

Ground target recognition using rectangle estimation

Christina Grönwall, Fredrik Gustafsson, Mille Millnert

Abstract—We propose a ground target recognition method based on 3D laser radar data. The method handles general 3D scattered data. It is based on the fact that man-made objects of complex shape can be decomposed to a set of rectangles. The ground target recognition method consists of four steps; estimation of the target's 3D size and orientation, segmentation of the target into parts of approximately rectangular shape, identification of segments that contain the main parts of the target and matching the of target with CAD models.

The core in this approach is rectangle estimation. The performance of the rectangle estimation method is evaluated statistically on simulated data. A case study on tank recognition is shown, where 3D data from three fundamentally different types of laser radar systems are used.

Index Terms—Rectangle estimation, laser radar, automatic target recognition

I. INTRODUCTION

A. Ground target recognition using 3D imaging laser radar

Laser radar systems have been investigated over several decades primarily for military applications [19, 25, 26]. The high resolution in angle-angle-range makes 3D imaging possible and due to the short wavelength, in general 0.5-10 μm , detailed range images of objects and background can be obtained. Due to the high resolution, even at km distances, details of a target can be resolved. This can be used for automatic target recognition (ATR). For example, if main parts of a tank (the barrel and turret) can be extracted, the hypothesis that the target is a tank is strengthened. Further, if articulated parts of a target can be identified, the target recognition can be simplified as the degrees of freedom reduce.

In this paper we propose a ground target recognition method based on 3D laser radar data. The method handles general 3D scattered data. It is based on the fact that man-made objects of complex shape can be decomposed into a set of rectangles. The method consists of four steps; 1) estimation of the target's 3D size and orientation, 2) segmentation of the target into parts of approximately rectangular shape, 3) identification of segments that contain the main parts of the target and 4) matching the target with library models.

From a computer vision perspective, this sequential processing of data is not optimal. An advantage is that even if a matching model cannot be found, we can report the estimated size and orientation and possibly some identified features. Further, when performing matching, the list of possible models has been limited.

C. Grönwall is with the Swedish Defence Research Agency, Dept. of Laser Systems, Linköping, Sweden. E-mail: christina.gronwall@foi.se. Her former surname was Carlsson.

F. Gustafsson and M. Millnert are with the Dept. of Electrical Engineering, Linköping University, Linköping, Sweden. E-mail: {fredrik,mille}@isy.liu.se.

B. The ATR framework

The framework of the target recognition method proposed in this paper is described in [2, 3, 15]. The framework is a query-based multi sensor information system for ground target recognition. Based on an operator's query, the system selects proper sensor data and analysis algorithms to perform the task. Once the target is detected, a four-step target recognition process is performed. The recognition is based on infrared, visual and laser radar data. First the sensor data is analyzed to estimate target attributes, for example position, dimensions and temperature. The attributes from different algorithms are then fused. Based on the attribute fusion, models of typical military vehicles are selected and the models are matched with sensor data. The model library contains wire-frame CAD models with thermal and visual textures. The results from the model matching are then subject to model match fusion and finally the most likely match results are presented to the operator. The method described in this paper is used both in the attribute estimation and in the model matching.

C. Outline

In the next section, we review some of the ATR work based on laser radar data and methods for rectangle estimation. In Section III, the rectangle estimation method is described and analyzed. In Section IV, the segmentation of objects with complex shape is described. In Section V, we propose a ground target recognition method based on rectangle estimation and in Section VI it is applied to tank recognition. The results and future work are discussed in Section VII and in Section VIII we conclude this paper.

II. RELATED WORK

A. Vehicle recognition using laser radar

Several ATR methods or systems for recognition of military ground vehicles based on laser radar data have been proposed over the years [3, 10, 30, 31, 33, 35]. During the last years, also ATR of civilian personal cars, mainly for traffic monitoring, have been proposed [16, 28, 34].

The approaches are applied to data of different resolution and different perspectives of the target. In [3, 28, 33]–[35], low resolution data is considered. A typical data set contains up to a few hundreds samples on a target, while [28] handles very low-resolution data (approx. 1.5 points/m²). In [10, 16, 30, 31], there are typically several hundreds of samples on the target. Typically, the data is collected in a forward-looking perspective, while in [3] and [28] were down-looking perspective data is considered. Often data is obtained using a scanning laser radar system, which results in irregularly sampled data. In [30, 31, 35], the laser sensor works in staring

mode, which gives regularly sampled data. Further, in [16, 30], data is collected from several views, which results in data that is less self-occluding.

In most cases, the ATR process is divided into two steps. Usually the first step consists of fast feature extraction or silhouette calculations [10, 31, 33, 34]. The feature extraction can retrieve geometrical properties of the target [3], lower-dimensional properties [35] or more abstract features like spin image representation [16, 30] (see [20] for description of spin images). The first step is used to reduce the list of potential targets. Then, the remaining targets are subject to 3D matching with library models, which are represented by CAD models [3, 10], some representation generated from CAD models [31, 33, 35] or 3D scatter data [10, 30, 34]. The ATR approach [35] is further evaluated in [18]. In [28], learning is used for the recognition. The methods in [10, 16, 30] can handle partly occluded targets. The problem with partly occluded targets is discussed, for example, in [3].

B. Rectangle estimation for complex shape analysis

When analyzing an object with complex shape, registered in 2D by passive imaging or projection of 3D data, the orientation can be estimated by rectangle fitting. An iterative approach is proposed in [12]. In [9, 32, 36], non-iterative approaches to rectangle estimation is used to find good initial values for further processing. The objects that are characterized are asteroids [36], buildings [32] and vehicles [9], respectively. In [32, 36], eigenvalue calculations are used to estimate the orientation of the object. After that, a rectangle that bounds the object samples [36] or is optimal in second order moment [32] is calculated. In [9], a rectangle that bounds the object data is estimated by solving an optimization problem, which is described further in Section III.

III. RECTANGLE ESTIMATION

A. Definition

The current approach for rectangle estimation has been described independently under the name *Rotating Calipers* [29] and in [8, 9]. This rectangle estimation approach is more general than the methods based on principal axis estimation [32, 36], as there is no demand that the orientation scatter matrix must be positive definite.

