Ground target recognition based on imaging LADAR point cloud data

Xiaofeng Li (李晓峰)*, Jun Xu (徐 军), Jijun Luo (罗积军), Lijia Cao (曹立佳), and Shengxiu Zhang (张胜修)

Xi'an Research Institute of High-tech, Xi'an 710025, China

**Corresponding author: xiaofeng_li2006@126.com*

Received August 15, 2011; accepted October 24, 2011; posted online April 18, 2012

We illustrate an approach to statistical model and sequential hypothesis designed to the automatic target recognition (ATR) problem for active imaging LADAR. The key to this approach is using multihypothesis sequential tests to reduce the number of target hypotheses under consideration as more observed data are processed. The approach is potentially useful when sensor data are plentiful but computation time and processing capability are constrained. We experimentally demonstrate that the proposed recognition approach can resolve the military ground vehicle recognition problem of active imaging LADAR with a high recognition rate.

OCIS codes: 100.3008, 280.3640.
doi: 10.3788/COL201210.S11002.

The battlefield scenario continues to grow in complexity as the use of high-resolution sensors and precision strike weapons has forced the increased use of concealment and camouflage technology to improve weapon survivability. Thus, automatic target recognition (ATR) capability is becoming increasingly important to the Defense Community. There are two fundamental challenges in ATR for military applications. The first is the massive data loads generated by modern high-resolution sensors, which can quickly escalate to unmanageable proportions. The second challenge is camouflage, concealment, and deception, this further increases the complexity of ATR and motivates the need for robust techniques[1].

A promising approach to tackle these challenges is LADAR imagery[2]. Active imaging LADAR typically provides both intensity and range images. Intensity images give an indication of target material. The range images provide explicit three-dimensional (3D) information about targets. This 3D data contains geometric invariant properties of targets and theory avoids the types of distortions and ambiguities created by two-dimensional (2D) sensors. Additionally, active imaging LADAR provides foliage and camouflage penetration. This simplifies foreground and background clutter removal, facilitates target detection and segmentation, and yields higher recognition rates and lower false alarms rates. There is considerable interest in developing robust 3D ATR technology as LADAR sensors become more widely available, and many approaches have been developed for the purposed of target recognition based on LADAR data. Recognition methods using LADAR intensity information or the combined range and intensity information, can be found for example in Refs. [3, 4], and here we focus on the methods only using LADAR range data which usually represented as "point cloud".

Previous works on automatic target recognition in LADAR is to convert the 3D space to a more simple 2D coordinate space. These recognition algorithms take advantage of the relatively large amount of processing tools in the 2D image analysis field. Vasil *et al.*[5] used spin images for 3D LADAR data recognition. In spin-image-based representations every point in a cloud will be associated with a representative 2D image that projects surrounding points with a locally referenced coordinate system. The efficiency of their approach is limited however, as it involves analyzing and classifying data through several layers of filters and iterative processes in a point by point basis. Grönwall *et al.*[6−8] used general 3D point scatterer in their work, which focuses on recognizing ground vehicles. Their approach is based on the assumption that man-made objects of complex shape can be decomposed into a set of rectangles. Many other LADAR range data based object recognition approaches are reviewed in Refs. [6, 9].

Recently, statistical approaches are usually found in the literature dealing with ATR based on LADAR range imagery[10−14]. Rather than treating data sets as the outcome of deterministic measurement processes, statistical approaches treat measurements as random variables under particular distributions in order to model observation uncertainty. The most important advantage of the statistical approaches is that once the prior knowledge and the assumptions about the data are given, the statistical approaches always give consistent and concise solutions to problems of detection, classification and parameter estimation by application of basic principles of statistical inference.

One such approach revolves around the use of a minimum probability of error decision rule to compute the likelihood that a set of 3D observations (such as point cloud) arose from any one of a number of known possible targets. Under this decision rule, however, likelihoods must be calculated for every component point in the cloud against each candidate target hypothesis. Applying this scheme to the practical application of recognizing real-world vehicles, an individual point-wise likelihood calculation necessitates a surface integral over a non-trivial figure. The number of these costly computations scale linearly with each data point processed and each target hypothesis tested. Thus, it is very desirable to reduce the number of times new likelihoods are evaluated.

