Machine vision for various manipulation tasks

Yukiyasu Domae

Advanced Technology R&D Center, Mitsubishi Electric Corp.

AIST

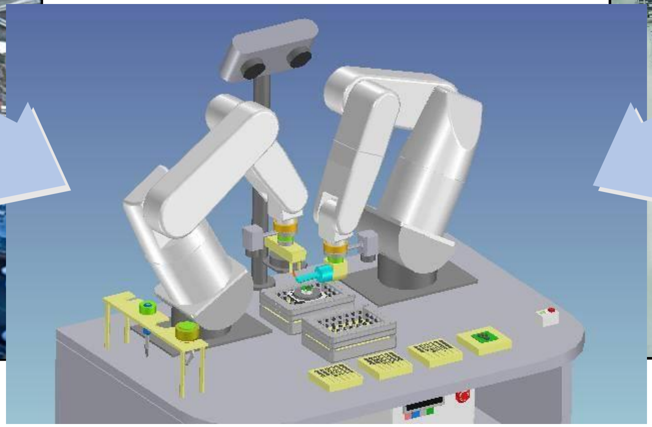
1. Machine Vision problems and algorithms in Factory Automation

Trend on Factory Automation

Megatrend：Increase of aging population, labor shortage

From mass production to variable production

Line production



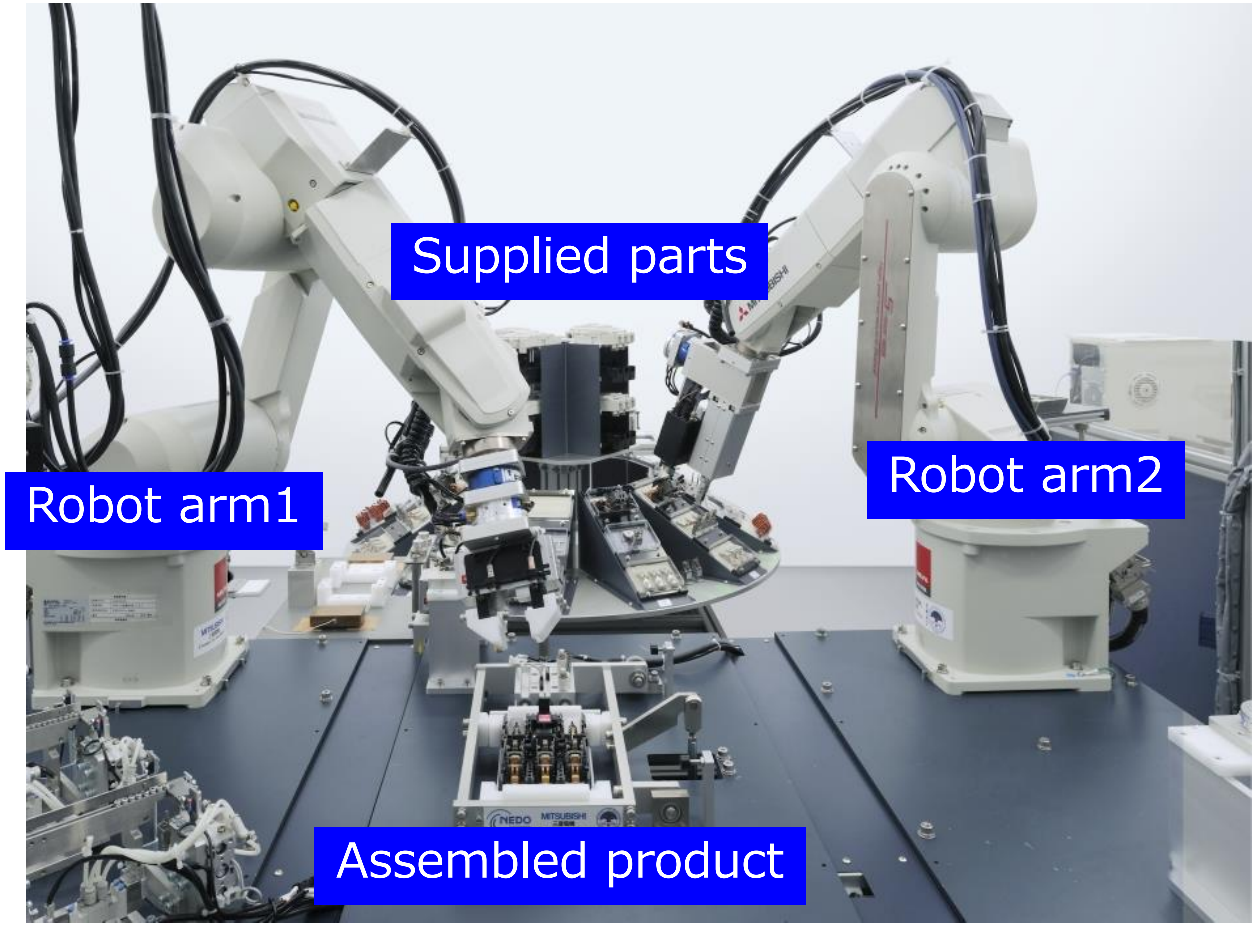
Cell production by hand

Cell production robot system

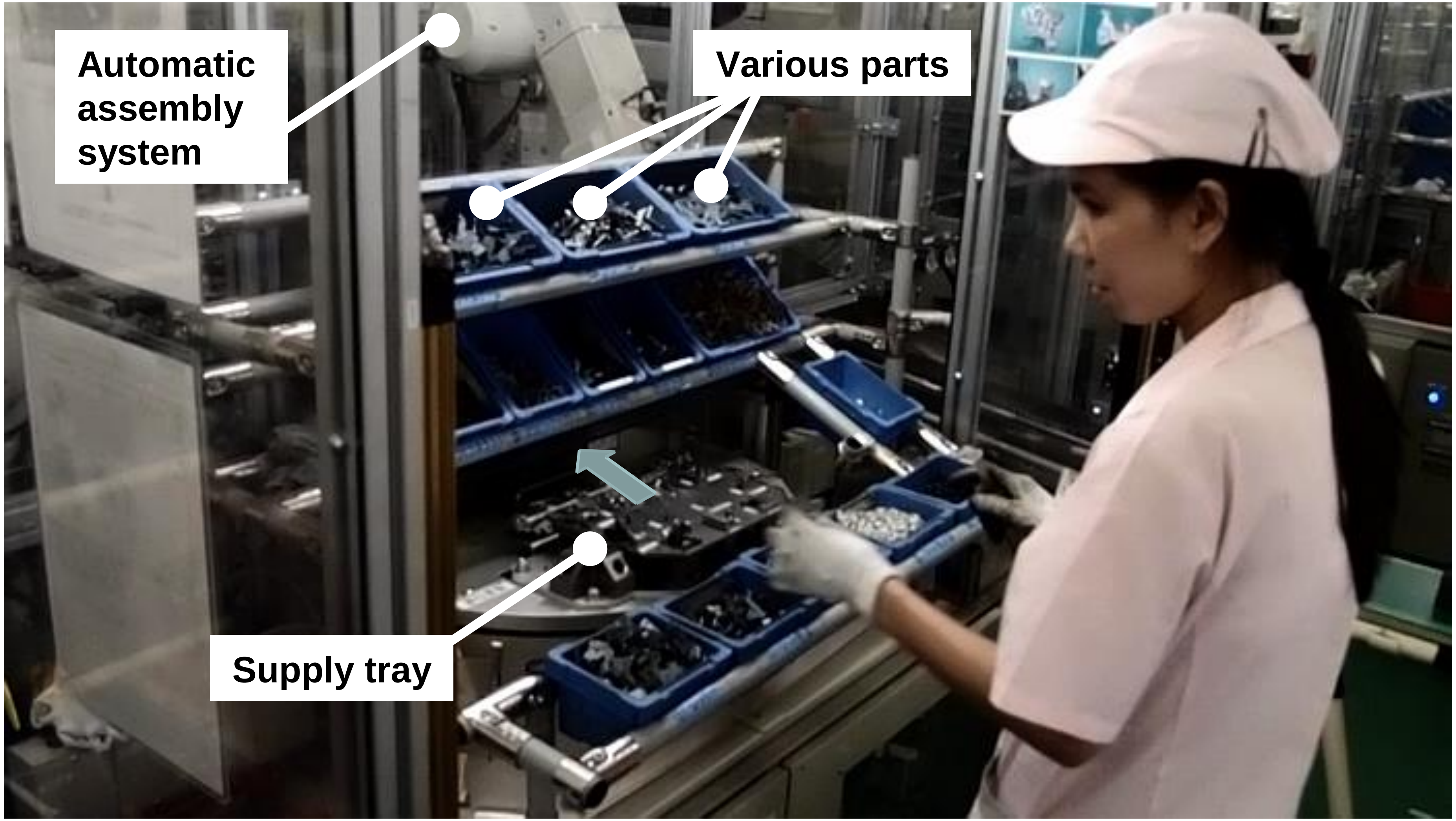
- High productivity
- High reliability
- High flexibility

Cell production robot system

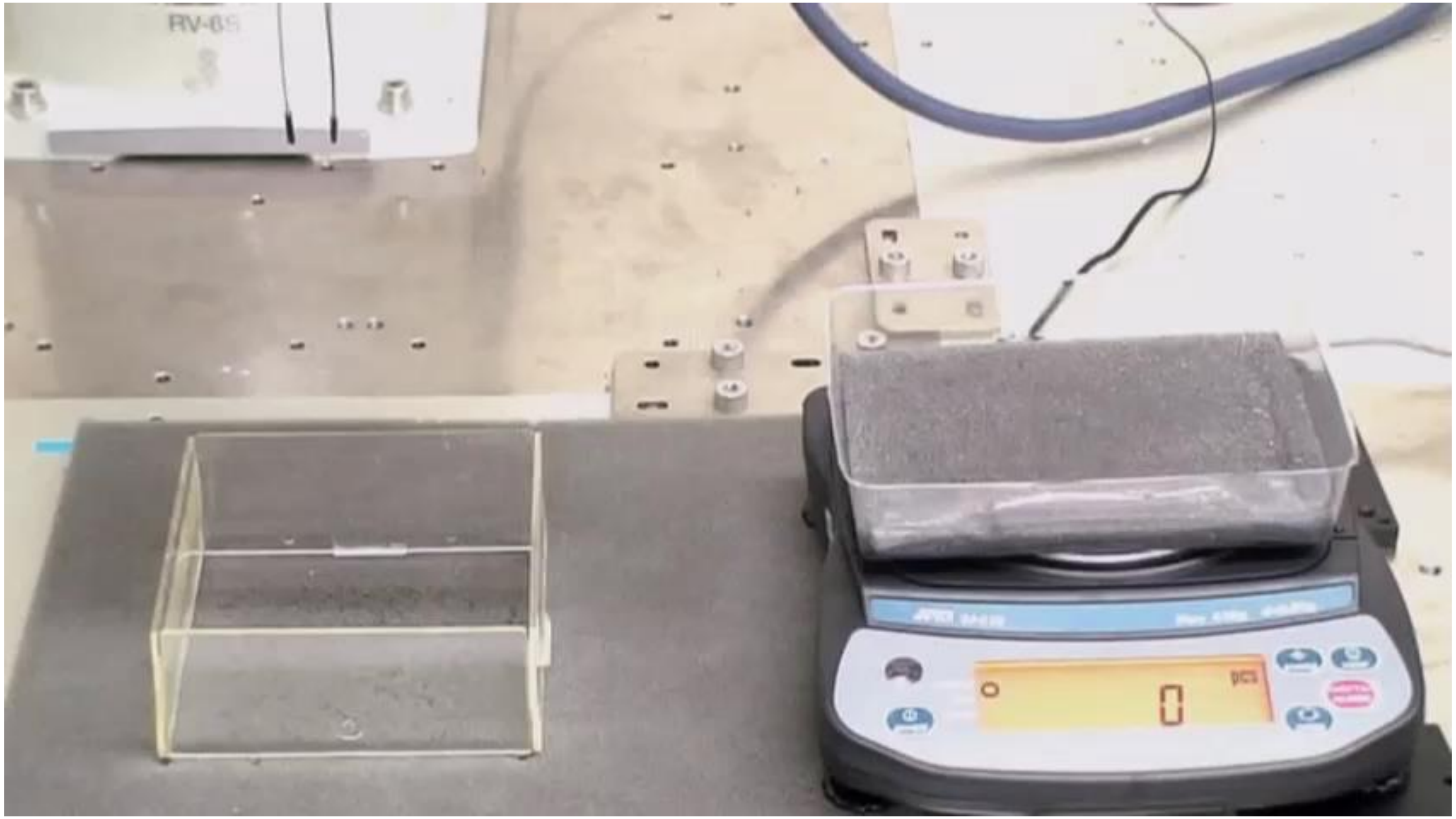
From parts supply to all assembly for products