We describe the rectangle estimation problem as an optimization problem. A straight line in two dimensions is described as $n_1x + n_2y - c = 0$, where the normal vector $\mathbf{n} = (n_1, n_2)^T$ defines the slope of the line and c the distance to origin and (x, y) is measurement data known to be on the object, possibly contaminated with noise. The points $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ are inside the rectangle or on one of the sides of the rectangle if

$$\text{Side 1} : n_1x_i + n_2y_i - c_1 \geq 0, \quad i = 1, \dots, N \quad (1a)$$

$$\text{Side 2} : -n_2x_i + n_1y_i - c_2 \geq 0, \quad i = 1, \dots, N \quad (1b)$$

$$\text{Side 3} : n_1x_i + n_2y_i - c_3 \leq 0, \quad i = 1, \dots, N \quad (1c)$$

$$\text{Side 4} : -n_2x_i + n_1y_i - c_4 \leq 0, \quad i = 1, \dots, N \quad (1d)$$

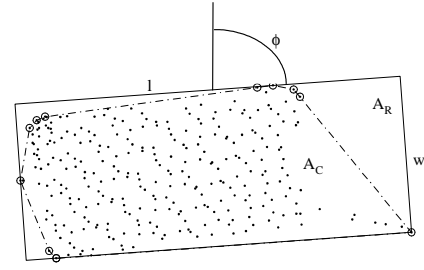


Fig. 1. Illustration of the parameters estimated in the rectangle estimation. A set of samples (dots), the convex hull (dashed line) and the estimated rectangle (solid line). The samples belonging to the convex hull are encircled. The parameters are length (l), width (w), orientation (ϕ), convex hull area (A_C) and rectangle area (A_R).

where $\mathbf{n}^T \mathbf{n} = 1$. If we introduce $X_i = (x_i, y_i)$ and the rotation matrix

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

we can formulate the rectangle estimation problem as a minimization problem, where the rectangle's area is the objective function:

$$\min (c_3 - c_1)(c_4 - c_2) \quad (2)$$

subject to

$$\begin{aligned} X_{1,i} \mathbf{n} - c_1 &\geq 0, & i = 1, \dots, N \\ X_{1,i} R \mathbf{n} - c_2 &\geq 0, & i = 1, \dots, N \\ X_{1,i} \mathbf{n} - c_3 &\leq 0, & i = 1, \dots, N \\ X_{1,i} R \mathbf{n} - c_4 &\leq 0, & i = 1, \dots, N \\ \mathbf{n}^T \mathbf{n} &= 1. \end{aligned}$$

Based on the estimates of \mathbf{n} and $c_j, j = 1, \dots, 4$, the rectangle's length, l , width, w , area, A_R , and orientation, ϕ , are calculated, as illustrated in Figure 1.

Problem (2) is not convex, as the objective function and the last constraint are not convex, but it is proven in [9, 23] that there exists a unique solution. There is a constraint that limits the number of possible orientations of the rectangle, see Theorem 1.

Theorem 1 (Minimal rectangle): The rectangle of minimum area enclosing a convex polygon has a side colinear with one of the edges of the polygon.

Proof: See [11]. The proof is also performed in [8, 9, 23]. ■

Using this theorem, we can limit the number of possible orientations of the rectangle, only rectangles that have one side colinear with one of the edges of the convex hull (that is a convex polygon) have to be tested.

In [9] and [29] (almost similar) algorithms are given for calculation of (2) in linear time, i.e., $O(N_v)$ where N_v is the number of vertices in the convex polygon. Further, the convex hull can be calculated in $O(N \log N)$ time if data is unsorted and in $O(N)$ time if data is sorted (N is the number of samples). In [8] a sorting algorithm for scanned laser radar data is proposed, whose execution time is linear in the number of samples. The implementation [9] is based on that four samples shall span the rectangle, one sample for each side, i.e., we have $N_v \geq 4$.

B. Performance

The performance of the estimation method (2) is investigated in Monte Carlo simulations. The performance is evaluated in terms of correctness in estimates of $\theta = (l, w, \phi, A_R)$. Further, the ratio between the convex hull's area and the rectangle's area, A_C/A_R , is studied. We start with random placement of N samples in (x, y) , where $x \in \mathcal{U}(-l^0/2, l^0/2)$ and $y \in \mathcal{U}(-w^0/2, w^0/2)$, respectively, where $\mathcal{U}(\cdot)$ is the uniform distribution. These samples are considered noise free. Random errors, Gaussian distributed with zero mean and equal variance $\sigma_{e_x}^2 = \sigma_{e_y}^2$ are added to $(x, y)_i, i = 1, \dots, N$. The noise is generated separately for x and y . The parameters θ are estimated using (2) on the perturbed data set. The statistical properties of the estimates are studied by the mean squared error (MSE) and bias, which are averaged over 100 sets. The MSE and the bias for parameter θ_j are defined as

$$\begin{aligned} \text{MSE}(\hat{\theta}_j) &= E(\hat{\theta}_j - \theta_j^0)^2 + E^2(\hat{\theta}_j - \theta_j^0) \\ &= \text{Var}(\hat{\theta}_j) + \text{bias}^2(\hat{\theta}_j), \end{aligned}$$

where θ_j^0 is the true, but unknown, parameter and $\hat{\theta}_j$ is the estimate. The properties of the area ratio A_C/A_R is studied using mean and standard deviation. The properties of the estimates are studied as a function of the number of samples, N , and signal to noise ratio (SNR). SNR is defined as

$$\text{SNR} = \min\left(\frac{r(x)}{\sigma_{e_x}}, \frac{r(y)}{\sigma_{e_y}}\right), \quad (4)$$

where $r(x)$ is the range in data, $r(x) = x_{\max} - x_{\min}$.

1) *Length, width and area estimates*: In Figure 2, the MSE in length estimate is shown for the case $l^0/w^0 = 2/1$. We can note a "knee" in the graph. For low SNR the dominating statistical distribution is the distribution of the noise, i.e., the Gaussian distribution. For high SNR, the dominating statistical distribution is the distribution of the samples, i.e., the uniform distribution. For lower SNR, more samples are needed to reach the uniform distribution as the dominating one. Similar results were obtained for $l^0/w^0 = 3$ and $l^0/w^0 = 4$. Similar results were also obtained for width and area estimates, see [14]. The length, width and area estimates contain bias. It is shown in [14, 22] that $\text{bias}(l) = 2l_0/(N+1)$, $\text{bias}(w) = 2w_0/(N+1)$ and $\text{bias}(A_R) = 4NA_0/(N+1)^2$.

