Spin-image surface matching based target recognition in laser radar range imagery*

Wang Li(王 丽)†, Sun Jian-Feng(孙剑峰), and Wang Qi(王 骐)

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

(Received 1 February 2010; revised manuscript received 16 March 2010)

We explore the problem of in-plane rotation-invariance existing in the vertical detection of laser radar (Ladar) using the algorithm of spin-image surface matching. The method used to recognize the target in the range imagery of Ladar is time-consuming, owing to its complicated procedure, which violates the requirement of real-time target recognition in practical applications. To simplify the troublesome procedures, we improve the spin-image algorithm by introducing a statistical correlated coefficient into target recognition in range imagery of Ladar. The system performance is demonstrated on sixteen simulated noise range images with targets rotated through an arbitrary angle in plane. A high efficiency and an acceptable recognition rate obtained herein testify the validity of the improved algorithm for practical applications. The proposed algorithm not only solves the problem of in-plane rotation-invariance rationally, but also meets the real-time requirement. This paper ends with a comparison of the proposed method and the previous one.

Keywords: Ladar, automatic target recognition, spin-image, statistical correlation coefficient

PACC: 4230S, 4230V, 4262

1. Introduction

Laser radar (Ladar) can acquire multiform real-time three-dimensional (3D) images such as intensity, range and Doppler images through multiple accurate measurements of echo intensity, range and velocity, and so on. On account of the range imagery including more information, target recognition becomes easier and the accuracy of automatic target recognition (ATR) is increased. The problem of ATR in Ladar range imagery has been an active topic of research for many years.[1−4]

Template matching is one of the most popular methods for ATR. We simply describe it here for easy understanding: silhouette features are extracted from the scene target first, and then compared with the known model library to identify and classify the specific target. Generally speaking, the primary problem of template matching is to represent a model target using a unique signature encoded mainly by a shape-based two-dimensional (2D) representation. Recent algorithms in which the shape-based representation is used are contour-based algorithm proposed by Shantaram and Hanmandlu,[5] Han *et al.*,[6] surface signatures by Yamany and Farag,[7,8] Zhou,[8] and spin-image algorithm by Andrew Johnson.[9] Researchers at Massachusetts Institute of Technology Lincoln Laboratory (MIT) introduced Johnson's spin-image surface matching to perform both target detection and recognition. The algorithm simplifies the problem of three-dimensional (3D) ATR into a 2D problem. Therefore, the processing of the algorithm can benefit from excellent achievements in the field of relatively mature 2D image processing. Using the spin-image algorithm, we have solved the problem of out-of-plane rotation-invariant recognition by extracting top half range image of target.[10−12] However, the procedures of recognition and evaluation of spin-image matching are very complicated and time-consuming, which is contrary to the real-time recognition requirement in practice. Therefore, a real-time algorithm is needed that can recognize targets from Ladar range data. This paper focuses on expediting the procedure of the spin-image algorithm.

In this paper, a novel evaluation procedure of recognition performance, named statistical correlation coefficient method (SCCM) is proposed. First, spin-images of models and scene targets are created, sepa-

*Project supported by the National Natural Science Foundation of China (Grant No. 60901046).

†Corresponding author. E-mail: wanglijy@yahoo.cn

© 2010 Chinese Physical Society and IOP Publishing Ltd

rately. Then scene spin-images are correlated to model library spin-image stack, and as a result, a series of correlation coefficients are obtained. Finally, the correlation coefficients greater than or equal to a threshold are counted to show recognition results. The final recognition performance using SCCM is consistent with the result obtained by the original method, which indicates the effectiveness of the new method. At the same time, the new method is proved to be high efficient and suitable for real-time target recognition when we explore the problem of in-plan rotation invariance for the vertical detection of Ladar.

2. Proposed algorithm

In this section, first we describe surface shape by a collection of oriented 3D points which are related to image possessing of the global properties of the surface in object-centred local coordinate systems. Secondly, by matching images, we determine correspondences between surface points, which is the so-called surface matching. Finally, for scenes targets, we propose SCCM to simplify the evaluation procedure and shorten computational time.

