma_u}\right), \quad (20)$$

The specific implementation steps of the SGJF algorithm are summarized as follows:

1. Input low-res range image  $D$  and high-res intensity image  $I$ .
2. Calculate the magnification according to the sizes of  $I$  and  $D$  and then calculate the guide image  $R_g$  using the Bicubic.
3. Set the filter window size  $\omega$  and weight parameters  $\sigma_d$ ,  $\sigma_c$  and  $\sigma_u$ .
4. Calculate the filter kernel  $w_{pq}$  of pixel  $p$  according to Equation (20).
5. Calculate the range value  $\tilde{D}(p)$  at pixel  $p$  according to Equation (18).
6. Repeat steps 4 and 5 for all pixels to obtain a high-res range image  $\tilde{D}$ .

Figure 1 shows the SR results of the Middlebury Art dataset with an SR factor of 4. The edges of the super-resolved image obtained by Bicubic are blurred. The super-resolved image obtained by MRF is less affected by the texture of the intensity image because the global optimization algorithm averages the local error to the entire image. However, it has fuzzy restoration details. The super-resolved image obtained by JBF is affected by the intensity image and blurring occurs at the slender edges caused by ‘texture transfer’. The super-resolved image obtained by SGJF has clear image edges and is less affected by the intensity image texture.



**Figure 1.** Visual comparison of super-resolved images of the Middlebury Art dataset. (a) Original high-res intensity image. (b) Original high-res range image. Super-resolved image by (c) Bicubic, (d) Markov random field (MRF), (e) joint bilateral filter (JBF) and (f) self-guided joint filtering (SGJF).



## 2.2. Recognition Algorithm

Target recognition based on ladar range images generally includes two processes of model acquisition and scene recognition and four steps of noise suppression, image segmentation, feature description and feature classification [23]. The model refers to the pre-defined 3D information of the target object to be recognized (i.e., the range images with known classification results) and the scene refers to the measurement data of the ladar (i.e., the range images to be classified). Generally, the range images acquired by ladar always contain various noises, which seriously affect the accuracy of the recognition algorithm. Therefore, the first step in target recognition is to suppress noise. Recognition algorithms based on local features typically require no image segmentation, whereas the recognition algorithms based on global features must undergo image segmentation. Feature description is intended to describe the geometric features of the target object in 3D space through mathematical language. Moreover, the feature classification refers to the classification of the geometric features described by mathematical language through algorithms to finally achieve target recognition.

### 2.2.1. Pre-Processing

Pre-processing includes noise suppression and image segmentation. Since the acquired range images are prone to produce some holes or obvious noises at the edges and areas with low reflectance of the target object, it is necessary to fill in these missing areas and eliminate outliers before processing the image. We used a statistical-based method [24] to complete the filling process by replacing the zero and extreme pixels with the statistical mode of the surrounding 25 pixels. This method can return sharper edges than using statistical mean values. Notably, we only corrected the outliers and did not change other known pixels, thereby avoiding blurring of the image edges.

After eliminating the outliers, we removed the ground and partial noise by fitting the plane. The least square method was initially used to fit the plane function of the ground through the outermost ring range data of the range image. The ground and partial noise could then be removed by giving up the range data of the plane and its vicinity. After that, a binary operation was used to extract the region of the target object to complete the image segmentation. However, when there was strong noise in the range image, the above operation was not enough to remove all the noise. It was also necessary to clean up the remaining noise spots by the method of extraction of connected components to obtain a clean binary image.

To improve the robustness, we acquired three range images for each pose of the target object and pre-processed them separately. The final binary image was obtained by voting after accumulating the three pre-processed images.

### 2.2.2. Combined Moment Invariants

Moment invariant as a global feature description vector has important applications in the field of ladar range image target recognition [25]. In general, lower order moments can reflect the overall characteristics of the image and higher order moments can reflect the detailed characteristics of the image. We used the combined moment invariants composed of Hu's seven moment invariants [26] constructed by the second- and third-order central moments and three of the affine moment invariants [27] constructed from algebraic invariant theory as the feature descriptors. The combined moment invariants have translation, rotation and scale invariance and are written as

$$\phi_1 = \eta_{20} + \eta_{02}, \quad (21)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \quad (22)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \quad (23)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (24)$$

$$\begin{aligned} \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ &+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right], \end{aligned} \quad (25)$$

$$\phi_6 = (\eta_{20} - \eta_{02})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \quad (26)$$

$$\begin{aligned} \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ &+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right], \end{aligned} \quad (27)$$

$$\phi_8 = \frac{1}{\eta_{00}^4}(\eta_{20}\eta_{02} - \eta_{11}^2), \quad (28)$$

$$\phi_9 = \frac{1}{\eta_{00}^2}(\eta_{30}^2\eta_{03}^2 - 6\eta_{30}\eta_{12}\eta_{21}\eta_{03} + 4\eta_{30}\eta_{12}^3 + 4\eta_{03}\eta_{21}^3 - 3\eta_{21}^2\eta_{12}^2), \quad (29)$$

$$\phi_{10} = \frac{1}{\eta_{00}^7}(\eta_{20}\eta_{21}\eta_{03} - \eta_{20}\eta_{12}^2 - \eta_{11}\eta_{30}\eta_{03} + \eta_{11}\eta_{21}\eta_{12} + \eta_{02}\eta_{30}\eta_{12} - \eta_{02}\eta_{21}^2), \quad (30)$$

where  $\eta_{pq}$  is the normalised center moment. To adjust the distribution range of the ten feature description vectors, we represented them in logarithm as

$$\mu_i = |\log|\phi_i||, \quad i = 1, 2, \dots. \quad (31)$$

### 2.2.3. BPNN

All of the features of all of the models described by combined moment invariants were then used to train the classifier for scene recognition. BPNN as a classifier is currently widely used to classify ladar range image features [28]. It uses gradient descent to calculate the correction amount of network connection weights and continuously corrects network weights from back to front until the network output error reaches a predetermined threshold.