2) *Orientation estimates*: In the Monte Carlo simulations of the orientation estimate, the squared bias level is 10-100 times lower than the MSE. Further, there is no obvious structure in the bias plots. This means that $\text{MSE}(\hat{\phi}) \approx \text{Var}(\hat{\phi})$ for all SNR values and N and that the orientation estimate is unbiased, see further evaluations in [14]. Figure 3 shows the MSE of the orientation estimate.

3) *Area ratio A_C/A_R* : For the area ratio A_C/A_R , the mean and standard deviation are studied, see Figures 4-5. For noise free data we have $1/2 \leq A_C/A_R \leq 1$, where the lower limit is reached for three samples ($N = 3$). The upper limit is reached when there is an infinite number of samples. For a low SNR and a large number of samples the shape of convex hull will approach an ellipse, i.e., $A_C/A_R \rightarrow \pi/4 (\approx 10^{-0.1})$. The knee in the graph in Figure 4 indicates when the

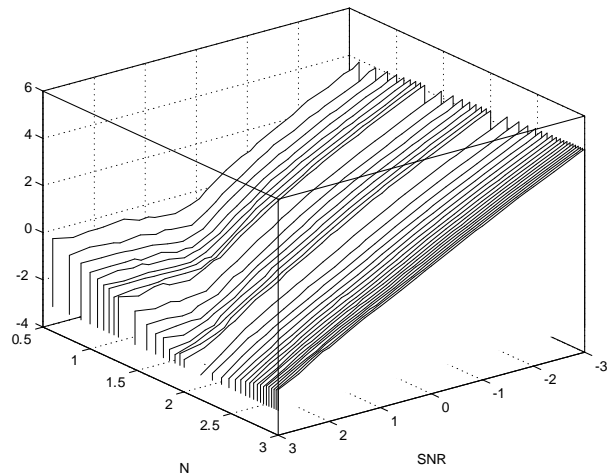


Fig. 2. MSE of length estimate, as a function of number of samples N and SNR. Logarithmic scale on axes.

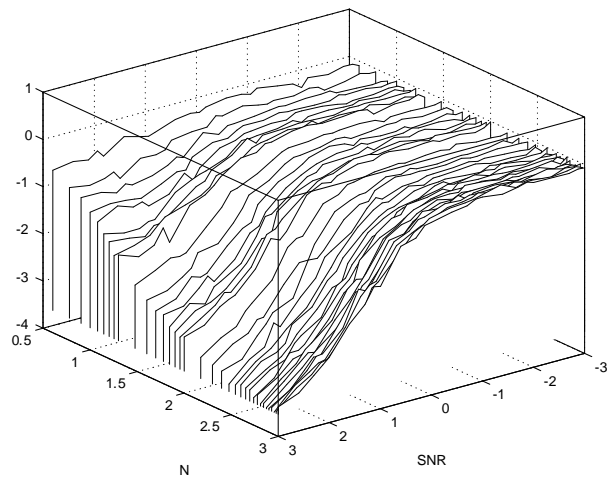


Fig. 3. MSE of orientation estimate, as a function of number of samples N and SNR. Logarithmic scale on axes.

convex hull approaches an ellipse. For small samples sets, both high and low SNR, the standard deviation of the estimate is approximately 10% and when the number of samples increases the standard deviation decreases to 2-4%.

IV. SEGMENTATION OF COMPLEX SHAPES

Man made objects, like vehicles and buildings, are in certain projections of rectangular shape. When the objects are of more complex shape, they can usually be decomposed into a set of rectangles. In this section, we describe an approach to decompose a complex shape to a set of rectangles. The approach have similarities with [32], a main difference is that it handles irregularly sampled data.

This method works on 2D data retrieved from projections of 3D data. If the current data set is not approximately similar to a rectangle, the data set is considered to describe a complex shape and it will be subject to segmentation. We split the object recursively by sliding a splitting line that is parallel to first the primary and then to the secondary axis of the rectangle. The

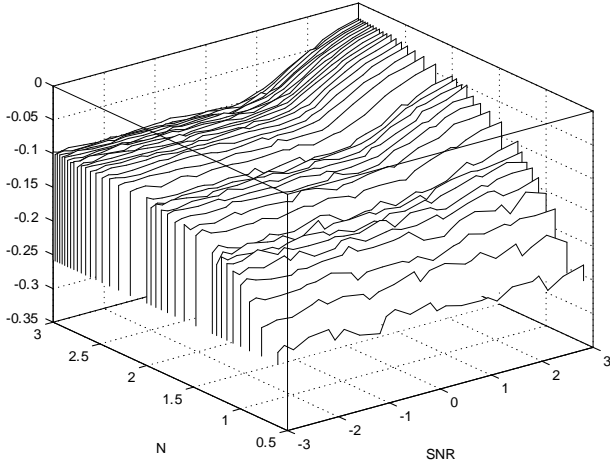


Fig. 4. Mean of area ratio A_C/A_R , as a function of number of samples N and SNR. Logarithmic scale on axes.

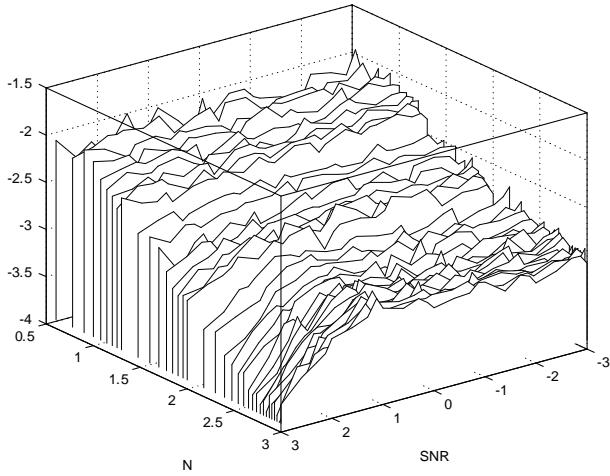


Fig. 5. Standard deviation of area ratio A_C/A_R , as a function of number of samples N and SNR. Logarithmic scale on axes.

data set will be traversed a certain distance δ in each iteration. Tests have shown that δ should be of the same magnitude as the searched subparts of the object. The two subsets of the object (part) that have the smallest total area are selected for segmentation. The result of the segmentation is stored in a binary tree T . In a tree, each terminating node (leaf), t , contains indices to either a rectangle-like part of the object or a part that cannot be further split.