Difficult task: parts supply



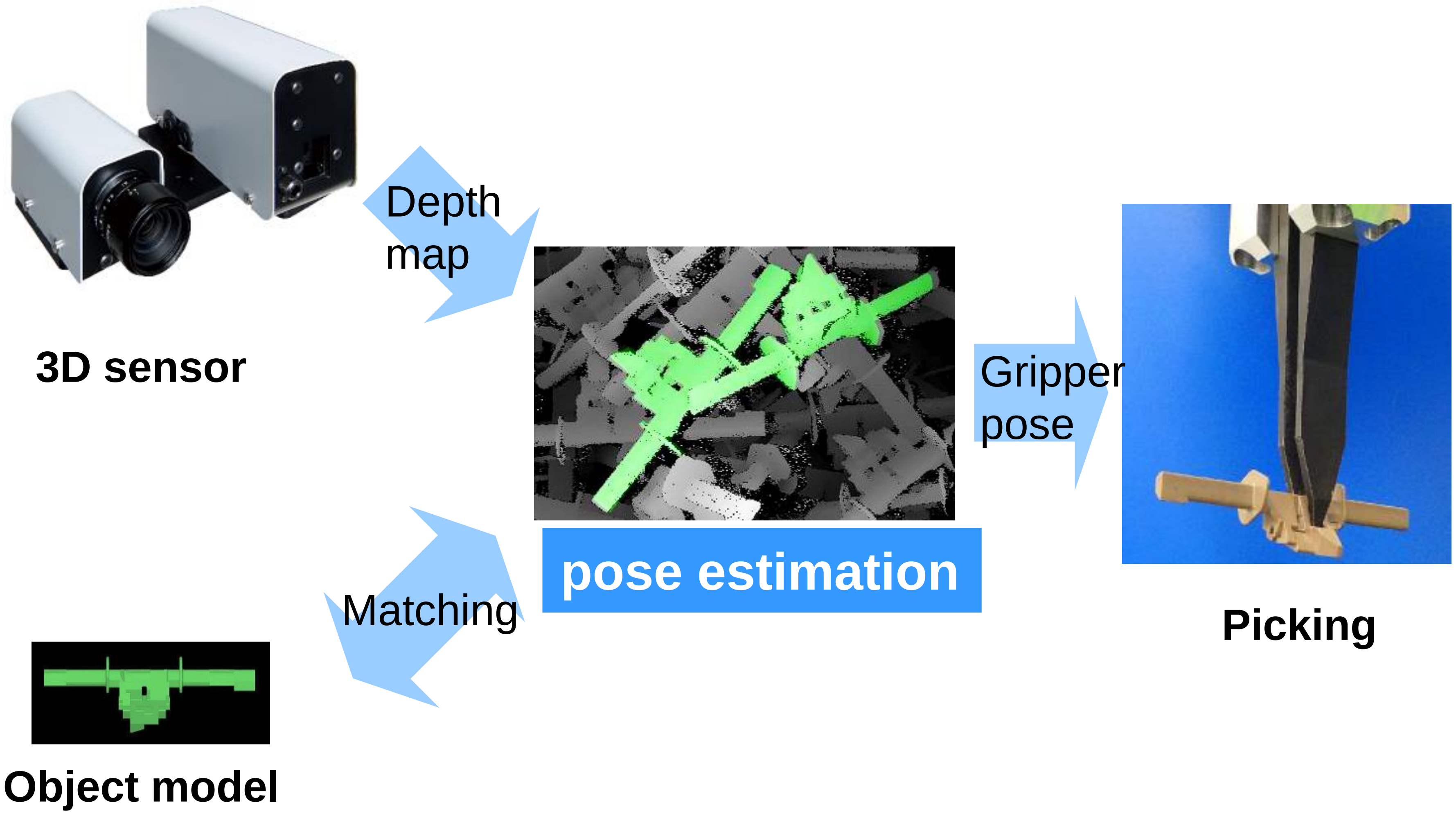
Developed technology



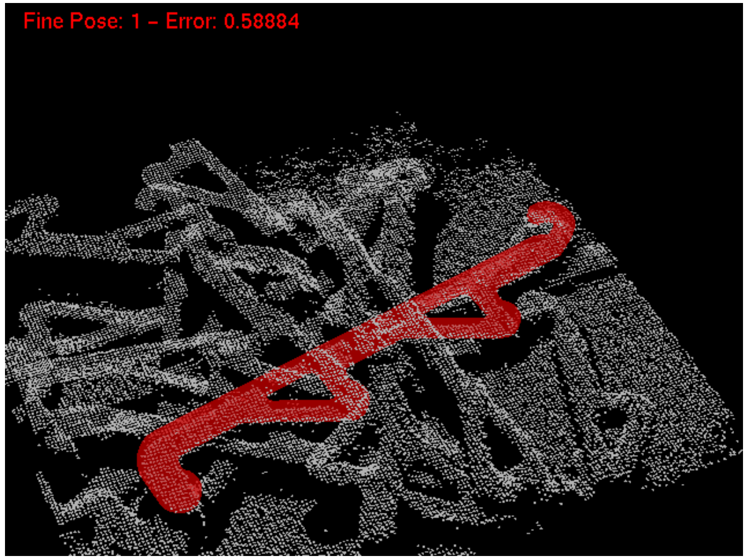
Fast picking from cluttered scenes

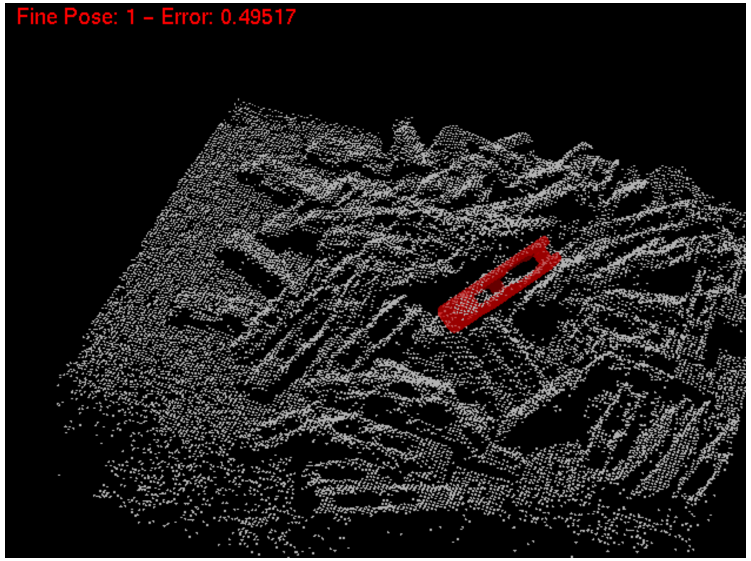
- R&D100 award 2014 winner -

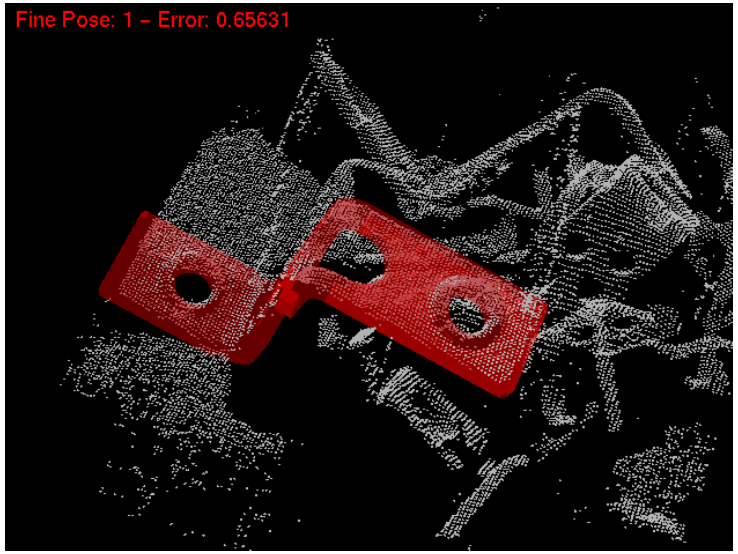
Basic recognition theory for picking

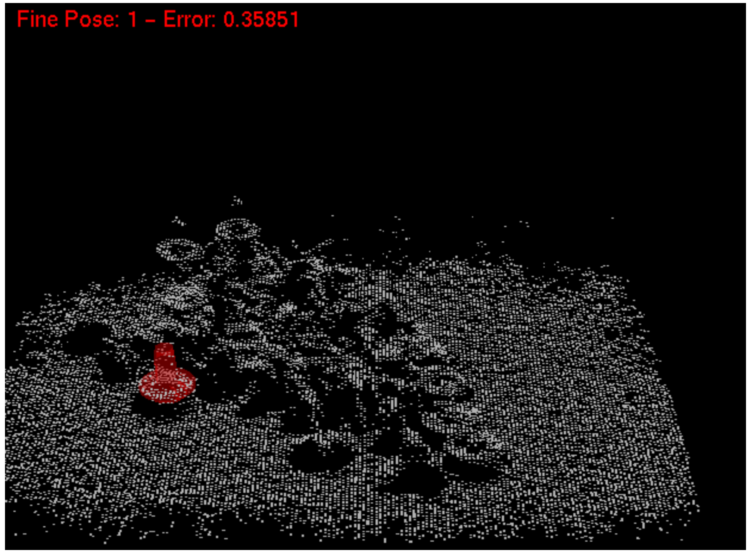


Examples of parts pose estimation





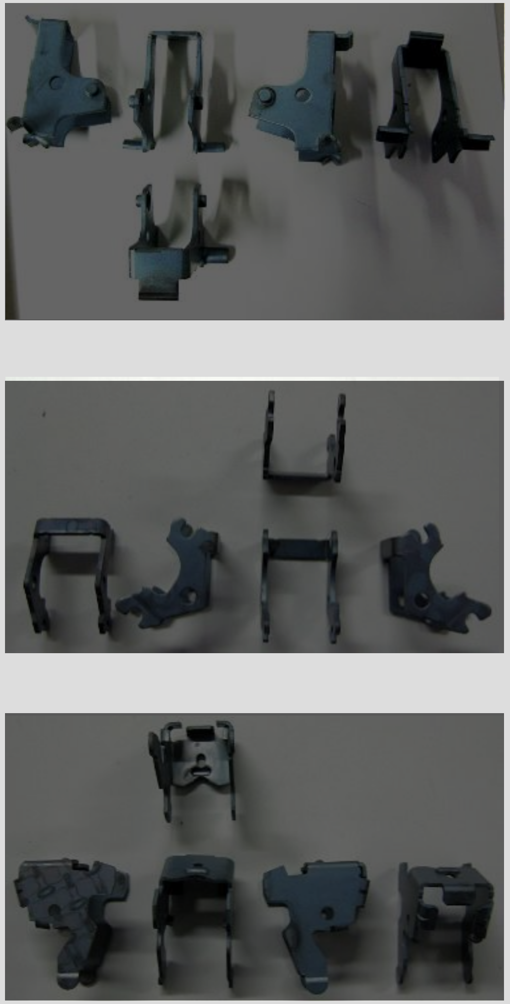




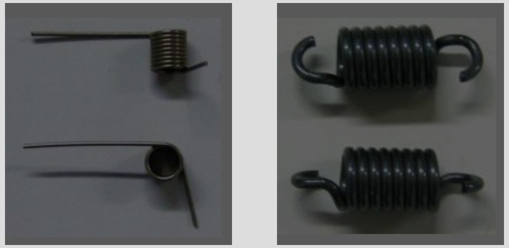
Applicable objects of pose estimation

Plane

Box

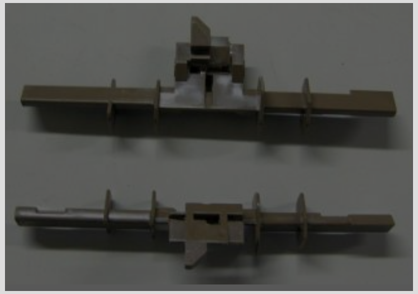


Spring



Cylinder

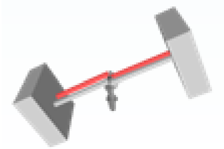
Others



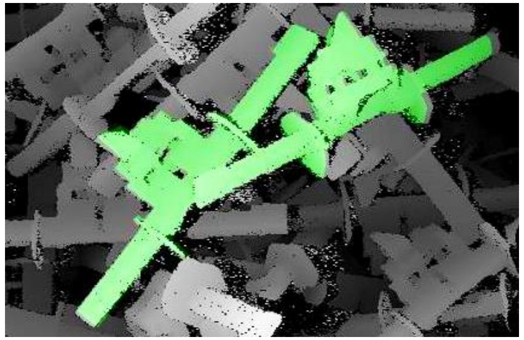
What's the problem for manipulation?

Need object pose estimation?

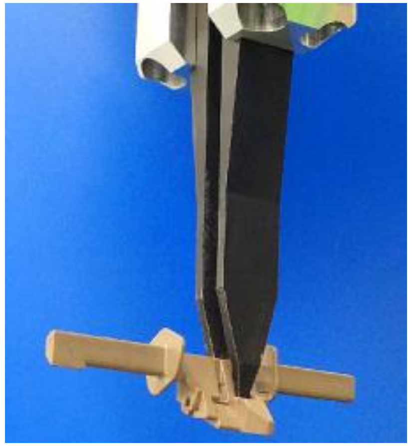
Need regrasping?
How to regrasp?

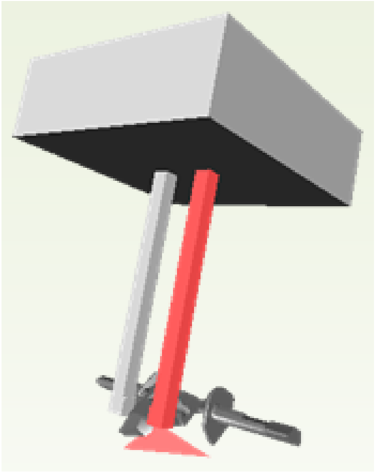


?



?





Finding

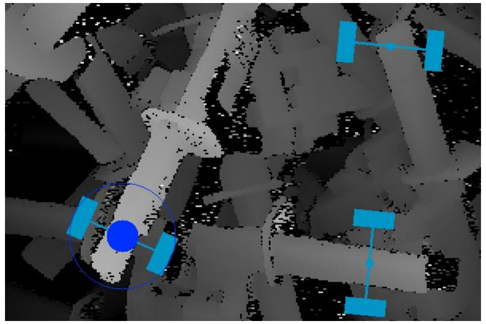
Grasping

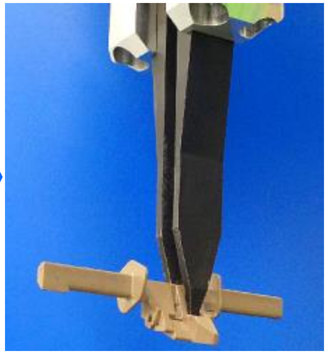
Kitting

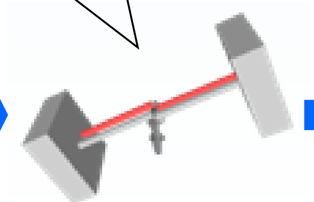
Practical approach for various object manipulation

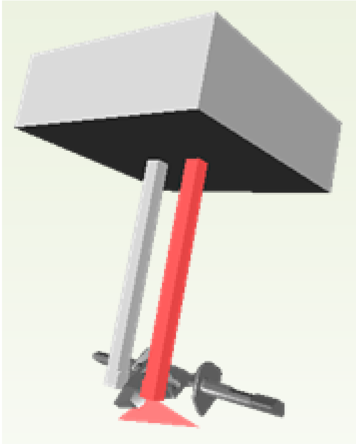
Fast graspability evaluation for object isolation