One way of reducing the number of likelihoods computed is to use sequential hypothesis testing. Many authors have reported on direct multi-hypothesis extensions of the basic SPRT. Unfortunately, it can also be shown that this type of optimality cannot extend to problems with more than two hypotheses. That is, no sequential test can minimize the ASN, subject to upper bounds on the probabilities of error, simultaneously across all true hypotheses. An approach to applying sequential multihypothesis tests was elucidated by Baum *et al.*[15], wherein they studied a generalization of sequential probability ratio test (SPRT) to more than two hypotheses that is motivated by Bayesian setting. This multihypothesis SPRT (MSPRT) was shown to be asymptotically optimal relative not only to the expected sample size but also to any positive moment of the stopping time distribution, when the error probabilities or, more generally, risks associated with incorrect decisions are small[16,17].

In this letter, we use the statistical approaches and MSPRT algorithm with the aim of reducing computation. The focus of our work is on using 3D active imaging LADAR point cloud data, and developing 3D ATR algorithm for military ground vehicle recognition.

Let the surface of an target denoted $S \subset \mathrm{R}^3$. A point $\boldsymbol{X}^* = [X_1^*, X_2^*, X_3^*]^{\mathrm{T}} \in S$ is randomly selected for measurement according to the density $p_{\boldsymbol{X}^*|\boldsymbol{\Theta},S}(\boldsymbol{X}^*|\theta, S)$ when the relative pose between the target and LADAR is θ. Let $\boldsymbol{X} = [X_1, X_2, X_3]^{\mathrm{T}}$ denote the measurement location of the surface point $\boldsymbol{X}^*$, and model the observation as $\boldsymbol{X} = \boldsymbol{X}^* + \boldsymbol{N}$, where $\boldsymbol{N}$ is an additive noise term that follows a multivariate Gaussian distribution with zero mean and covariance matrix $\sum_{\theta,S}$. The conditional probability density function for $\boldsymbol{X}$ is[12]

$$p_{\boldsymbol{X}|\boldsymbol{\Theta},S,\boldsymbol{X}^*}(x|\theta, S, x^*) = \frac{1}{(2\pi)^{3/2}\left|\sum_{\theta,S}\right|^{1/2}} \cdot \exp\left[-\frac{1}{2}(x-x^*)^{\mathrm{T}}\sum\nolimits_{\theta,S}^{-1}(x-x^*)\right]. \quad (1)$$

Given these conditional densities, the probability density function for an observation $\boldsymbol{X}$ given that surface S is measured with relative pose θ is

$$p_{\boldsymbol{X}|\boldsymbol{\Theta},S}(x|\theta, S) = \int_{x^*\in S} p_{\boldsymbol{X}|\boldsymbol{\Theta},S,\boldsymbol{X}^*} \cdot (x|\theta, S, x^*) p_{\boldsymbol{X}^*|\boldsymbol{\Theta},S}(x^*|\theta, S)\mathrm{d}x^*. \quad (2)$$

This expression reflects the fact that a given observation x could have resulted from any point x^* on the surface S, the likelihood of which depends on how likely it is that x^* was measured and that additive noise would have resulted in the corresponding observation x. The integral of all such points x^* gives the likelihood that x could have arisen from a measurement of S along θ.

Given a set of points and corresponding relative pose $\mathcal{X} = [(\boldsymbol{X}_k, \theta_k)]_{k=1}^K$, where the measured points are conditionally independent given the object from which they were measured. Classification is performed by finding the target $m \in \{1, 2, \cdots, M\}$ that most likely resulted in the measured point-cloud. The possibility that the data were drawn from the mth distribution is referred to as hypothesis m, denoted H_m. The possibility that the data was drawn according to some other unknown distribution is referred to as the null hypothesis, denoted H_0.