An indication that node t needs further splitting is the dissimilarity of the bounding rectangle's area and the area of the convex hull of the samples stored in node t . The area ratio is similar to the Hausdorff measure used in [32]. Let $A_R(t)$ denote the bounding rectangle's area for the samples in node t and $A_C(t)$ the area of the convex hull for the samples in node t . The area ratio for t is defined as

$$M(t) = \frac{A_C(t)}{A_R(t)}, \quad (5)$$

where $1/2 \leq M(t) \leq 1$. If $M(t)$ is smaller than a threshold

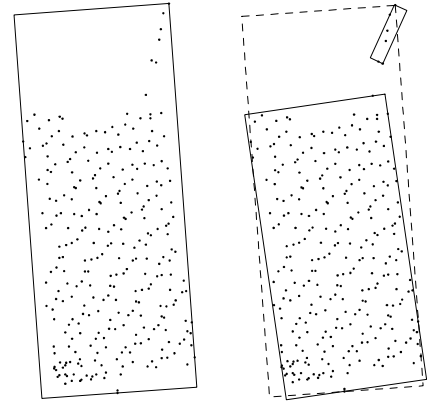


Fig. 6. Example of splitting of node t . Left: node t , $M(t)=0.79$. Right: after splitting, the dashed rectangle is the rectangle for node t . Upper node is t_L , $M(t_L)=0.94$, and lower node is t_R , $M(t_R)=0.27$.

τ , the data set stored in node t is considered not being of rectangular shape. Thus, the contents in t will be split to t_L and t_R (i.e., left and right leaf in the binary tree). The segmentation algorithm can be summarized in six steps:

- 1) Calculate $M(t)$, see (5).
- 2) Calculate SNR (4) and select τ from a table.
- 3) If $M(t) < \tau$, proceed below. Otherwise, terminate.
- 4) Split node t into t_L and t_R . Do one separation for each increment δ .
- 5) Select the t_L and t_R that have the smallest total rectangle area.
- 6) Check area ratios $M(t_L)$ and $M(t_R)$:
 - a) if $M(t_L) \geq \tau$ or $M(t_R) \geq \tau$, save t_L or t_R , respectively, and terminate.
 - b) if $M(t_L) < \tau$ or $M(t_R) < \tau$, segment t_L or t_R , respectively, further.

The threshold τ is based on statistics of the number of samples and SNR (4) in the current data set (see Figures 4-5, Section III), where the noise variance $\sigma_{e_x}^2$ is given by the measurement system model. Segmentation is only performed if $N \geq 8$. An example of segmenting is shown in Figure 6. On rare occasions the samples are distributed such that the convex hull contains less than four edges. Then the bounding rectangle cannot be calculated. The bounding rectangle for the contents in node t_L (or t_R) will then be approximated by its upper bound:

$$A_R(t_L) \leq A_R(t) - A_R(t_R)$$

and the orientation will be estimated using principal component analysis.

V. APPLICATION TO GROUND TARGET RECOGNITION

A. Introduction

In this section, we apply the rectangle estimation and segmentation approach to recognition of ground targets. The main steps of the method are described here and are illustrated in the next section.

We assume that a vehicle viewed in different projections can be approximated by a set of rectangles and that in some

views the rectangles will describe the main parts of the target. When a target is measured with a laser radar, we can derive a 3D view of the object. This means that data can be projected to an arbitrary view. On the other hand, a laser beam does not penetrate dense materials like metal surfaces. Thus, we only collect data from the parts of the object that are visible from the laser radar's perspective (so-called self-occlusion). Further, in this application we cannot assume that the vehicle is placed in a certain pose relative to the sensor and we cannot assume any certain orientation or articulation of the vehicle.

The object recognition algorithm consists of four steps:

- 1) Estimate the target's 3D size and orientation using the rectangle estimation method described in Section III.
- 2) Segment the target into parts of approximately rectangular shape using the method described in Section IV. The main parts of the object are stored in (some of) the terminating leaves.
- 3) Check the terminal nodes for possible target parts by simple geometric comparisons. One node can belong to several classes.
- 4) Match the entire object with a wire-frame model. The model's main parts are rotated to the estimated orientations.

By using a large segmentation step ($\delta = 1$ meter) the typical main parts of a vehicle can be detected. The mean of A_C/A_R is used as threshold (τ).

B. 3D orientation estimation

We first study the object in top view and then rotate to side and front/back views. The 3D orientation consists of five steps:

1. Transform data to top view perspective.
2. Estimate a rectangle based on top view data (x, y) using (2). The main directions of the target are given by the orientation of the rectangle. The yaw angle is given by the orientation of the rectangle's main axis.
3. Project the data set into the direction (x', y') , where x' is parallel to the main and y' is parallel to the secondary axis.
4. Estimate a rectangle based on side view data (x', z) . The pitch angle is given by the orientation of this rectangle.
5. Estimate a rectangle based on back/front view data (y', z) . The roll angle is given by this rectangle's orientation.

C. Target segmentation and node classification

The target is segmented in each view and in horizontal and vertical directions, respectively. This results in six descriptions of the target, stored in six binary trees T_1, \dots, T_6 . Depending on the sample density, what parts of the target that are registered and the correctness in the 3D orientation estimation, some terminating leaves will contain the main parts of the target while other do not have a clear geometrical interpretation. The leaves of the six trees are searched for typical features, like barrel and turret, using geometric rules of length, width, height and distance between the part's inertia and the main part's inertia. The geometric rules are given by the model library.

D. Matching

The 3D data of the target will be matched with a low-resolution CAD model. The distance between the target samples and the model facets is calculated using the bidirectional Hausdorff distance [6]. If the target's main parts have been identified the model's parts are rotated to the estimated orientations. Otherwise, the target will be matched with the model in default orientation.

The matching score is calculated using the relative mean squared error (RE) [7]. Let $(x, y, z)_i$ define target sample i and $(x', y', z')_i$ the projection on the closest model facet. The RE is defined as

$$RE = \frac{H((x, y, z), (x', y', z'))}{S(x, y, z)}, \quad (6)$$

where $H((x', y', z'), (x, y, z))$ is the MSE from the one-sided Hausdorff calculation and $S(x, y, z)$ is the spread in data estimated by

$$S(x, y, z) = \frac{1}{N} \sum_{i=1}^N \|(x, y, z)_i - \bar{\mu}\|_2^2,$$

where $(x, y, z)_i$, $i = 1, \dots, N$, is the perturbed data set and $\bar{\mu} = (\bar{\mu}_x, \bar{\mu}_y, \bar{\mu}_z)$ is the estimated mean value. The RE is always nonnegative and for good initial fits of model and target, $H((x, y, z), (x', y', z')) < S(x, y, z)$ [7], thus $0 \leq RE < 1$.

