3D content-based model matching using geometric features

Christina Grönwall, Fredrik Gustafsson

Division of Automatic Control
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: http://www.control.isy.liu.se
E-mail: stina@isy.liu.se, fredrik@isy.liu.se

31st January 2006

Report no.: LiTH-ISY-R-2726
Submitted to JASP

Abstract

We present an approach that utilizes ef cient geometric feature extraction and a matching method that takes articulation into account. It is primarily applicable for man-made objects. First the object is analyzed to extract geometric features, dimensions and rotation are estimated and typical parts, so-called functional parts, are identi ed. Examples of functional parts are a box's lid, a building's chimney, or a battle tank's barrel. We assume a model library with full annotation. The geometric features are matched with the model descriptors, to gain fast and early rejection of non-relevant models. After this pruning the object is matched with relevant, usually few, library models. We propose a sequential matching, where the number of functional parts increases in each iteration. The division into parts increases the possibility for correct matching result when several similar models are available. The approach is exemplified with an vehicle recognition application, where some vehicles have functional parts.

Keywords: geometric features, early rejection, content-based matching, least squares

3D content-based model matching using geometric features

Christina Grönwall

Dept. Laser Systems
FOI, Linköping, Sweden
christina.gronwall@foi.se

Fredrik Gustafsson

Dept. Electrical Engineering
Linköpings Universitet
Linköping, Sweden

We present an approach that utilizes efficient geometric feature extraction and a matching method that takes articulation into account. It is primarily applicable for man-made objects. First the object is analyzed to extract geometric features, dimensions and rotation are estimated and typical parts, so-called functional parts, are identified. Examples of functional parts are a box's lid, a building's chimney, or a battle tank's barrel. We assume a model library with full annotation. The geometric features are matched with the model descriptors, to gain fast and early rejection of non-relevant models. After this pruning the object is matched with relevant, usually few, library models. We propose a sequential matching, where the number of functional parts increases in each iteration. The division into parts increases the possibility for correct matching result when several similar models are available. The approach is exemplified with an vehicle recognition application, where some vehicles have functional parts.

Key words: geometric features, early rejection, content-based matching, least squares

1 Introduction

A common problem in computer vision is to find whether an unknown man-made object matches a model in a model library. The object is registered with a sensor system and the data set describing the object is contaminated with noise and other uncertainties. The object's pose cannot be controlled during the registration and the object may contain articulated parts. For example, vehicles can easily change appearance by opening of a door, adding of load etc., and boxes may have the lid in various positions. Buildings may be of complex shape and contain a chimney or a tower. In these cases, it is important that the object and model are placed in the same pose before matching to avoid that several poses have to be tested for each object-model combination. Applications where this occur are in traffic monitoring, industrial applications, urban structure modelling and military target recognition.

One approach is to analyze data before querying the model library. This can both reduce the number of poses that need to be tested for each object-model combination and the number of models that are relevant for matching. We present an approach, primarily applicable for man-made objects, that utilizes efficient geometric feature extraction of the object and a matching method that takes articulation into account.

First the object is analyzed to extract geometric features. Its dimensions (length, width and height) and rotation are estimated and typical parts, here called *functional parts*, are identified. Examples of functional parts are a box's lid, a building's chimney, a vehicle's door, or the barrel of a battle tank. We assume a model library with full annotation. The geometric features are matched with the model descriptors, to have fast and early rejection of non-relevant models. In the next step the object is matched with relevant library models, usually a low number of models. The estimation of dimensions and rotations of each part simplifies the model matching as the degrees of freedom reduce. The distance between the object and the model is minimized in least squares (LS) sense. Functional parts may be subject to articulation and this is taken into account in the matching. We propose a sequential matching, where the number of functional parts increases in each iteration. The division into parts increases the possibility for correct matching result, when several similar models are available.

In this paper, the object data is a 3D point scatter retrieved with a sensor system. The paper focuses on the interaction with the model data base, i.e., the model selection using model descriptors and model matching of articulated objects. The extraction of geometric features is described shortly. In Section 2 we survey previous work. In Section 3, we describe an approach to least squares fitting of two point scatters including articulation and in Section 4 this is applied to fitting of a point scatter with a wire-frame model. The proposed approach is applied to vehicle recognition in Section 5. A discussion is found in Section 6 and in Section 7 the paper is summarized and concluded.