Regrasp motion planning



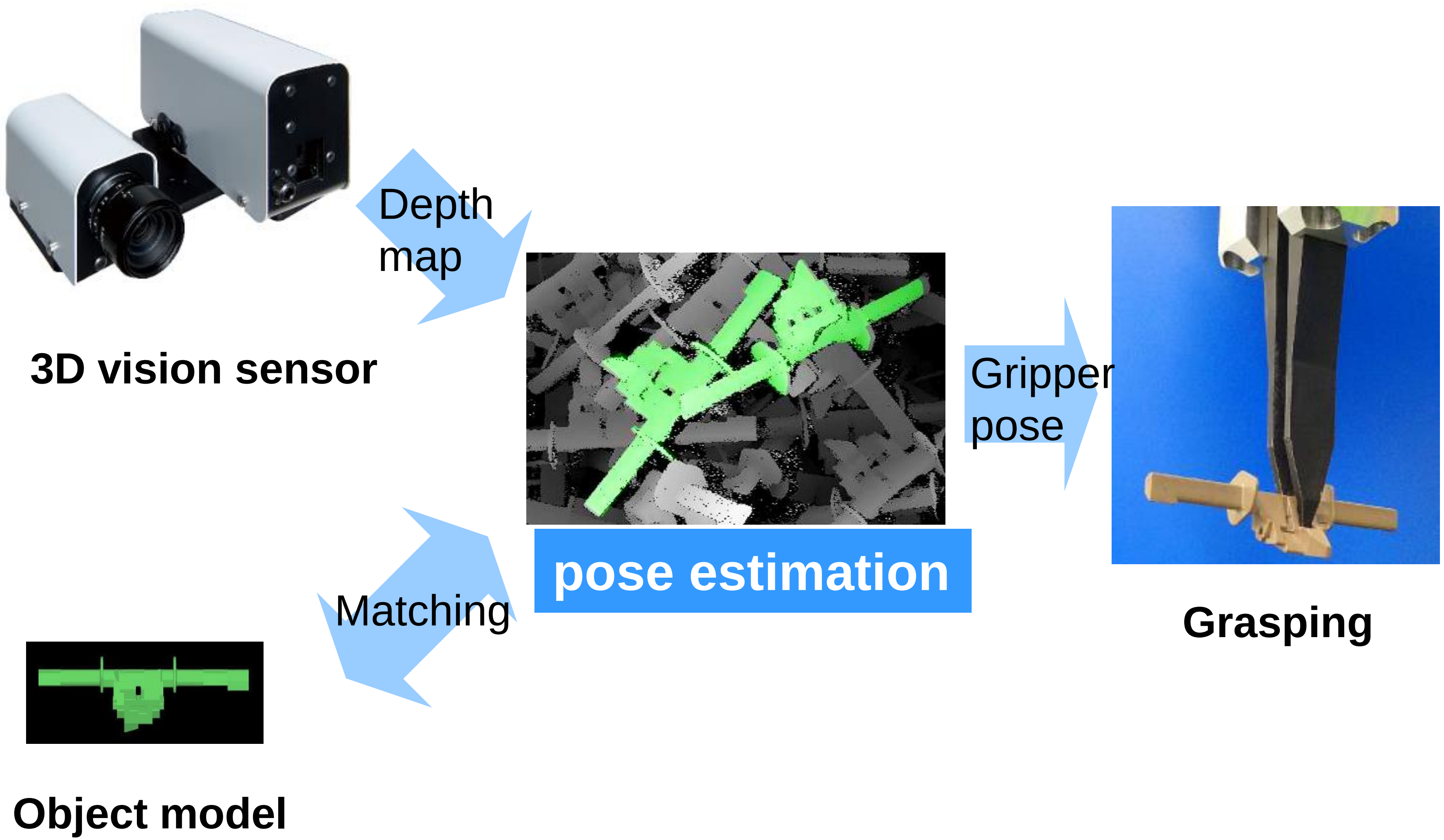




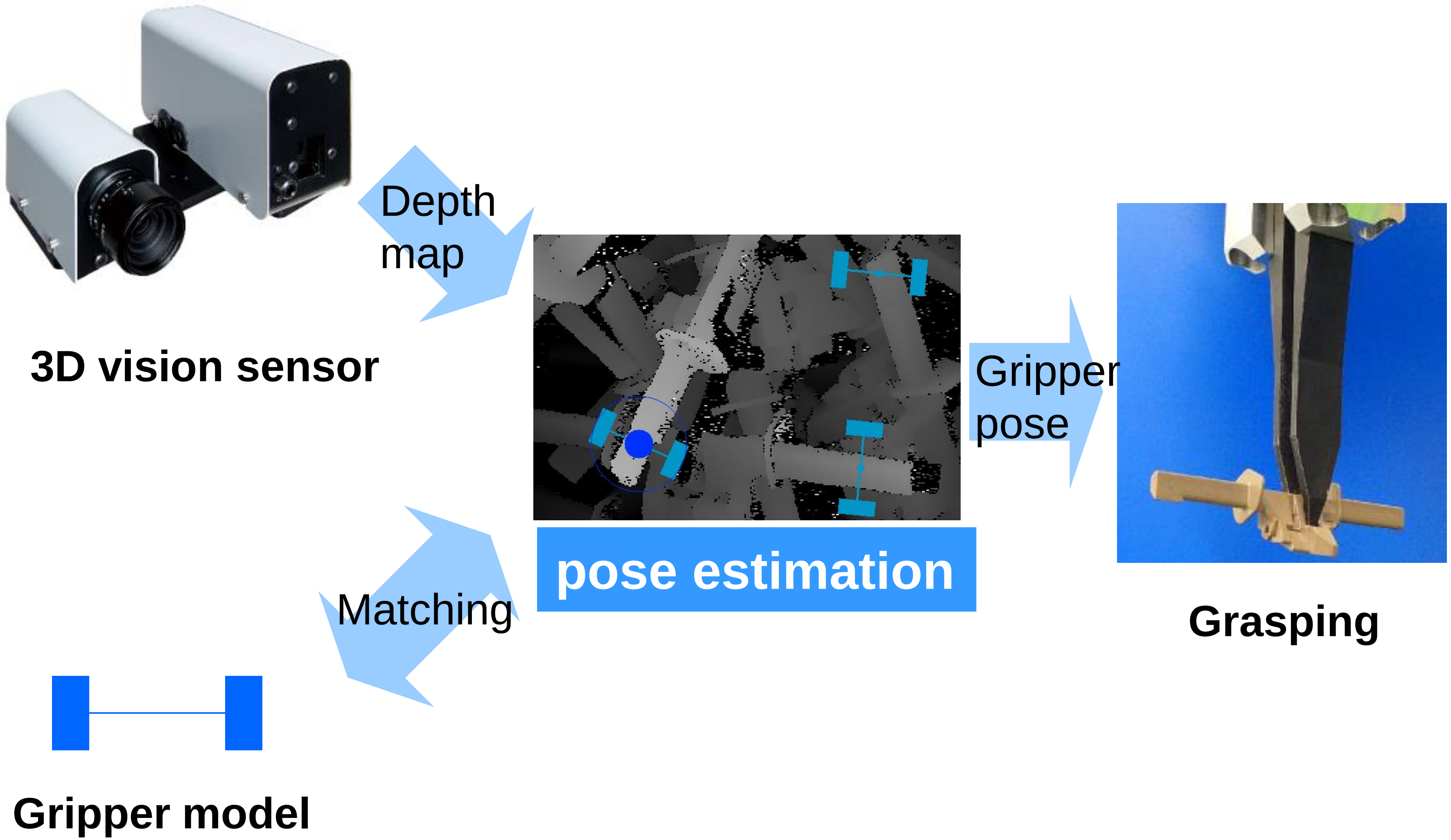


Finding **Grasping** **Placing** on a planer table **Regrasping** **Kitting**

Basic theory of pose estimation

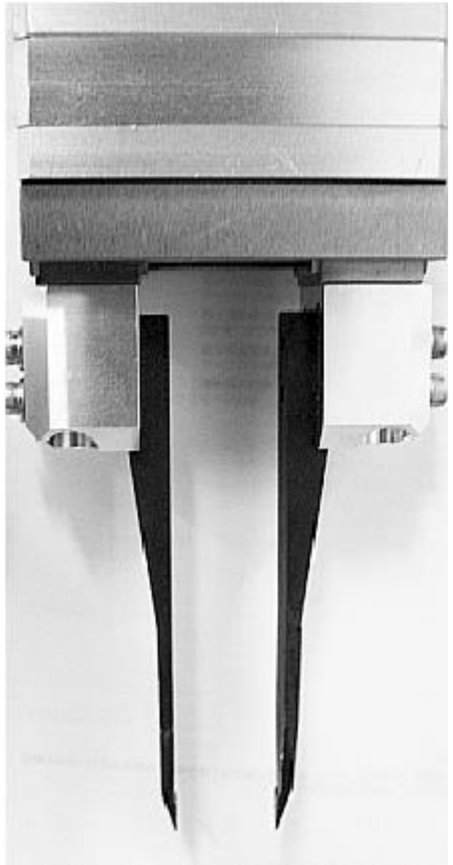


Basic thery of picking recognition



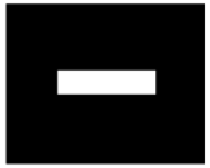
Y.Domae, H.Okuda, et.al. “Fast Graspability Evaluation on a Single Depth Maps for Bin-picking with General Grippers”, ICRA2014

Fast graspability evaluation

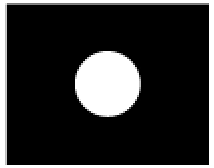


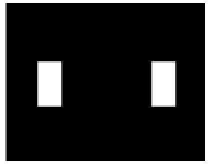
2-finger

For each grippers,
we define
its contact and collision region
using two masks



Contact region





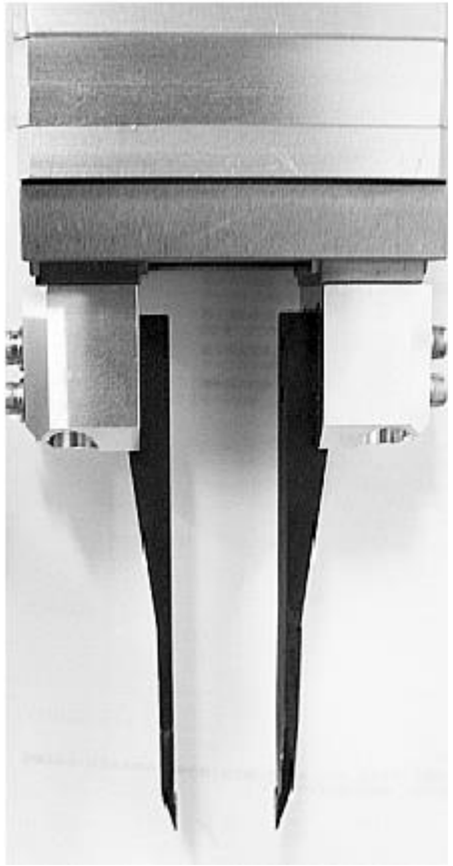
Collision region

Vacuum

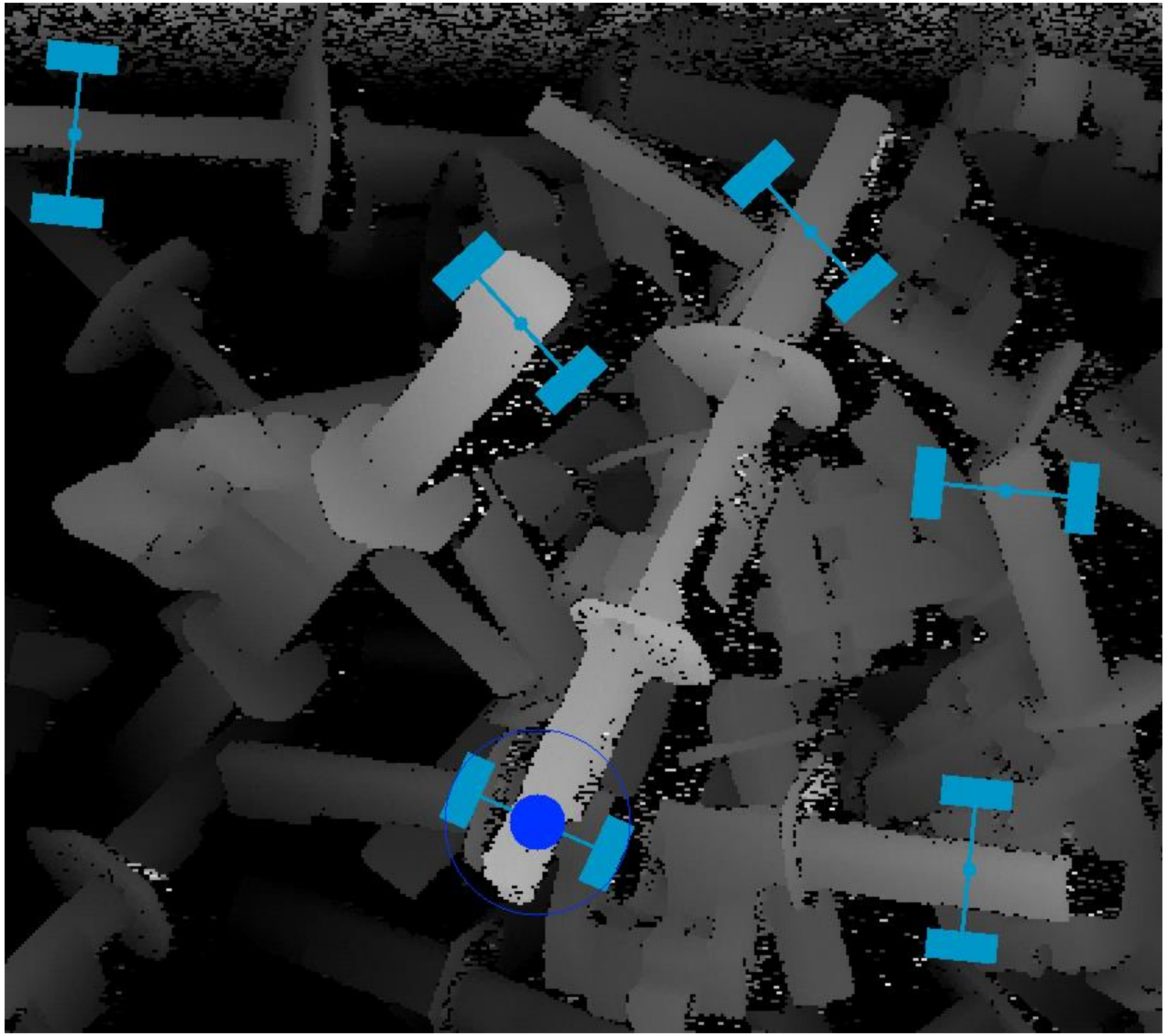
Y.Domae, H.Okuda, et.al. “Fast Graspability Evaluation on a Single Depth Maps for Bin-picking with General Grippers”, ICRA2014