Generally, BPNN includes at least three layers: an input layer, a hidden layer and an output layer. The input layer to the output layer is one-way propagation and the relationship between the input data and the output result is non-linear. The three-layer BPNN used in the experiment can be expressed as [29]

$$g_k(X) \equiv z_k = f\left[\sum_{j=1}^{n_H} w_{jk}f\left(\sum_{i=1}^d w_{ji}x_i + w_{j0}\right) + w_{k0}\right], \quad (32)$$

where  $f$  is the transfer function,  $x$  is the input vector,  $z$  is the output vector,  $w$  is the weight of each layer and  $d$  and  $n_H$  denote the dimensions of the input vector and the hidden layer, respectively.

## 3. Experiments and Results

This section is divided into four parts. Section 3.1 introduces the overall steps of the experiments and the platform used to acquire range images. Sections 3.2–3.4 are three independent experiments and each experiment is described in the order of experiment purpose, organization method and result analysis.

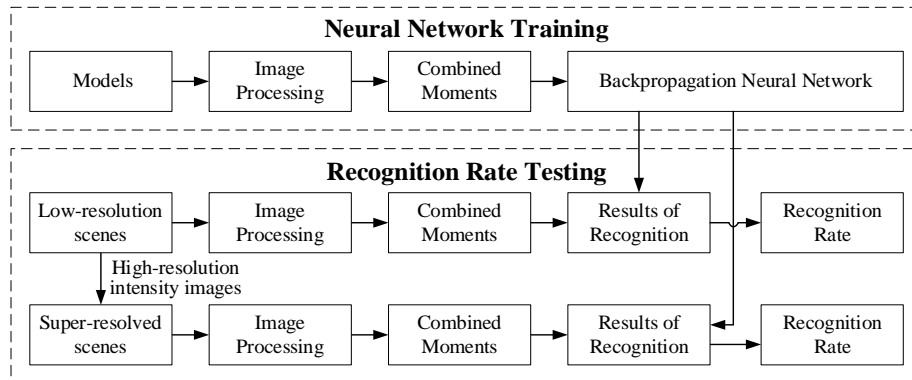
### 3.1. Experimental Steps and Platform

#### 3.1.1. Experimental Steps

Figure 2 illustrates the operating steps of the designed SR recognition system for range images, which is divided into neural network training and recognition rate testing. For training neural networks, the noise of models was initially removed through image pre-processing and converted into binary images with target shape information. The combined moments were then used as the feature descriptor to calculate the feature vectors of the binary image. Finally, the BPNN was trained by these feature vectors as the feature classifier for scene recognition. In the process of recognition rate testing,



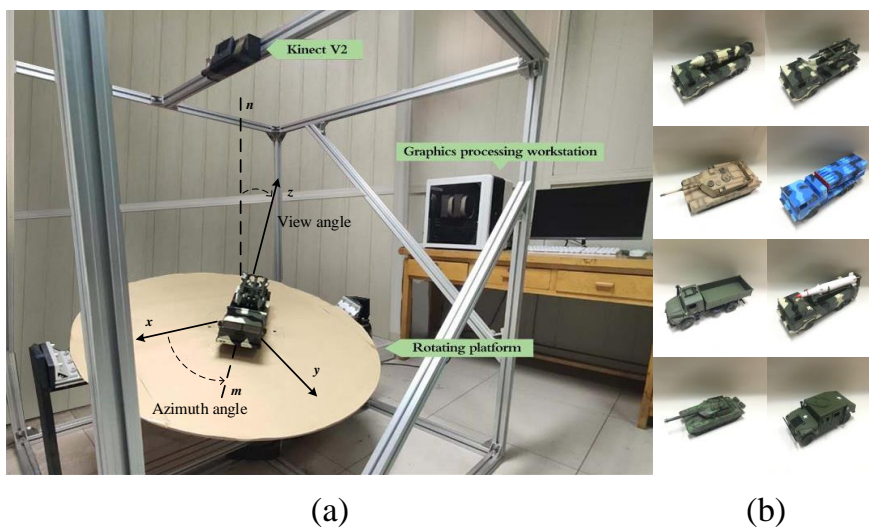
the low-res scenes and their super-resolved images were calculated in the same way to obtain feature vectors. These were then inputted into the trained BPNN to obtain the classification results before calculating the recognition rate. Finally, the effect of the SR algorithms on recognition was evaluated by comparing the two sets of recognition rates.



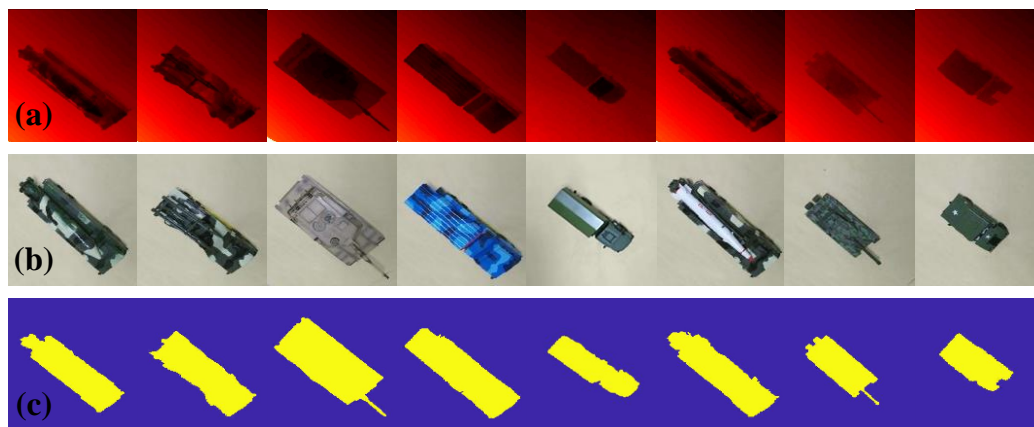
**Figure 2.** Block diagram for the super-resolution (SR) recognition system.