2 Previous Work

We review earlier work concerning size and orientation estimation, segmentation of complex shapes and functional part identification, LS fitting of 3D point scatters and, finally, matching of wire-frame models with 3D point scatters.

The orientation of objects, registered in 2D by passive imaging or projections of 3D data, can be estimated by rectangle fitting. An iterative approach is proposed in [3] and in [2, 8, 13, 16] non-iterative approaches are described. The objects being characterized are asteroids [16], buildings [13] and vehicles [2, 8], respectively. Eigenvalue calculations are used to estimate the orientation of the object [13, 16]. After that, a rectangle that bounds the object samples [16] or is optimal in second order moment [13] is calculated. In [2], a rectangle that bounds the object data is estimated and in [8] the orientation is given by the direction of point pairs and the dimensions by density calculations.

Segmentation of complex shapes into rectangular functional parts is treated in [5, 13]. The segmentation [13] is based on interval division, maximizing of each rectangle, and minimization of overlap between rectangles. In [5], the segmentation is based on search for rectangles similar to the searched functional parts. The search is constrained on the data noise and number of samples.

The problem with articulated functional parts of vehicles and vehicle recognition is treated in a general way in [5, 8, 10, 12, 14]. In [5, 8, 10] the functional parts are identified and the model is oriented according to the estimated orientations of the functional parts. In [14], each model is stored with several articulations and fast matching (geometric hashing) is used. The functional parts are identified using combinations of geometric rules and cuboid fitting [8], rectangle fitting [5], a hypothesis testing [10], or division into surface patches [14]. The dimension estimate and functional part identification is used for retrieval of relevant models in the model library [14]. In [12], the functional parts are identified using spin image technique where the functional parts are stored in a library of spin images.

The problem of relating two three-dimensional data sets under the presence of noise has been subject to much attention [1, 4, 6, 9, 11]. The "absolute orientation problem" of finding the least squares solution of a rotation and translation rigid body transformation is addressed in [1]. A refinement of [1], that avoids the reflections problems in [1] and also handles scaling transformation, is presented in [11]. In [9], a method based on a mixed least squares - total least squares solution is proposed, assuming noise to be present in both point sets to be fitted. The methods in [1, 9, 11] all have the disadvantage of requiring both data sets to have the same number of points, and that the point correspondence between the sets is known a priori. By contrast, [4] solve the "approximate geometric pattern matching problem", based on approximately minimizing the directed Hausdorff distance from the pattern set to the background set using rigid body transformations. The points sets can be of different sizes and point correspondence is not assumed. An approach to fit a 3D point scatter to a face model using projections of the point scatter is presented in [6].

We will extend the work in [6] to a general case with several functional parts and a model library containing several models. The initial estimate of the object's dimensions, rotation and identification of functional parts will follow [5].

3 Matching articulated point sets

We first present the global LS fitting problem for two point scatters with point correspondences, earlier presented in [1, 9, 11]. The problem is then extended to modular LS fitting where the object's articulation is treated. Modular fitting for two parts is earlier presented in [6].

3.1 Global LS fitting

Assume that we have two 3D point sets $P = (p_1, p_2, ..., p_N)^t$ ($N \times 3$) and $Q = (q_1, q_2, ..., q_N)^t$ ($N \times 3$) that are related by $Q = (PR + T) + E$, where R is a rotation matrix, T is a translation vector and $E = (e_1, e_2, ..., e_N)^t$ is noise. The (noise free) model is represented by Q and the noisy object by P. We assume that $e_i, i = 1, ...N$, has zero mean, equal variance and that the elements in E are independently and identically distributed (iid). We will find R and T that in least squares sense minimize the estimation error

$$V = \min_{R,T} \quad \|Q - (PR + T)\|_2^2 \qquad (1a)$$

$$\text{subject to} \quad RR^t = I \qquad (1b)$$

$$\det R = 1, \qquad (1c)$$

where V is the mean square error (MSE), $\|\cdot\|_2$ is the Euclidian norm, $(\cdot)^t$ is matrix transpose, and I is the identity matrix. Define the regressor ϕ and the parameter vector θ

$$\begin{aligned} \phi^t &= \begin{pmatrix} P & \mathbf{1}_{N\times 1} \end{pmatrix} \\ \theta &= \begin{pmatrix} R^t & T \end{pmatrix}^t, \end{aligned}$$

where $\mathbf{1}_{N\times 1} = (1, 1, ..., 1)^t$. The minimization problem (1) can then be written

