This is the accepted version of a paper presented at 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, September 28-October 2, 2015.

Citation for the original published paper:
A Principle of Minimum Translation Search Approach for Object Pose Refinement.
http://dx.doi.org/10.1109/IROS.2015.7353776

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-47942
A Principle of Minimum Translation Search Approach for Object Pose Refinement

Rasoul Mojtahedzadeh and Achim J. Lilienthal
Center of Applied Autonomous Sensor Systems (AASS), Örebro University, Sweden

Abstract—The state-of-the-art object pose estimation approaches represent the set of detected poses together with corresponding uncertainty. The inaccurate noisy poses may result in a configuration of overlapping objects especially in cluttered environments. Under a rigid body assumption the inter-penetrations between pairs of objects are geometrically inconsistent. In this paper, we propose the principle of minimum translation search, PROMTS, to find an inter-penetration-free configuration of the initially detected objects. The target application is to automate the task of unloading shipping containers, where a geometrically consistent configuration of objects is required for high level reasoning and manipulation. We find that the proposed approach to resolve geometrical inconsistencies improves the overall pose estimation accuracy. We examine the utility of two selected search methods: A-star and Depth-Limited search. The performance of the search algorithms are tested on data sets generated in simulation and from real-world scenarios. The results show overall improvement of the estimated poses and suggest that depth-limited search presents the best overall performance.

I. INTRODUCTION

The problem of object detection and pose estimation is an essential task in robotic systems. In real-world applications of robotic systems a scene composed of cluttered objects is a challenging environment for object detection and pose estimation algorithms. An example of a highly cluttered environment is a shipping container filled with goods (e.g., carton boxes) that could come in random configurations (see a few examples in Fig. 1). A complete and accurate estimation of the poses of the objects is of great importance especially for high level reasoning [1] and motion planning for manipulation of the objects. State-of-the-art object pose estimation methods [2] [3] represent the uncertainty in their estimations. The performance of the search algorithms are tested on data sets generated in simulation and from real-world scenarios. The results show overall improvement of the estimated poses and suggest that depth-limited search presents the best overall performance.

Fig. 1: Two example snapshots of the real-world configurations of the carton boxes inside shipping containers.

In this work, we assume an object detection and pose estimation module providing object properties, shapes and poses, as input data. In addition, a database of 3D models of the objects is assumed to be available. The method we present here considers objects with convex polyhedron shapes and assumes that there is only uncertainty in the estimated poses (i.e., shapes are assumed to be detected correctly). However, our method can be easily extended for objects with concave shapes by decomposing the shapes into a set of connected convex polyhedrons [4]. We call a set of poses of the objects a configuration; in an inter-penetration-free configuration there is no overlapping pair of objects.

We pose the problem of obtaining an inter-penetration-free configuration from a geometrically inconsistent set of poses as a search problem. Since the desired configuration is a set of new poses of the objects, it is reasonable to perform the search in the state space of the poses. It should be noted that the type of search we propose in this work differs from the search for the initial poses that an object pose estimation algorithm performs, i.e., we do not estimate the initial poses per se. Our ultimate goal is to refine a set of estimated poses in order to obtain a geometrically consistent configuration of objects. An object pose estimation, on the other hand, is a method to search for the poses to best fit object models to sampled data (e.g., a point cloud) of a given scene; the best fit, however, not necessarily results in a geometrically consistent configuration.

The main contributions of this paper are three-fold: 1) We propose search methods based on a principle of minimum translation when resolving inter-penetration inconsistencies to limit the search space; 2) we study the application of a fast algorithm for computing the inter-penetration between a

1http://www.roblog.eu
pair of convex objects, and 3) the results show empirically that PROMTS leads to increased object pose accuracy in a large set of simulation and real world experiments.

The paper is organized as follows. Related work is discussed in Section II. Section III outlines our proposed search algorithms to find a collision-free configuration of the initially overlapping objects. Experimental results are reported in Section IV, while section V concludes the paper.

II. RELATED WORK

A number of approaches have been proposed to estimate the poses of objects from 2D images [5] [6] [7] and 3D sampled points [8] [9] [2]. The main focus of the proposed approaches for object pose estimation is to obtain accurate object poses, while geometrical consistency of the estimated poses has received less attention. For example, Lim et al. [10] propose a fine pose estimation method to fit 3D models of IKEA furniture in the images. They use a database of the 3D models and define a multi-criteria score function to find the best fit for the 3D models in an image. As their results show, the error in the pose of the fit 3D model may result in inter-penetration with the environment. For instance, in Figure 9 in [10], the fit 3D model of a bookcase considerably intersects with the ground floor due to the error in its estimated pose. Such geometrically inconsistent situations can be resolved, for example, by a collision detection algorithm (that is expected to push the bookcase up) resulting in a higher accuracy of the estimated pose.