The matching score can be improved by least squares fitting [5]. In this approach we minimize the distance between the targets samples and their projected samples, i.e.,

$$\min_{R, T} \sum_{i=1}^N \|(x, y, z)_i - ((x', y', z')_i R) + T\|, \quad (7)$$

where R is the rotation matrix and T the translation.

VI. CASE STUDY: TANK RECOGNITION

In this section the steps of a target recognition process is shown in five examples. The examples show registrations of T72 tanks performed with three fundamentally different types of laser radar systems. Two of the laser radars register both 3D and reflectance, but the reflectance data is not used in this paper. All targets are placed in open terrain on grass fields and no occluding objects were present.

A. The data sets

The first examples, target A, B and C, are recorded with a helicopter-borne down-looking scanning laser radar¹ [17]. The helicopter was flying at 25 m/s at an altitude of 130 meters above ground. The scanning laser radar operates in the near infrared (NIR) at 1.06 μm with 0.1 mJ/pulse and a sampling rate of 7 kHz. The footprint on ground is approximately 0.14 m and the distance between samples approximately 0.3 m along the scanning lines and 0.5 m between the scanning lines. The measurement uncertainty is approximately 0.1 meters in x , y and z . The field of view is 20° (degrees) perpendicular to the

¹The TopEye system from TopEye AB, see www.topeye.com.

flight direction. The scanning constitutes a zigzag pattern on the ground and the resulting data is in point scatter format containing 3D position and reflected intensity in each sample, i.e., the data is an unordered set of samples (x, y, z, r) . The measurement model of this system is given in [13].

Target D is recorded with another scanning laser radar system². It operates at $1.5 \mu\text{m}$ with a sampling rate of 2 kHz. The footprint on the target is approximately 0.015-0.02 m and the distance between samples is approximately 0.3 m both along and between the scanning lines. The maximum field of view is $40^\circ \times 40^\circ$ degrees. The resulting data, after post processing, is an unordered set of samples (x, y, z, r) . The measurement uncertainty is approximately 0.015 m in x and y and 0.02 m in z (depth). The laser radar system was placed 5 m above ground and approximately 190 m from the target, to constitute forward-looking perspective.

Target E is recorded with a horizontally looking, ground-based range scanning system, i.e., a gated viewing system [24]. In gated viewing, a camera is time controlled with respect to a pulsed illuminating source. The gated viewing laser is an experimental system working at 532 nm with 63 mJ/pulse and a range gate of 40 ns (corresponding to a depth resolution of 6 meters). For every laser pulse approximately six meters of the terrain is illuminated. By sliding of the time gate, a sequence of 2D images is obtained. Using the method described in [4], the set of 2D intensity images are transformed into a regular grid with range information (3D data of the scene). The measurement uncertainty is approximately 0.02 m in x and y and 0.04 m in z (depth). The system is ground-based and the target is registered at a distance of approx. 2 km in side-looking perspective. The field of view is approximately $0.5^\circ \times 0.5^\circ$.

B. Preprocessing

We assume that the target area is detected [2, 3, 15]. An area of approximately 15×15 meters containing the target is selected. In this case, where the targets are placed in open terrain, we use step 4-5 in the 3D orientation estimation algorithm (Section V-B) to estimate the ground's slope. We compensate for the slope and the ground and the targets are separated using height difference.

C. 3D orientation estimation

The 3D orientation and size estimates for target A-E are shown in Figures 9-13. The estimated dimensions are shown in Table I, the true orientations are not known. For target B, C and E the barrel is not pointing straight forward, this affects the length and width estimates. The length estimates for the complete target (with barrel) are within $\pm 10\%$ of the true value and the length estimates for the target's main part (without barrel) are within $\pm 6\%$ of the true value. The width estimates for the complete target (with barrel) are within $\pm 37\%$ of the true value and the width estimates for the target's main part (without barrel) are within $\pm 14\%$ of the true value. The reduction in length and width estimation errors are due to the

Target	#Samples	Length (m)	Width (m)	Height (m)
A w. barrel	129	8.69 (-0.96)	3.58 (+0.06)	2.25 (-0.24)
A no barrel	126	6.67 (-0.46)	3.58 (+0.06)	2.25 (-0.24)
B w. barrel	191	8.98 (-0.67)	4.81 (+1.29)	2.46 (-0.03)
B no barrel	185	6.96 (-0.17)	4.01 (+0.49)	2.46 (-0.03)
C w. barrel	287	9.59 (-0.06)	3.86 (+0.34)	2.38 (-0.11)
C no barrel	281	7.25 (+0.12)	3.55 (+0.03)	2.38 (-0.11)
D w. barrel	770	8.84 (-0.81)	3.26 (-0.26)	2.57 (+0.08)
D no barrel	756	7.13 (± 0)	3.26 (-0.26)	2.57 (+0.08)
E w. barrel	1156	9.07 (-0.58)	3.55 (+0.03)	2.42 (-0.07)
E no barrel	1139	7.23 (+0.10)	3.55 (+0.03)	2.42 (-0.07)

TABLE I

ESTIMATED DIMENSIONS OF THE TARGETS, ESTIMATION ERRORS IN PARENTHESIS. THE TRUE VALUES (FROM CAD MODEL) ARE LENGTH WITH BARREL POINTING FORWARD: 9.65M, LENGTH WITHOUT BARREL 7.13 M, WIDTH 3.52 M AND HEIGHT 2.49 M.

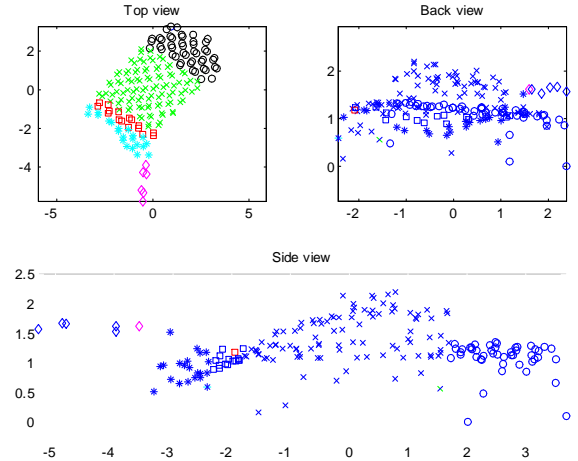


Fig. 7. Result of segmentation of target B in side view, short side segmentation. The data is divided into five segments, where one is identified as a barrel (marked with rhombs). Axes in meters.

removal of the articulated barrel. The height values for both the complete target and the main part are within $\pm 10\%$ of the true values.