### 3.1.2. Image Acquisition Platform

An image acquisition platform (Figure 3a) was designed to acquire range and intensity images of eight target objects (Figure 3b). This platform was equipped with a Kinect V2, which combined a Time-of-Flight (ToF) camera and an RGB camera. Moreover, it acquired range and intensity images of the target objects at any attitude angle (azimuth angle  $0^{\circ}$ – $360^{\circ}$  and view angle  $0^{\circ}$ – $80^{\circ}$ ). The acquired images were divided into training and test sample sets in which the former was used to simulate the model library to train the BPNN and the latter was used to simulate the scene library to test the recognition rate. We defined the acquired range image, with a spatial resolution of  $320 \times 320$  pixels, as the original high-res range image (OHRRI) and the intensity image aligned with the OHRRI as the original high-res intensity image (OHRRI) in the experiments. In addition, given that temperature could have a profound influence on the credibility of the range data acquired by the Kinect V2 ToF camera, all data of the experiment were acquired after the Kinect V2 was run for 30 min to avoid temperature effects [30]. Figure 4 shows the range and intensity images of the eight target objects and the binary image obtained after pre-processing.



**Figure 3.** Experimental setup. (a) Image acquisition platform. (b) Target objects.



**Figure 4.** Acquired and pre-processed target object images. (a) Range images. (b) Intensity images. (c) Binary images.

Table 1 shows the parameters of the SR algorithms. For an objective comparison, the values of the common parameters of JBF and SGJF were set to be the same and the value of  $\omega$  was different according to the SR factor as  $s$ . Bicubic used the function `imresize()` in MATLAB. For BPNN, the input layer node was set to 10 to correspond to 10 moment invariant feature description vectors and the output node was set to 8 to correspond to 8 target objects. The transfer function of its first layer was the ‘Tansig’ function and the second layer was the ‘Purelin’ function. The number of iterations was set to 500. Considering that different training sample sets contained different numbers of training samples, the number of hidden layer neurons was also different when training BPNN to avoid underfitting or overfitting. The selection of all parameters represented the best result of multiple tests.

**Table 1.** SR algorithm parameters.

MRF	Value	JBF	SGJF	Value
$\lambda_S$	0.20	$\sigma_d$		3
$\lambda_N$	0.20	$\sigma_c$		0.10
$\sigma_I$	0.05		$\sigma_u$	0.15
$\sigma_g$	0.05			3 ( $s = 2$ )
$t_{se}$	0.70	$\omega$		5 ( $s = 4$ )
				9 ( $s = 8$ )

### 3.2. Arbitrary Attitude Angle Recognition

In the case where the SR algorithm was not used, the recognition effect of the recognition system on scenes with any attitude angle within the acquisition range of the platform was tested and the influence of the model library size on the recognition rate was analyzed. A test sample set for simulating a scene library was firstly established. At  $10^\circ$  intervals, range images of each target object at view angles of  $5^\circ$  to  $75^\circ$  and azimuth angles of  $5^\circ$  to  $175^\circ$  were acquired; thus, the test sample set contained 1152 range images. Next, training sample sets for simulating model libraries were established. Still at  $10^\circ$  intervals, range images of each target object at view angles of  $10^\circ$  to  $80^\circ$  and azimuth angles of  $0^\circ$  to  $180^\circ$  were acquired and one image when the view angle was  $0^\circ$ ; thus, a total of 153 range images of each target object were acquired. Different numbers of range images were finally extracted as training samples to establish the training sample sets with different sizes. According to the view angle, the number of range images taken at different intervals were as follows: (1) three at  $40^\circ$ , (2) five at  $20^\circ$  and (3) nine at  $10^\circ$ . For azimuth, the numbers of range images taken at different intervals were as follows: (1) three at  $90^\circ$ , (2) four at  $60^\circ$ , (3) seven at  $30^\circ$ , (4) 10 at  $20^\circ$  and (5) 19 at  $10^\circ$ . Combining the view and azimuth angles, 15 training sample sets with different sizes were obtained. Table 2 shows the recognition rates

of the test sample set by 15 BPNNs trained by the training sample sets and the number of samples contained in each training sample set is shown in parentheses.

**Table 2.** Recognition rates on different training sample sets.

View Interval	Azimuth Interval				
	90° (3)	60° (4)	30° (7)	20° (10)	10° (19)
40° (3)	87.50% (56)	87.76% (72)	89.41% (120)	92.18% (168)	93.40% (312)
20° (5)	92.80% (104)	94.36% (136)	95.05% (232)	95.66% (328)	96.18% (616)
10° (9)	94.79% (200)	96.18% (264)	96.88% (456)	97.22% (648)	97.22% (1224)

From the table, the recognition rate of the test sample set by the BPNN was typically positively related to the training sample set size. However, when the training sample set had fewer samples selected at the view angle (or azimuth angle), the recognition rate was low even if many samples were selected in another direction. By contrast, when the training sample set size increased to a certain number, the recognition rate grew slowly or even stopped. If we wanted to maintain the recognition rate above 95%, a minimum of 232 training samples were required and if we wanted to maintain the recognition rate above 90%, a minimum of 104 training samples were required.