A sequential test procedure is a pair (N_S, ϕ_S), where $N_S(X_1, X_2, \cdots)$ is called a stopping rule, and ϕ_S is a terminal decision function, maps observations $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_K$ to one of the M distributions. For the two-hypothesis case, the sequential probability ratio test (SPRT) has been shown to minimize the average sample number (ASN) over all tests that have class conditional probabilities of error no greater than some specified ε_1 and ε_2, regardless of which hypothesis is true. At each stage, the SPRT calculates the likelihood ratio of all available data and compares the result against two threshold values. If the likelihood ratio is smaller than the minimum threshold or greater than the maximum threshold, the corresponding hypothesis is selected. If the ratio lies between the two thresholds, the data are taken to be ambiguous, and additional data is collected. The test then repeats with additional data[13].

Like the SPRT, the multihypothesis SPRT (MSPRT) compute the likelihood of the entire data collection under each hypothesis at every stage. By dropping unlikely hypotheses from further consideration, the number of hypotheses, and thus the amount of processing for each new data sample, decreases with each stage.

Suppose that when H_m is true, we use the data model described, the log-likelihood of surface S_m is

$$L_m^k(S_m; \mathcal{X}) = \sum\nolimits_{k=1}^{K} \log[p_{\boldsymbol{X}|\boldsymbol{\Theta},S}(x_k|\theta_k, S_m)], \quad (3)$$

where S_m is the surface of target m.

Let $F_m^k(l)$ denote the cumulative distribution function (CDF) of $L_m^k(S_m; \mathcal{X})$ under the assumption that H_m is true. Optimal hypothesis testing algorithm are based on the fact that when H_m is not true, we expect the log-likelihood hypothesis $L_m^k(S_m; \mathcal{X})$ to be small. Assuming $F_m^k[L_m^k(S_m; \mathcal{X})] = \alpha$, there is a 100α percent probability that by pure chance alone an even less likely sequence of data would be observed if H_m were in fact true. This can form the basis of a discrimination test for whether or not to drop H_m at stage N.

So the whole process of recognition algorithm is as follows[13].

Step 1: define α to be the largest tolerable false rejection rate for the problem.

Step 2: initialize the stage number $k = 0$ and let $M^0 = \{1, 2, \cdots, M\}$ to be set of target classes initially under consideration.

Step 3: increment k by a pre-specified value, collect observation $\mathbf{X}_k$ and compute $L_m^k(S_m; \mathcal{X})$(Eq. (3)) for each $m \in M$.

Step 4: let $M^k = \{m \in M^{k-1} \,|\, F_m^k[L_m^k(S_m; \mathcal{X})] > \alpha\}$ be the set of target classes still under consideration after stage k.

Step 5: if $|M^k| > 0$, report.

(a) $\widehat{m} = \arg\max_{m\in M^n} L_m^k(S_m; \mathcal{X})$ as the most likely hypothesis found at the end of stage k;

(b) $F_m^k[L_m^k(S_m; \mathcal{X})]$ as the significance of that hypothesis;

(c) M^k as the set of feasible alternatives.

Step 6: if $|M^k| = 0$, report that all of the known target hypothesis have been rejected.

Step 7: if $|M^k| > 1$ go to step 3, otherwise terminate.

For the study of our ATR algorithm, we used the Georgia Tech LADAR simulator[18] to generate the raw point cloud data. We use eight specific of interest, consisting of military ground vehicles ranging from tanks to camion and armored car. Computer Aided Design (CAD) models of the specific targets are shown in Fig. 1. The model library contains two multi-vehicle targets-classes, namely tanks and Armored Personnel Carriers (APCs). The tank class is composed of the M1-A1, M60 and WMCVUN. The APC target class includes the M2-A2, GMC, Hemtt, Willys and Hummer vehicles. For each target, there are 72 orientations, corresponding to the aspect angles of 0°, 5°, ⋯ , 355° in azimuth.

Based on the above CAD models, a target model library was constructed to simulate an ideal 3D LADAR point clouds data of each target. Each of these eight targets were generated by LADAR simulator at 10 different ranges: 100, 200, ⋯, 1 000 m. At each range, point clouds were acquired for 4 different target poses: front (0°), front left corner (45°), left side (90°), and back left corner views (135°). Measurement points were collected all around each model, expect for the undersides where is not observed by the sensor. Examples of Simulated point clouds from CAD models (shown in Fig. 1) at 200 meters and 45° poses were shown in Fig. 2.

Fig. 1. Target CAD models.