The presence of other objects – either fixed (e.g., a wall) or movable – nearby the target object corresponds to additional geometrical constraints which requires extra analysis. Aldoma et. al [11] propose an approach for verifying 3D models of objects (hypothesis verifying) in cluttered scenes. The method determines object and pose instances according to a global optimization paradigm by minimizing a cost function which encompasses geometrical cues. Grundmann et al. [12] propose a probabilistic approach, Rule Set Joint State Update (RSJSU), to estimate the poses of a set of objects simultaneously using the full joint posterior. They assume independence between prior belief, measurement and prediction models to approximate the full state. The results of the proposed method were presented on tabletop scenarios with only one object, though. Our work differs from both hypothesis verifying and full state estimation as we do not perform global pose estimation. In a closely related work, Wong et. al [13] propose collision-free state estimation where they attempt to solve a constrained optimization problem in order to find a feasible collision-free configuration. They assume that all the objects are resting stably on a 2D surface (i.e., no object is on top of another object). In their method the projections of the objects onto the 2D surface create a set of boundaries, and the inter-penetrations between the boundaries are resolved through optimization. However, their method is not applicable in our problem where goods are usually stacked on top of each other with arbitrary configurations inside shipping containers.

It is worth mentioning that the collision resolvers of Physics Engines (e.g. see [14]) are based on dynamic collision detection where impulse forces are used to simulate the trajectories of two objects after their dynamic collision. Such impulse force based algorithms when initialized with a static configuration of overlapping objects result in a spread of objects far from their initially estimated poses. Contrary, the PROMTS approach proposed in this paper attempts to resolve all initial inter-penetrations between objects with minimum change in their initially estimated poses.

III. PROMTS APPROACH

The methodology we propose in this paper consists of two major parts: representation and computation of the inter-penetrations between pairs of overlapping convex polyhedrons, and a solver that resolves these inter-penetrations through search algorithms. A naive approach that one may utilize to resolve the inter-penetrations is to perform a complete search in the state space of six degrees of freedom (6-DOF) poses of the objects to find a geometrically consistent solution. However, as the number of objects increases, the naive search becomes infeasible since the branch factor of the search space grows exponentially. In order to reduce the search space and yet retain the most promising solutions, we propose a principle of minimum translation search, and we call it PROMTS. Although PROMTS is not tied to specific search algorithms, in this work, we employ A-star and Depth Limited search methods [15] to find a geometrically consistent configuration of objects with minimum change compared to their initially estimated poses.

A. Inter-penetration Computation

The inter-penetration between two overlapping polytopes can be represented by another polytope that contains the overlapping space. Although this representation is a precise description of the inter-penetration space, the computation of the overlapping polytope, especially in 3-dimensional space, is computationally expensive [16]. Another representation of
the overlapping space between two polytopes is an inter-penetration vector such that translating one of polytopes by the vector will resolve the inter-penetration between the polytopes; the length of the vector is called depth of penetration (DOP), and the direction of the vector identifies a separating axis.

A fast method to compute the inter-penetration vector is based on the separating axis theorem (SAT) [16], which is widely used in computer graphics and physics simulations for collision detection. SAT is a corollary of the separating hyperplane theorem [17], which is an essential theory in convex set analysis. The separating hyperplane theorem states that for two convex sets \( A \) and \( B \), either the two sets are overlapping or there exists at least one separating hyperplane \( P \) such that \( A \) is on one side of \( P \) and \( B \) is on the other side. The normal of a separating hyperplane is called a separating axis for the two convex sets.

