Team NimbRo Picking at ARC 2017: Fast Learning Semantic Perception and Coordinating Two Arms

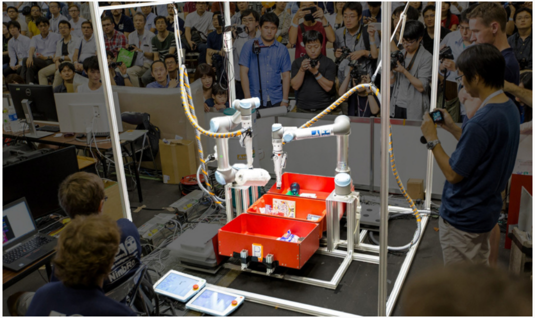


Image by Amazon Robotics

amazon robotics CHALLENGE

Max Schwarz,
Christian Lenz,
Germán Martín García,
Seongyong Koo,
Arul Selvam Periyasamy,
Michael Schreiber,
and Sven Behnke

UNIVERSITÄT BONN AIS

Computer Science Institute VI
Autonomous Intelligent Systems

This talk: Focus on Lessons Learned from APC 2016 and ARC 2017

More details also in interactive presentation / paper:

Fast Object Learning and Dual-arm Coordination for Cluttered Stowing, Picking, and Packing
Max Schwarz, Christian Lenz, Germán Martín García, Seongyong Koo, Arul Selvam Periyasamy, Michael Schreiber, and Sven Behnke

ICRA 2018, **session WeA@L.6**

What did we learn from APC 2016?

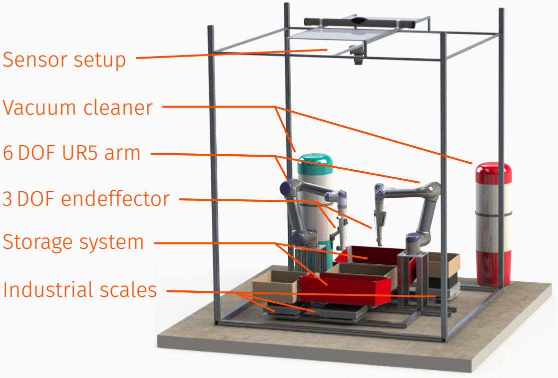
NimbRo Picking APC 2016:
stow: 2nd place
pick: 3rd place

- Hybrid grasping strategies work well: Suction + Pinch grasping (Delft, MIT, PFN)
- Complex grasping actions can be performed using keyframe interpolation techniques (NimbRo)
- Stationary sensor setups are faster (Delft, ...)
- Measuring weight is valuable
- Speed! If you fail, just retry.

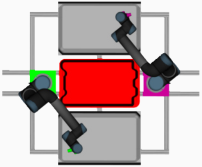
Rule changes for ARC 2017

- Unknown objects
  → need fast semi-automatic capture & training
- Pack three boxes in parallel
  → multiple arms
- No deep shelf bins
  → linear actuator not required