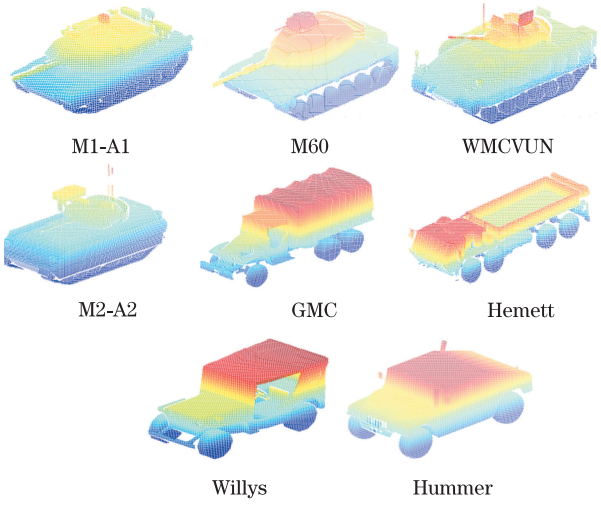

Fig. 2. LADAR point clouds simulated from CAD models.

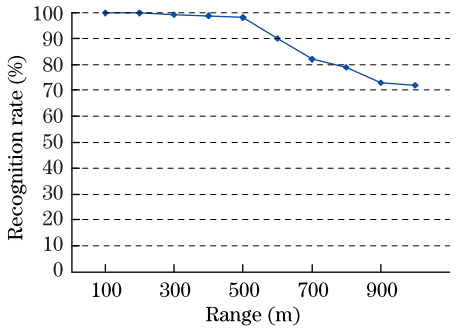

Fig. 3. Average recognition rate versus range.

In order to avoid redundant computation and expedite experimental runs, all the log-likelihoods corresponding to every point-cloud/target-hypothesis pair were computed offline prior to the experiments. Sample means and variances were also pre-computed from random 1 000-point sampling of the likelihoods corresponding to the correct classification under every target hypothesis. These likelihoods sets were shuffled point-wise at the beginning of each experimental trial to simulate the unordered nature of collected data. A 5% significance level is set for most of the experiments as that was the common value found in the literature.

In the experiment, target truth was known prior to data collection. We assume the entire target could be seen. As the number of measured points is limited by the LADAR system and operational conditions, a LADAR system working at close range can acquire more points on target than when working at large distances. Thus, the number of points on target is lower in longer range.

The average recognition rate versus range is presented in Fig. 3. Figure 3 shows that the proposed algorithm successfully identified nearly all targets in all poses up to 500 m. The performance starts to drop past 500 m and reaches 72% at 1 000 meters. This due to rapidly decreasing the number of measured points will result in lower recognition accuracy. These results demonstrate that the ATR algorithm works very well with simulated data.

Table 1 presents the recognition confusion matrix obtained from the comparison of model library to each of target scenes for all ranges combined. The left column represents the input target vehicle and the top row represents a target scene to model library comparison. For instance the first row contains the comparison between a M1-A1 scene measurement and the model library. The recognition confusion matrix in Table 1 resembles an identity matrix, which would be the ideal result. Analyzing the failure cases demonstrates that the system performance degrades gracefully with decreasing data quality. Even when the system does not select the right vehicle, it selects objects that are similar in shape most of the time.

To evaluate the performance of our algorithm, we propose to compare it with the spin-image surface-matching algorithm presented in Ref. [5]. The experiment results are summarized in Table 2. As we can notice, the average recognition time per model of our method is approximately 3 s on a Windows XP workstation with CPU speed of 2.8 GHz and 1 GB of memory. To compare with, the algorithm presented in Ref. [5] lasted 137 s under the

Table 1. Recognition Confusion Matrix

	M1-A1	M60	WMCVUN	M2-A2	GMC	Hemtt	Willys	Hummer	Unknown	Total Pass	Total Scene	Recognition Rate (%)
M1-A1	**33**	3	2	2	0	0	0	0	0	33	40	82.5
M60	4	**32**	3	0	0	0	0	0	1	32	40	80
WMCVUN	1	2	**35**	2	0	0	0	0	0	35	40	87.5
M2-A2	3	1	1	**34**	0	0	0	0	1	34	40	85
GMC	0	0	0	0	**37**	0	2	0	1	37	40	92.5
Hemtt	0	0	0	0	0	**39**	0	0	1	39	40	97.5
Willys	0	0	0	0	2	0	**34**	4	0	34	40	85
Hummer	0	0	0	0	1	0	3	**35**	1	35	40	87.5