Recognition examples

We compute the graspability map by filtering a depth map with the masks



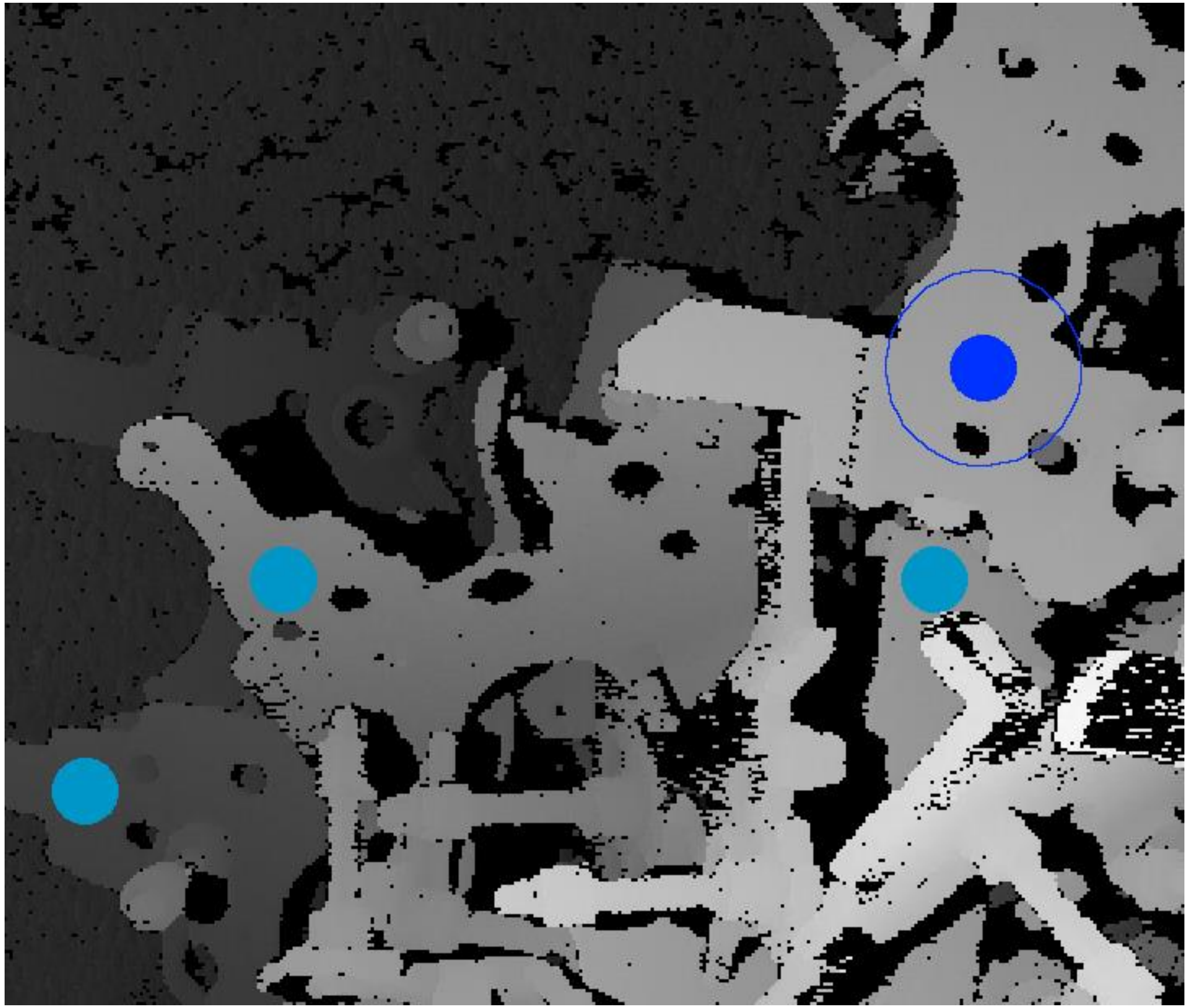
2-finger



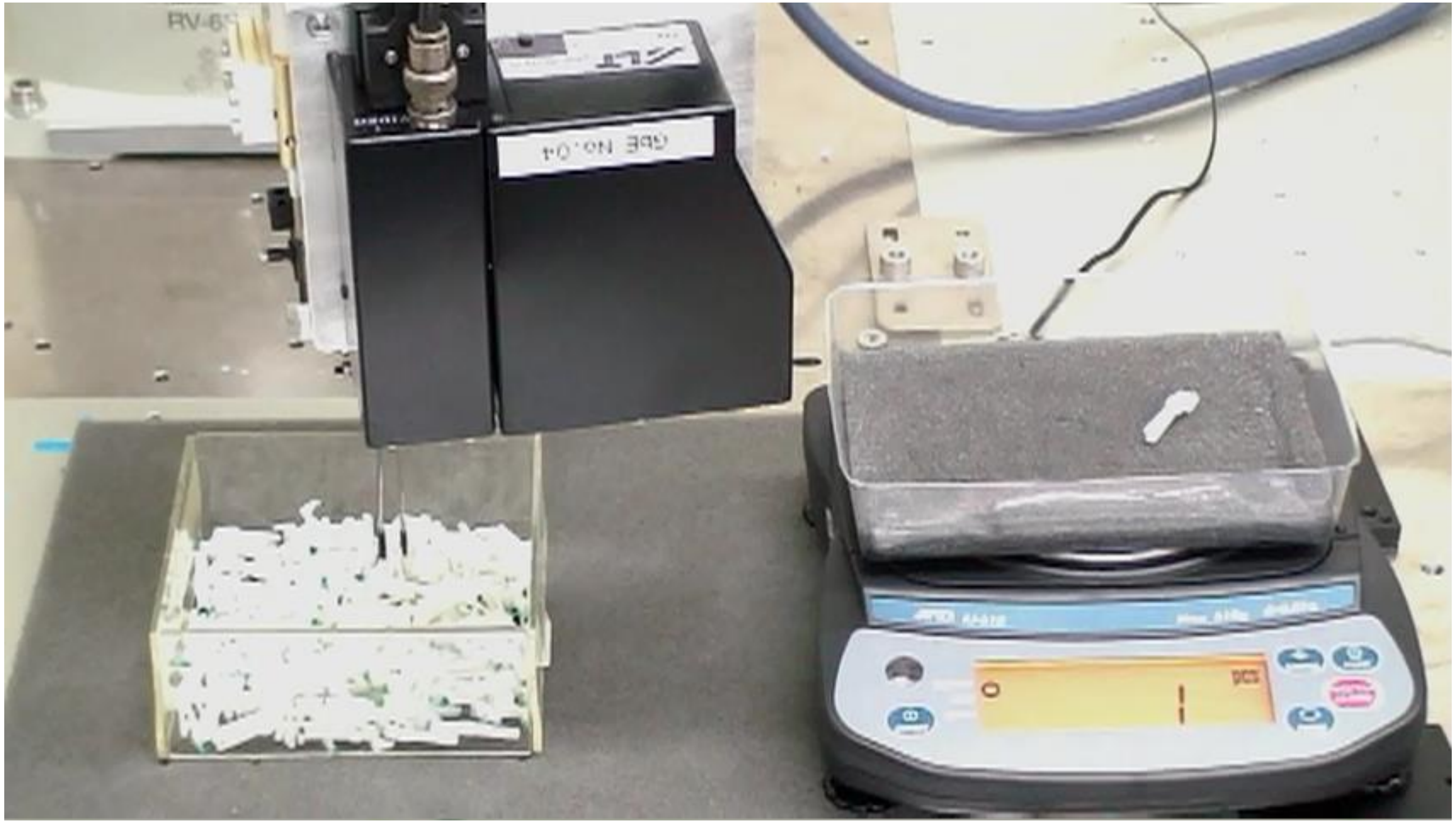
Recognition examples

We compute the graspability map by filtering a depth map with the masks

Vacuum



Demo

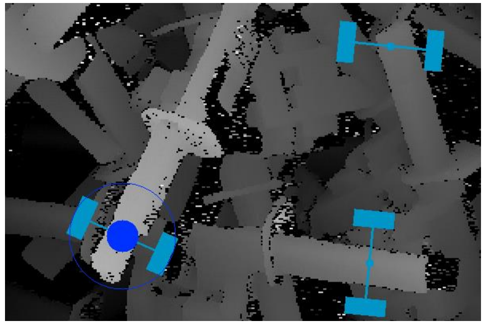


Y.Domae, H.Okuda, et.al. “Fast Graspability Evaluation on a Single Depth Maps for Bin-picking with General Grippers”, ICRA2014