For two non-overlapping convex polytopes, \( A \) and \( B \), if \( L \) is a separating axis along the unit vector \( \hat{l} \), then the orthogonal projections of \( A \) and \( B \) onto \( L \) result in two non-overlapping intervals (see Fig. 2). In other words, if there exists at least one axis on which the orthogonal projections of two convex polytopes have non-overlapping intervals, then the two polytopes are separated. On the other hand, if \( A \) and \( B \) are two overlapping convex polytopes, in order to separate them with minimum translation, we compute the orthogonal projections of \( A \) and \( B \) onto all their fundamental axes and select the axis on which the overlapping interval (DOP) is minimum; the vector along this axis with DOP length is called minimum translation vector (MTV). In 3-dimensional space, for each pair of convex polyhedrons, \( A \) and \( B \), the set of fundamental axes, \( L\text{Set} \), contains all the normals of the faces as well as all possible cross products between the edges of \( A \) and the edges of \( B \) [16]. In Fig. 3, the set of fundamental axes for computing overlapping intervals of two polytopes, \( A \) and \( B \) is depicted for the 2D case. The procedure of computing MTV and DOP for two convex polyhedrons is presented in Algorithm 1.

For a configuration of more than two objects in which there is at least one overlapping pair of objects, resolving the inter-penetration may result in a new set of overlapping objects. Fig. 4 illustrates a configuration of three movable polytopes, \( A \), \( B \) and \( C \), and a fixed object, \( W \). (a) Initial state, \( s_1 \) with an inter-penetration between \( A \) and \( B \) which generates two possible actions, \( a(s_1) = \{a^1_A, a^1_B\} \). (b) Taking action \( a^1_A \) translates \( A \) by MTV resulting in a new inter-penetration between \( A \) and \( W \). Since \( W \) is fixed, there is only one possible action in \( s_2 \), \( a(s_2) = \{a^2_A\} \), where taking \( a^2_A \) goes back to \( s_1 \), and hence this path in the search will not be expanded further. (c) Taking action \( a^2_B \) translates \( B \) by negative MTV which results in a new inter-penetration between \( B \) and \( C \). Since \( C \) is a movable object, there are two possible actions in \( s_3 \), \( a(s_3) = \{a^3_B, a^3_C\} \). (d) Taking \( a^3_C \) results in \( s_4 \) which is an inter-penetration-free configuration, i.e., a goal state.

In order to find a configuration with a sum of zero inter-
which the total depth of penetrations is zero. We propose the
In our problem, a goal state is a configuration of objects in
of taking action
a
successors of states are expanded. This limits the number
the practical application of the algorithm. In our problem,
the state space of MTVs can grow exponentially as the
successors of states are expanded. This limits the number
of objects and the inter-penetrations that the proposed A-
search algorithm approximately satisfies the two criteria: maximizing geometrical consistency
and minimizing the sum of translations required to reach a
goal state.
2) Depth Limited Search: Although the A-star search algorithm with an approximated consistent heuristic function
could find a near to optimum solution (if there exists any),
the time complexity of A-star in a large state space limits
the practical application of the algorithm. In our problem,
the state space of MTVs can grow exponentially as the
successors of states are expanded. This limits the number
of objects and the inter-penetrations that the proposed A-
search algorithm is capable to deal with in a reasonable time.
This limitation motivated us to examine the utility of a
second search algorithm that only explores a branch of the
state space and finds suboptimal but geometrically consistent
solutions. A suboptimal solution is a sequence of translation
actions that results in an inter-penetration-free configuration
while the total cost of taking the actions is not necessarily
minimum; we employ the depth-first search [15] algorithm
for this purpose. A suboptimal solution just refers to how
an inter-penetration-free configuration is reached but we
are mainly interested in the final configuration and not so
much in the sequence of actions. On the other hand, as we
mentioned earlier in Section III-A, there could be configu-
ations of overlapping objects for which no goal state exists
(i.e., there exists no inter-penetration-free configuration).

Algorithm 1: Computation of MTV and DOP

Data: Vertices and LSet of two convex polyhedrons, A and B
Result: MTV and DOP of A and B
DOP ← inf;
MTV ← 0;
for each axis L in LSet do
  project vertices of A and B on L;
  compute each projection interval on L;
  if two intervals intersect then
    d ← the length of intersection;
    if d < DOP then
      DOP ← d;
      MTV ← MTV · ı̂;
  end
else
  DOP ← 0;
end
return

Fig. 5: Two simulated configurations with random poses of
objects inside shipping containers.

The heuristic function in Eq. 2 is not consistent in
general. However, we observed that in many cases when the
minimum required cost to reach the goal state is equal to
all the translation actions that must be taken to resolve the
inter-penetrations the heuristic function in Eq 2 becomes
consistent with good results, and hence we use it as an
approximation for a consistent heuristic function.