### 3.3. SR Recognition of Scene-Model with the Same Resolution

#### 3.3.1. Experiment without Noise

The SR recognition of low-res scenes without noise was tested and the resolution of their super-resolved scene was the same as the resolution of the model. The down-sampling factor was defined as  $d$  and the SR factor as  $s$ . Firstly, OHRRI and OHRI of each target object (Figure 5a) were acquired. OHRRI were then down-sampled by factors  $d$  of 2, 4 and 8 to simulate low-res scenes (Figure 5b). Finally, the four SR algorithms described above were applied to low-res scenes and  $s$  was the same as  $d$  (Figure 5c). OHRI were used as guidance for SR algorithms requiring intensity image guidance.

The binary image could well represent the contour of the target object and the effect of the down-sampled images slowly worsened as the factor  $d$  increased. For super-resolved images, the difference was not very obvious when the factor  $s$  was small. When the factor was large, the edge of the Bicubic became blurred, MRF was difficult to see the details, JBF had obvious texture transfer and SGJF was less affected by the texture of intensity and the details were clear.

To test the recognition rate, the training sample sets were established by OHRRI and the test sample sets were established by down-sampled images and their SR images. Compared with Section 3.2, the attitude angle of the acquired image was simplified to only acquire images with a view angle of 40°, whereas the acquisition method of the azimuth angle was unchanged. For the training sample set, extracting the azimuth angle in the same way as in Section 3.2 established five training sample sets with different sizes and the number of range images they contained was 24, 32, 56, 80 and 152, respectively. For the test sample set, each target object first acquired 18 images. These were then down-sampled by factors  $d$  of 2, 4 and 8 to establish three down-sampling test sample sets. Next, the four SR algorithms were applied to the three down-sampling test sample sets with the same  $s$  as  $d$ . According to the different SR algorithms and factors  $s$ , 12 SR test sample sets were established. Therefore, a total of 15 test sample sets were finally established and each of them contained 144 (18 × 8) images. Figure 6 shows the recognition results of the 15 test sample sets by five BPNNs trained by the training sample sets.

Figure 6 shows that the recognition rate of all SR test sample sets was higher than that of the down-sampling test sample sets, especially when the factor was large. It represented that the SR algorithm helped improve the recognition rate of the acquired low-res scenes; a high factor increased the improvement. However, for each SR algorithm, the recognition rate of the large factor sample set was lower than that of the small factor sample set. Specifically, the SGJF test sample set often

maintained a high recognition rate. The JBF test sample set had a lower recognition rate than others when the factor was small. However, as the factor increased, its decrease was smaller. The MRF test sample set had the highest recognition rate when the factor was small and was only lower than the SGJF test sample set when the factor was large. The recognition rate of the Bicubic test sample set decreased the most with the increase in the factor.

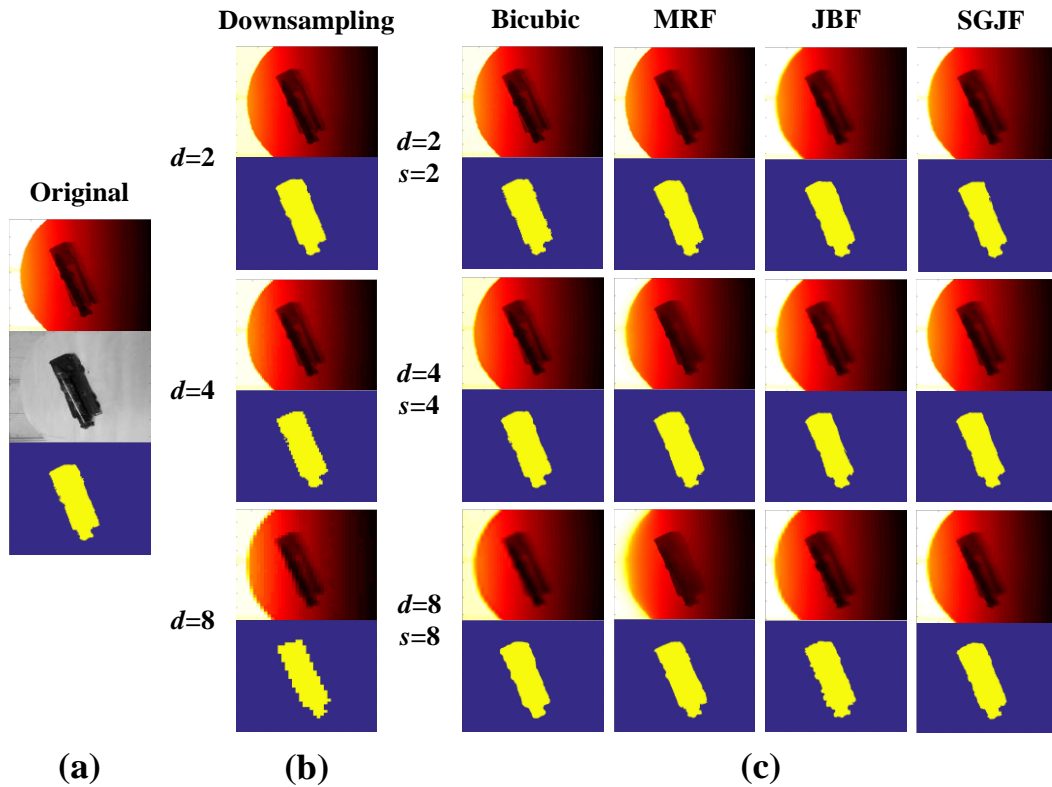


Figure 5. Explanation of the experiment phenomenon. (a) Original images. (b) Down-sampled images. (c) Super-resolved images.