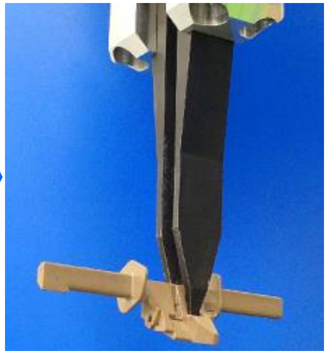
Regrasp method

Pick-and-place
for object isolation

Regrasp motion planning

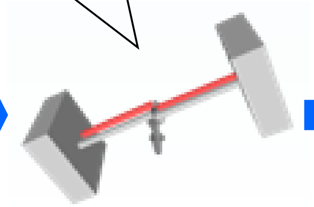


Finding

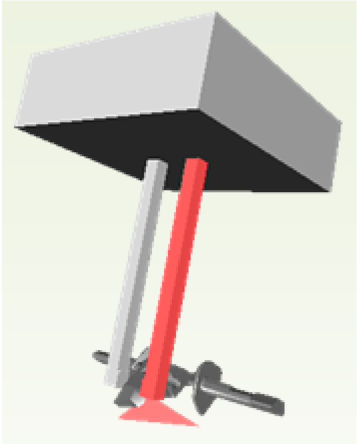


Grasping

Placing
on
a planer table

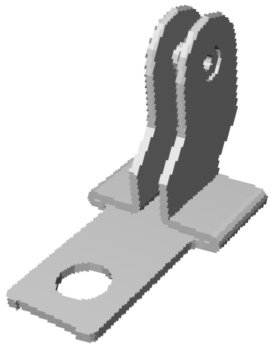


Regrasping



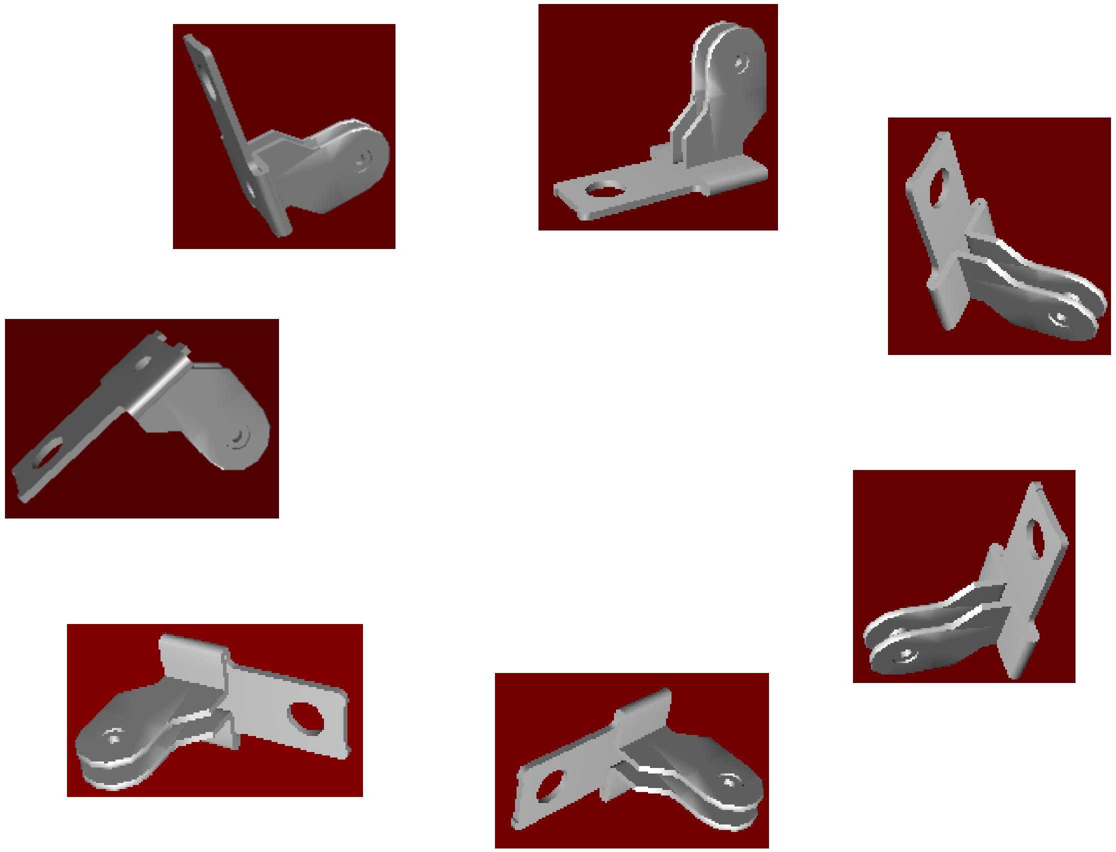
Kitting

Regrasp motion planning



Object model

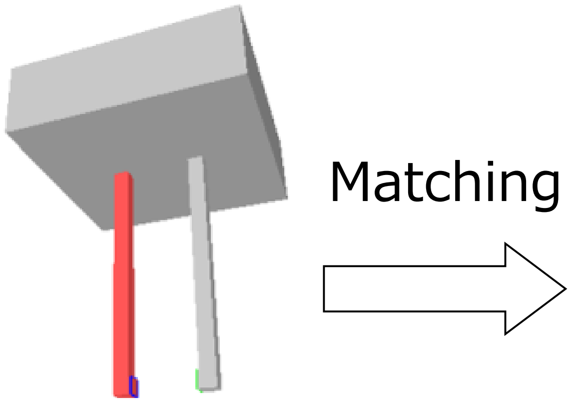
Physical simulation



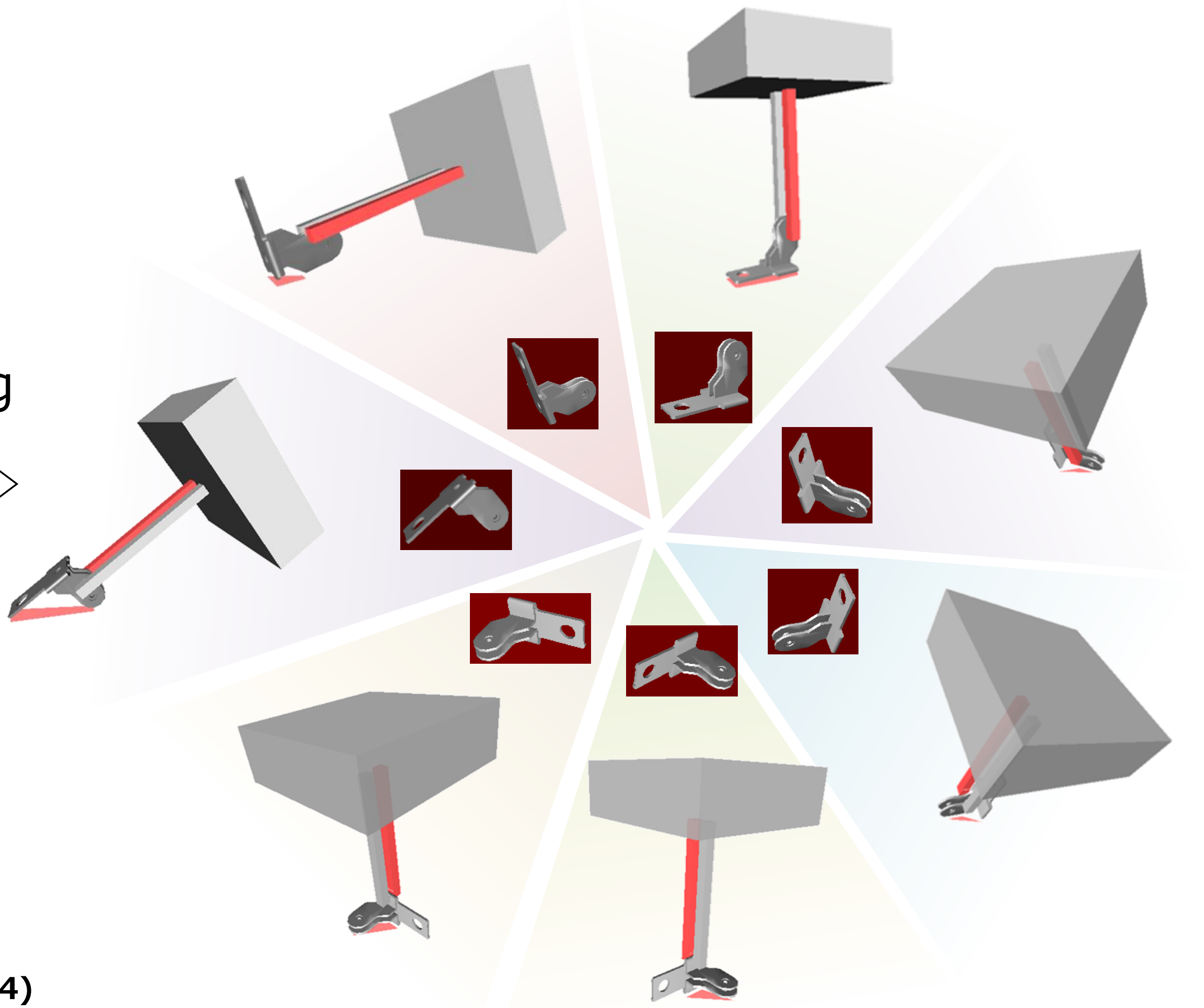
Stable poses on a planer table

Robot Programming for Assembly Tasks (ICRA2014)
http://gamma.cs.unc.edu/MPIR/

Regrasp motion planning



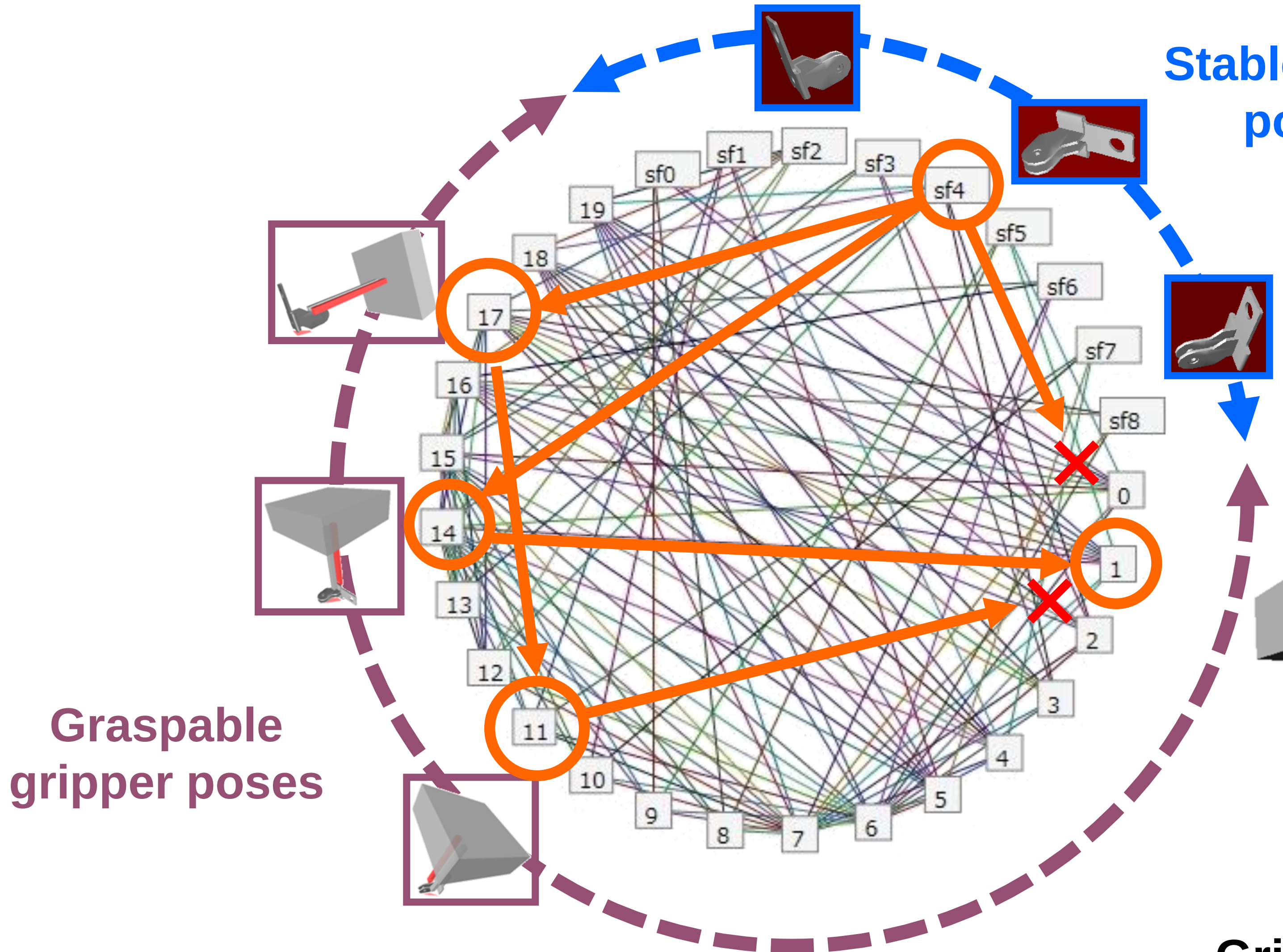
Gripper model



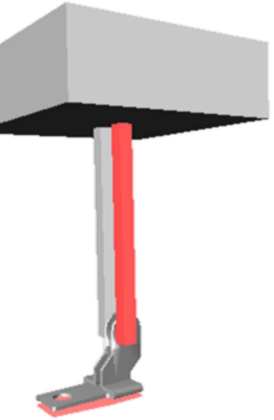
Graspable gripper poses

Robot Programming for Assembly Tasks (ICRA2014)
http://gamma.cs.unc.edu/MPIR/

Regrasp motion planning



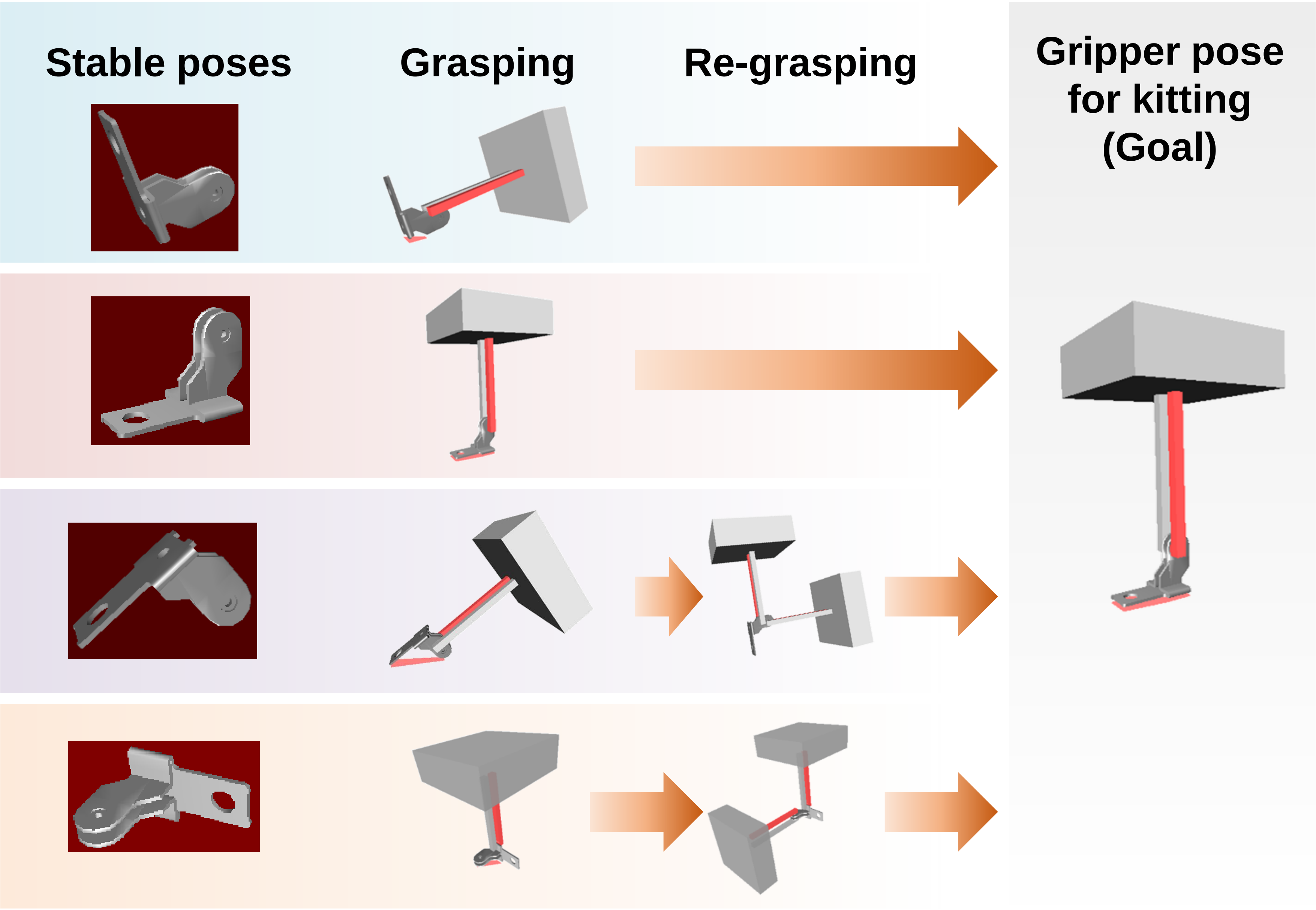
State transition graph



Gripper pose for kitting

Robot Programming for Assembly Tasks (ICRA2014)
http://gamma.cs.unc.edu/MPIR/

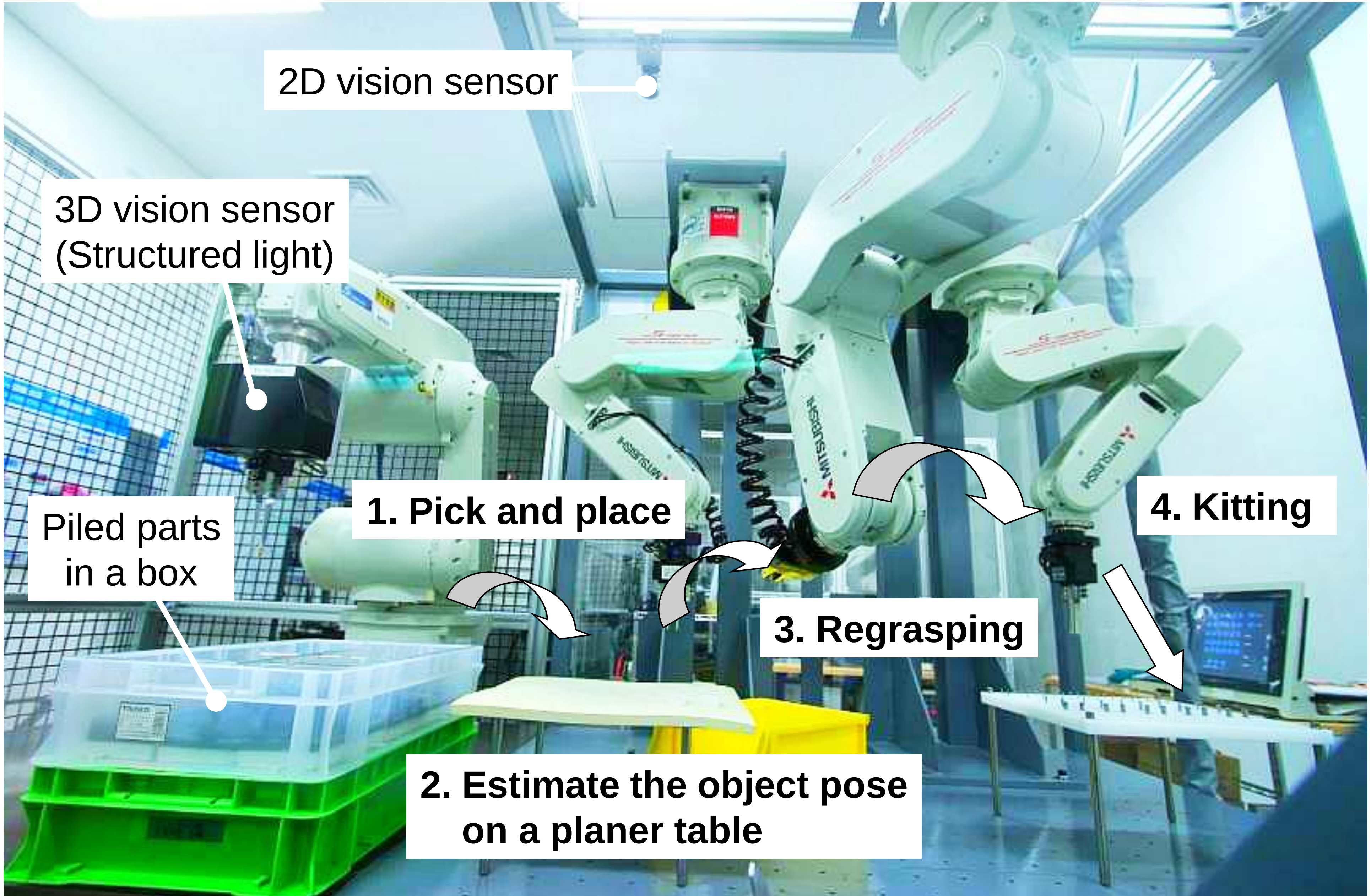
Planning examples



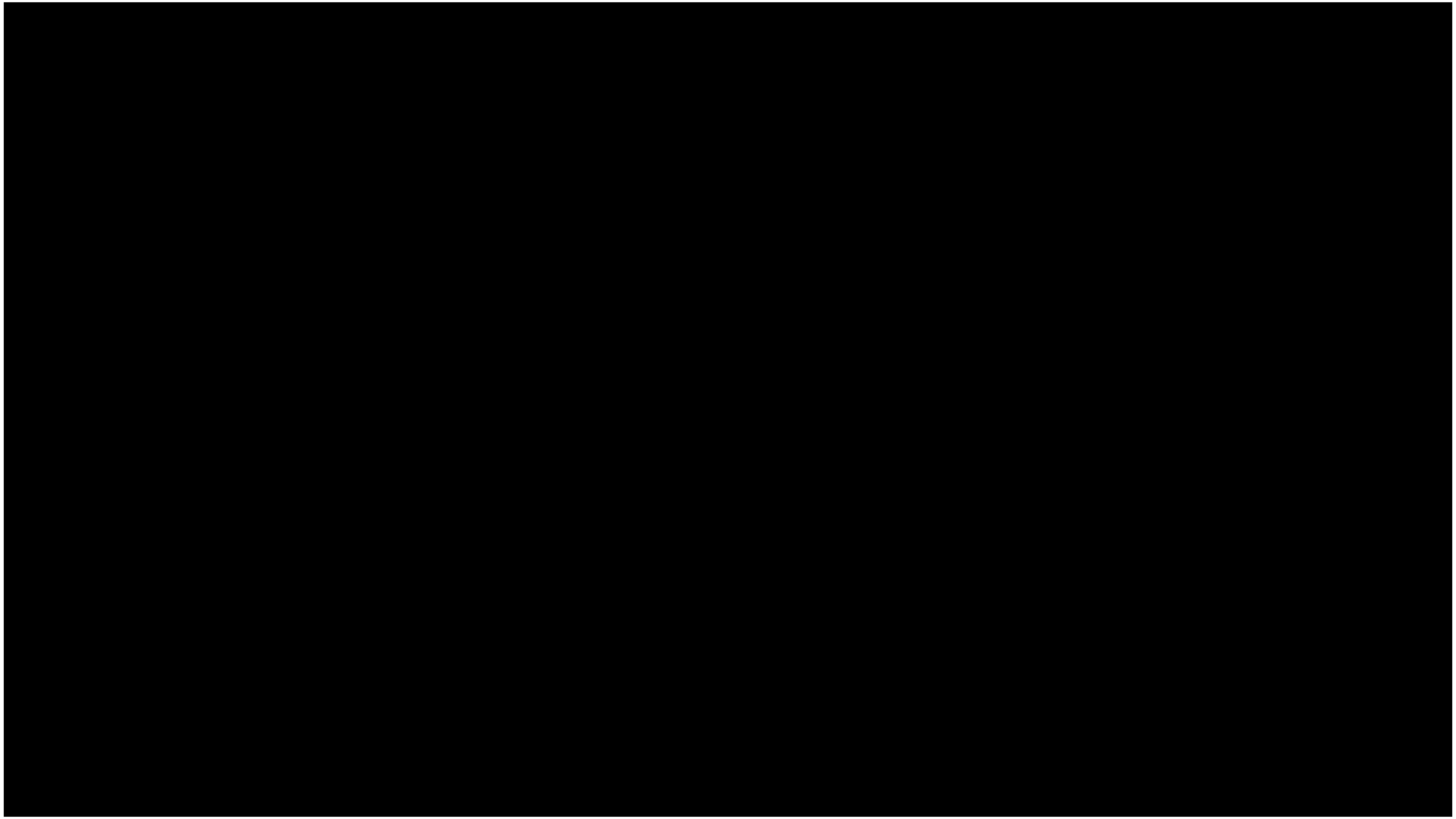
Example of mid-level planning for manipulation

A mid-level planning system for object reorientation
https://www.youtube.com/watch?v=X8Ltgs7ppsk