2.1. Overview of spin-image surface matching

A fundamental component of surface matching representation is an oriented point used to create a spin-image, a 3D point with an associated direction which is in the direction of the surface normal.[13] The first spin-image parameter is α, the perpendicular distance relative to $\boldsymbol{n}$, and the second parameter is β, the signed perpendicular distance to the plane P (see Fig. 1). Given an oriented point basis O, we can define a spin-map function S_0, which projects 3D point $\boldsymbol{x}$ to the 2D coordinate of a particular basis (p, n) as follows:[9]

$$S_{\mathrm{O}}(x) \to (\alpha, \beta) = \Big(\sqrt{\|\boldsymbol{x}-\boldsymbol{p}\|^2 - (\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{p}))^2}, \boldsymbol{n} \cdot (\boldsymbol{x}-\boldsymbol{p})\Big). \quad (1)$$

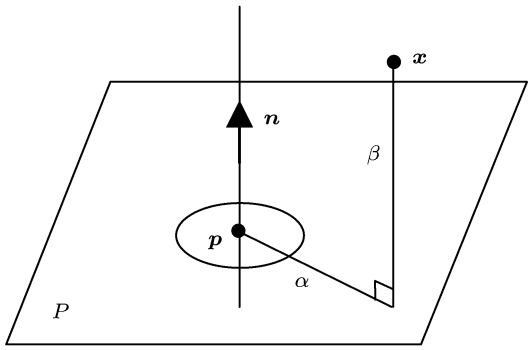


Fig. 1. Illustration of spin-image parameters.

By projecting each 3D oriented point to α–β parameter space, a 2D point set is formed at each oriented point. Herein, we bilinearly interpolate a 2D point set to create a spin-image. Using the spin-image generation criteria, the 3D data are gridded in spin parameter space by the following equations

$$i = \left\lfloor \frac{\alpha}{b_{\mathrm{s}}} \right\rfloor \quad \text{and} \quad j = \left\lfloor \frac{\beta_{\max} + \beta}{b_{\mathrm{s}}} \right\rfloor, \quad (2)$$

where $\lfloor f \rfloor$ is the floor operator which rounds f down to the nearest integer, and b_{s} is the bin size. Then each point in a 2D point set contributes to the four surrounding bins by bilinear interpolation which helps to reduce the error brought from the girding process. Bilinear interpolation weights are calculated according to[9]

$$a = \alpha - ib_{\mathrm{s}} \quad \text{and} \quad b = \beta - \beta_{\max} - jb_{\mathrm{s}}. \quad (3)$$

The contribution of a point to a spin-image is calculated by the bilinear interpolation and then all 3D points are binned to form a spin-image. The procedure to create a 2D array representation of a spin-image is described visually in Fig. 2.

For template matching, the best way to compare linearly related images is to normalize linear correlation coefficient. For two spin-images M and N with K bins each, the linear correlation coefficient $R(M, N)$ is defined as[9,10]

$$R(M,N) = \frac{K\sum m_i n_i - \sum m_i \sum n_i}{\sqrt{(K\sum m_i^2 - (\sum m_i)^2)(K\sum n_i^2 - (\sum n_i)^2)}}, \quad (4)$$

where R is between -1 (anti-correlated) and 1 (completely correlated). Each scene spin-image is correlated to all the model spin-images, resulting in a distribution of similarity measure values. When comparing two

spin-images M and N, the similarity measure C is defined as

$$C(M,N) = (\text{atanh}(R(M,N)))^2 - \lambda\left(\frac{1}{H-3}\right), \tag{5}$$

where H is the number of overlapping pixels used in the computation of R. The correspondences to model spin-image comparison, obtained for each scene spin-image, are filtered using similarity measure threshold. The remaining correspondences are further filtered and then grouped by geometric consistency in order to compute plausible transformations that align the scene to the model data sets. The initial scene-to-model alignment is verified and refined using a modified version of the Iterative Closest Point (ICP) algorithm.[10,14−18] The purpose of ICP algorithm is to find the best matches between model and scene.

