Journal of Russian Laser Research, Volume 30, Number 6, 2009

MODEL-BASED ALGORITHM FOR ROTATION INVARIANT TARGET RECOGNITION USING LASER RADAR RANGE IMAGERY

Wang Li,* Sun Jianfeng, and Wang Qi

National Key Laboratory of Tunable Laser Technology
Harbin Institute of Technology
Harbin, Heilongjiang 150001, China

*Corresponding author e-mail: wanglijy@yahoo.cn

Abstract

The out-of-plane rotation invariance is demonstrated for a recognition system for laser radar (ladar) range imagery. The key of our method is to transform an original ladar range image of scene data into a height-range image. We demonstrate experimentally that our method provides the out-of-plane rotation invariance property and preserves in-plane rotation invariant characteristics at the same time. We also show experimentally the relationship between the recognition rate and depression angles.

Keywords: ladar, automatic target recognition, spin image, rotation invariance, height-range imagery.

1. Introduction

Coherent laser radars (ladars) are capable of collecting intensity, range, and Doppler images by raster scanning of the field of view. Since the range imagery includes more information, the target recognition becomes easier, and the accuracy of automatic target recognition (ATR) increases [1]. We carried out imaging experiments using a coherent CO_2 imaging ladar and obtained the intensity and range images that can be used for recognizing a target [2]. The reason why a CO_2 laser is chosen as the radiation source of the ladar system can be explained by the fact that a CO_2 laser with a wavelength of 10.6 μm possesses two obvious advantages, namely, it has good atmospheric transmission characteristics and it is eye-safe for all possible exposures [3, 4].

In most pattern-recognition methods based on ladar range imaging, the problem of out-of-plane rotation invariance is a puzzle, and it must somehow be solved. Such pattern-recognition problems arise in many fields. In the early years, the representation schemes included generalized cylinders, planar-face-surface curvatures, and superquadric and splash representations. Early works mainly dealt with polyhedral objects, which means segmenting of curved surfaces into planar surfaces. However, the planar patch is not the most suitable representation for free-form surfaces. Other recent surface representations include the surface-point signature, the harmonic shape image, the 3D-shape contexts, and harmonic shape contexts [5–11]. However, the most published approaches to solving this problem either assumed high-angular resolution and high range-precision range images or were based on creating a model library using real 3D objects rather than the real range image. Due to the nature of the specific imagery we are

Manuscript submitted by the authors in English first on August 3, 2009 and in final form on November 10, 2009.
1071-2836/09/3006-0615

interested in, these approaches are not suitable for our problem [12–14]. Thus, we use Johnson's spin-image surface-matching algorithm [15] and recognition method presented in [16] to study the possibility of out-of-plane rotation-invariant recognition of ladar range imagery. The cited methods were first introduced to deal with the in-plane rotation-invariance problem. The algorithm reduces the problem of 3D ATR to a 2D problem and encodes the target signature using a shape-based 2D representation. In reality, a spin image is a 2D parametric space histogram that includes the majority of the shape information in the 3D-scene target. Moreover, the processing of spin images can benefit from excellent achievements in the field of relatively mature 2D-image processing. Even so, the original spin image algorithm is not out-of-plane rotation invariant.

In this paper, we improve the algorithm and make it suitable for our problem. When an airborne ladar detects a ground-based object, the whole or partial top half of the object is always located in the field of view, and the area of the top-half imagery is changeable with the depression angle. So the top-half range image can be extracted using geometrical relations from the original range imagery, which is obtained directly by the ladar. The spin images corresponding to the top-half image are created and then compared with the spin images of the model library. Our recognition method has out-of-plane rotation invariance and preserves in-plane rotation invariance at the same time.

This paper consists of two parts.

In the first part, we describe the spin-image surface matching algorithm. In the second part, we transform the original ladar range imagery to the height-range imagery and successfully demonstrate our recognition system using eight scene data, and finally we analyze the relationship between the depression angle and recognition rate in the tests.

2. Overview of Spin-Image Surface Matching Algorithm

In this section, first we describe the surface-shape representation by a collection of oriented 3D points that have associated images capturing the global properties of the surface in the object-centered local coordinate system. Second, in view of matching images, we determine the correspondence between the surface points, which provide the surface matching.