General parts feeding system



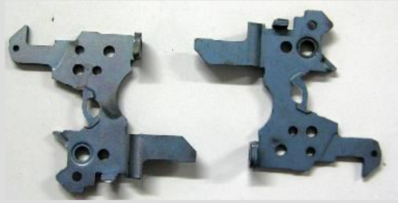
Demo



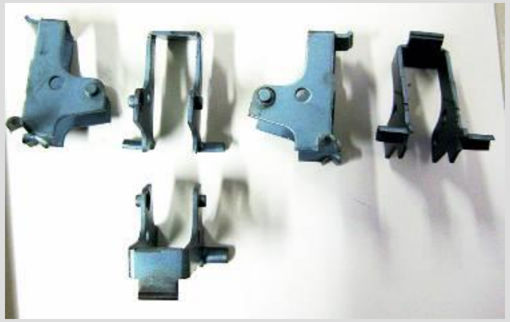
出典：一般形状部品の多品種供給の自動化
日本ロボット学会誌，Vol33, No.5, pp.387-394, 2015

Applicable objects(ALL)

Plane



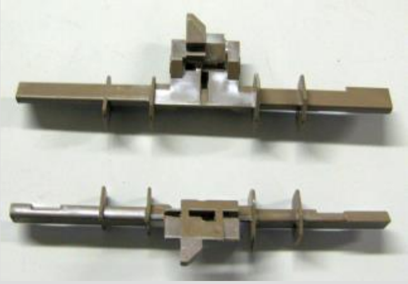
Box



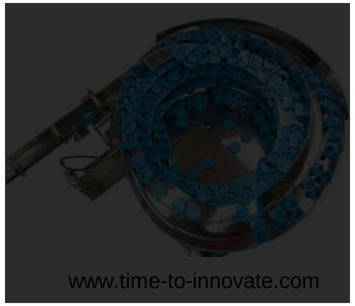
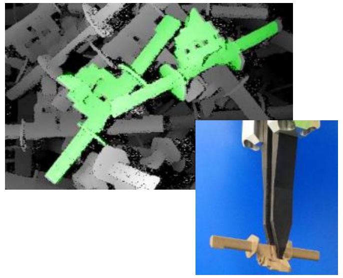
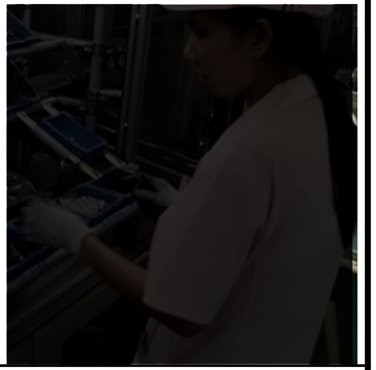
Spring

Cylinder

Others

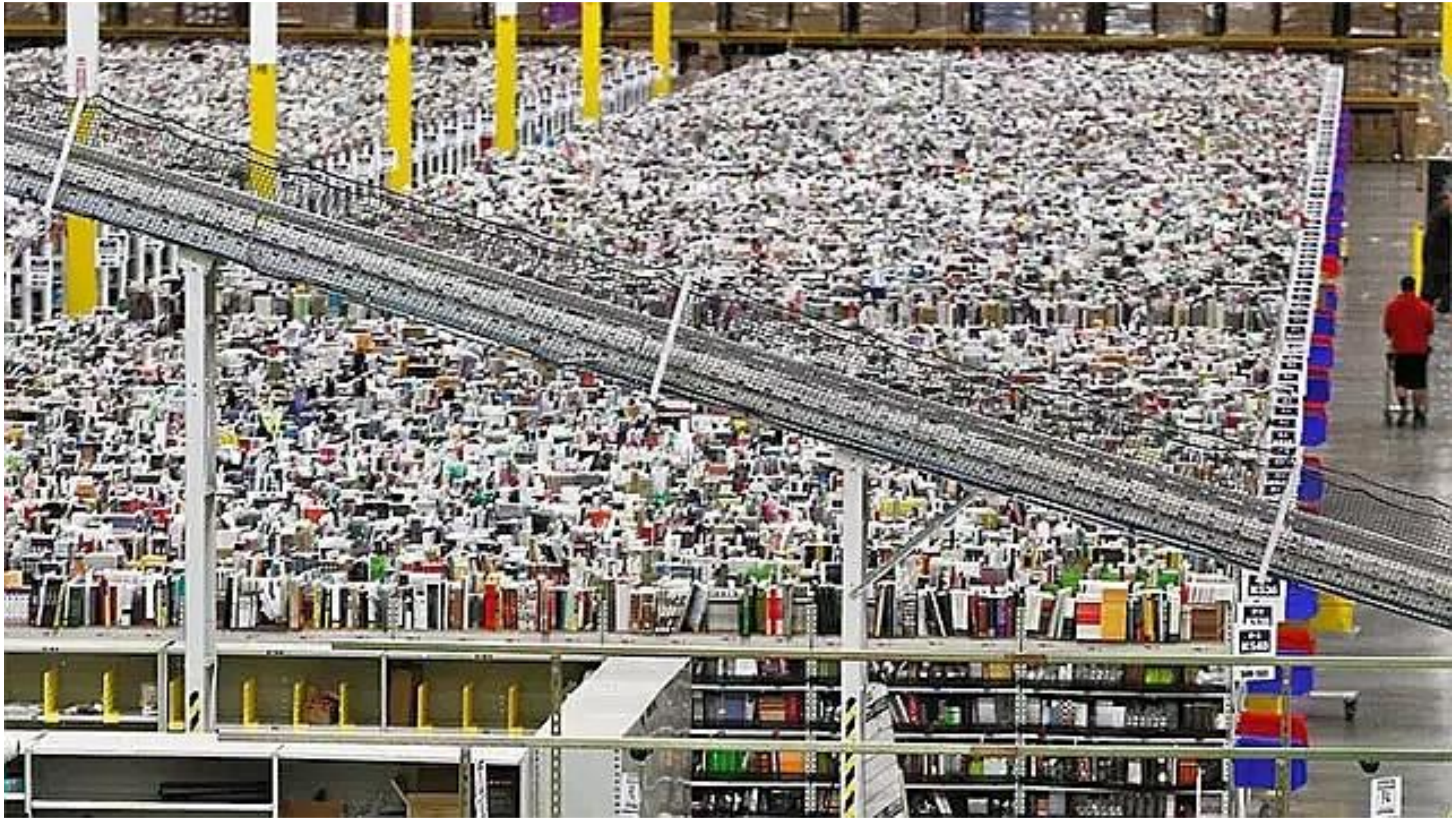


Comparisons

	Our method 	Parts feeders 	Traditional robot system 	Manual labor 
Arbitrary part shapes	OK	NG	NG	OK
Cycle time	3～5 seconds	1～2.5 seconds	3.5～10 seconds	1～3 seconds
Lead time for product change	2～3 days for robot programming	1 month for H/W renewal	2 weeks for S/W renewal	1 hour for starting 2 weeks for mastership

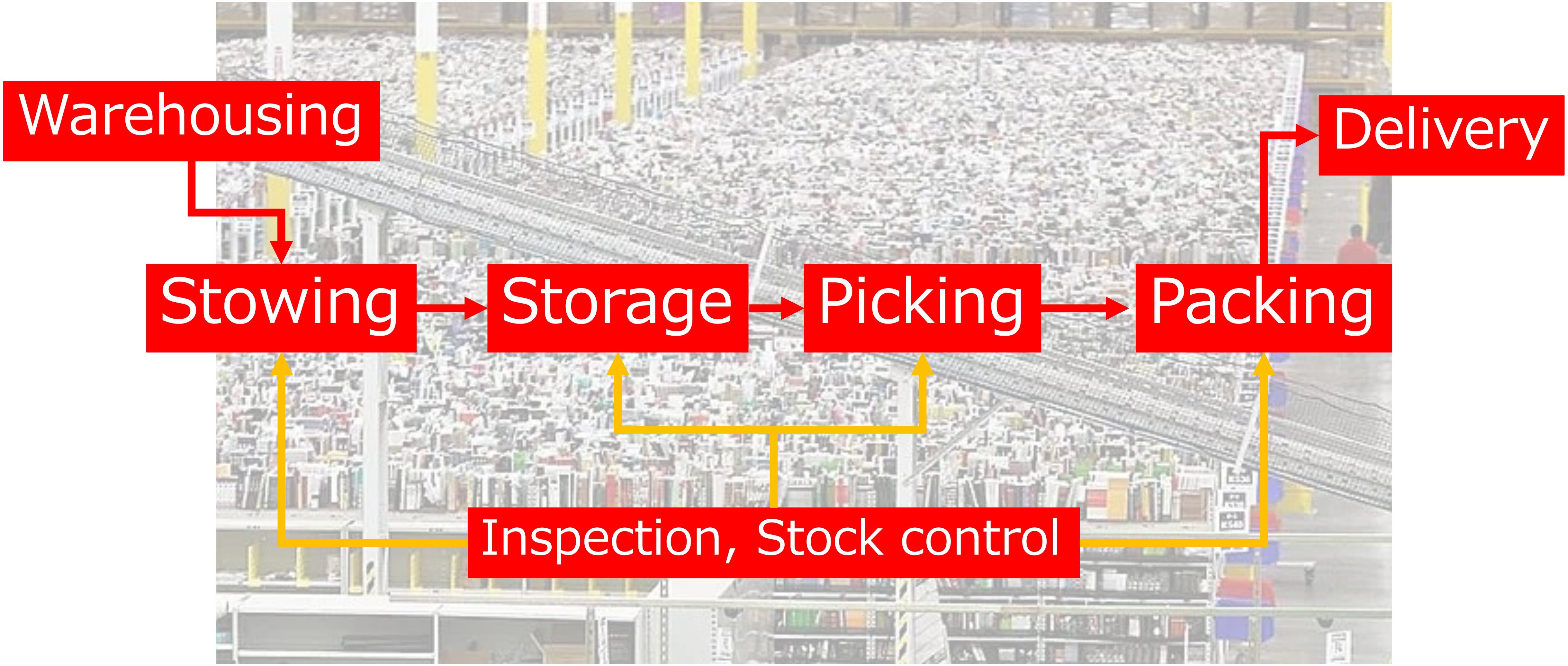
2. Machine vision problems in Warehouse Automation

Inside look at an Amazon warehouse



BGR: **An inside look at an Amazon warehouse**
http://bgr.com/2012/11/30/amazon-warehouse-an-inside-look/

Warehouse working processing



出典：BGR: An inside look at an Amazon warehouse
http://bgr.com/2012/11/30/amazon-warehouse-an-inside-look/

Trend on warehouse automation

- From mass logistics to variable logistics -

High flexibility is also needed to warehouse.