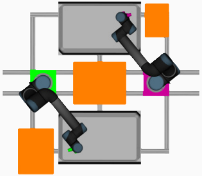
System Design



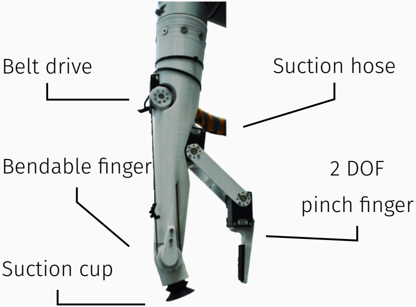
Stow setup with tote



Pick setup with boxes

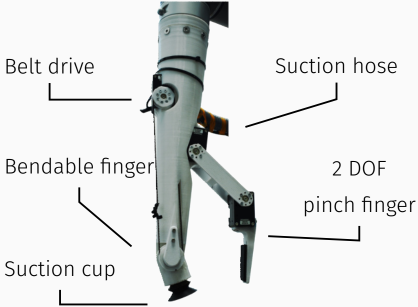


Endeffector Design - Suction



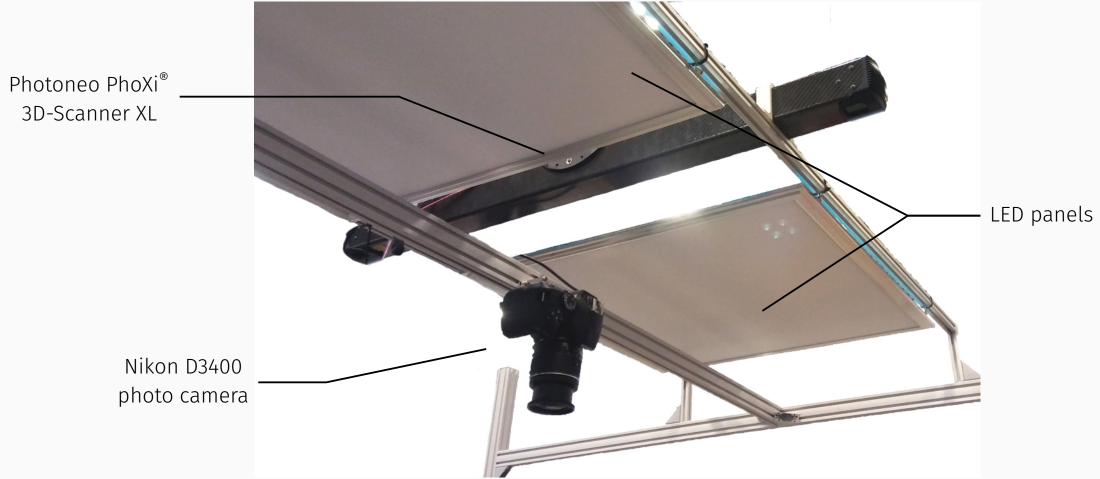
This endeffector design allows us to grasp items using suction...

Endeffector Design - Pinch Grasp



… and perform pinch grasps with both fingers.

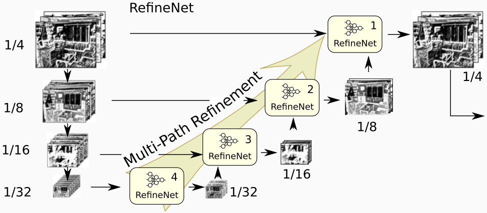
Sensor Setup



Object Perception

Semantic Segmentation

A state-of-the-art semantic segmentation method is used to perceive objects.



RefineNet: Multi-Path Refinement Networks

for High-Resolution Semantic Segmentation

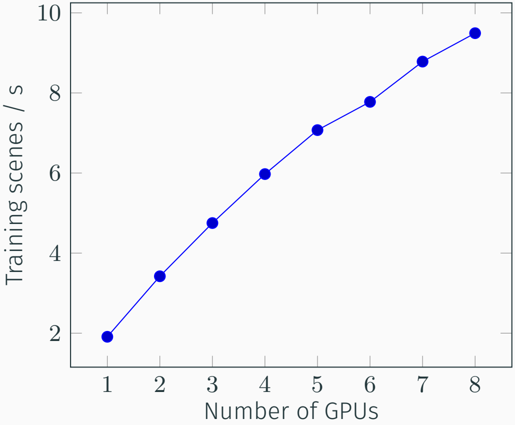
Guosheng Lin, Anton Milan, Chunhua Shen, Ian Reid

CVPR 2017

Data Capture, Scene Synthesis & Training

We capture new objects using a turntable and generate synthetic scenes on top of annotated dataset frames. Training is performed in $\approx 30\,\text{min}$ on four Titan X cards.

Segmentation results



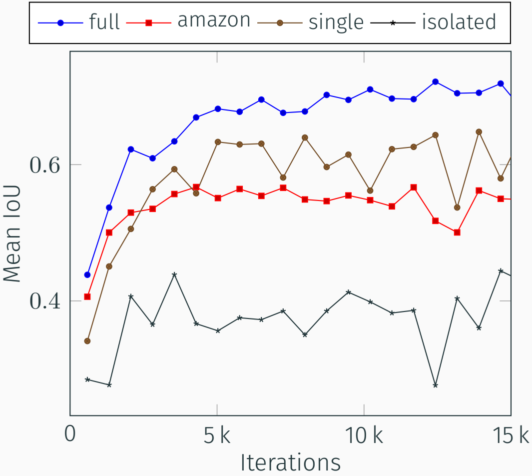


Figure 1: Segmentation experiments. Left: Training image throughput depending on the number of GPUs. Right: Test set IoU during training.