2.1. Spin-Image Fundamentals

A fundamental component of the surface matching representation is the point of orientation (oriented point) used to create a spin image, i.e., a 3D point associated with the normal to its surface [17, 18]. An oriented point defines a five-degrees-of-freedom basis $(\mathbf{p}, \mathbf{n})$ using the tangent plane P through $\mathbf{p}$ oriented perpendicularly to the unit normal $\mathbf{n}$. Figure 1 demonstrates the concept of surface matching. A spin image is created first by constructing a local coordinate system at each oriented point. Using this local coordinate system, the position of all other points on the surface can be encoded by two parameters. The first spin-image parameter is α, the distance perpendicular to $\mathbf{n}$, and the second parameter is β, a signed distance perpendicular to the plane P. The parameter α is positive, and the parameter β can be positive or negative. Given an oriented-point basis O, we can define a spin-map function S_0, which projects a 3D point $\mathbf{x}$ to the 2D coordinate of a particular basis $(\mathbf{p}, \mathbf{n})$ as follows [15]:

$$S_0(\mathbf{x}) \to (\alpha, \beta) = \left(\sqrt{\|\mathbf{x} - \mathbf{p}\|^2 - (\mathbf{n}(\mathbf{x} - \mathbf{p}))^2}, \mathbf{n}(\mathbf{x} - \mathbf{p}) \right). \quad (1)$$

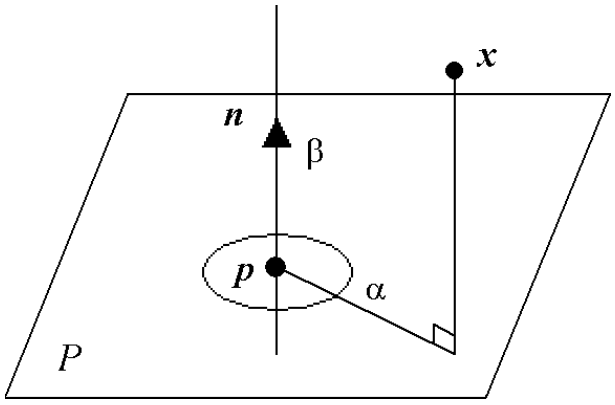

Fig. 1. Illustration of generation of the spin-image parameters.

The application of the function $S_0(x)$ to all oriented points in a 3D-point cloud provides a set of 2D points in the $\alpha - \beta$ space. The first step in building a spin image is to determine which points can contribute to the spin image. This is accomplished by gridding the 3D data in the spin-parameter space and applying the spin-image generation criteria to the list of gridded points. Based on the (α, β) spin coordinates associated with all points, the data is gridded according to

$$i = \left\lfloor \frac{\alpha}{b_s} \right\rfloor, \qquad j = \left\lfloor \frac{\beta_{\max} + \beta}{b_s} \right\rfloor, \tag{2}$$

where $\lfloor f \rfloor$ is the floor operator which rounds f down to the nearest integer, and b_s is the bin size. The contribution of a point is bilinearly interpolated to the four surrounding bins in a 2D array. By spreading the contribution of the point in a 2D array, the bilinear interpolation provides further reduction of the effect of variations in the 3D point position on the 2D array. The bilinear-interpolation weights are calculated according to

$$a = \alpha - ib_s, \qquad b = \beta - \beta_{\max} - jb_s. \tag{3}$$

A point's contribution to the spin image S_i resulting from the bilinear interpolation is calculated as [10,15]

$S_i(i,j) = S_i(i,j) + (1-a)(1-b)$	$S_i(i+1,j) = S_i(i,j) + a(1-b)$
$S_i(i,j+1) = S_i(i,j+1) + (1-a)b$	$S_i(i+1,j+1) = S_i(i+1,j+1) + ab$

(4)

Finally, all 3D points are rearranged, according to Eq. (4), to the form S_i. The procedure of creating a 2D-array representation of the spin image is shown in Fig. 2.

2.2. 3D-Surface Matching

The surface-matching algorithm [15, 16] uses a scene data set along with a spin-image model library. The spin-image model library contains the ideal 3D ladar signatures of each target derived from CAD models [19]. Each 3D ladar model data set also has its associated spin-image database, with a corresponding spin image for each model 3D point. The scene spin image is compared with each model's spin-image database within the model library. For template matching, the best way of comparing is to relate the images linearly with the help of a normalized linear correlation coefficient. Given two spin images M and N with K bins each, the linear correlation coefficient $R(M,N)$ is defined as [15]

$$R(M,n) = \frac{K\sum m_i n_i - \sum m_i \sum n_i}{\sqrt{\left[K\sum m_i^2 - \left(\sum m_i\right)^2\right]\left[K\sum n_i^2 - \left(\sum n_i\right)^2\right]}}, \tag{5}$$

where R changes between -1 (anticorrelated) and 1 (completely correlated). Each scene spin image is compared with all model spin images, which provides a distribution of similar measured values. When