The solution that A-star search finds (if there exists any)
is a sequence of actions such that their execution results in a
transition from start to goal state with a minimum total cost
taking actions. Since a goal state is an inter-penetration-
free configuration, and the total cost (translations) of reaching
the goal state in many cases is minimum, a solution
returned by the A-star search algorithm approximately satisfies
the two criteria: maximizing geometrical consistency and
minimizing the sum of translations required to reach a
goal state.

2) Depth Limited Search:

In order to find an inter-penetration-free configuration of
objects (i.e., a goal state) with minimum translations in the
initially estimated poses, a complete search on the state space
of MTVs is required. We explore two search algorithms, A-
and depth limited search to find a goal state.

We represent a state, s, as a new set of poses, P(s), of the
objects, and we define a set of actions for each state, a(s).
Given a state, s, for each pair of overlapping objects (i.e., ∀
DOP_{ij} ∈ s such that DOP_{ij} ≠ 0) we define two possible
actions, a_i, a_j ∈ a(s), such that
• a_i translates i-th object by the MTV_{ij};
• a_j translates j-th object by the negative MTV_{ij}.

If a static object (e.g., a wall) overlaps with a movable
object, only the action that translates the movable object is
considered in the search (see Fig. 4).

1) A-star Search: A-star is a heuristic graph search algo-
rithm which is optimal if the heuristic function is consist-
tent [15]. If h(s) is the heuristic function that estimates the
cost to reach the goal state from s and c(s, a', s') is the cost
of taking action a' ∈ a(s) to go from s to the successor state
s', then the heuristic function h(.) is said to be consistent if,

$$h(s) \leq c(s, a', s') + h(s')$$ (1)

In our problem, a goal state is a configuration of objects in
which the total depth of penetrations is zero. We propose the

penetrations between objects, one solution that we propose
is to utilize search in the state space of MTVs to find a
goal state satisfying a set of predefined constraints. We note
that depending on the initial configuration of objects and
the structure of the environment, there might be no solution
resulting in an inter-penetration-free configuration.

B. Search Algorithms

In order to find an inter-penetration-free configuration of
objects (i.e., a goal state) with minimum translations in the
initially estimated poses, a complete search on the state space
of MTVs is required. We explore two search algorithms, A-
and depth limited search to find a goal state.

We represent a state, s, as a new set of poses, P(s), of the
objects, and we define a set of actions for each state, a(s).
Given a state, s, for each pair of overlapping objects (i.e., ∀
DOP_{ij} ∈ s such that DOP_{ij} ≠ 0) we define two possible
actions, a_i, a_j ∈ a(s), such that
• a_i translates i-th object by the MTV_{ij};
• a_j translates j-th object by the negative MTV_{ij}.

If a static object (e.g., a wall) overlaps with a movable
object, only the action that translates the movable object is
considered in the search (see Fig. 4).

1) A-star Search: A-star is a heuristic graph search algo-
rithm which is optimal if the heuristic function is consist-
tent [15]. If h(s) is the heuristic function that estimates the
cost to reach the goal state from s and c(s, a', s') is the cost
of taking action a' ∈ a(s) to go from s to the successor state
s', then the heuristic function h(.) is said to be consistent if,

$$h(s) \leq c(s, a', s') + h(s')$$ (1)

In our problem, a goal state is a configuration of objects in
which the total depth of penetrations is zero. We propose the

penetrations between objects, one solution that we propose
is to utilize search in the state space of MTVs to find a
goal state satisfying a set of predefined constraints. We note
that depending on the initial configuration of objects and
the structure of the environment, there might be no solution
resulting in an inter-penetration-free configuration.

B. Search Algorithms

In order to find an inter-penetration-free configuration of
objects (i.e., a goal state) with minimum translations in the
initially estimated poses, a complete search on the state space
of MTVs is required. We explore two search algorithms, A-
and depth limited search to find a goal state.

We represent a state, s, as a new set of poses, P(s), of the
objects, and we define a set of actions for each state, a(s).
Given a state, s, for each pair of overlapping objects (i.e., ∀
DOP_{ij} ∈ s such that DOP_{ij} ≠ 0) we define two possible
actions, a_i, a_j ∈ a(s), such that
• a_i translates i-th object by the MTV_{ij};
• a_j translates j-th object by the negative MTV_{ij}.

If a static object (e.g., a wall) overlaps with a movable
object, only the action that translates the movable object is
considered in the search (see Fig. 4).