$$V = \min_{R,T} \quad \left\|Q - \phi^t\theta\right\|_2^2 \qquad (2a)$$

$$\text{subject to} \quad RR^t = I \qquad (2b)$$

$$\det R = 1. \qquad (2c)$$

3.2 Modular LS fitting

Suppose that we have a method to identify the functional parts of the object, that we can divide the data set into these parts and that it is possible to do the same division with the model data. Let us assume that the object has a main part and on that part, an articulated part is placed and on that second part a third, articulated part is placed, etc. The general fitting problem, with model and object data divided into J parts, can be expressed

$$\begin{aligned} Q_1 &= P_1 R_1 + T_1 + E_1 \\ Q_2 &= (P_2 R_1 + T_1) R_2 + T_2 + E_2 \\ &\vdots \\ Q_J &= P_J R_J + T_J + E_J \\ R_J &= R_1 R_2 \cdots R_{J-2} R_{J-1} \\ T_J &= T_{J-2} R_{J-1} + T_{J-1}, \end{aligned}$$

where the elements in $E = [\ E_1 \quad E_2 \quad \cdots \quad E_J\ ]^t$ are iid with zero mean and equal variance. Part P_j contains N_j samples, $N_1 + N_2 + \cdots + N_J = N$. Define ϕ^t and θ as

$$\begin{aligned} \phi^t &= \begin{pmatrix} P_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & P_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & P_J & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix}, \\ \theta &= \begin{pmatrix} R_1^t & \cdots & R_J^t & T_1 & \cdots & T_J \end{pmatrix}^t, \end{aligned}$$

and the minimization problem can again be written

$$V = \min_{R_1,\cdots,R_J,T_1,\cdots,T_J} \left\| Q - \phi^t \theta \right\|_2^2 \tag{3a}$$

$$\text{subject to} \quad R_j R_j^t = I, \quad j = 1, ..., J \tag{3b}$$

$$\det R_j = 1, \quad j = 1, ..., J \tag{3c}$$

4 Fitting point set with face model

In most cases, two point sets with point correspondence are not available [4, 6]. Instead we have a point scatter describing the object and the model is a face model, denoted $\mathcal{M}$. It is then possible to fit the object samples with their projections on the closest facets. After initial projection and fit, a new projection and fit can be performed and the iterations continue until the MSE varies only slightly.

Define P as the point set describing the object and Q as the point set describing the model, where Q is the projection of the elements in P to the closest model facet, i.e.,

$$P = \text{Proj}\left(Q|\mathcal{M}\right).$$

If the orthogonal projection of an element in P is not on a facet, the projected sample is set to the closest facet edge. An outlier rejection is also necessary; elements in Q that have too long distances to the corresponding samples in P will be rejected. The outlier distance is user defined. The iterative algorithm for fitting of a 3D point set with a face model, when the number of functional parts is fixed, is summarized in Algorithm 1.

Algorithm 1 (Fitting of point scatter and facets)

1. Estimate the object's orientation, including orientation of functional parts, and place the model in similar position. This gives the initial rotations $R_1^0, \cdots, R_J^0$ and translations $T_1^0, \cdots, T_J^0$.
2. For iteration k, calculate the projected samples of $P_j^k, j = 1, ..., J$ on the model $\mathcal{M}$, $Q_j^k =$Proj$\left(P_j^k|\mathcal{M}\right)$, to get point correspondences.
3. Reject outlier elements in Q_j^k and their corresponding elements in $P_j^k, j = 1, ..., J$.
4. Estimate rotations $R_1^k, \cdots, R_J^k$ and translations $T_1^k, \cdots, T_J^k$ and calculate the MSE of the estimation error, $V^k\left(\mathcal{M}\right)$, see (3).
5. If $\tau < V^k\left(\mathcal{M}\right)/V^{k-1}\left(\mathcal{M}\right) \leq 1$, terminate. Otherwise, continue to iteration $k+1$. The threshold τ is user-defined.

In the general case we have a model library with M face models $\mathcal{M}^m, m = 1, ..., M$. The problem of finding the correct model for the object can be expressed

$$\arg\min_{\mathcal{M}^m} V\left(\mathcal{M}^m, J\right). \tag{4}$$

This is a nonlinear optimization problem, due to the projections, which means that good initialization is necessary. The initialization gives us the initial rotation R_0 and the initial translation T_0. The separation into functional parts is controlled by the model with which the object is matched. If the model contains one or more parts, the matching will be performed with the parts articulated according to estimates from data.