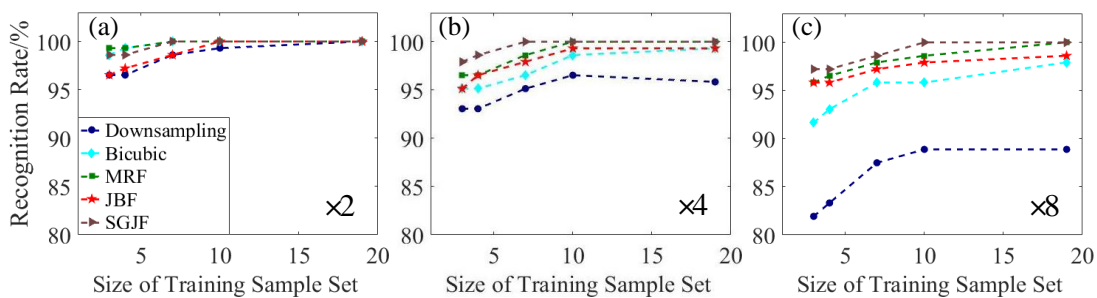
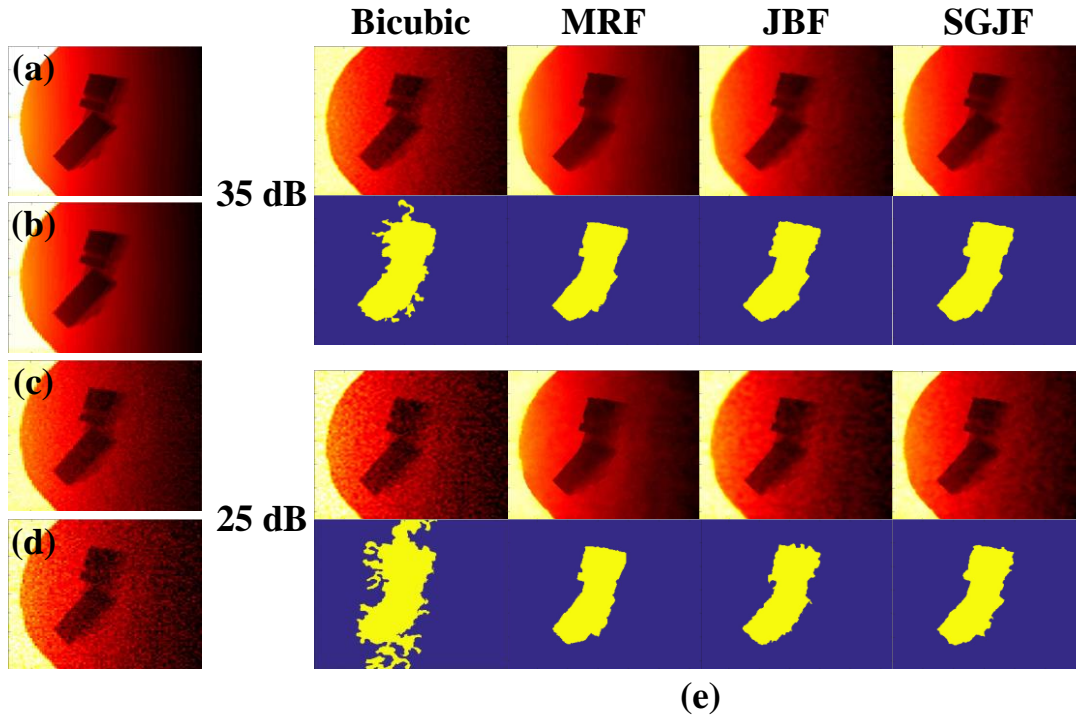


Figure 6. Comparison of recognition rates of different SR algorithms and factors without noise. (a)  $d = s = 2$ . (b)  $d = s = 4$ . (c)  $d = s = 8$ .

### 3.3.2. Experiment with Noise

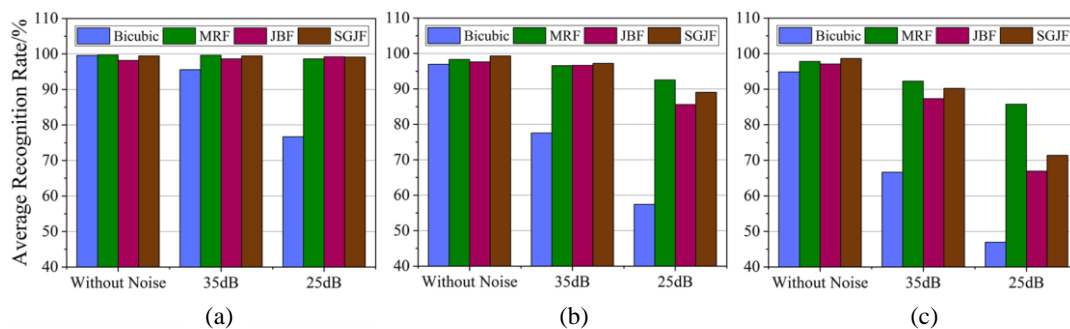
In practical applications, the ladar range image often contains various noises, mainly Gaussian noise. To test the influence of noise on the SR algorithm in the system, this section used the same experimental procedure as Section 3.3.1. The difference was that before running the SR algorithm, Gaussian noise with a signal-to-noise ratio (SNR) of 35 dB and 25 dB was first added to the down-sampled image. Figure 7 shows the super-resolved images and their binary images when the factor was 4. The Bicubic algorithm was incapable of suppressing noise and the noise seriously affected the extraction of binary images. The other three algorithms obviously suppressed the noise, where the

MRF super-resolved range image was the clearest and its binary image obtained smooth edges even when the noise intensity was large. Both JBF and SGJF super-resolved range images were somewhat affected when the noise intensity was large. In addition, the JBF super-resolved range image was also affected by the intensity image texture. These findings were all reflected in the binary image.



**Figure 7.** Explanation of the experiment phenomenon. (a) Original high-res range image (OHRRI). (b) Down-sampled image at  $d = 4$ . (b,c) with 35 dB noise. (b,d) with 25 dB noise. (e) Super-resolved images of (c,d) and their binary images.

