

Object Detection and Pose Estimation for Robotic Manipulation using Physics Simulation and Monte-Carlo Tree Search

Chaitanya Mitash, Kostas Bekris, Abdeslam Boularias
Rutgers, the State University of New Jersey

Motivation

- ▶ Provide a dataset for the task of robotic grasping and manipulation in a typical constrained warehouse environment
- ▶ A difficult environment:
 - ▶ Inconsistent and indirect lighting conditions induce noise in sensors
 - ▶ Objects are often partially occluded by other objects or the shelving unit
 - ▶ Narrow bins allow only very limited robotic mobility
- ▶ However, the task is of great importance to industrial applications
- ▶ Developing high accuracy pose estimation algorithms is of utmost importance
- ▶ **Objective:** provide a representative dataset that *allows researchers to determine effects of specific environmental variables on pose estimation algorithms*

Amazon Picking Challenge (APC)



Figure : The 25 APC objects categorized by broad object class

- ▶ Shelf-picking challenge simulating a warehouse environment
- ▶ Tight, cluttered workspaces bins within the shelving unit
- ▶ 25 target objects stocked in 12 bins, 1-4 objects per bin
- ▶ Target items show variety of forms and traits

Data Collection



Figure : Our data collection hardware and environment

- ▶ Collection performed on Motoman Dual-arm SDA10F humanoid robot using Kinect v1 RGBD camera
- ▶ Camera located on wrist of robot, position calibrated prior to data collection
- ▶ 10,368 RGBD images collected involving 24 target objects, each shown in:
 - ▶ 12 different bin locations
 - ▶ 3 clutter states per bin (1, 2, 3 items)
 - ▶ 3 viewpoints per clutter state
 - ▶ 4 frames per viewpoint
- ▶ 6D Ground truth annotated by hand, including all transformations between object, camera, and robot base

Dataset Features

- ▶ Large dataset covering wide variety of objects with 6D ground truth object pose annotations
- ▶ Good sample coverage of probable object poses within the shelves
- ▶ Provides researchers the ability to determine the effects of immediate clutter on pose estimation
 - ▶ Same ground truth target object pose is given with varying degrees of clutter within the bin
- ▶ Allows opportunity to analyze accuracy based on the camera's viewpoint
 - ▶ For each shelf configuration, samples are generated from 3 different viewing positions
- ▶ Provides the ability to determine the effects of slight sensor noise on deterministic algorithms
 - ▶ For each shelf configuration at each viewpoint 4 samples are generated