4.1 Penalty on number of functional parts

There is a trade-off between the number of samples, uncertainties in measurement data and the model complexity. The fundamental cost of parameters can be expressed by the final prediction-error criterion (FPE) [7]:

$$F\left(\mathcal{M}, J\right) = \frac{1 + d_{\mathcal{M}}/N}{1 - d_{\mathcal{M}}/N} V\left(\mathcal{M}, J\right), \tag{5}$$

where the object and model is divided into J parts and $d_{\mathcal{M}} = \dim \theta^{\mathcal{M}}$. Note the (obvious) criterion of number of samples: $N > d_{\mathcal{M}}$, which prevents over-fitting. $F(\mathcal{M}, J)$ decreases as N increases. A lower fitting error $V(\mathcal{M}, J)$ gives a lower $F(\mathcal{M}, J)$, thus, good initial fit of Q and P and a low noise level in the measurement data P gives a lower $F(\mathcal{M}, J)$.

4.2 Modular matching algorithm

We propose a sequential approach for inclusion of functional parts. The separation of the object and the model into main part (P_1, Q_1) and subparts $(P_2, ..., P_J, Q_2, ..., Q_J)$ is application dependent. This sequential fitting can be used for selection of correct model when several similar models are available. The fitting at iteration j gives good initialization for iteration $j + 1$. The algorithm for model matching for articulated objects is summarized in Algorithm 2.

Algorithm 2 (General modular matching)

1. Estimate the object's orientation, including orientation of functional parts, and place the model in similar position.
2. For iteration $j = 1, 2, ...$, separate the object data set and the model into the main parts and subparts according to definitions applicable for the current problem.
3. Calculate the best fit for j articulations according to (4) using Algorithm 1.
4. Calculate the FPE, $F(\mathcal{M}, j)$, according to (5).
5. If $F(\mathcal{M}, j) \geq F(\mathcal{M}, j-1)$, terminate. Otherwise, continue to iteration $j + 1$.
6. If $F(\mathcal{M}, j) \geq F(\mathcal{M}, j-1)$, the correct division into parts is found for $J = j - 1$. The estimates of rotation and translation for each parts are found.

5 Case study: vehicle recognition

5.1 Introduction

We exemplify the approach for an application on recognition of military vehicles, where some have functional parts. In this case we have a 3D registration of a vehicle and the problem is to recognize the correct model in the data base. In the data base, the models are stored as wire-frame face models, which come from CAD libraries or are generated from earlier registrations. The number of data base models selected for matching with the object, i.e., the probe, is reduced sequentially, see Algorithm 3.

Algorithm 3 (Model association)

1. Estimate the object's dimensions and orientations in 3D. Select models of the library where the dimensions are correct, within a tolerance.
2. Identify the object's functional parts, their dimensions and orientations. Select the models from the previous step that contain those functional parts.
3. For each model remaining after step 2, perform matching using Algorithm 2. First global matching will be performed ($j = 1$) and the modular matching up to $j = J$, where there are $J - 1$ functional parts.
4. The model that fits best to the object is the model with lowest value on $V(\mathcal{M}, j)$ and where $F(\mathcal{M}, j) < F(\mathcal{M}, j-1)$.

The method used for dimension and orientation estimate and functional parts identification is described shortly in Section 5.3 and further in [5]. The results of steps 1–2 in Algorithm 3, the selection of relevant models, are shown in Section 5.4 and the results of the model matching (steps 3–4 in Algorithm 3) are shown in Section 5.5.

The goal with identification and fitting of functional parts for vehicles is to simplify the model matching. If the parts are identified we can match with libraries regardless of the relative position of the functional parts. Different configurations of a vehicle can be handled in a structural way. If the functional parts of a tank (the barrel and turret) can be extracted, the hypothesis that the object is a tank is strengthened. When the object's functional parts can be identified, the recognition can be simplified as the degrees of freedom reduce. Further, for a tank the orientation of the barrel indicates the tank's intention, which can be useful in security or military applications.

5.2 Data sets and data base contents

We will use data sets from four types of vehicles to illustrate the early rejection and the efficiency of modular matching. In this case the objects are registered with a scanning laser radar, although other 3D registering systems are applicable. The data set consists of two tanks (one T72 and one M60), one anti-aircraft gun (ZSU23) and one armored personal carrier (MTLB). The estimated dimensions and the identified functional parts are shown in Table 1. The uncertainties in dimension estimates are based on estimates of measurement noise variance, the number of samples and the performance of the dimension and orientation estimator.