The five BPNNs trained by the training sample set in Section 3.3.1 were still used to test the recognition rate and the acquisition of the test sample set also followed the steps in Section 3.3.1, except for adding noise with a different SNR. According to the result of Section 3.3.1, the recognition rate increased with the training sample set size and the recognition rate curves of different SR algorithms basically did not cross. Therefore, the average recognition rate (i.e., the average of the recognition rates of five different BPNNs) was used to describe the recognition results of the system under noisy conditions, as shown in Figure 8.



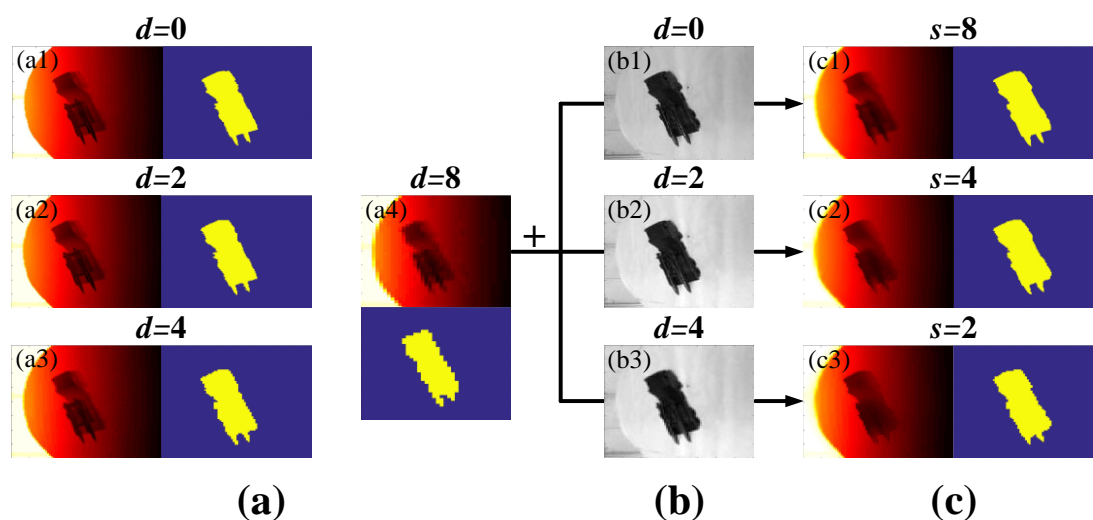
**Figure 8.** Comparison of the average recognition rates of different SR algorithms and factors with noise. (a)  $d = s = 2$ . (b)  $d = s = 4$ . (c)  $d = s = 8$ .



Overall, a larger factor hastened the decrease in the average recognition rate and 25 dB decreased faster than 35 dB. Specifically, the Bicubic algorithm was extremely sensitive to noise and was unsuitable for use when the noise intensity and factor was large. The MRF algorithm performed best in noise and the recognition rate of its test sample set only decreased slowly with the increase in noise intensity and factor. However, it needed to adjust more parameters and the calculation was more complicated. The JBF and SGJF algorithms were also robust to noise. When the factor or noise intensity was low, they were less affected. However, when the factor and noise intensity increased simultaneously, the average recognition rate dropped significantly and the gap between the JBF and SGJF algorithms also increased. Taken together, the average recognition rate of the SGJF algorithm was significantly smaller than the MRF algorithm only when the noise intensity and factor were at the maximum, whilst it was superior or similar to other algorithms in other cases. Therefore, the SGJF algorithm had a very good comprehensive recognition performance.

### 3.4. SR Recognition of Scene-Model with Multiple Resolutions

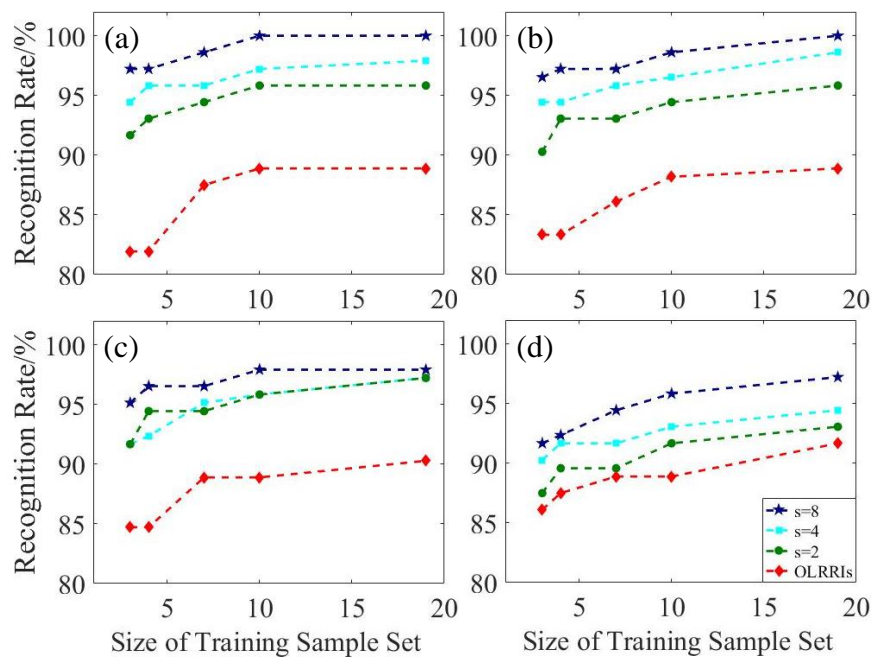
Recognizing multi-resolution (multi-res) super-resolved scene libraries by multi-res model libraries without noise was tested where the resolution of the super-resolved scene library may have been lower than, equal to or higher than the model library. Firstly, OHRRI and OHRRII of each target object (Figure 9(a1,b1)) were acquired. OHRRII were then down-sampled by factors  $d$  of 2, 4 and 8 (Figure 9(a2–a4)) and the down-sampled image with  $d = 8$  was called the original low-res range image (OLRRI). Meanwhile, the down-sampled images of OHRRII with factors  $d$  as 2 and 4 (Figure 9(b2,b3)) together with OHRRII were used as the guide images. Finally, the SGJF algorithm with a very good comprehensive recognition performance was applied to OLRRI and the factors  $s$  were 2, 4 and 8 (Figure 9(c1–c3)). The smaller the factor  $d$  of the down-sampled image and the larger the factor  $s$  of the SR image, the clearer the image and the smoother the edges of their binary image.