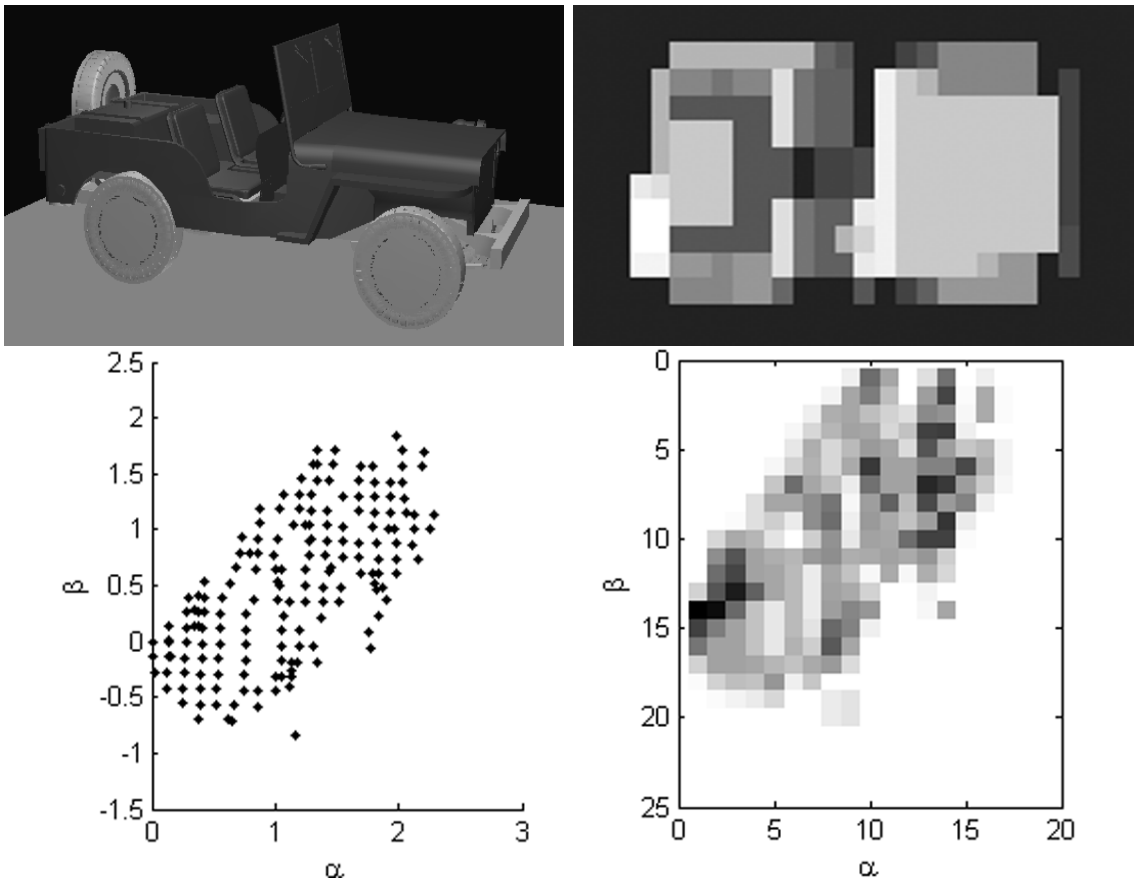

Fig. 2. Example of the spin-image creation of a jeep.

comparing two spin images M and N, where S is the number of overlapping pixels used in the computation of R, the similarity measure C is defined as

$$C(M,N) = [\operatorname{arctanh}(R(M,N))]^2 - \lambda \frac{1}{S-3}\,. \tag{6}$$

The correspondences obtained through comparison of each scene spin image with the model spin images are filtered using the similarity-measure threshold. This process is repeated for the rest of the scene spin images, resulting in a wide distribution of the similarity measures. Given a new distribution of the similarity measures, the second-similarity threshold is applied to remove unlike correspondences. The remaining correspondences are further filtered and then grouped based on geometric consistence in order to compute plausible transforms that align the scene to the model data sets. Finally, the initial scene-to-model alignments are verified and refined using a modified version of the iterative closest point algorithm (ICP) [19]. The goal of using the ICP is to find the best match between the model and the scene by eliminating incorrect matches. Figure 3 is a detailed block diagram of the surface matching procedure.

The target recognition can be measured using the recognition goodness of fit value V_{GOF} between the