Daifuku: Shuttle track
http://www.daifuku-logisticssolutions.com/image.jsp?id=2234

Roogato: Amazon using robot automation: Says human employees still needed
http://roogato.com/amazon-using-robot-automation-says-human-employees-still-needed/

Autonomous mobile for flexibility

Tablet monkeys : Amazon Warehouse Robots
https://www.youtube.com/watch?v=quWFjS3Ci7A

Probrem : Picking various items

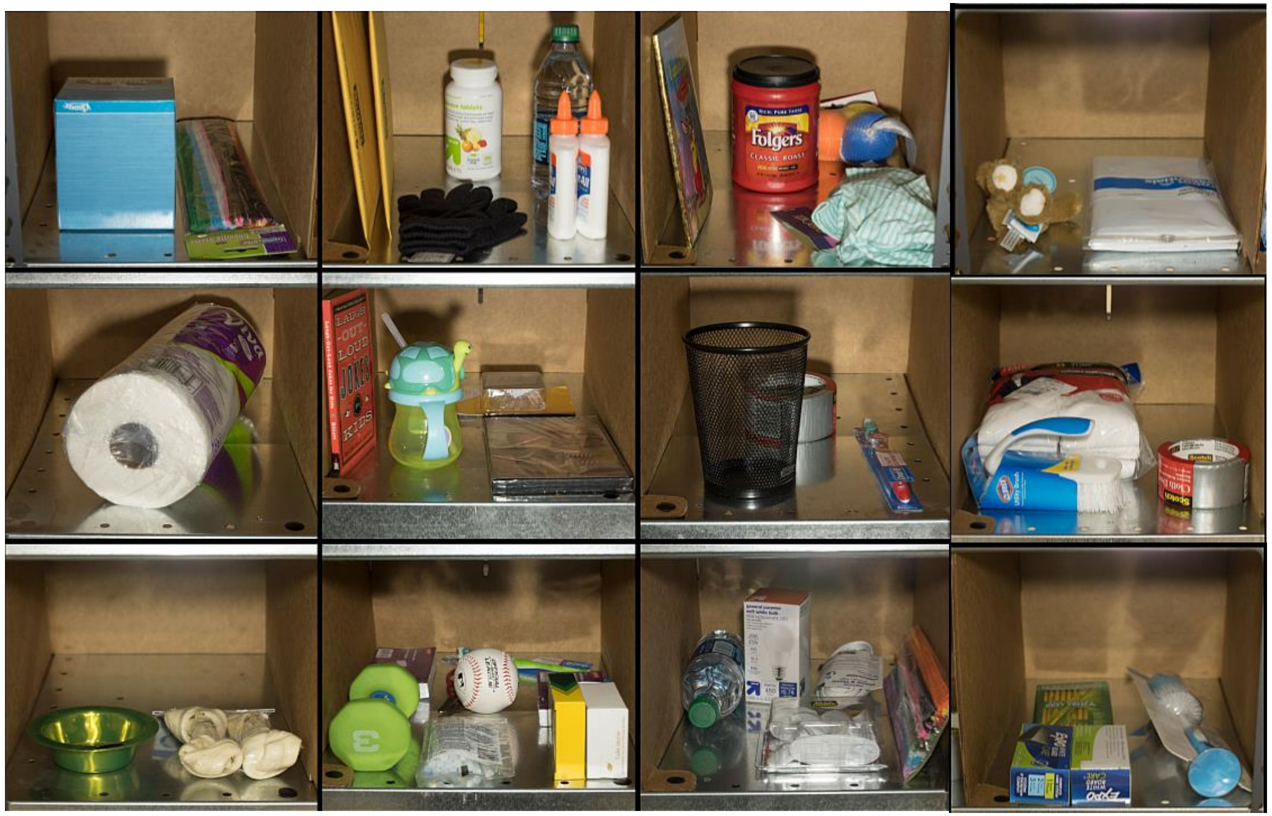
Picking is by hand in warehouses

Amazon Picking Challenge
http://amazonpickingchallenge.org/

CG demo of mobile manipulator

I AM ROBOTICS: Swift – Mobile Picking Robot
https://www.iamrobotics.com/

Problem as pattern recognition



Amazon Picking Challenge 2016 例題

1．Item classification　→　Specific object recognition

2．Picking　→　pose estimation, grasp planning

Problem as pattern recognition

1． various types of items
- Measurement of translucent, transparent, shiny, brack items




- Recognition of flexible items


2． Cluttered scene
- Overlap, hidden


Amazon Picking Challenge (APC)

Amazon Picking Challenge 2015

Target items in APC 2015

Team C^2M (Chubu, Chukyo, Mitsubishi)

CHUBU UNIVERSITY AICHI JAPAN

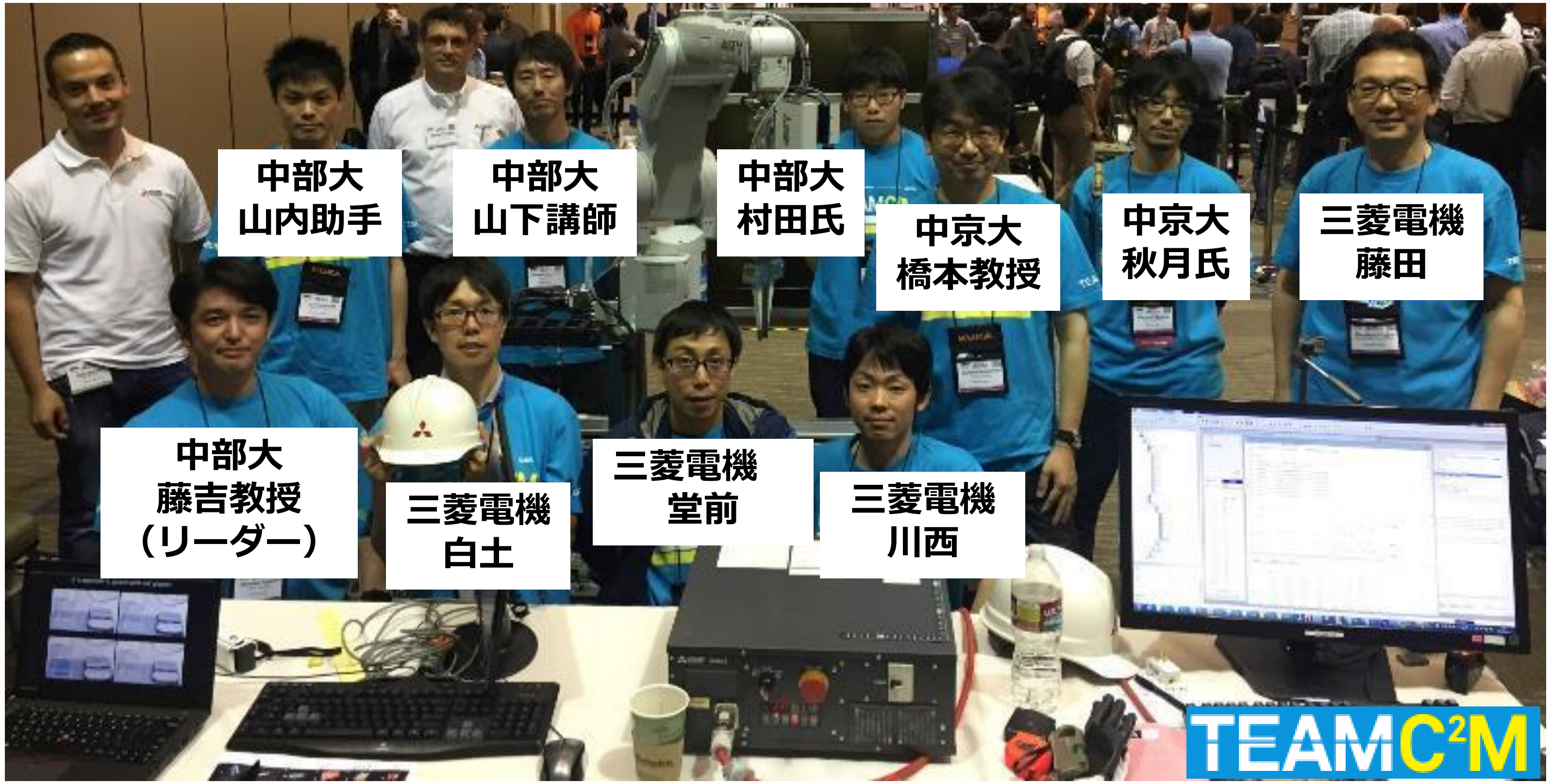
MPRG
MACHINE PERCEPTION AND ROBOTICS GROUP

CHUKYO UNIVERSITY

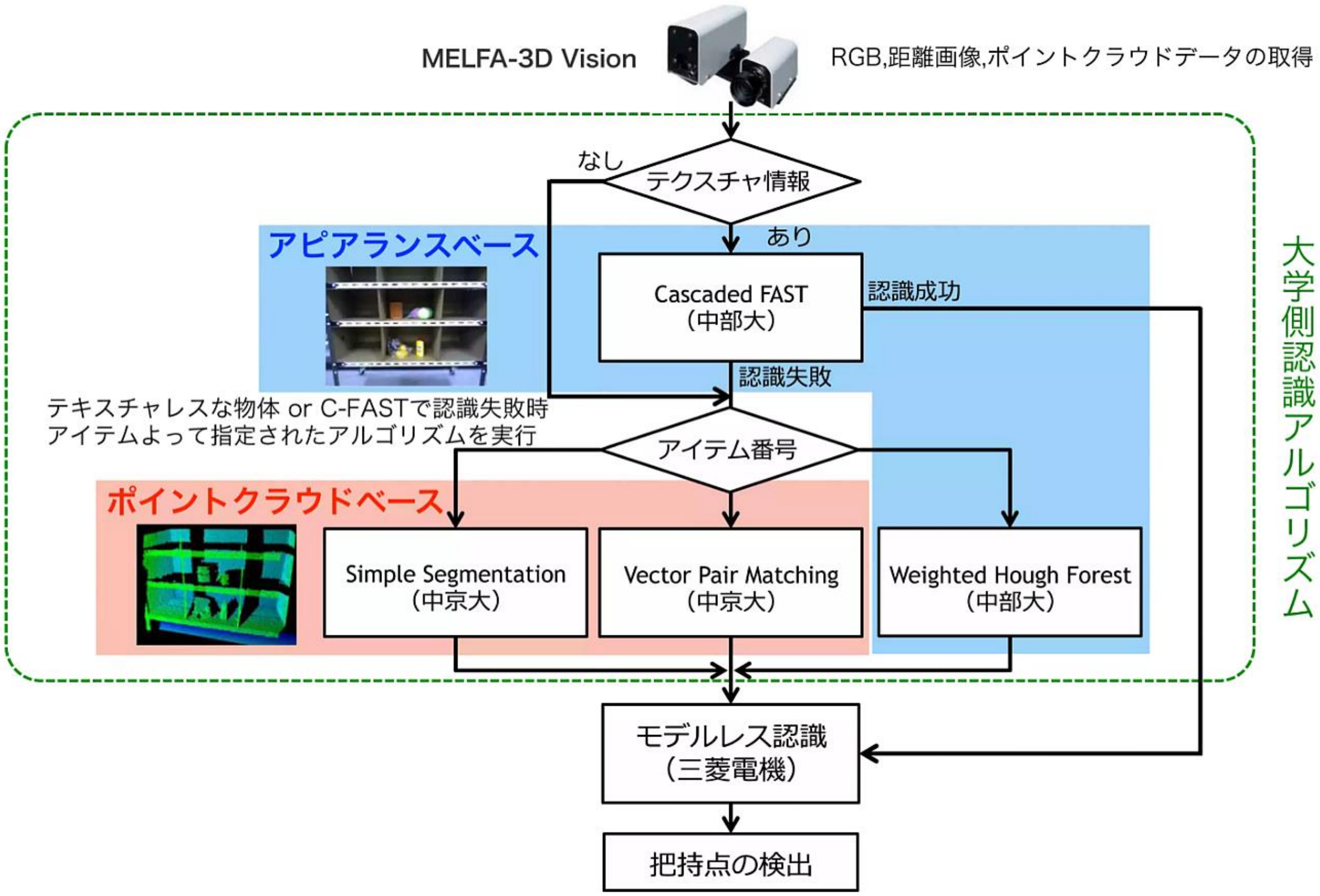
Intelligent Sensing Laboratory
Hashimoto Lab., Chukyo University

MITSUBISHI ELECTRIC

Advanced R&D Center

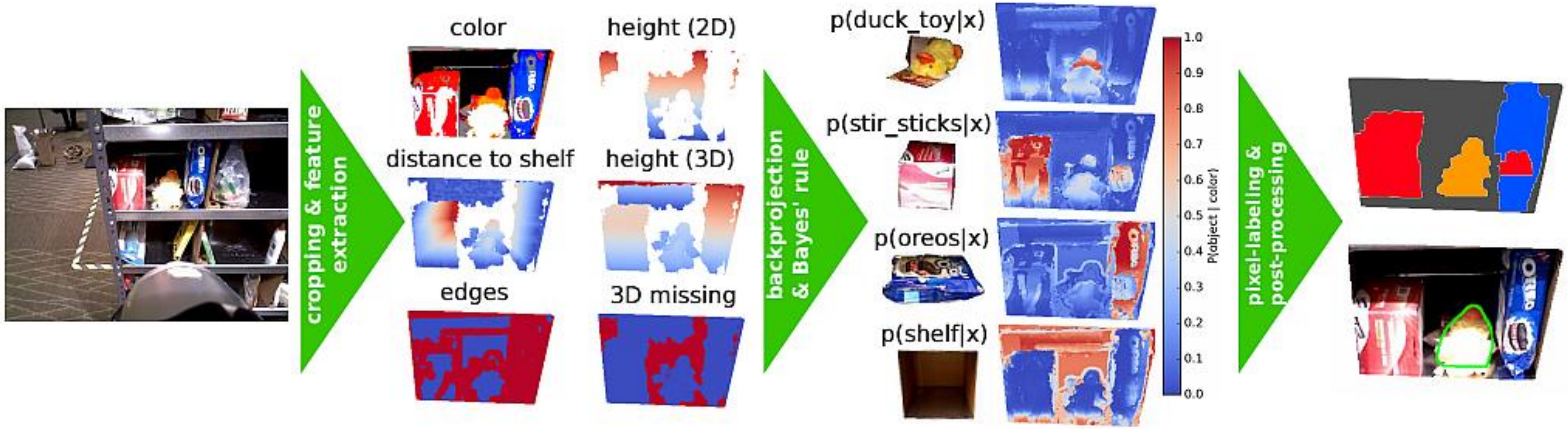


Recognition approach of Team C^2M



mrpg: Amazon Picking Challenge 2015: Team C^2M
http://mprg.jp/research/amazon_picking_challenge_j