In the model library, i.e., the data base, there are 16 models of common tanks, an anti-aircraft gun, howitzers, armored personal carriers and smaller personal carriers. All models are represented by wire-frame face models of low resolution, typically there are 500-800 facets. We assume that the model's properties are stored for all models, i.e.,

Object	Dimension estimates			Part id.	
	Length	Width	Height	Turret	Barrel
T72	9.07 ± 0.60	3.55 ± 0.75	2.42 ± 0.75	Y	Y
M60	9.52 ± 0.90	3.56 ± 1.35	3.09 ± 1.35	Y	Y
ZSU23	5.53 ± 0.90	3.00 ± 1.35	3.37 ± 1.35	Y	N
MTLB	6.44 ± 0.60	3.07 ± 0.75	2.00 ± 0.75	N	N

Table 1: Estimated dimensions (in meters) and identified functional parts of the four objects. The estimated three standard deviation distances are given as dimension uncertainties.

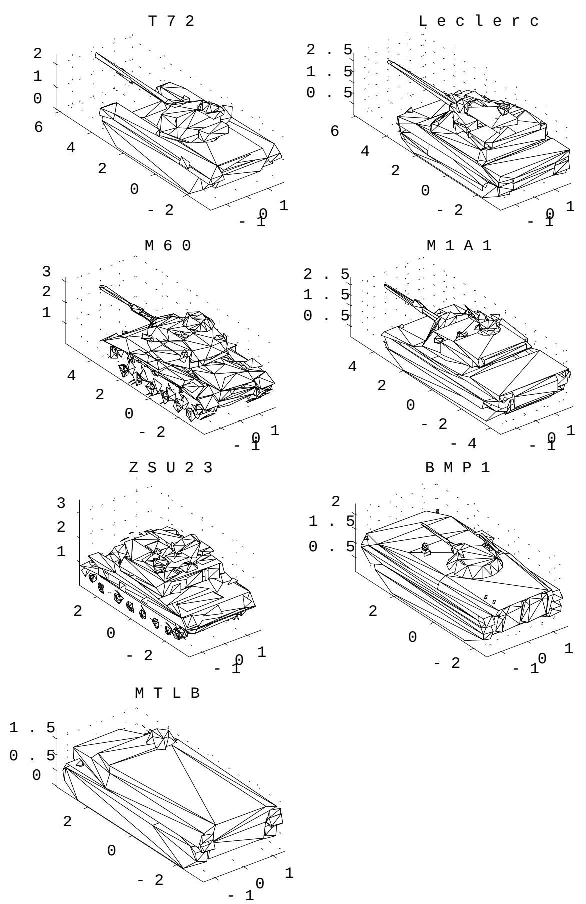

Figure 1: The models in low resolution, axes in meters.

we assume full annotation. The descriptors are dimensions (length, width and height), allowed orientations, and presence of functional parts (barrel and turret). The T72 and M60 models contain both turret and barrel, the ZSU23 model contains a turret but not a barrel and the MTLB model has neither a turret nor a barrel, see Figure 1.

5.3 Initialization and functional part identification

Approached for separation of complex objects and functional part identification have been presented earlier in the literature [5, 8, 10, 12, 13, 14]. The method used in this case study is further described in [5]. The method handles general 3D scattered data. It takes advantage of the 3D structure and that the dimensions are known in laser radar data. The estimation of initial position and segmentation into functional parts is based on the assumption that man-made objects, like vehicles and buildings, in certain projections are of rectangular shape. A man-made object of complex shape can be decomposed into a set of rectangles and in some views the rectangles will describe the functional parts of the object. When an object 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 object is placed in a certain pose relative to the sensor and we cannot assume any certain orientation or articulation of the object.

The initialization algorithm consists of three steps; 1) Estimate the object's 3D size and orientation using the rectangle estimation. 2) Segment the object data into parts of approximately rectangular shape. The functional parts can be found in some of the rectangles. 3) Identify the functional parts by simple geometric comparisons and estimate their dimensions and orientations. An example of initialization and identification of functional parts for the T72 object is shown in Figure 2.

5.4 Model pruning using descriptor match

We first compare the estimated features of the objects with the model's descriptors. In Table 2 the impact of comparing estimated object features with models' descriptors is shown. The number of models subject to model matching reduces vastly in this step. The remaining model for the T72 object is the T72 model, for the M60 object the remaining models are M60, T72, Leclerc and M1A1. For the ZSU23 object the remaining models are ZSU23 and T72 and for the MTLB object the remaining models are MTLB and BMP1.