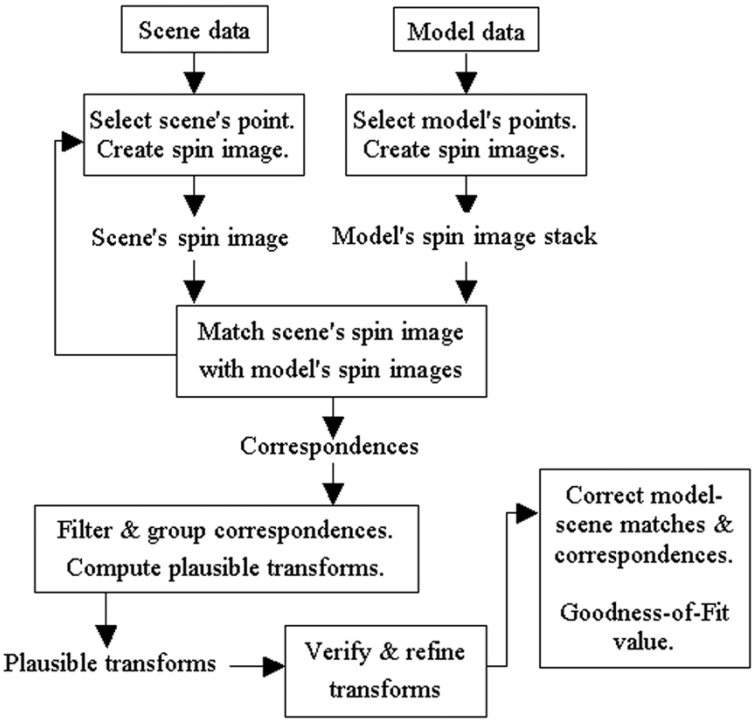

Fig. 3. Block diagram of the surface matching.

scene and the model, which is defined as [16]

$$V_{\mathrm{GOF}}(s,m) = \frac{\varepsilon^2 N_{\mathrm{pt}}}{\mathrm{MSE}}\,, \tag{7}$$

where s and m represent the scene and the model, respectively, ε is the fraction of overlapping between s and m determined by a process of verification, N_{pt} is the number of plausible pose transforms, and MSE is the mean-square error determined by the ICP. In order to express clearly the recognition performance of the surface matching, V_{GOF} is normalized to all the V_{GOF} values in the model library. If the scene s correctly matches the model i in the model library including T models, the normalized V_{GOF} is defined as [16]

$$\bar{V}_{\mathrm{GOF}}(s,m_i) = \frac{V_{\mathrm{GOF}}(s,m_i)}{\sum\limits_{j=1}^{T} V_{\mathrm{GOF}}(s,m_j)}, \tag{8}$$

where $\bar{V}_{\mathrm{GOF}}$ ranges from 0 to 1, and the sum of $\bar{V}_{\mathrm{GOF}}$ for all the models is equal to unity. When the scene matches none of the models, the sum of the $\bar{V}_{\mathrm{GOF}}$ is equal to zero.

3. Out-of-Plane Rotation-Invariant Object Recognition

In this section, the original ladar range imagery is transformed to the height-range imagery only, including the top-half range information on the object. Then our recognition system is successfully

demonstrated for eight scenes with targets in arbitrary depression and azimuth angles. In the end, we analyze the relation between the depression angle and the recognition rate in all tests.

3.1. Generation of Height-Range Imagery

Given a range imagery containing a target with arbitrary depression angle obtained directly by the ladar, the original range imagery is transformed into a height-range imagery carrying information on only the top-half of the target, which is described visually in Fig. 4. Height-range values are calculated according to the relation

$$H = L \sin A. \tag{9}$$

In the height-range imaging, we assume that the ground level height is zero and set an appropriate threshold height according to the target type. Employing this threshold, some points that cannot contribute to the recognition performance are filtered, and the others located on the top half are kept. Two noiseless range images for a car and a bus are shown in Figs. 5 and 6 (left) for the depression angles of 65°and 75°, respectively, and the arbitrary azimuths. Figures 5 and 6 (right) show the transformed height-range images corresponding to the above-mentioned range images.

3.2. Creation of the Model Library

We use eight specific simulated targets to build a model library consisting of ground vehicles ranging from cars to trucks and scrapers. Figure 7 displays their 3D models. Since the depression and azimuth angles are arbitrary, the range image of a model target should contain all top-half information. Figure 8 shows a set of simulated noiseless range images for every model. These images were generated for the case of orthographic ladar view, and contain 32×64 pixels each.

Based on the these models, a target model library was constructed to simulate an ideal 3D ladar spatially-resolved signature of each target. The simulated ideal target models are then converted to the spin images. A measured spin image should then be compared to every spin image of the model library to recognize and identify the scene target.

3.3. Experimental Confirmation

Due to limitations caused by the experimental conditions, it is very difficult to obtain the real range images of a target at arbitrary depression angles. Thus, we simulated noise scene data required in the test, using real ladar range images, which were acquired by a pulse CO_2 laser heterodyne detecting ladar. The laser transmitter of the ladar system is a pulse CO_2 laser with a wavelength of 10.6 μm, a peak power of 1.5 kW, and a repetition rate of 20 kHz. The space resolution of the ladar system is 32×64 pixels. Figure 9 shows a block diagram of the pulse CO_2 laser heterodyne detecting system.