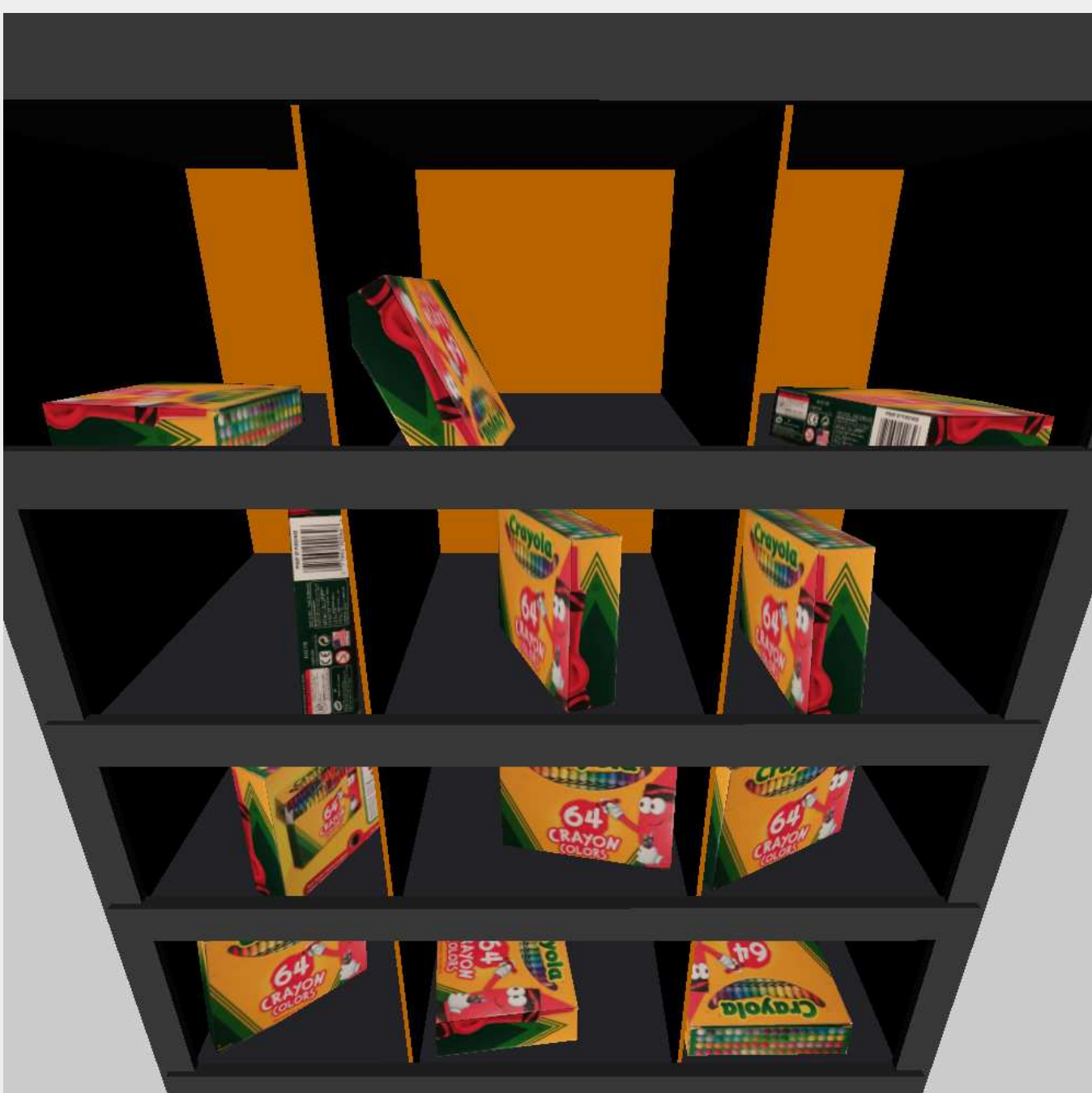


Figure : Variety of coverage in our dataset of ground truth poses for a single object



Figure : Examples from our dataset showing the same target object pose (the "duck") in varying degrees of clutter within the bin

Pose Estimation Evaluation

- ▶ Dataset is readily used to compare pose estimation algorithms for this difficult task
- ▶ Acceptable accuracy thresholds are at the discretion of the user
- ▶ May use the metrics provided by the dataset to determine a series of improvements in order to deal with difficulties of the task

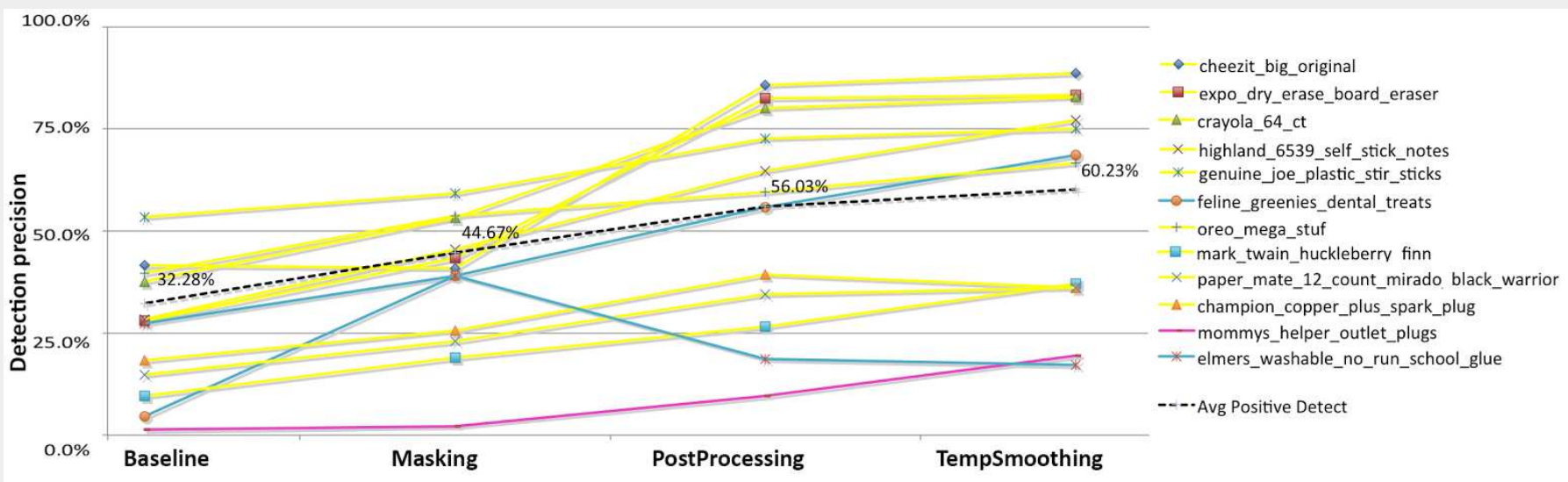


Figure : Evaluation made possible several improvements to an example pose estimation algorithm