Table 2. Comparison of ATR Performance between Two Methods

Method	Average Recognition Rate (%)								Average Recognition Time per Model (s)
	M1-A1	M60	WMCVUN	M2-A2	GMC	Hemtt	Willys	Hummer	
Proposed Algorithm	82.5	80	87.5	85	92.5	97.5	85	87.5	3.02
Algorithm in Ref. [5]	85	76	87	91	96	99	81	92	137.5

same conditions. In our algorithm, the log-likelihoods and sample means and variances were computed offline, as the algorithm in Ref. [5] needs to create spin-image and matching making its recognition time more slowly. Note that the approach in Ref. [5] considers all point-clouds and all hypotheses, making its average recognition rate a bit higher than our algorithm in some scenes.

In conclusion, we present a sequential hypothesis testing algorithm based on statistical approaches for the recognition of targets in LADAR point clouds data. The proposed recognition algorithm performance is successfully demonstrated on eight 3D models of potential targets. 8 military ground vehicles are acquired during experiment with LADAR Simulator. 3D point clouds data of targets are acquired at ranges from 100 to 1 000 m in 4 different poses at every range. It is experimentally shown that the proposed algorithm can reach a high recognition rate in fewer computations. The target recognition performance of our method is successfully demonstrated on simulated point cloud data; however, the performance needs to be validated with real scene data if the condition will be satisfied in the future. The statistical approach needs to be improved in order to increase the recognition rate for the whole algorithm.

This work was supported by the National Natural Science Foundation of China under Grant No. 60874093.

References

1. S. Roy, S. Se, and V. Kotamraju, Proc. SPIE **7696,** 76960M (2010).
2. E. Bovenkamp and K. Schutte, Proc. SPIE **7684,** 76840Z (2010).
3. N. R. Pal, T. C. Cahoon, and J. C. Bezdek, IEEE Trans. Fuzzy Systems **9,** 44 (2001).
4. M. R. Stevensa, M. Snorrasona, and S. Amphayb, Proc. SPIE **4729,** 133 (2002).
5. A. N. Vasile and R. M. Marino, Lincoln Laboratory J. **15,** 61(2005).
6. C. Gröwall, F. Gustafsson, and M. Millnert, IEEE Trans. Image Processing **15,** 3401 (2006).
7. C. Gröwall, "Ground Object Recognition Using Laser Radar Data: Geometric Fitting, Performance Analysis, and Applications", PhD Thesis, (Linköings Universitet, 2006).
8. C. Gröwall, G. Tolt, and D. Karlsson, Proc. SPIE **7696,** 76960N (2010).
9. R. D. Nieves and W. D. Reynolds, Proc. SPIE **7684,** 76840Y (2010).
10. M. D. Devore and X. Zhou, Proc. SPIE **6234,** 623407 (2006.)
11. X. Zhou, "Statistical Model-based Object Recognition from Three-Dimensional Point-Cloud Data" PhD Thesis, (University of Virginia, 2008).
12. X. Zhou and M. D. Devore, IEEE Trans. Aerospace And Electronic Systems **44,** 1416 (2008).
13. I. O. Reyes, M. D. Devore, and P. A. Beling, Proc. SPIE **7696,** 769615 (2010).
14. M. D. Devore and A. J. Bateman, Proc. SPIE **7696,** 76960R (2010).
15. C. W. Baum and V. V. Veeravalli, IEEE Trans. Information Theory **40,** 1994 (1994).
16. V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli, IEEE Trans. Information Theory **45,** 2448 (1999).
17. V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli, IEEE Trans. Information Theory **46,** 1366 (2000).
18. J. H. Dixon, "LADAR Simulator User Manual", http://users.ece.gatech.edu/∼jdixon/ladar_sim/ (October 12, 2010).