In order to determine the recognition performance, multiple noise scenes were analyzed. The measured scene data, each containing a target instance, were obtained at arbitrary depression and azimuth angles. All scene data with 32×64 pixels are low-space-resolution images. If we can recognize the target with a high recognition rate from such a low-space-resolution image, our method would appear to be more applicable to high-space-resolution range images. Eight scene data were transformed via the method described in Sec. 3.1 to height-range images shown in Fig. 10. The spin-image algorithm takes the scene data set and creates a spin-image database based on subsampling of the points. The sampling ranges from 20 to 50% of all scene data points. The scene data points are not judiciously picked – the points

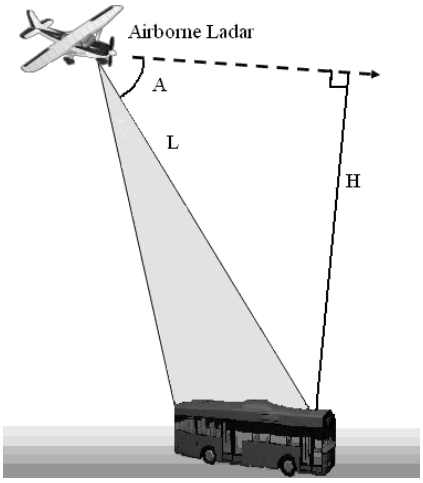

Fig. 4. Transform from the range value to the height-range value.

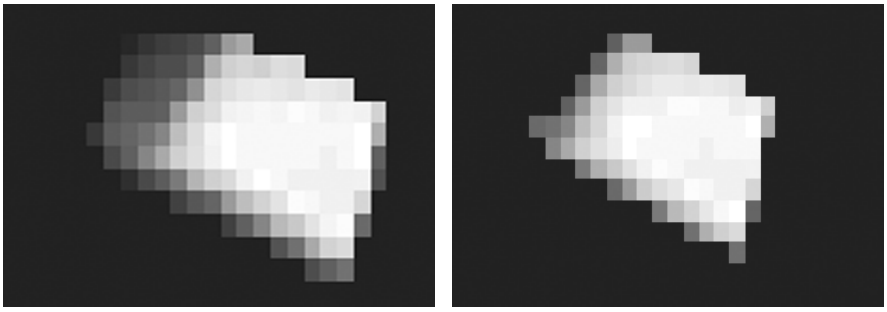

Fig. 5. Noiseless range image for a car (left) and the corresponding height-range image (right).

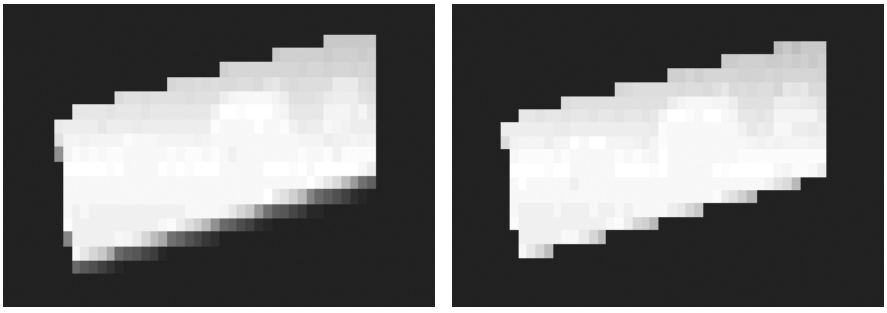

Fig. 6. Noiseless range image for a bus (left) and the corresponding height-range image (right).