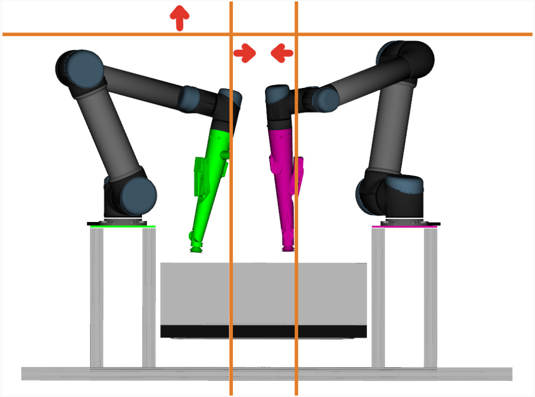
Grasp Generation

- Object contours are extracted from segmentation.

- Object contours are extracted from segmentation.
- 2D grasp points with maximum distance to the contour are found.
- 6D grasp poses are calculated from depth and local surface normals.

Motion Generation and Dual-arm Coordination

Nullspace-optimizing IK

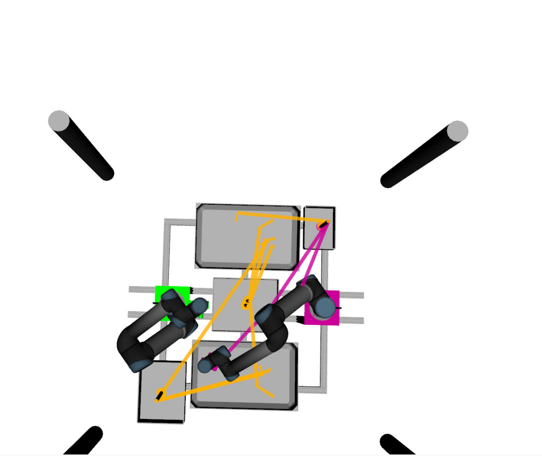


- Secondary objective optimized during IK: Keep wrist as high as possible and away from the robot base
- For suction grasps, consider only 5D poses

$\Rightarrow$ Reach any visible suction pose in the bin without arm$\leftrightarrow$bin collisions.

Dual-arm Coordination

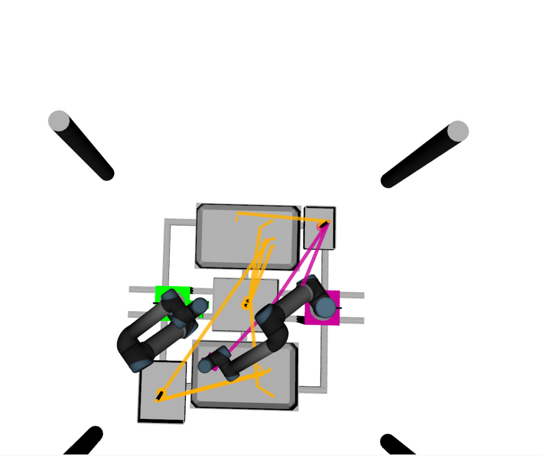
- Manipulation tasks are defined by line segments between endeffector waypoints.



Green & purple: Arm activities. Yellow: Next tasks.

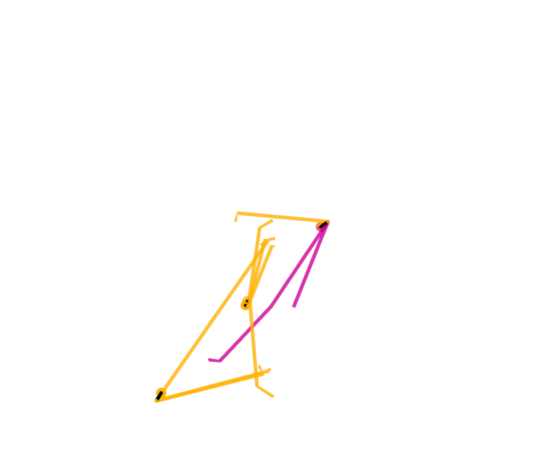
Dual-arm Coordination

- Manipulation tasks are defined by line segments between endeffector waypoints.
- Line segments are projected in 2D for collision checking.