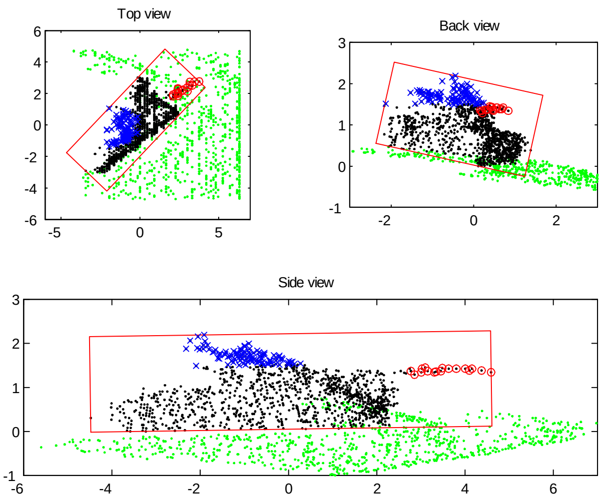

Figure 2: Initialization of the T72 object. The rectangles show the dimensions and orientation estimate. The identified functional parts are turret (x) and barrel (o). Axes in meters.

Object	No. of	correct	dim.	Incl.	part id.
	≥ 1	≥ 2	3	≥ 2	≥ 3
T72	14	3	2	2	1
M60	15	13	4	7	4
ZSU23	16	14	5	7	2
MTLB	13	6	2	-	-

Table 2: The number of remaining models after comparision of object features and model descriptors.

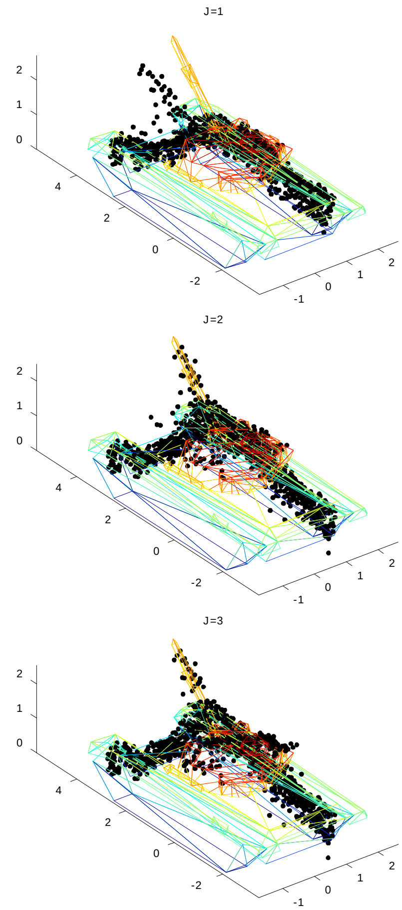

Figure 3: Matching sequence for the T72 object, axes in meters. Top: J=1, MSE=6.73, 15 iterations in Algorithm 1, FPE=6.80. Middle: J=2, MSE=4.41, 8 iterations in Algorithm 1, FPE=4.51. Bottom: J=3, MSE=4.21, 3 iterations in Algorithm 1, FPE=4.34.

5.5 Modular matching

We continue with matching for the M60, ZSU23 and MTLB objects using Algorithm 2, as more than one model turned out to be relevant for these objects. In Algorithm 1 (called by Algorithm 2), we set the outlier distance to 1.5 meters and threshold $\tau = 0.99$. In the first iteration of Algorithm 2 global fitting is performed and in the second and third iteration modular fitting is performed. In the second iteration we set the chassis to the main part and the articulated part contains turret and barrel (when applicable). For the tank objects there is a third iteration, where the chassis is defined as the main part, the turret is the first articulated part and the barrel is the second articulated part. The fitting sequence for the T72 object is shown in Figure 3.

For M60, ZSU23 and MTLB objects we show the iterations of Algorithm 2 in Tables 3-5. The tables show the ben-

Object	Model			
	M60	T72	Leclerc	M1A1
J=1, V	3.61 (20)	3.96 (13)	3.48 (10)	4.45 (8)
J=1, F	3.76	4.12	3.62	4.62
J=2, V	3.25 (3)	3.69 (3)	3.27 (3)	4.03 (3)
J=2, F	3.51	3.99	3.48	4.36
J=3, V	3.00 (3)	3.48 (2)	3.11 (3)	3.87 (2)
J=3, F	3.39	3.92	3.50	4.35

Table 3: Match results for the M60 object, separation into J parts. MSE, with the number of iterations in Algorithm 1 in parenthesis, and FPE are shown.