D. Target segmentation and node classification

In Figures 7-8 the segmentations of target B are shown. For this target the main parts of a tank were identified in the side view projection. In Figures 9-13 segmentation results of all targets are shown. It can be noted that for target A the turret is not identified. This is probably due to a combination of few samples on the turret and the pitch orientation of barrel. In both side and back/front view the turret and barrel are segmented as one part and thus not identified.

E. Matching

In the information system [2, 3, 15], matching is only performed with models of similar dimensions. To test this approach, matching is performed with several models that contain turret and barrel. In the model library five tanks, four armored personal carriers (APC), one howitzer and one multipurpose vehicle contain these subparts. A common target

²The 3D-ILRIS system from OpTech Inc., see www.optech.on.ca.

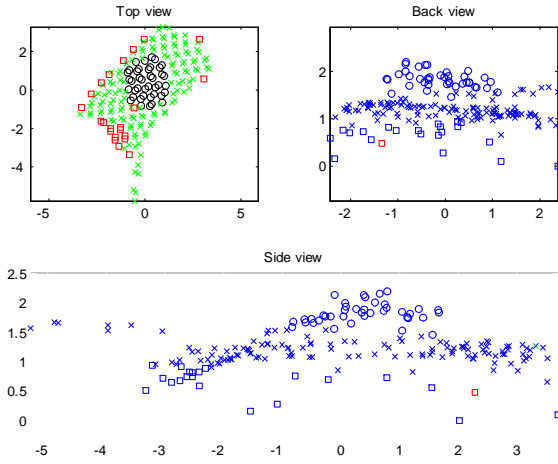


Fig. 8. Result of segmentation of target B in side view, long side side segmentation. The data is divided into three segments, where one is identified as a turret (marked with circles). Axes in meters.

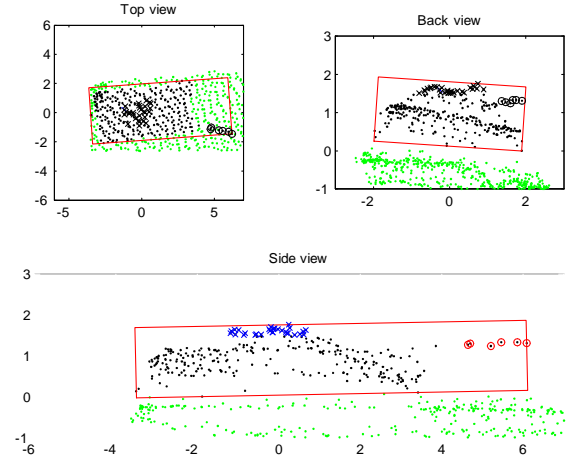


Fig. 11. Result after node classification, target C. The rectangles show the estimated size and orientation. Identified barrel samples are marked with 'o' and turret samples with 'x'. Axes in meters.

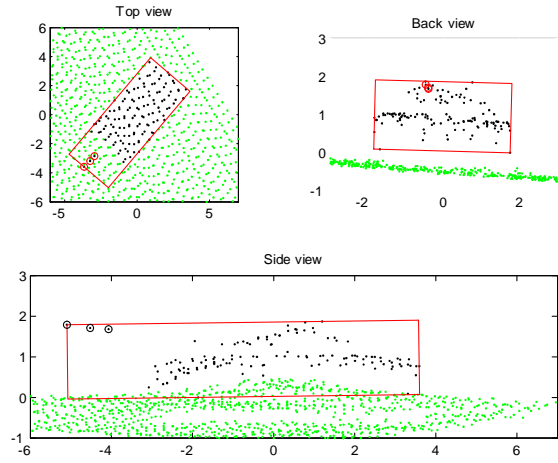


Fig. 9. Result after node classification, target A. The rectangles show the estimated size and orientation. Identified barrel samples are marked with 'o'. Grey marks ground samples and black target samples. Axes in meters.

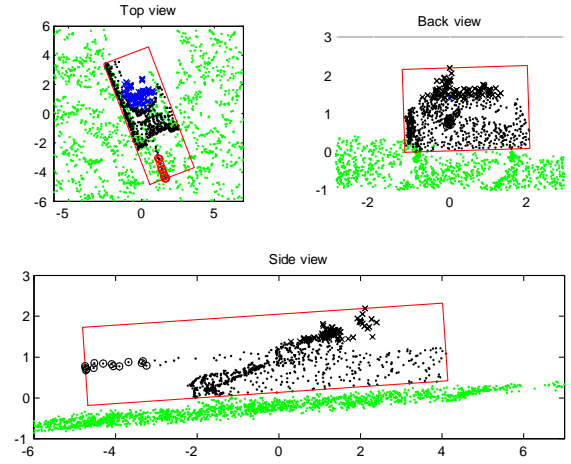


Fig. 12. Result after node classification, target D. The rectangles show the estimated size and orientation. Identified barrel samples are marked with 'o' and turret samples with 'x'. Axes in meters.

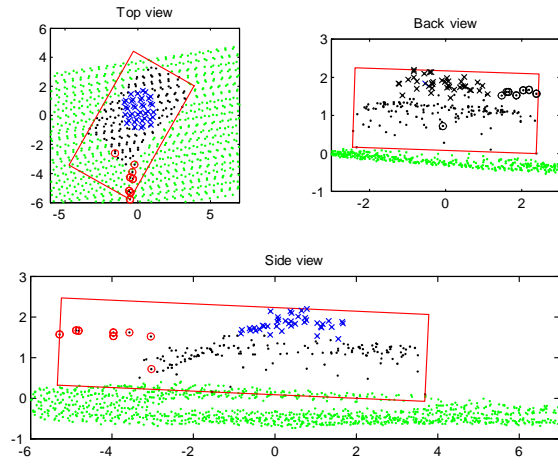


Fig. 10. Result after node classification, target B. The rectangles show the estimated size and orientation. Identified barrel samples are marked with 'o' and turret samples with 'x'. Axes in meters.

model library is used [1], where each model is described by its 3D structure (face/wire-frame models). The highest matching scores (lowest RE values) comes from matching of the T72 data with models of T72 and T80. A T80 has a shape that is very similar to a T72. Good estimates of orientation and articulation give quite good matching results even when parts of the target are missing. Least squares fitting (7) improved the results somewhat, see Table II and Figure 14.