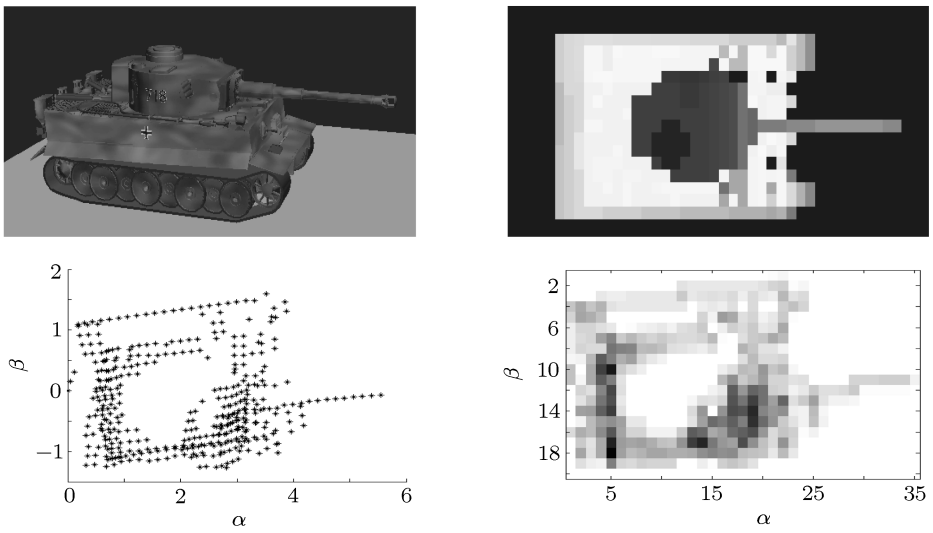


Fig. 2. Example spin-image from a tank.

Target recognition can be measured using Recognition Goodness of Fit value (V_{GOF}) between the scene and the model. V_{GOF} is defined as[12]

$$V_{\text{GOF}}(s,m) = \frac{\varepsilon^2 \cdot N_{\text{pt}}}{MSE}, \tag{6}$$

where s and m represent a scene and a model respectively, ε is the fraction of overlap between s and m, N_{pt} is the number of plausible pose transformations, and MSE is the mean-squared error determined by ICP algorithm. 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 model i in a model library including T models, then normalized V_{GOF} is defined as

$$\overline{V}_{\text{GOF}}(s,m_i) = \frac{V_{\text{GOF}}(s,m_i)}{\sum_{j=1}^{T} V_{\text{GOF}}(s,m_j)}, \tag{7}$$

where $\overline{V}_{\text{GOF}}$ ranges from 0 to 1, and the sum of $\bar{V}_{\text{GOF}}$ for all the models is 1. When the scene matches none of the models, the sum of the value of $\bar{V}_{\text{GOF}}$ equals 0.

2.2. Proposed 3D surface matching

The spin-image surface matching system is very useful for target recognition in Ladar range imagery. Figure 3 demonstrates the surface matching procedure in detail. However, it is very complicated and time-consuming in previous method in which correspondences are eliminated and verified using similarity measure

value and modified ICP. After the correlation coefficient yielded by Eq. (4), the sophisticated filtering and verifying procedures used in previous method can be simplified by introducing SCCM. Namely, the processes in the imaginary line frame in Fig. 3 are replaced by SCCM. The procedure of the SCCM is illustrated in Fig. 3.