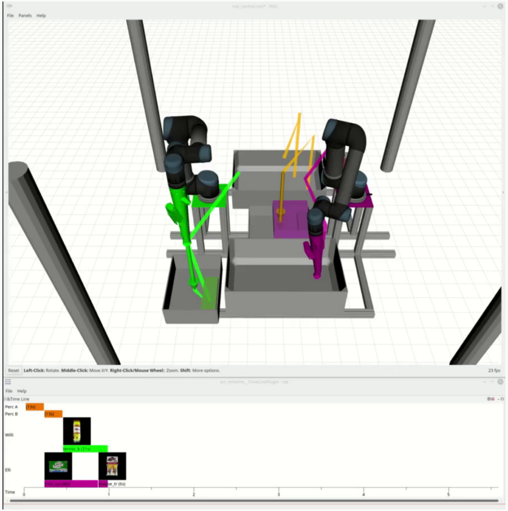
Green & purple: Arm activities. Yellow: Next tasks.

- Manipulation tasks are defined by line segments between endeffector waypoints.
- Line segments are projected in 2D for collision checking.
- Next tasks are assigned to a free arm if the minimum distance between all pairs of line segments is large enough.



Green & purple: Arm activities. Yellow: Next tasks.

Dual-arm Coordination



Collision-free task assignment:

Green & purple: Arm activities.
Yellow: unassigned task generated from latest perception result.

Timeline of system activities.

Experiences from ARC 2017

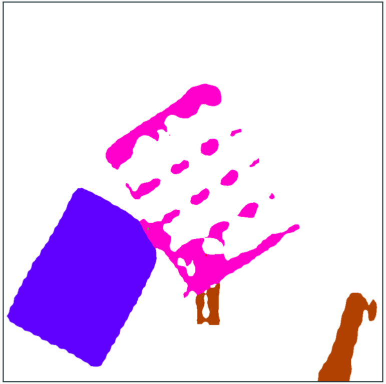
Final run

Highly successful: Stowed 14 out of 16 objects, picked 8 out of 9 objects ⇒ 235 points.

Failure 1: Could not pick last two objects from tote

Failure 2: Could not pick last object during pick phase

- Items correctly segmented
- Control module always selected the `kitchen_masher`
- Computed pregrasp pose collided with the bin
- No randomness involved



Failure mode a):

- Item undersegmented due to very sparse annotation
- Control module starts moving other objects to other bin

Final run: Picking Phase

Failure mode b):

- Item grasped, but fails weight check:
  `weight diff=0.059g, expected weight=0.086g`
- Noise problems with scales

Lessons Learned

- Motion generation
    - Full motion planning is not really necessary.
    - If using keyframe-based approach, make keyframe generation as robust as possible.
- Object perception
    - Deep Learning techniques are applicable in this setting.
    - Make sure you are training the correct objective!
- High-level control
    - Don’t get stuck in loops!
    - Separate verification method (weight) relaxes demands on segmentation accuracy and grasp precision.

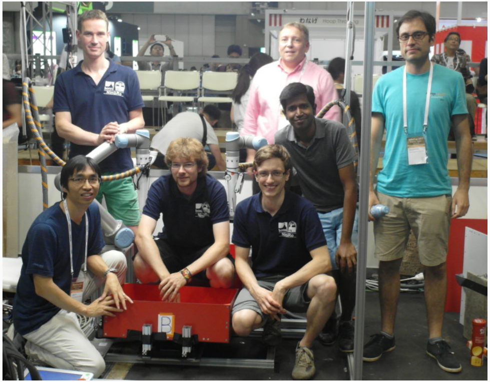
Results

- Robust and fast item perception
  in cluttered scenes.
- Perception pipeline can be quickly adapted to novel items.
- Robust grasp generation
  for a large variety of items.
- Planning & Coordination for dual-armed manipulation in shared workspace.
- 2nd place in the ARC 2017 Finals!

Amazon Robotics Challenge
(Final)

Rank	Team	Score
1	ACRV	272
2	NimbRo	235
3	Nanyang	225

Thank you



Michael Schreiber Sven Behnke Arul Selvam Periyasamy Germán Martín García
Seongyong Koo Christian Lenz Max Schwarz