1) A-star Search: A-star is a heuristic graph search algo-
rithm which is optimal if the heuristic function is consist-
tent [15]. If h(s) is the heuristic function that estimates the
cost to reach the goal state from s and c(s, a', s') is the cost
of taking action a' ∈ a(s) to go from s to the successor state
s', then the heuristic function h(.) is said to be consistent if,

$$h(s) \leq c(s, a', s') + h(s')$$ (1)

In our problem, a goal state is a configuration of objects in
which the total depth of penetrations is zero. We propose the

penetrations between objects, one solution that we propose
is to utilize search in the state space of MTVs to find a
goal state satisfying a set of predefined constraints. We note
that depending on the initial configuration of objects and
the structure of the environment, there might be no solution
resulting in an inter-penetration-free configuration.
In such cases an unlimited search algorithm may generate infinitely many intermediate states. In order to overcome this issue we propose to use the depth limited search (DLS) algorithm which is the depth limited version of depth-first search. Since backward translation represents a redundant state we preserve the visited states in memory to prevent the generation of successors that are already visited (e.g., see Fig. 4 where back translation of $A$ in $s_2$ results in a redundant state same as $s_1$ which is already visited).

### IV. Results

We examined the performance of the proposed approach, PROMTS, on both simulated and real-world data. Using scenarios generated in simulation enables us to create a large data set of different configurations of objects with their ground truth poses to capture the statistical properties of the approach. The real-world configurations were used to verify the approach on real data.

#### A. Simulated Configurations

We selected three categories of shapes of objects which are commonly used in shipping containers: box, cylinder and barrel. A total of 200 scenarios was generated in equally sampled five groups of $N = 10, 20, 30, 40, 50$ objects; the number of shapes in each scenario was equally likely drawn from the three categories with uniform random dimensions. In addition to $N$ objects, in each scenario (e.g., see Fig. 5) there are 6 fixed objects: left, right, back wall, floor and ceiling of the container as well as the ground plane that supports all other objects. For every scenario, we add Gaussian noise, $N(0, \sigma^2)$, to each component of the translation vectors and the Euler angles of the objects poses to generate a set of noisy poses. The noisy poses simulate the error in the estimated poses by an existing object detection module. We used standard deviations of 0.05m and 5 degrees for the translation and the rotation components respectively.

#### B. Real-World Configurations

Fig. 7 shows the real-world configurations of objects (i.e., carton boxes and cylinders) we used for verifying our proposed approach. A Microsoft Kinect captured a point cloud of the scene to which the 3D models of the objects were registered, and the poses were manually refined for obtaining the ground truth (see right column of Fig. 7). In order to examine the approach independent of and not tied-to any particular object pose estimation algorithm, we sample a set of noisy pose estimates from the ground truth. This means that we can expect the same results if noise that comes from the sensing and estimation process is distributed in the same way. We add Gaussian noise to each component of the translation vectors and the Euler angles of the poses, which may result in a configuration of overlapping adjacent objects.

### TABLE I: The success rate (see Section IV-C) of the proposed search algorithms with respect to the number of objects.

<table>
<thead>
<tr>
<th>No. Objects</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>A* Success Rate</td>
<td>80%</td>
<td>72.5%</td>
<td>35%</td>
<td>25%</td>
<td>2.5%</td>
</tr>
<tr>
<td>DLS Success Rate</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### TABLE II: The results of applying the proposed search methods on real-world configurations (see Fig. 7).

<table>
<thead>
<tr>
<th>Scene</th>
<th>A* DLS</th>
<th>A* DLS</th>
<th>A* DLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scene 1</td>
<td>Avg. ExT (ms)</td>
<td>9.53</td>
<td>0.69</td>
</tr>
<tr>
<td>Scene 2</td>
<td>Succ. Rate(%)</td>
<td>96.1</td>
<td>100</td>
</tr>
<tr>
<td>Scene 3</td>
<td>Avg. PER (%)</td>
<td>8.6</td>
<td>11.3</td>
</tr>
<tr>
<td></td>
<td>Avg. RPE (m)</td>
<td>0.511</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>Avg. IPE (m)</td>
<td>0.559</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>Avg. IOL (#N)</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>
Fig. 7: (left) Three real-world configurations of objects. (right) The ground truth 3D models and poses.

(see Fig. 8a for an example). We generated 1000 samples for each real-world configuration with a standard deviation of 0.05m for the translation noise, and a standard deviation of 5 degrees for the rotation noise.