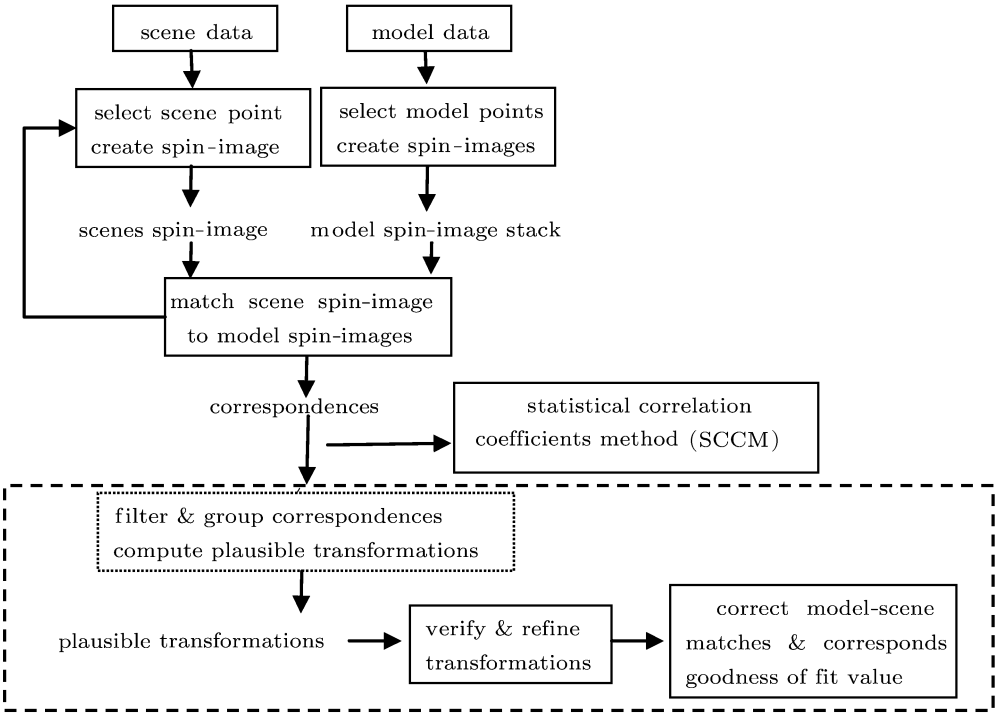


Fig. 3. Surface matching block diagram.

The procedures based on statistical correlation coefficient are shown in detail as follows: a) after spin-images of scene target and model library are built, each spin-image of scene target is compared with all spin-images of a model in the library, and a series of correlation coefficient values is obtained from Eq. (4) (Note that the number of correlation coefficient values is the same as the number of spin-images of the model.); b) if one or several correlation coefficients are over threshold H_s (Here, H_s is the minimum value representing correct matching between two spin images), the spin-image of the scene is correctly matched with that of the model as illustrated in Fig. 4(a); conversely, if none of them is over threshold H_s, the spin images of the scene and the model are not matched with each other (see Fig. 4(b)); c) the other spin-images of the scene target are compared with spin-images of the model, and the number of spin-images of the scene target is counted when the condition of being over the threshold H_s is satisfied; d) repeat the above processes and match spin-images of the scene target with spin-images of other models in the library. If the scene s correctly matches model i in a model library including T model, the normalized statistical correlation coefficient $\bar{R}$ is defined as

$$\bar{R}(s, m_i) = \frac{P_R(s, m_i)}{\sum_{j=1}^{T} P_R(s, m_j)}, \tag{8}$$

where $P_R(s, m_i)$ is the number of the correlation coefficient R which is over a threshold, $\bar{R}$ ranges from 0 to 1, and the sum of $\bar{R}$ for all the models is 1. When the scene matches none of the models, the sum of the $\bar{R}$ values equals 0. The higher the value of $\bar{R}$ for a scene-to-model contrast is thus the better the model matches the scene. Thus, the $\bar{R}$ value that falls on each model represents a confidence measure of how the model matches the scene.

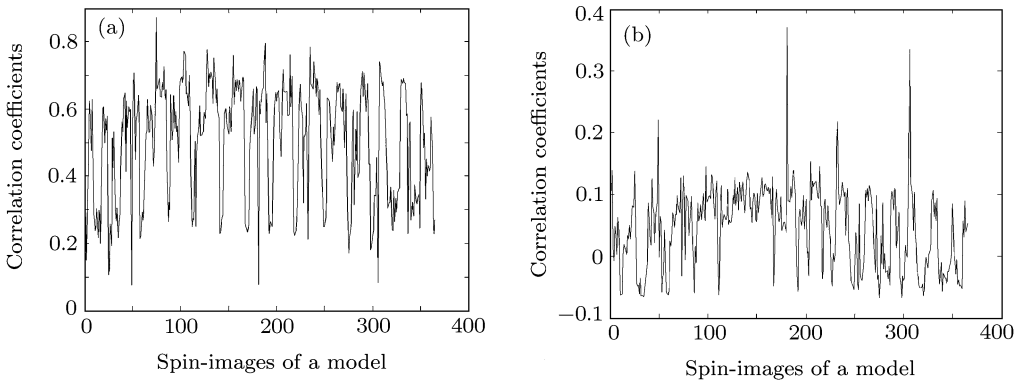


Fig. 4. Correlation of two spin-images from different scenes to spin-images of the same model: (a) Correct matching, (b) incorrect matching.