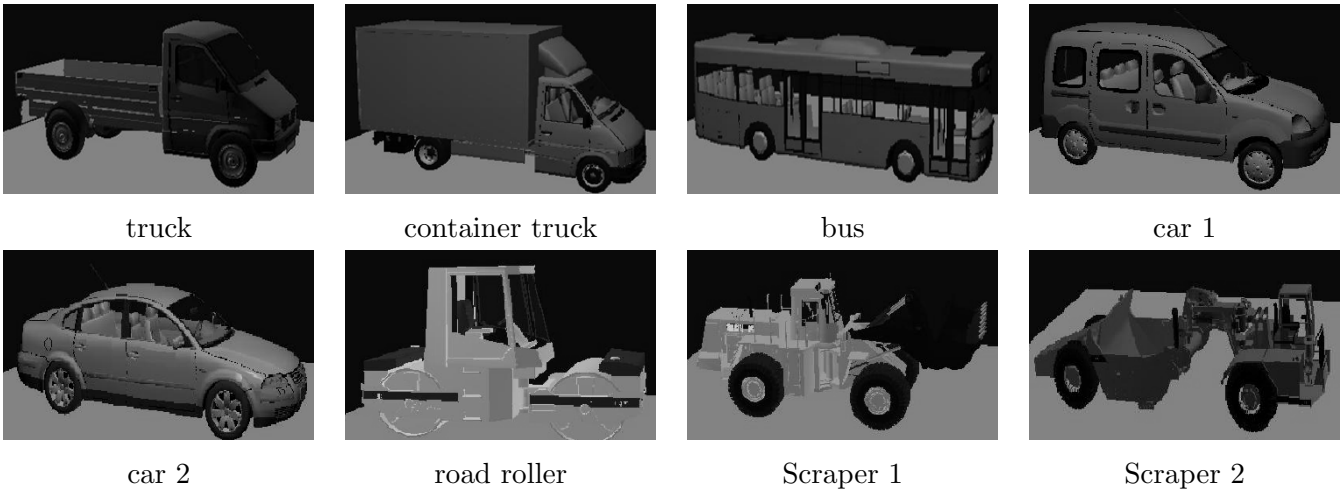

Fig. 7. 3D model library.

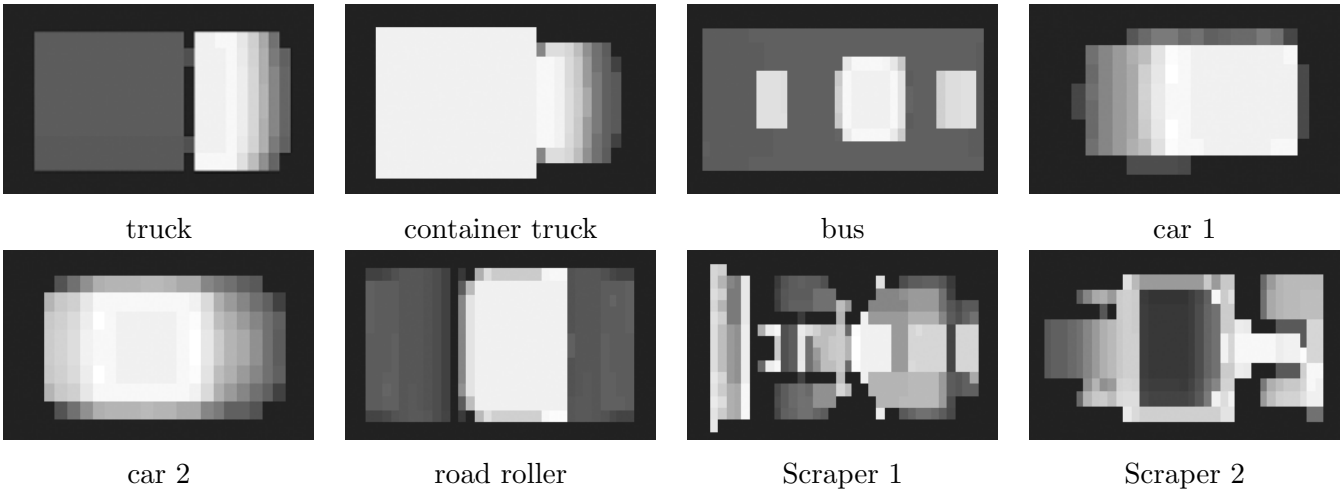

Fig. 8. Simulated noiseless range images.

are uniformly distributed across the given scene. Therefore, no feature extraction is performed to pick the spin-image points. The spin images of a scene target are then compared to the ones of each model target. Finally, series of $\bar{V}_{\mathrm{GOF}}$ values representing the recognition performance are obtained [16].

A confusion matrix is created using series of $\bar{V}_{\mathrm{GOF}}$ values, and each element of the matrix is obtained by comparing every scene target with all model targets. Table 1 denotes the recognition confusion matrix. The elements on the principal diagonal denote the confidence measurement $\bar{V}_{\mathrm{GOF}}$, and those on the off diagonals represent the errors. When the model target matches the scene target correctly, the highest $\bar{V}_{\mathrm{GOF}}$ value is obtained for this model target. However, for most of the remaining model targets, the $\bar{V}_{\mathrm{GOF}}$ values are equal to zero, because the recognition algorithm finds no match between the particular scene and the corresponding target. This means that the model target matches the scene target perfectly, which is an ideal result.