Object	Model	
	ZSU23	T72
J=1, V	4.29 (15)	7.77 (13)
J=1, F	4.47	8.09
J=2, V	4.00 (7)	9.53 (4)
J=2, F	4.33	10.31

Table 4: Match results for the ZSU23 object, object separation into J parts. MSE, with the number of iterations in Algorithm 1 in parenthesis, and FPE are shown.

efit of sequential matching and of the MSE value to identify the correct model. Further, the tables show that the FPE indicates the level articulation that is applicable in the fitting. The level of fit is represented by the MSE (V) and the level of split into part by the FPE (F). In Table 3-5 the values are normalized [5].

The M60 object is matched with four models that are fairly similar in shape, see Table 3. For the global match ($J = 1$) the Leclerc is the best fit and the M60 is second best. For modular match ($J = 2, 3$), the hypothesis that the object is a M60 is strengthened. For the other models there is less, or none, improvement when the match is performed for three parts. Even if the MSE is lower, the FPE value indicates that the current split of model and data is expensive.

In Table 4, the match results for the ZSU23 object are shown. In this case the hypothesis that the object is a ZSU23 is strengthened in the modular match, while the T72 hypothesis is attenuated. When the ZSU23 object is matched with the T72 model we can clearly see that the FPE value indicates the level articulation that is applicable. For the MTLB object a correct match was received, see Table 5.

Object	Model	
	MTLB	BMP1
J=1, V	1.54 (14)	3.40 (13)
J=1, F	1.57	3.46

Table 5: Match results for the MTLB object, global fitting (J=1). MSE, with the number of iterations in Algorithm 1 in parenthesis, and FPE are shown.

6 Discussion

The proposed approach is based on a quadratic minimum criteria, which makes it sensitive to outliers. Some robustness to outliers exists due to the outlier rejection in Algorithm 1. However, the approach presented here gives best results when outliers are not present in object data. This implies careful preprocessing of data where most outliers are rejected.

We use the threshold τ to decrease the number of iterations in Algorithm 1. This can be replaced by the expression $V^k(\mathcal{M}) \geq V^{k-1}(\mathcal{M})$, with the risk of introducing many iterations with small fitting adjustments.

The definitions of turret and barrel used in this case study are general and apply for several vehicle types in the data base. The geometric rules are described in [5]. These definitions can be replaced by type-specific definitions retrieved from the models, this will increase the number of matches in the first step.

The first step in this approach, the model pruning based on estimated object features, is fast. The estimation of dimensions and rotations, and the identification of functional parts is also fast [5]. The modular matching step is more time consuming and it is necessary to have early rejection of uninteresting models.

In the sequential fine tuning, the level of detail in the models can increase (i.e., increase of the number of facets) as the model is fine tuned. For example, after a few iterations in Algorithm 2, the low-resolution model is replaced by a high-resolution model.

Future work includes to compare this LS fitting approach with the more common iterative closest point (ICP) fitting method.

7 Conclusions

We presented an approach, primary applicable for man-made objects, that contains efficient geometric feature extraction of the object and matching that takes articulation into account. First the object's dimensions and rotation are estimated and the functional parts are identified. This careful analysis is used for early rejection of non-relevant model. In the next step the object is matched with relevant library models, usually only a few models remain. The distance between the object samples and the model is minimized in least squares sense. If functional parts have been identified, this and their articulations are taken into account in the matching. We proposed a sequential matching, where the number of functional parts increases in each iteration.

The impact of comparing estimated object features with models descriptors is shown. The number of models subject to model matching reduces vastly in this step.

To match with articulation, good initial fit is needed, this is acquired by the sequential process. The division into parts increases the possibility for correct matching result, when several similar models are available. After a few iterations the correct hypothesis is strengthened. For the other models there is less, or none, improvement when the match is performed for three parts.

The first step in this approach, the model pruning based on estimated object features, is fast. The estimation of dimensions and rotations, and the identification of functional parts is also fast. The matching is more time consuming, it is necessary to have early rejection of uninteresting models.

References

[1] K. S. Arun, T. S. Huang and S. D. Blostein, "Least-squares fitting of two point sets", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-9, no 5, pp. 698-700, 1987.