3. Experimental results

Due to the limitation of experiment conditions, it is very difficult to obtain real range images of targets in a random direction in plan. Thus we simulate scene noise data required in the test, based on real Ladar range image which is acquired by coherent imaging Ladar.[10] We use eight specific simulated targets to build a model library, consisting of military ground vehicles ranging from tank, to camion and armored car.[19,20] Figure 5 displays their 3D models. Figure 6 shows simulated non-noise range images which are 32×64 pixels and generated in the case of vertical detection.

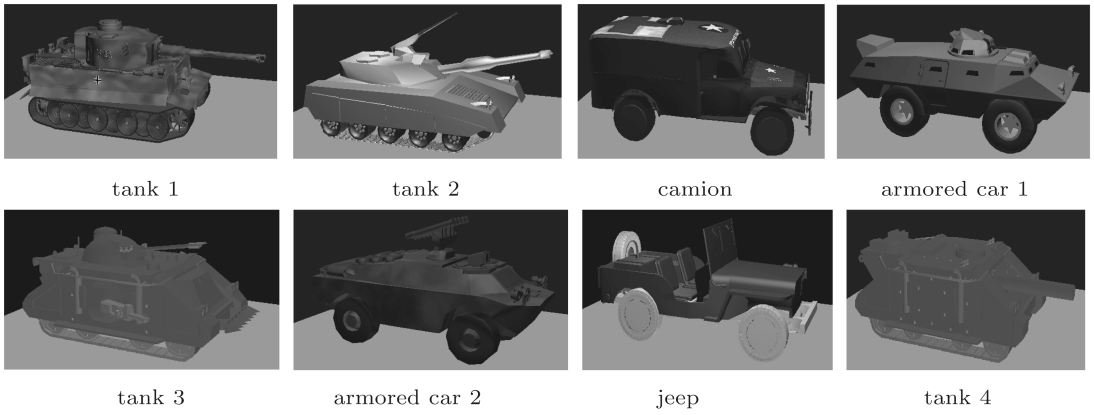


Fig. 5. 3D model library.

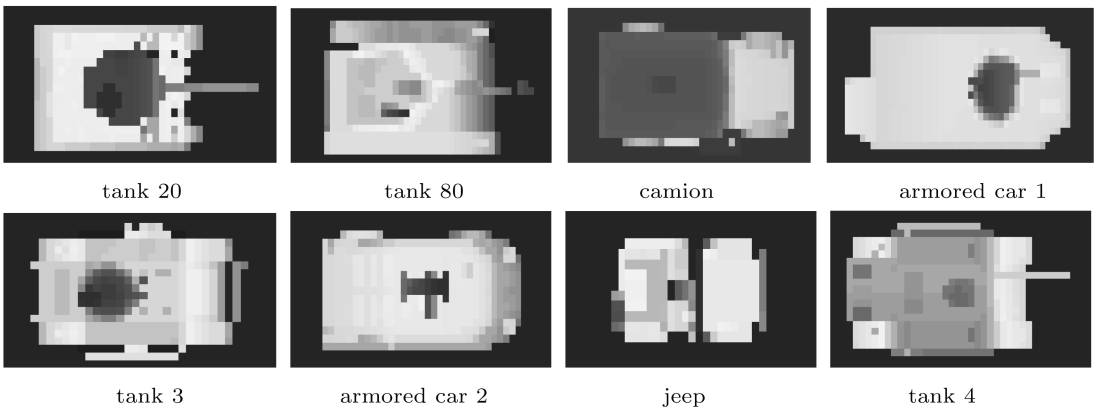


Fig. 6. Simulated non-noise range images.

Based on the above models, a target model library is constructed to simulate an ideal 3D Ladar spatially-resolved signature of each target. The simulated ideal target models are then represented as spin-images through the spin image representation method. The resulting model spin-image library is used to compare the models with a measured scene to recognize and identify the scene target.

In order to examine the recognition performance, multiple scenes are analysed. Figure 7 shows an orthographic projection for sixteen measured data sets. There exist the spatial noises in all of the measured data sets. The scene noise targets are in arbitrary orientation in plane.