**Figure 9.** Explanation of the experiment phenomenon. (a) OHRRI and its down-sampled images. (b) OHRRII and its down-sampled images. (c) Super-resolved images of the original low-res range image (OLRRI).

To test the recognition rate, training and test sample sets were established and the attitude angle of the image in each sample set was the same as Section 3.3.1. The images in the training sample sets contained four different resolutions, which were OHRRI and its down-sampled image with factors  $d$  of 2, 4 and 8 (i.e., OLRRI). Moreover, each resolution image established five training sample sets with different sizes. Therefore, a total of 20 training sample sets were obtained. In addition, the resolutions of the four test sample sets established were also different. They were composed of OLRRI and their

super-resolved images obtained by SGJF algorithm with factors  $s$  of 2, 4 and 8. Figure 10 shows the recognition results of the four test sample sets by 20 BPNNs trained by the training sample sets.



**Figure 10.** Comparison of the recognition rate of the multi-res training sample set with the multi-res test sample set without noise. The resolution of the training sample set was (a) OHRRI. (b) Down-sampled image of OHRRI by factor  $d$  of 2. (c) Down-sampled image of OHRRI by factor  $d$  of 4. (d) Down-sampled image of OHRRI by factor  $d$  of 8 (i.e., OLRRI).

Regardless of which resolution training sample set was used to train the BPNN, the higher the resolution of the test sample set (i.e., the greater  $s$ ), the higher the recognition rate. This finding was explained by the fact that the difference between different test samples increased as the resolution increased. Therefore, test samples with a higher resolution were easier to distinguish and recognize. Conversely, the results of each test sample set recognized by training sample sets with different resolutions were compared longitudinally. When the resolutions of the test and training sample sets were the same, the recognition rate was higher than when their resolutions were different.

According to the experimental results, the method of using the SR algorithm in actual scene recognition is obtained as follows. (1) When a low-res scene is acquired, if it can be super-resolved then it should be super-resolved. (2) If multiple SR factors are available to choose from, the largest one should be selected even if the resolution of the super-resolved scene is higher than that of the model in the model library. (3) If only one SR factor exists to choose from, but multiple model libraries are available with different resolutions, the model library with the same resolution as the super-resolved scene should be selected.

#### 4. Conclusions

This work proposed an SR algorithm for range images based on SGJF, which can reduce the effect of intensity image texture on super-resolved images by adding the range information of the range image as a coefficient of the filter. We also constructed an SR recognition system for range images, which acquired range and intensity images of eight military target objects through the image acquisition platform to establish training and test sample sets for the simulation model and scene libraries and used the Bicubic, JBF, MRF and SGJF algorithms to test the effect of the SR algorithm on the recognition of ladar range images. The recognition of scenes with arbitrary attitude angles by model libraries of different sizes was initially tested. The results demonstrated that the recognition rate was



positively correlated with the model library size. However, the recognition rate grew slowly or even stopped when it increased to a certain number. The recognition of the system under different noise conditions was then tested when the resolutions of the super-resolved scene and the model were the same. The results demonstrated that all the SR algorithms we tested can help improve the recognition rate of low-res scenes; moreover, a larger factor increases the improvement. The SGJF algorithm we proposed largely avoids the problem of ‘texture transfer’. It had the highest recognition rate when no noise occurred and showed noise robustness after adding noise. Therefore, the SGJF algorithm has a very good comprehensive recognition performance. Finally, the recognition of multi-res super-resolved scene libraries by multi-res model libraries was tested. The results demonstrated that regardless of the resolution of the model library, an increase in the resolution of the super-resolved scene also increased the recognition rate. In addition, using the model library with the same resolution as the super-resolved scene could obtain the highest recognition rate. According to the experimental results, suggestions on the use of SR algorithms in actual scene recognition are proposed. Our future work includes studying the application of the SR algorithm in the recognition of radar range images based on local features.

**Author Contributions:** Conceptualization: Y.Z. and J.L.; Data curation: Y.Z., S.H. and F.L.; Formal analysis: W.X., S.H. and W.L.; Funding acquisition: S.H.; Methodology: W.X.; Project administration: S.H.; Software: J.L.; Writing—original draft: Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the International Science and Technology Cooperation Project (2015DFR10830).