C. Evaluation

In our experiments, in order to fairly compare the proposed search algorithms, we limited the number of visited nodes for both DLS and A-star search algorithms to 5000 and set the maximum depth for DLS to 1000. We define the result of a search by A-star and DLS a successful search if a valid solution (i.e., an inter-penetration-free configuration) can be found given the limited depth and number of node visits. The success rate is the percentage of successful searches.

In order to evaluate pose accuracy, we define pose error reduction (PER) as the difference between the initial pose error (IPE) and the refined pose errors (RPE)

\[
\text{PER} = \frac{\text{IPE} - \text{RPE}}{\text{IPE}} \times 100\%
\]

where,

\[
\text{IPE} = \sum_{i=1}^{N} ||t_{d,i}^n - t_{g,i}^n||
\]

\[
\text{RPE} = \sum_{i=1}^{N} ||t_{r,i}^n - t_{g,i}^n||
\]

and \(N\) is the number of objects, \(t_{d,i}^n\), \(t_{r,i}^n\) and \(t_{g,i}^n\) are the translation vectors of the \(i\)-th object’s ground truth, detected (i.e., noisy) and refined (i.e., a goal state) poses respectively. A positive value of PER indicates a reduction in the refined poses with respect to the initially detected poses.

1) Simulated Configurations: In Fig. 6a and Fig. 6b, the average of PER for each search method is depicted with respect to the number of objects and initial inter-penetrations between objects respectively. Fig. 6c shows the average execution time for the search methods with respect to the number of objects. It can be seen that both search methods are approximately equally fast for scenarios with 10 objects. However, as the complexity of the scenarios increases with an increasing number of objects, the execution time for A-star rapidly increases, while the depth-limited search algorithm is able to resolve the inter-penetrations between objects in highly cluttered scenarios still in a reasonable time (less than 50 seconds on average for scenarios with 50 objects).

Table I depicts the success rate of the search algorithms with respect to the number of objects. While depth-limited search manages to find a goal state for all the simulated test scenarios, A-star with the proposed approximate heuristic...
shows a decreasing performance as the number of objects increases.

2) Real-World Configurations: Fig. 8 visualizes a typical search for an inter-penetration-free configuration of the real-world scenario shown in Fig. 7a. The initial state is depicted in Fig. 8a that introduces inter-penetrations between pairs of objects due to inaccuracies in the initial poses. In Fig. 8b, a goal state, where the inter-penetrations are resolved, is depicted. A few intermediate states are shown in Fig. 8c, Fig. 8e, Fig. 8f and Fig. 8d in the order in which they are evaluated by the proposed algorithm. In Table II the results of applying the proposed search methods to the real-world configurations are summarized. The first observation is that both search methods reduce the average pose error and result in configurations of objects which are geometrically consistent. It can be also seen that the proposed approach is computationally inexpensive (less than 100 mili-seconds) for real-world configurations where the number of visible objects to the perception module is less than 10. The success rate of A-Star search, as in simulation, is less than with DLS, which manages to successfully find inter-penetration-free configurations in all the trials. We also observe that the results obtained for the real-world data is consistent with that of simulated data.

V. SUMMARY AND CONCLUSION

In this paper we proposed PROMTS, a principle of minimum translation search approach to resolve the set of inter-penetrations between initially estimated 3D models and poses of objects. Our experimental evaluation shows that resolving the inter-penetrations not only represents a geometrically consistent model of the environments, but also reduces the total pose error. Our target application is to refine the poses of the detected objects inside shipping containers in the process of automating the task of unloading goods. We employ the separating axis theorem for fast computation of minimum translation vectors between pairs of overlapping objects. We examined the utility of two search methods, A-star with an approximate heuristic and depth limited search, to explore the state space of minimum translation vectors to find geometrically consistent poses. The approach was tested and verified on data sets generated from real-world and simulated configurations. From the results we observe that using the depth-limited search technique significantly prunes the state space to find a geometrically consistent solution. The results also suggest that a trade-off analysis between computational resources and the amount of resolved inter-penetrations with respect to the number of objects is necessary to select a proper search paradigm. Nevertheless, for our target application, depth-limited search algorithm showed the best overall performance. Future work concerns to evaluate other heuristic functions and graph search pruning techniques to improve the search time. Examining the search along fundamental axis in case no solution along the MTV directions found is another direction for the future work.

ACKNOWLEDGEMENT

This research has been partially supported by the RobLog project, funded by the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement ICT-270350.

REFERENCES