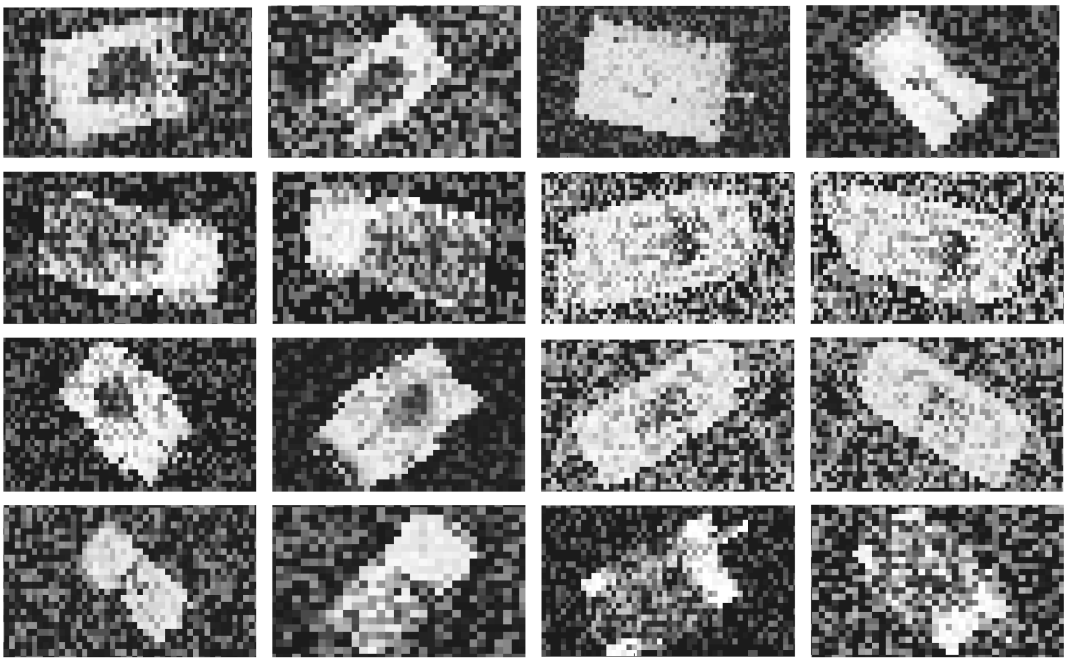


Fig. 7. Noise range images with arbitrary angles.

The 20%–50% disperse points on the measured scene target are selected to produce homologous spin images. Spin images of a scene target are compared with ones of each model target, and during the comparison, the number of correlation coefficient R exceeding threshold $H_s = 0.8$ is counted. After every scene target is compared with every model target, a series of $\bar{R}$ values representing recognition performance is obtained.

A confusion matrix is created using a series of $\bar{R}$ values, and each of the elements in the matrix is obtained from the comparison of every scene target with every model target. Table 1 denotes the recognition confusion matrix. Every column of the confusion matrix represents a comparison of a scene target with the model library. The elements on the principal diagonal denote the confidence measurement $\bar{R}$, and those off the principal diagonal represent the errors. When the model target matches the scene target perfectly, the highest $\bar{R}$ value will be obtained on the correct model target. However, for most of the remaining model targets, $\bar{R}$ values equal zero, because no match between the particular scene and the respective model target is found by the recognition algorithm. It means that the model target matches the scene target truth perfectly, and that is an ideal result.

From Table 1, we can see that the $\bar{R}$ values fall almost entirely on the correct targets at $\bar{R}$ levels exceeding 90% in eleven of the sixteen data sets. For the remaining five scenes, the correct targets are still assigned to the highest $\bar{R}$ values, but a portion of the $\bar{R}$ values fall on targets rather than the target truth. While the $\bar{R}$ values do not entirely fall on the correct target, the distribution of $\bar{R}$ values falls almost entirely on the correct target class. For example, for the tank 3-1 scene, the recognition algorithm is able to correctly identify the scene as a tank with 100% of $\bar{R}$ values and classify the target as a tank-3 with 84% of $\bar{R}$ values. For the armored car 2-1 scene, the algorithm is able to correctly identify the target as an armored car 2 with 89% of $\bar{R}$ values, with the remaining $\bar{R}$ values falling on armored car 1. Thus, the recognition algorithm classifies the scene as armored car with 100% of $\bar{R}$ values in the case.