The $\bar{V}_{\mathrm{GOF}}$ values fall almost entirely on the correct targets at $\bar{V}_{\mathrm{GOF}}$ levels exceeding 90% in five of the eight data sets. For the remaining three scenes, the correct targets are still assigned the highest values $\bar{V}_{\mathrm{GOF}}$, but there is a portion of the $\bar{V}_{\mathrm{GOF}}$ values that are not the true targets. While the $\bar{V}_{\mathrm{GOF}}$ values do not completely fall on the correct target, the distribution of $\bar{V}_{\mathrm{GOF}}$ values falls almost completely on the correct target class. For the car 1 scene, the recognition algorithm is able to identify correctly the scene as a car with a $\bar{V}_{\mathrm{GOF}}$ of 100% and classify the target as the car 1 with a $\bar{V}_{\mathrm{GOF}}$ of 86%. For the scraper 1 scene, the algorithm is able to identify correctly the target as a scraper 1 with a $\bar{V}_{\mathrm{GOF}}$ of 85%, with the remaining $\bar{V}_{\mathrm{GOF}}$ value falling on another scraper. Thus, the recognition algorithm correctly classifies the scenes with a $\bar{V}_{\mathrm{GOF}}$ of 100% in this case.

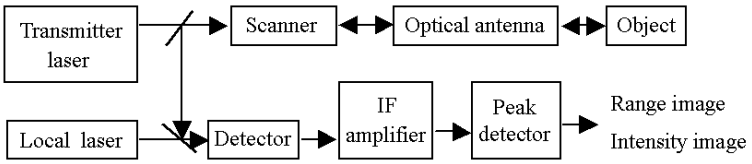

Fig. 9. Pulse CO_2 laser heterodyne detecting system.

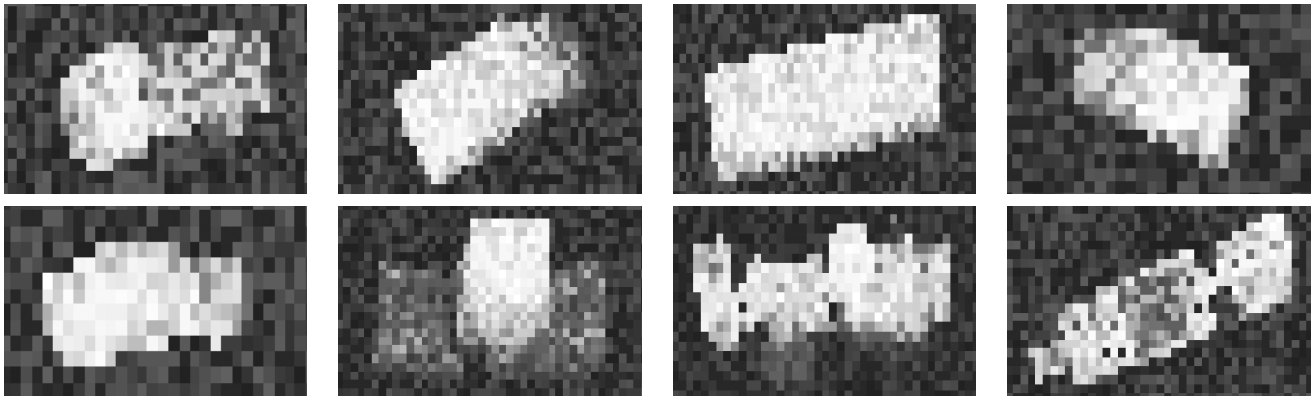

Fig. 10. Measured noise scene targets with arbitrary depression and azimuth angles.

Table 1. Recognition Confusion Matrix.

Field target data	Depression angle	Model							
		Truck	Container truck	Bus	Car 1	Car 2	Road roller	Scraper 1	Scraper 2
Truck	70°	**0.93**	0.07	0.00	0.00	0.00	0.00	0.00	0.00
Container truck	70°	0.05	**0.91**	0.04	0.00	0.00	0.00	0.00	0.00
Bus	75°	0.00	0.01	**0.99**	0.00	0.00	0.00	0.00	0.00
Car 1	65°	0.00	0.00	0.00	**0.86**	0.14	0.00	0.00	0.00
Car 2	25°	0.05	0.03	0.00	0.20	**0.72**	0.00	0.00	0.00
Road roller	60°	0.02	0.00	0.00	0.02	0.01	**0.95**	0.00	0.00
Scraper 1	45°	0.00	0.00	0.00	0.01	0.01	0.03	**0.85**	0.10
Scraper 2	80°	0.00	0.00	0.00	0.00	0.00	0.00	0.03	**0.97**

However, for the car 2 scene, only a $\bar{V}_{\mathrm{GOF}}$ of 72% falls on the true target, with the remaining $\bar{V}_{\mathrm{GOF}}$ falling on the other kind of targets. The reason for recognition performance's worsening, is that the depression angle for the car is just 25°, i.e., it is relatively small. The smaller the depression angle, the less the top-half information in the height-range image of a target. Thus, the recognition rate is

lower also. Figure 11 shows the $\bar{V}_{\mathrm{GOF}}$ as a function of the depression angle. First, $\bar{V}_{\mathrm{GOF}}$ increases with augmentation of the depression angle and then is kept on a certain level, which indicates no further improvement.