VII. DISCUSSION AND FUTURE WORK

The proposed method assumes that most parts of the object have been registered, which demands that the detection method(s) and the target-ground segmentation are stable. This is the case when targets are placed in open terrain, but not for partly occluded targets. Detection of partly occluded objects needs further research. A laser radar's penetration capability of sparse structures, like vegetation and camouflage nets, is quite

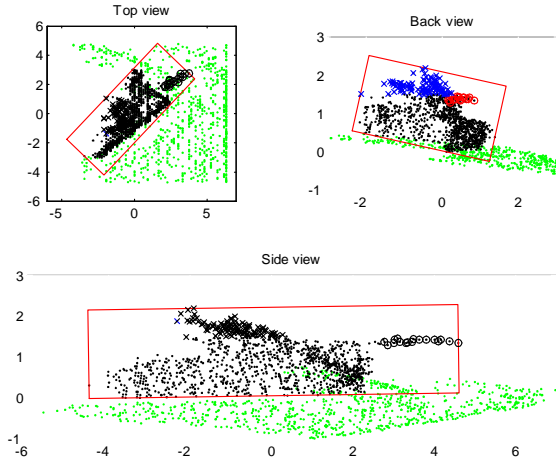


Fig. 13. Result after node classification, target E. The rectangles show the estimated size and orientation. Identified barrel samples are marked with 'o' and turret samples with 'x'. Axes in meters.

Model \ Target	A	B	C	D	E
T72 (tank)	0.0064	0.0075	0.0039	0.0390	0.0263
T80 (tank)	0.0103	0.0087	0.0061	0.0490	0.0368
Leclerc (tank)	0.0108	0.0098	0.0150	0.0423	0.0460
Leopard (tank)	0.0323	0.0289	0.0303	0.0664	0.0698
M1A1 (tank)	0.0261	0.0206	0.0174	0.0662	0.0538
BMP1 (APC)	0.0203	0.0311	0.0236	0.0575	0.0408
BTR80 (APC)	0.0343	0.0504	0.0368	0.0575	0.0492
M2A2 (APC)	0.0317	0.0435	0.0346	0.0643	0.0695
MTLB (APC)	0.0229	0.0385	0.0286	0.0855	0.0568
MI09 (how.)	0.0348	0.0294	0.0600	0.0633	0.1179
Hum-Tow (veh.)	0.1596	0.2768	0.2022	0.2884	0.4149

TABLE II

LEAST SQUARES FIT WITH WIRE-FRAME MODELS, RE VALUES GIVEN. THE THREE LOWEST RE VALUES FOR EACH TARGET ARE IN BOLD FACE.

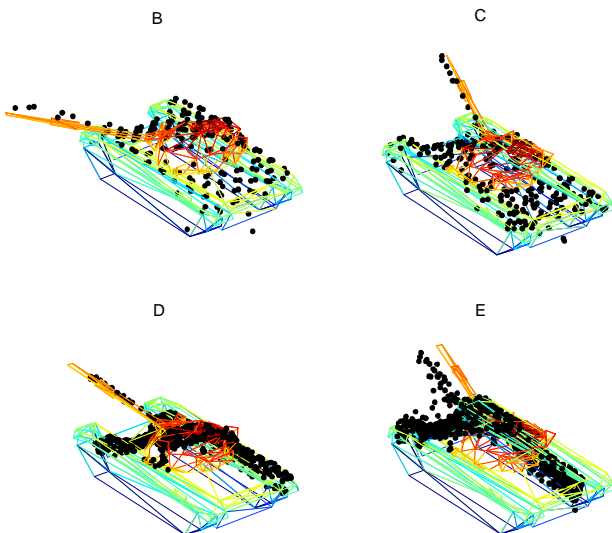


Fig. 14. Matching results, LS fit with the T72 model.

large [27, 30], which is promising from an ATR perspective. As data is a 3D scatter, there is some robustness in the method for objects with missing parts.

The rectangle estimation has quite large MSE and bias for small samples sets. This means that the estimation error of an articulated part (like a barrel) can be quite large. Further, to obtain good estimates of orientation and dimensions at least two sides of the target must be registered. To handle these problems iterative fitting approaches can be applied in the matching step. Application of an iterative fitting approach can also provide a method that can be used in target identification problems. The intensity values can also be used in this step [33].

We consider data as a 3D point scatter rather than a regular grid (a matrix). The reason for this is that 3D imaging systems may not collect data in matrix format in one single frame but from multiple views. Also, the spatial resolution is often rather low and we may introduce further uncertainties in data by resampling to matrix format.

The proposed method for 3D size and orientation estimation is fast but not minimum variance. It can be used to get good starting values for more accurate, iterative methods that use both object and surrounding background data [9]. Alternatively, the 3D size and orientation estimates can be used as starting values for more advanced target recognition methods, e.g., [2] and [21].

In the future, we will study detection methods for partly occluded objects. We will also apply iterative approaches in the matching step, to tackle the problems with unsatisfactory initial fits, small data sets and non-consecutive data sets.

VIII. CONCLUSIONS

In this paper an approach to ground target recognition has been proposed. The method is based on general 3D scattered data and can handle arbitrary perspectives of the target. The object recognition algorithm consists of four steps; estimation of the target's 3D size and orientation, segmentation of the target into parts of approximately rectangular shape, identification of segments that contains the main parts of the target and finally, matching the target with CAD models.

The core in this approach is rectangle estimation. The proposed rectangle estimation method is minimum variance in orientation estimates but the length and width estimates contains bias. The target recognition approach was tested on five data sets of ground targets. The sets contained data from tanks and the number of samples on the targets varied from 129 to 1156 samples. The targets were registered in down-looking, forward-looking and side-looking perspective. The estimated dimensions were in most cases within $\pm 10\%$ of the true values. In the segmentation and node classification, the turret was identified in all five cases while the barrel was identified in four cases. In the matching step, the five targets were correctly recognized and the matching results improved somewhat by least squares fitting.

ACKNOWLEDGMENTS

The authors appreciate that the data sets were made available and preprocessed. We acknowledge TopEye AB, Pierre

Andersson and Tomas Chevalier, FOI Laser Systems.