Table 1. Recognition confusion matrix corresponding to $\bar{R}$.

			model							
		angular diversity	tank 1	tank 2	camion	armored car 1	tank 3	armored car 2	jeep	tank 4
field target data	tank 1-1	160°	0.94	0.06	0.00	0.00	0.00	0.00	0.00	0.00
	tank1-2	320°	0.91	0.07	0.00	0.00	0.02	0.00	0.00	0.00
	tank 2-1	190°	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
	tank 2-2	230°	0.04	0.95	0.00	0.00	0.01	0.00	0.00	0.00
	camion-1	20°	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	camion-2	200°	0.00	0.00	0.97	0.01	0.00	0.02	0.00	0.00
	armored car 1-1	20°	0.00	0.00	0.00	0.90	0.00	0.10	0.00	0.00
	armored car 1-2	340°	0.00	0.00	0.00	0.86	0.00	0.14	0.00	0.00
	tank 3-1	40°	0.01	0.00	0.00	0.00	0.84	0.00	0.00	0.13
	tank 3-2	140°	0.04	0.02	0.00	0.00	0.89	0.00	0.00	0.05
	armored car 2-1	150°	0.00	0.00	0.00	0.11	0.00	0.89	0.00	0.00
	armored car-2	30°	0.00	0.00	0.00	0.08	0.00	0.92	0.00	0.00
	jeep-1	45°	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	jeep-2	135°	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	tank 4-1	145°	0.03	0.02	0.00	0.00	0.07	0.00	0.00	0.88
	tank 4-2	210°	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.93

Although the available data and variety of targets are quite limited in this study, the confusion matrix shows that the proposed recognition algorithm can identify the correct target by assigning the largest $\bar{R}$ values to all sixteen recognition tests. Thus, we can conclude that the recognition algorithm is able to correctly classify the scene sets with 99% of $\bar{R}$ values. In addition, no matter in what direction in plan the simulated scene target is, the SCCM can recognize the target correctly, which means the new method solves the problem of the rotation invariance in plan.

In order to compare the recognition performance, $\bar{V}_{\mathrm{GOF}}$ values of the original algorithm are calculated simultaneously using the same measured scene targets and model library. Table 2 presents recognition confusion matrix composed of $\bar{V}_{\mathrm{GOF}}$ values. The $\bar{V}_{\mathrm{GOF}}$ values fall almost entirely on the correct targets at $\bar{V}_{\mathrm{GOF}}$ levels exceeding 90% in ten of the sixteen data sets. For the remaining six scenes, the correct target is still assigned to the highest $\bar{V}_{\mathrm{GOF}}$ value, but a portion of the $\bar{V}_{\mathrm{GOF}}$ values fall on targets instead of the target truth. While the $\bar{V}_{\mathrm{GOF}}$ values do not entirely fall on the correct target, the distribution of $\bar{V}_{\mathrm{GOF}}$ values falls almost entirely on the correct target class. The matrix in Table 2 resembles the matrix in Table 1, which means that the recognition performance obtained using $\bar{R}$ is the same as the one obtained by $\bar{V}_{\mathrm{GOF}}$.

Table 2. Recognition confusion matrix corresponding to $\bar{V}_{\mathrm{GOF}}$.

			model							
		angular diversity	tank 1	tank 2	camion	armored car 1	tank 3	armored car 2	jeep	tank 4
field target data	tank 1-1	160°	0.96	0.04	0.00	0.00	0.00	0.00	0.00	0.00
	tank1-2	320°	0.93	0.05	0.00	0.00	0.02	0.00	0.00	0.00
	tank 2-1	190°	0.05	0.85	0.00	0.00	0.04	0.00	0.00	0.06
	tank 2-2	230°	0.04	0.92	0.00	0.00	0.03	0.00	0.00	0.01
	camion-1	20°	0.00	0.00	0.98	0.00	0.00	0.00	0.02	0.00
	camion-2	200°	0.00	0.00	0.97	0.01	0.00	0.02	0.00	0.00
	armored car 1-1	20°	0.00	0.00	0.00	0.96	0.00	0.04	0.00	0.00
	armored car 1-2	340°	0.00	0.00	0.00	0.89	0.00	0.11	0.00	0.00
	tank 3-1	40°	0.05	0.00	0.00	0.00	0.85	0.00	0.00	0.10
	tank 3-2	140°	0.04	0.02	0.00	0.00	0.87	0.00	0.00	0.07
	armored car 2-1	150°	0.00	0.00	0.00	0.09	0.00	0.91	0.00	0.00
	armored car-2	30°	0.00	0.00	0.00	0.16	0.00	0.84	0.00	0.00
	jeep-1	45°	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	jeep-2	135°	0.03	0.00	0.00	0.00	0.09	0.00	1.00	0.00
	tank 4-1	145°	0.05	0.02	0.00	0.00	0.05	0.00	0.00	0.88
	tank 4-2	210°	0.03	0.00	0.00	0.00	0.07	0.00	0.00	0.90