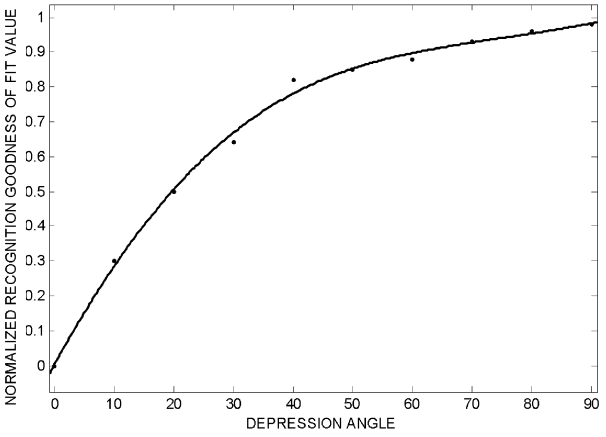

Fig. 11. The $\bar{V}_{\mathrm{GOF}}$ value as a function of the training-set size.

From Fig. 11, we can see that at depression angles of 40° and 50°, the $\bar{V}_{\mathrm{GOF}}$ values are equal to 80% and 86%, respectively. That is to say, when the depression angle is greater than 40°, the recognition rate is high and acceptable for applications.

4. Conclusions

In this paper, we develop a recognition system for ladar imagery target recognition. The output is invariant to out-of-plane rotations. The success of the method is based on the extraction of a height-range image from the original range image. Such a modified recognition approach is attractive and preferable to others because it keeps all features that the original spin-image surface-matching algorithm possesses and, in addition, provides the user with the possibility of resolving both in-plane and out-of-plane rotation-invariant recognitions. Automatic target recognition was successfully shown for eight measured noise scenes with targets, which were measured at arbitrary depression and azimuth angles. Finally, it was experimentally shown that, when the depression angle is greater than 40°, a high recognition rate is achieved, which is acceptable for applications.

References

1. Nikhil R. Pal, Tobias C. Cahoon, Jim C. Bezdek, and Kuhu Pal, *IEEE Trans. Fuzzy Syst.*, **9**, 44 (2001).
2. Jianfeng Sun and Qi Wang, *Laser Phys.*, **19**, 502 (2009).
3. S. Pershin, *SPIE*, **2222**, 392 (1994).
4. D. Sliney, *Opt. Photon. News*, **8**, 31 (1997).
5. S. M. Pershin, A. F. Bunkin, V. A. Lukyanchenko, and R. R. Nigmatullin, *Laser Phys. Lett.*, **4**, 808 (2007).
6. B. Bhanu, *IEEE Trans. Pattern Anal. Machine Intell.*, **6**, 340 (1984).
7. F. Solins and R. Bajcsy, *IEEE Trans. Pattern Anal. Machine Intell.*, **12**, 131 (1990).
8. F. Stein and G. Medioni, *IEEE Trans. Pattern Anal. Machine Intell.*, **14**, 125 (1992).
9. S. M. Yamany and A. Farag, in: *Proceedings of the Seventh IEEE International Conference on Computer Vision*, IEEE Computer Society, Washington, DC (1999), Vol. 2, p. 1098.
10. D. Zhang and M. Hebert, in: *Proceedings of the Seventh IEEE International Conference on Computer Vision*, IEEE Computer Society, Washington, DC (1999), Vol. 2, p. 524.
11. A. Frome, D. Huber, R. Kolluri, and Bulow, "Recognizing objects in range data using regional point

descriptors," in: *Proceedings of the European Conference on Computer Vision*, Springer Verlag, Berlin/Heidelberg (2004), Vol. 3, p. 224.
12. Jacques G. Verly, Richard L. Delanoy, and Dan E. Dudgeon, *Opt. Eng.*, **31**, 2540 (1992).
13. Lena Klasen, Pierre Andersson, Hakan Larsson, et al., *Proc. SPIE*, **5412**, 321 (2004).
14. Cem Ünsalan, *Pattern Recogn. Lett.*, **28**, 49 (2007).
15. A. Johnson, "A representation for 3D surface matching," Ph.D. Thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh (1997).
16. Alexandru N. Vasile and Richard M. Marino, *Lincoln Lab. J.*, **15**, 61 (2005).
17. Niloy J. Mitra and An Nguyen, in *Proceedings of 19th ACM Symposium on Computational Geometry, June 8-10, 2003, San Diego, CA, USA*, ACM, New York (2003), p. 322.
18. Zheng You Zhang, *Int. J. Computer Vision*, **13**, 119 (1994).
19. Paul J. Best and Neil D. Mckay, *IEEE Trans. Pattern. Anal. Mach. Intell.*, **14**, 239 (1992).
20. G. Bouchette, P. Iles, C. English, et al., *Proc. SPIE*, **6566**, 656619 (2007).