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

## References

- Hua, T.; Dai, K.; Zhang, X.J.; Yao, Z.C.; Zhai, Y.; Wang, H.J. Optimal VMD-Based Signal Denoising for Laser Radar via Hausdorff Distance and Wavelet Transform. *IEEE Access* **2019**, *7*, 167997–168010. [[CrossRef](#)]
- Xia, W.Z.; Han, S.K.; Cao, J.; Yu, H.Y. Target recognition of log-polar ladar range images using moment invariants. *Opt. Laser. Eng.* **2017**, *88*, 301–312. [[CrossRef](#)]
- Roldan, I.; del-Blanco, C.R.; de Quevedo, A.D.; Urzaiz, F.I.; Menoyo, J.G. DopplerNet: A convolutional neural network for recognising targets in real scenarios using a persistent range-Doppler radar. *IET Radar Sonar Nav.* **2020**, *14*, 593–600. [[CrossRef](#)]
- Teimouri, N.; Dyrmann, M.; Jorgensen, R.N. A Novel Spatio-Temporal FCN-LSTM Network for Recognizing Various Crop Types Using Multi-Temporal Radar Images. *Remote Sens.* **2019**, *11*, 990. [[CrossRef](#)]
- Du, J.; Zhou, H.X.; Qian, K.; Tan, W.; Zhang, Z.; Gu, L.; Yu, Y. RGB-IR cross input and sub-pixel upsampling network for infrared image super-resolution. *Sensors* **2020**, *20*, 281. [[CrossRef](#)]
- Hou, H.; Andrews, H. Cubic splines for image interpolation and digital filtering. *IEEE Trans. Acoust. Speech Signal Process.* **1978**, *26*, 508–517.
- Stark, H.; Oskoui, P. High-resolution image recovery from image-plane arrays, using convex projections. *J. Opt. Soc. Am. A* **1989**, *6*, 1715. [[CrossRef](#)]
- Chang, H.; Yeung, D.Y.; Xiong, Y.M. Super-resolution through neighbor embedding. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004.
- Konno, Y.; Tanaka, M.; Okutomi, M.; Yanagawa, Y.; Kinoshita, K.; Kawade, M. Depth map upsampling by self-guided residual interpolation. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 1394–1399.
- Eichhardt, I.; Chetverikov, D.; Janko, Z. Image-guided ToF depth upsampling: A survey. *Mach. Vision. Appl.* **2017**, *28*, 267–282. [[CrossRef](#)]
- Kopf, J.; Cohen, M.; Lischinski, D.; Uyttendaele, M. Joint bilateral upsampling. *ACM Trans. Graph.* **2007**, *26*, 96. [[CrossRef](#)]
- Min, D.; Lu, J.; Do, M.N. Depth video enhancement based on weighted mode filtering. *IEEE T. Image Process.* **2012**, *21*, 1176–1190.
- Diebel, J.; Thrun, S. An application of Markov random fields to range sensing. In *Advances in Neural Information Processing Systems (NIPS 2005)*; Platt, J., Ed.; MIT Press: Cambridge, MA, USA, 2005; pp. 291–298.

14. Ferstl, D.; Reinbacher, C.; Ranftl, R.; Matthias, R.; Horst, B. Image guided depth upsampling using anisotropic total generalized Variation. In Proceedings of the 2013 IEEE International Conference on Computer Vision (ICCV), Sydney, Australia, 1–8 December 2013; pp. 993–1000.
15. Yang, W.M.; Zhang, X.C.; Tian, Y.P.; Wang, W.; Xue, J.H. Deep learning for single image super-resolution: A brief review. *IEEE Trans. Multimed.* **2019**, *21*, 3106–3121. [[CrossRef](#)]
16. Zhang, J.; Shao, M.H.; Yu, L.L.; Li, Y.S. Image super-resolution reconstruction based on sparse representation and deep learning. *Signal Process Image* **2020**, *87*, 115925.
17. Vo, T.H.; Lee, G.S.; Yang, H.J.; Kim, S.H. Pyramid with Super Resolution for In-the-Wild Facial Expression Recognition. *IEEE Access.* **2020**, *8*, 131988–132001. [[CrossRef](#)]
18. ElSayed, A.; Kongar, E.; Mahmood, A.; Sobh, T. Unsupervised face recognition in the wild using high-dimensional features under super-resolution and 3D alignment effect. *Signal Image Video Process.* **2018**, *12*, 1353–1360. [[CrossRef](#)]
19. Park, J.; Kim, H.; Tai, Y.W.; Brown, M.S.; Kweno, I. High quality depth map upsampling for 3D-TOF cameras. In Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 1623–1630.
20. Vedaldi, A.; Fulkerson, B. VLFeat: An Open and Portable Library of Computer Vision Algorithms. Available online: <http://www.vlfeat.org/> (accessed on 28 August 2019).
21. Chen, J.; Tang, C.K.; Wang, J. Noise brush: Interactive high quality image-noise separation. *ACM Trans. Graph.* **2009**, *28*, 146.
22. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision (ICCV'98), Washington, DC, USA, 4–7 January 1998; pp. 839–846.
23. Xia, W.Z.; Han, S.K.; Cao, J.Y.; Wang, L.; Zhai, Y.; Cheng, Y. Target recognition of ladar range images using slice image: Comparison of four improved algorithms. *Opt. Eng.* **2017**, *56*, 073107. [[CrossRef](#)]
24. Sanford, K. Smoothing Kinect Depth Frames in Real-Time. Available online: <http://www.codeproject.com/Articles/317974/KinectDepthSmoothing> (accessed on 16 September 2018).
25. Liu, Z.J.; Li, Q.; Xia, Z.W.; Wang, Q. Target recognition of ladar range images using even-order Zernike moments. *Appl. Opt.* **2012**, *51*, 7529–7536.
26. Hu, M.K. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* **1962**, *8*, 179–187.
27. Suk, T.; Flusser, J. Affine moment invariants generated by graph method. *Pattern Recognit.* **2011**, *44*, 2047–2056. [[CrossRef](#)]
28. Liu, Z.J.; Li, Q.; Xia, Z.W.; Wang, Q. Target recognition for small samples of ladar range image using classifier ensembles. *Opt. Eng.* **2012**, *51*, 087201. [[CrossRef](#)]
29. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*, 2nd ed.; Wiley-Interscience: New York, NY, USA, 2000.
30. Wasenmüller, O.; Stricker, D. Comparison of Kinect V1 and V2 depth images in terms of accuracy and precision. *Lect. Notes Comput. Sci.* **2017**, *10117*, 34–45.