[2] C. Carlsson and M. Millnert, "Vehicle Size and Orientation Estimation using Geometric Fitting", in Proc. SPIE, 2001, vol. 4379, pp.412–423.

[3] J. De Geeter, H. Van Brussel, J. De Schutter and M. Decréton, "A smoothly constrained Kalman filter", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-19, no. 10, pp. 1171 - 1177, Oct. 1997.

[4] M. T. Goodrich, J. S. B. Mitchell, M. W. Orletsky, "Approximate Geometric Pattern Matching Under Rigid Motion", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-21, no. 4, pp. 371-379, 1999.

[5] C. Grönwall, F. Gustafsson, M. Millnert, "Ground target recognition using rectangle estimation", Dept. of Electrical Eng., Linköpings Universitet, Linköping, Sweden, March 2005, Tech. Report no. LiTH-ISY-R-2684. Submitted to Trans. IEEE Image Proc.

[6] C. Grönwall, P. Andersson, F. Gustafsson, "Least squares fitting of articulated objects", Workshop on advanced 3D imaging for safety and security, San Diego, CA, USA, 25 June 2005, In Proc. CVPR, 2005.

[7] L. Ljung: System Identification - Theory For the User, 2nd ed, PTR Prentice Hall, Upper Saddle River, N.J., 1999.

[8] J. Neulist and W. Armbruster, "Segmentation, classification and pose estimation of military vehicles in low resolution laser radar images", in Proc. SPIE, 2005, vol. 5791, pp. 218-225.

[9] J. A. Ramos and E. I. Verriest, "Total Least Squares Fitting of Two Point Sets in m-D", Proc. 36th conf. Decision and Control, pp. 5048-5053, Dec. 1997.

[10] E. Sobel, J. Douglas and G. Ettinger, "3D LADAR ATR based on recognition by parts", in Proc. SPIE, 2003, vol. 5094, pp. 29-40.

[11] S. Umeyama, "Least-Squares Estimation of Transformation Parameters Between Two Point Patterns", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-13, no. 4, pp. 376-380, 1991.

[12] 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.

[13] S. Vinson and L. D. Cohen, "Multiple Rectangle Model for Buildings Segmentation and 3D Scene Reconstruction", in Proc. ICPR , 2002, pp. 623-626.

[14] M. R. Wellfare and K. Norris-Zachery , "Characterization of articulated vehicles using ladar seekers", in Proc. SPIE, 1997, vol. 3065, pp. 244-254.

[15] Q. Zheng, S. Z. Der and H. I. Mahmoud., "Model-based Target Recognition in Pulsed Ladar Imagery", IEEE Trans. Image Processing, 2001, vol. 10, no. 4, pp. 565-572, Apr. 2001.

[16] D. Q. Zhu and C.-C. Chu, "Characterization of irregularly shaped bodies",in Proc. SPIE, 1995, vol. 2466, pp. 17-22.

Avdelning, Institution
Division, Department

Division of Automatic Control
Department of Electrical Engineering

Datum
Date

2006-01-31

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ ____________

Rapporttyp
Report category

☐ Licentiatavhandling
☐ Examensarbete
☐ C-uppsats
☐ D-uppsats
☒ Övrig rapport
☐ ____________

ISBN
—

ISRN
—

Serietitel och serienummer
Title of series, numbering

ISSN
1400-3902

LiTH-ISY-R-2726

URL för elektronisk version

http://www.control.isy.liu.se

Titel
Title

3D content-based model matching using geometric features

Författare
Author

Christina Grönwall, Fredrik Gustafsson

Sammanfattning
Abstract

We present an approach that utilizes ef cient geometric feature extraction and a matching method that takes articulation into account. It is primarily applicable for man-made objects. First the object is analyzed to extract geometric features, dimensions and rotation are estimated and typical parts, so-called functional parts, are identi ed. Examples of functional parts are a box's lid, a building's chimney, or a battle tank's barrel. We assume a model library with full annotation. The geometric features are matched with the model descriptors, to gain fast and early rejection of non-relevant models. After this pruning the object is matched with relevant, usually few, library models. We propose a sequential matching, where the number of functional parts increases in each iteration. The division into parts increases the possibility for correct matching result when several similar models are available. The approach is exemplified with an vehicle recognition application, where some vehicles have functional parts.

Nyckelord
Keywords

geometric features, early rejection, content-based matching, least squares