The comparison indicates that both $\bar{V}_{\mathrm{GOF}}$ of the previous algorithm and $\bar{R}$ of the SCCM can represent the recognition performance. However, we recommend using $\bar{R}$ instead of $\bar{V}_{\mathrm{GOF}}$, because $\bar{R}$ can greatly simplify recognition procedures and shorten the computational time. According to our calculation, the total computing time of the SCCM is decreased by 30%–50% in comparison with the previous method. The exact percentage shortened by the new method needs to be examined by more tests.

4. Conclusions

In this paper, a novel method named SCCM is proposed to simplify the evaluating system of the spin-image surface matching algorithm. A target model library, composed of eight military ground vehicles, is compared with eight simulated random-noise range images with targets in arbitrary orientation in plane. The recognition result obtained by the simplified algorithm shows that the in-plane rotation-invariance problem for vertical detection of Ladar is resolved reasonably and a balance is achieved between the recognition efficiency and the recognition rate for practical applications. The new method meets the requirement of real-time target recognition along with a high recognition rate.

References

[1] Jacques G V, Richard L D and Dan E D 1992 *Opt. Eng.* **31** 2540

[2] Wang M, Hu X F and Wu X P 2006 *Acta Phys. Sin.* **55** 5135 (in Chinese)

[3] Gong Y J and Wu Z S 2009 *Acta Phys. Sin.* **58** 6227 (in Chinese)

[4] Nikhil R P, Tobias C C, Jim C B and Kuhu P 2001 *IEEE Trans. Fuzzy Syst.* **9** 44

[5] Shantaram V and Hanmandlu M 2002 *Proc. Int. Conf. on Information Technology* p. 274

[6] Zhao R, Liang Z C, Han B, Zhang H C, Xu R Q, Lu J and Ni X W 2009 *Chin. Phys.* B **18** 1877

[7] Gao Y Q, Zhu B Q, Liu D Z, Liu X F and Lin Z Q 2009 *Chin. Phys.* B **18** 215

[8] Zhou G Q 2009 *Chin. Phys.* B **18** 2779

[9] Andrew E J 1999 *IEEE Trant. Patt. Anal. Mach. Intell.* **21** 433

[10] Wang L, Sun J F and Wang Q 2009 *J. Russ. Laser Res.* **30** 615

[11] Zhao W Q, Chen S S and Feng Z D 2006 *Acta Phys. Sin.* **55** 3363 (in Chinese)

[12] Alexandru N V and Richard M M 2004 *Proc. of SPIE* **5426** 67

[13] Niloy J M and Nguyen A 2003 *Proc. of ACM Symp. on Computational Geometry & Applications* p. 322

[14] Paul J B and Neil D M 1992 *IEEE Trant. Patt. Anal. Mach. Intell.* **14** 239

[15] Liao T H and Gao Q 2006 *Chin. Phys.* **15** 347

[16] Pang C Y, Zhou Z W and Guo G C 2006 *Chin. Phys.* **15** 3039

[17] Zhang C, Bai L F and Zhang Y 2007 *Acta Phys. Sin.* **56** 3227 (in Chinese)

[18] Zheng Y Z 1994 *Int. J. Comput. Vision* **13** 119

[19] Sun J F, Li Q, Lu W and Wang Q 2007 *Chin. Opt. Lett.* **5** 118

[20] Li Q, Wang Y Z, Wang Q and Li Z Q 2005 *Acta Opt. Sin.* **25** 581 (in Chinese)