□

REFERENCES

- [1] Wire-frame/face 3D Models, <http://www.facet3dmodels.com>.
- [2] J. Ahlberg, et al., "Automatic target recognition on a multi-sensor platform", in *Proc. SSAB*, 2003, pp. 93-96.
- [3] J. Ahlberg, et al., "Ground Target Recognition in a Query-Based Multi-Sensor Information System", *Integrated Computer-Aided Engineering Journal*, submitted Dec. 2004.
- [4] P. Andersson et al., "Long Range Gated Viewing and Applications to Automatic Target Recognition", in *Proc. SSAB*, 2003, pp. 89-92.
- [5] K. S. Arun, et al., "Least-squares fitting of two point sets", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no 5, pp. 698-700, Sep. 1987.
- [6] N. Aspert, et al., "Mesh: Measuring error between surface using the Hausdorff distance", in *Proc. ICME*, 2002, vol. 1, pp. 705 - 708.
- [7] L. Brieman, et al., *Classification and regression trees*, Monterey: Wadsworth and Brooks, 1984, Chapter 8.3.
- [8] C. Carlsson, "Vehicle Size and Orientation Estimation Using Geometric Fitting", Dept. of Electrical Eng., Linköping University, Linköping, Sweden, Jun. 2000, Licentiate Thesis no. 840.
- [9] C. Carlsson and M. Millnert, "Vehicle Size and Orientation Estimation using Geometric Fitting", in *Proc. SPIE*, 2001, vol. 4379, pp.412-423.
- [10] C. E. English, et al., "Development of a practical 3D automatic target recognition and pose estimation algorithm", in *Proc. SPIE*, 2004, vol. 5426, pp. 112-123,
- [11] H. Freeman and D. Shapira, "Determining the Minimum-Area Encasing Rectangle for an Arbitrary Closed Curve", *Communications of the ACM*, 1975, Vol. 18, No 7, pp. 409-413.
- [12] J. De Geeter, et al., "A smoothly constrained Kalman filter", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-19, no. 10, pp. 1171 - 1177, Oct. 1997.
- [13] C. Grönwall, et al., "Performance analysis of measurement error regression in direct-detection laser radar imaging", in *Proc. ICASSP*, 2003, vol. VI, pp.545-548
- [14] Internal report = this paper with extra information in appendices.
- [15] T. Horney, et al., "An information system for target recognition", in *Proc. SPIE*, 2004, vol. 5434, pp. 163-175.
- [16] D. Huber, et al., "Parts-based 3D object classification", in *Proc. CVPR*, 2004, vol. 2 , pp. II-82 - II-89.
- [17] E. J. Huising and L. M. Gomes Pereira, "Errors and Accuracy Estimates of Laser Data Acquired by Various Laser Scanning Systems for Topographic Applications", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 53, pp. 245-261, 1998.
- [18] B. Hutchinson, et al., "Simulationbased analysis of range and cross-range resolution requirements for the identification of vehicles in lidar imagery", *Opt. Eng.*, vol. 42, no. 9, pp. 2734-2745, Sep. 2003.
- [19] A. V. Jelalian, *Laser Radar Systems*, Norwood, MA: Artech House, 1992.
- [20] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-21, no. 5, pp. 433 - 449, May 1999.
- [21] L. Klasén, "Image Sequence analysis of Complex Objects. Law Enforcement and Defence Applications", Dept. of Electrical Eng., Linköping University, Linköping, Sweden, Dissertation no. 762, 2002.
- [22] E. L. Lehmann and G. Casella, *Theory of point estimation. 2nd ed.*, New York: Springer Verlag, 2001.
- [23] H. Pirzadeh, "Computation Geometry with the Rotating Calipers", Faculty of Graduate Studies and Research, McGill University, Canada, Master Thesis, Nov. 1999.
- [24] O. Steinvall, et al., "Gated viewing for target detection and recognition", in *Proc. SPIE*, 1999, vol. 3707, pp. 432-448.
- [25] O. Steinvall, et al., "Laser Based 3D imaging. New capabilities for Optical Sensing", FOI Sensor Technology, FOI, Linköping, Sweden, Tech. Rep. FOI-R-0856-SE, Apr. 2003.
- [26] O. K. Steinvall, et al., "3D laser sensing at FOI: overview and a system perspective", in *Proc. SPIE*, 2004, vol. 5412, pp. 294-309.
- [27] O. K. Steinvall, et al., "Characterizing targets and backgrounds for 3D laser radar", presented at SPIE Remote Sensing Europe, London, UK, 2004
- [28] C. K. Toth, et al., "Vehicle recognition from lidar data", in *Proc. of ISPRS working group III/3 workshop*, 2003, pp. 162-166.
- [29] G. Toussaint, "Solving Geometric Problems with the Rotating Calipers", in *Proc. IEEE MELECON*, 1983.
- [30] A. N. Vasile and R. Marino, "Pose-independent automatic target detection and recognition using 3D LADAR data", in *Proc. SPIE*, 2004, vol. 5426, pp. 67-83.
- [31] J. G. Verly and R. L. Delanoy, "Model-based automatic target recognition (ATR) system for forwardlooking groundbased and airborne imaging laser radars (LADAR)", *Proc. IEEE.*, 1996, vol. 84, no. 2, pp. 126 - 163.
- [32] S. Vinson and L. D. Cohen, "Multiple Rectangle Model for Buildings Segmentation and 3D Scene Reconstruction", in *Proc. ICPR* , 2002, pp. 623-626.
- [33] M. R. Wellfare and K. Norris-Zachery , "Characterization of articulated vehicles using lidar seekers", in *Proc. SPIE*, 1997, vol. 3065, pp. 244-254.
- [34] T. Yano, et al., "Vehicle identification technique using active laser radar system", in *Proc. MFI*, 2003, pp.275 - 280.
- [35] Q. Zheng, et al., "Model-based Target Recognition in Pulsed Ladar Imagery", *IEEE Trans. Image Processing*, 2001, vol. 10, no. 4, pp. 565-572, Apr. 2001.
- [36] D. Q. Zhu and C.-C. Chu, "Characterization of irregularly shaped bodies", in *Proc. SPIE*, 1995, vol. 2466, pp. 17-22.



Christina Grönwall is a PhD student at... Fredrik Gustafsson is professor... Mille Millnert is professor...