Figure : An example detection of the "duck" from our dataset

Dataset Comparison

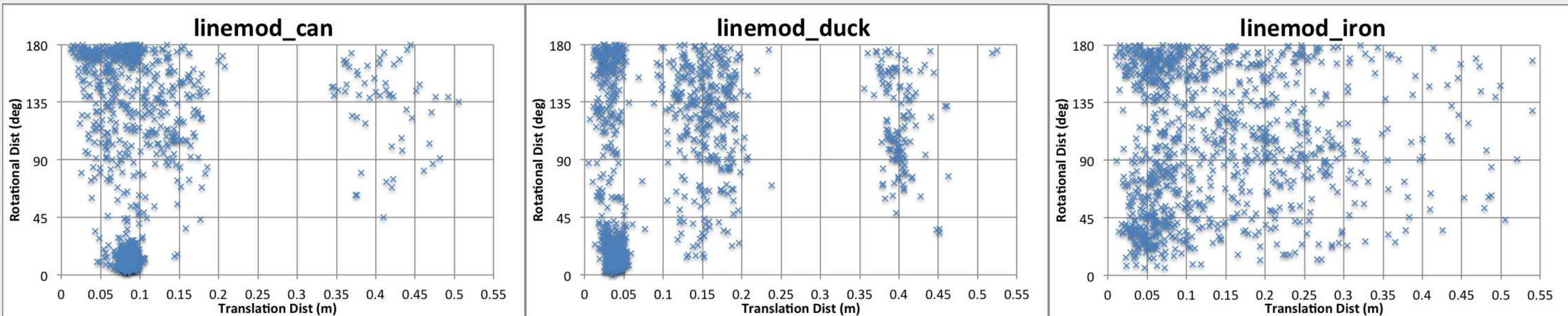


Figure : Example objects from LINEMOD dataset evaluated using our example algorithm

- ▶ LINEMOD dataset of 10,000+ RGBD images with ground truth 6D pose
- ▶ Shows target items on an AR board, in a heavily cluttered environment
- ▶ Samples generated from a variety of viewpoints around and above the object
- ▶ Can evaluate rotational, translational accuracy using the provided object models

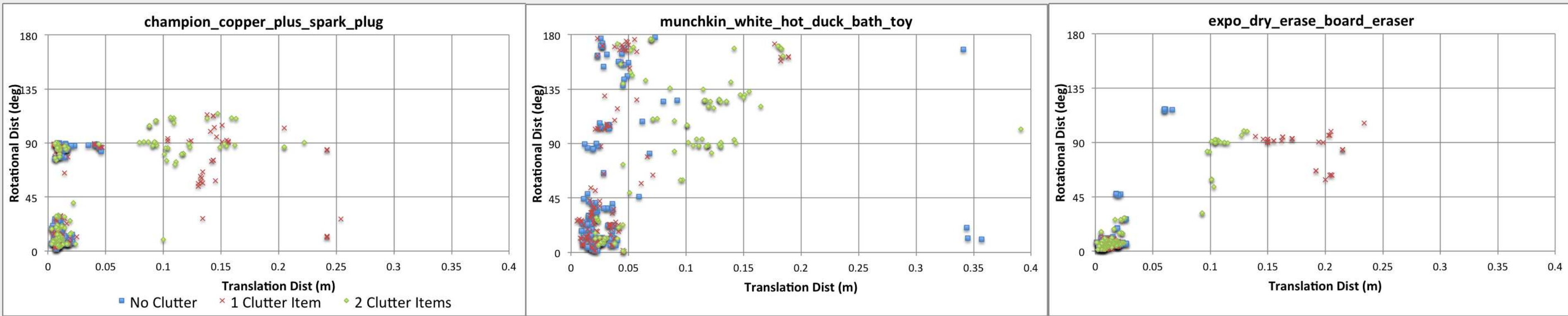


Figure : Objects from our dataset evaluated using the same algorithm; Colors indicate clutter conditions

- ▶ Our similar setup: object models provided, 10,000+ RGBD images
- ▶ Able to extract effects of immediate clutter on accuracy
- ▶ Can also determine robustness to noise in the sensor samples
- ▶ Evaluation with our dataset allows researchers to identify certain weak points in pose estimation algorithms on this difficult task