Recognition approach of Team RBO(2015 winner)



- Feature extraction from RGBD images
- Item classification for each pixels by using Bayes rule
- Pixel labeling and item segmentation

RBO: Probabilistic Multi-Class Segmentation for the Amazon Picking Challenge
http://www.robotics.tu-berlin.de/fileadmin/fg170/Publikationen_pdf/Jonschkowski-16-Technical-Report.pdf

Recognition trends in APC 2015

Team	Sensor	Perception	Motion Planing
RBO	3D imaging on Arm, Laser on Base, Pressure sensor, Force-torque sensor	Multiple features (color, edge, height) for detection and filtering 3D bounding box for grasp selection	No
MIT	Both 2D and 3D imaging on Head and Arm	3D RGB-D object matching	No
Grizzly	2D imaging at End-effector, 3D imaging for head, and laser for base	3D bounding box segmentation and 2D feature based localization	Custom motion planning algorithm
NUS Smart Hand	3D imaging on Robot	Foreground subtraction and color histogram classification	Predefined path to reach and online cartesian planning inside the bin using MoveIt.
Z.U.N.	(respondent skipped response)	(respondent skipped response)	MoveIt RRT Planning for reaching motion and use pre-defined motion inside bin
C^2M	3D imaging on End-effector and force sensor on arm	RGB-D to classify object and graspability	No
Rutgers U. Pracsys	3D imaging on Arm	3D object pose estimation	Pre computed PRM paths using PRACSYS software & grasps using GraspIt
Team K	3D imaging on Arm and Torso	Color and BoF for object verification	No
Team Nanyang	3D imaging on End-effector	Histogram to identify object and 2D features to determine pose	No
Team A.R.	3D imaging on End-effector	Filtering 3D bounding box and matching to a database	No
Georgia Tech	3D imaging on Head and Torso	Histogram data to to recognize and 3D perception to determine pose	Pre-defined grasp using custom software and OpenRave
Team Duke	3D imaging on End-effector	3D model to background subtraction and use color / histogram data.	Klamp't planner to reaching motion
KTH/CVAP	3D/2D imaging on head, Tilting laser on Torso and Laser on Base	Matched 3D perception to a stored model	Move to 6 pre-defined working pose and use MoveIt to approach and grasp object

1．Senor
- Cheap and low resolution sensor
- 2D and 3D imaging
- Laser scanner

2．Classification
- RGBD-feature-based (color, edge, etc)
- Pose estimation (bounding box, graspability)

→ No deep learning

Lessons from the Amazon Picking Challenge
http://www.mathpubs.com/detail/1601.05484v2/Lessons-from-the-Amazon-Picking-Challenge

Picking demo of APC 2015

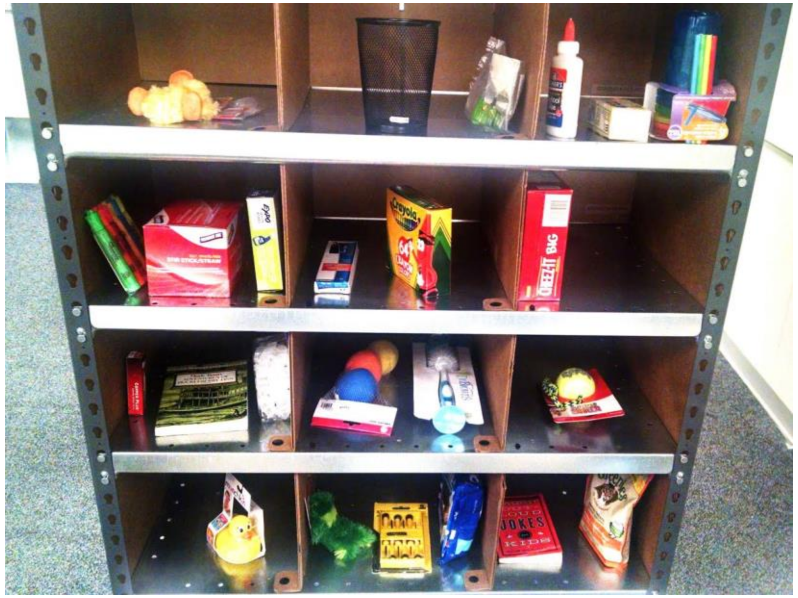
Picking Robot: Mitsubishi Electric Corp.
https://www.youtube.com/watch?v=AEKEce_ZKgg

Amazon Picking Challenge 2016

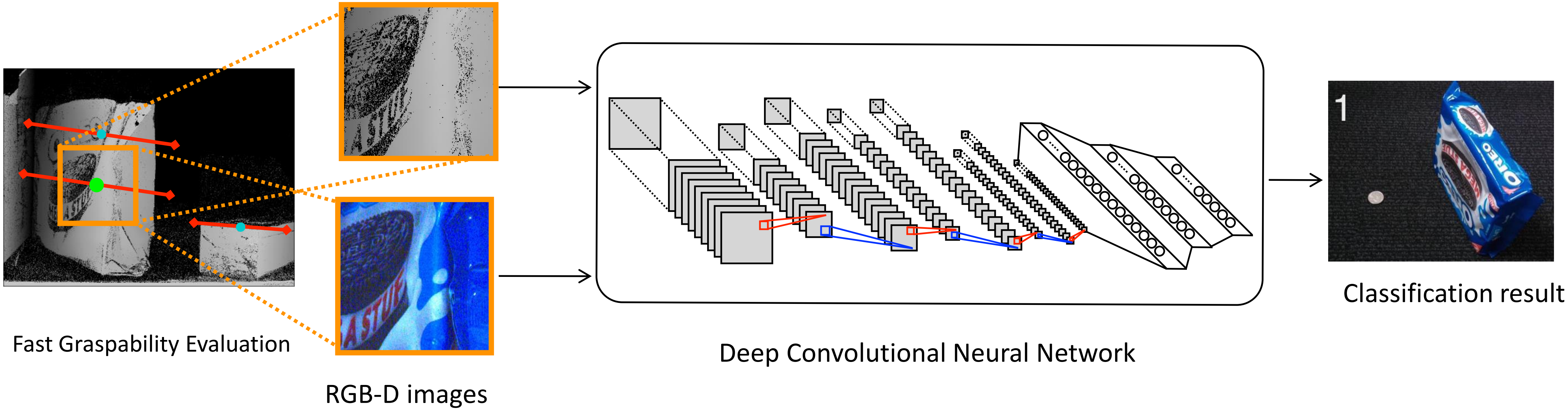
1． Additional items
25 items → 38 items

2． Additional task : Stow

Stowing → Storage → Picking



Recognition trends in APC 2016



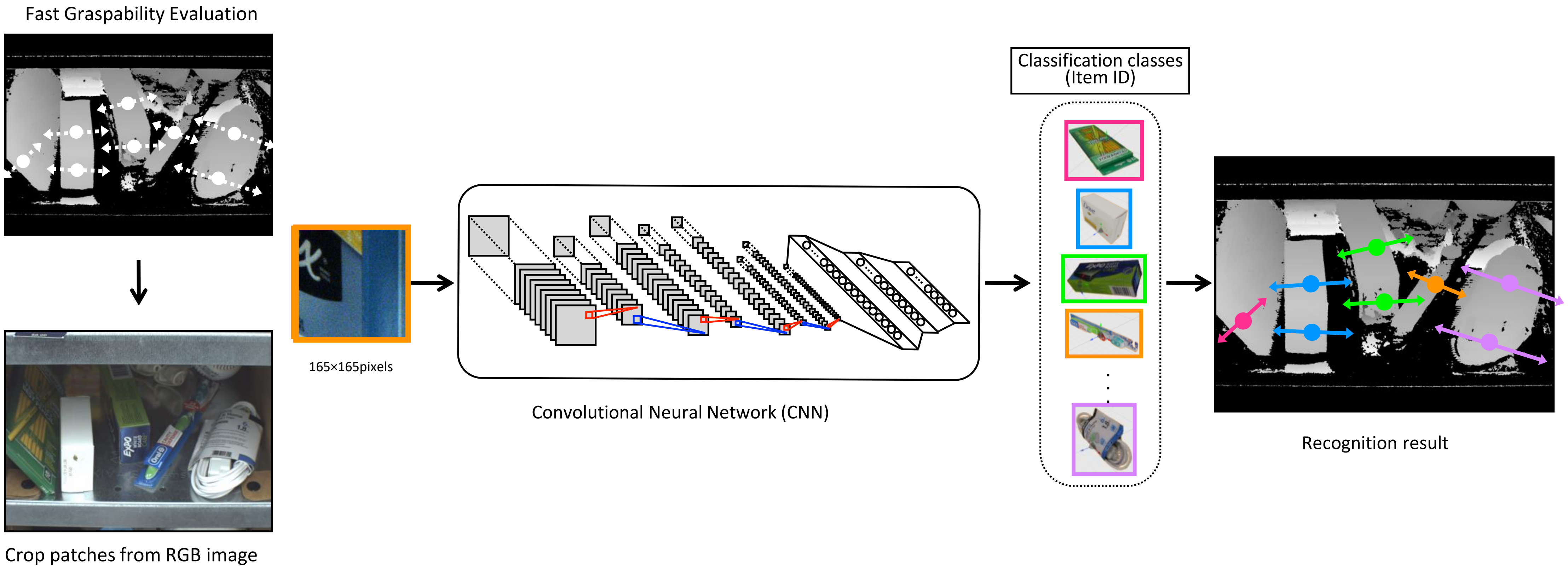
Recognition approach of team C^2M in APC 2016.

Almost all teams used Deep Learning for item recognition!

Demo by Team Delft (2016 winner)

Amazon Picking Challenge 2016 - Team Delft picking
https://www.youtube.com/watch?v=3KlzVWxomqs

Recognition strategy by Team C^2M in APC 2016



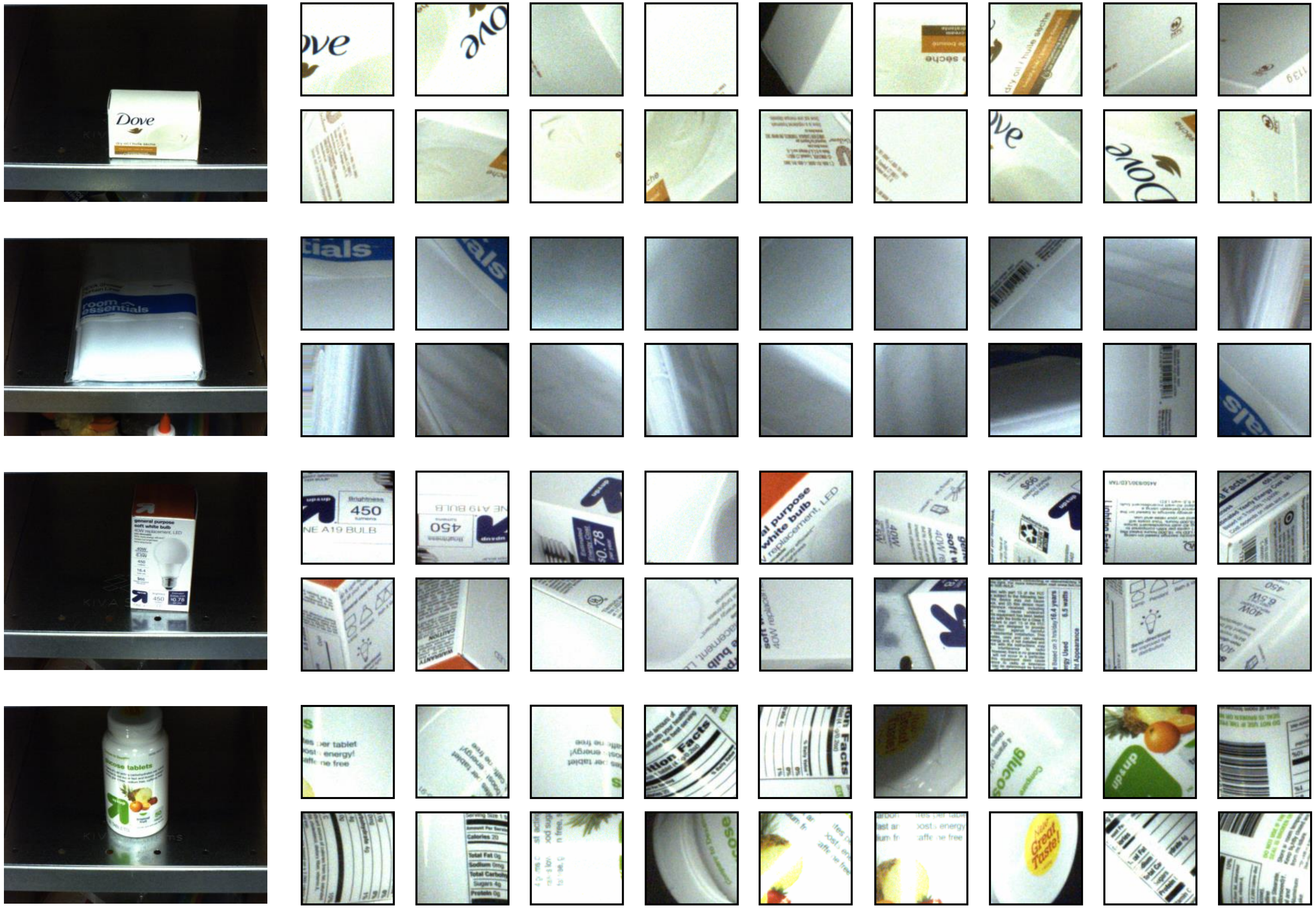
1. Patches are cropped by using Fast Graspability Evaluation
2. Convolutional Neural Network

   Input: RGB patches

   Output: Item ID (and Gripper pose)

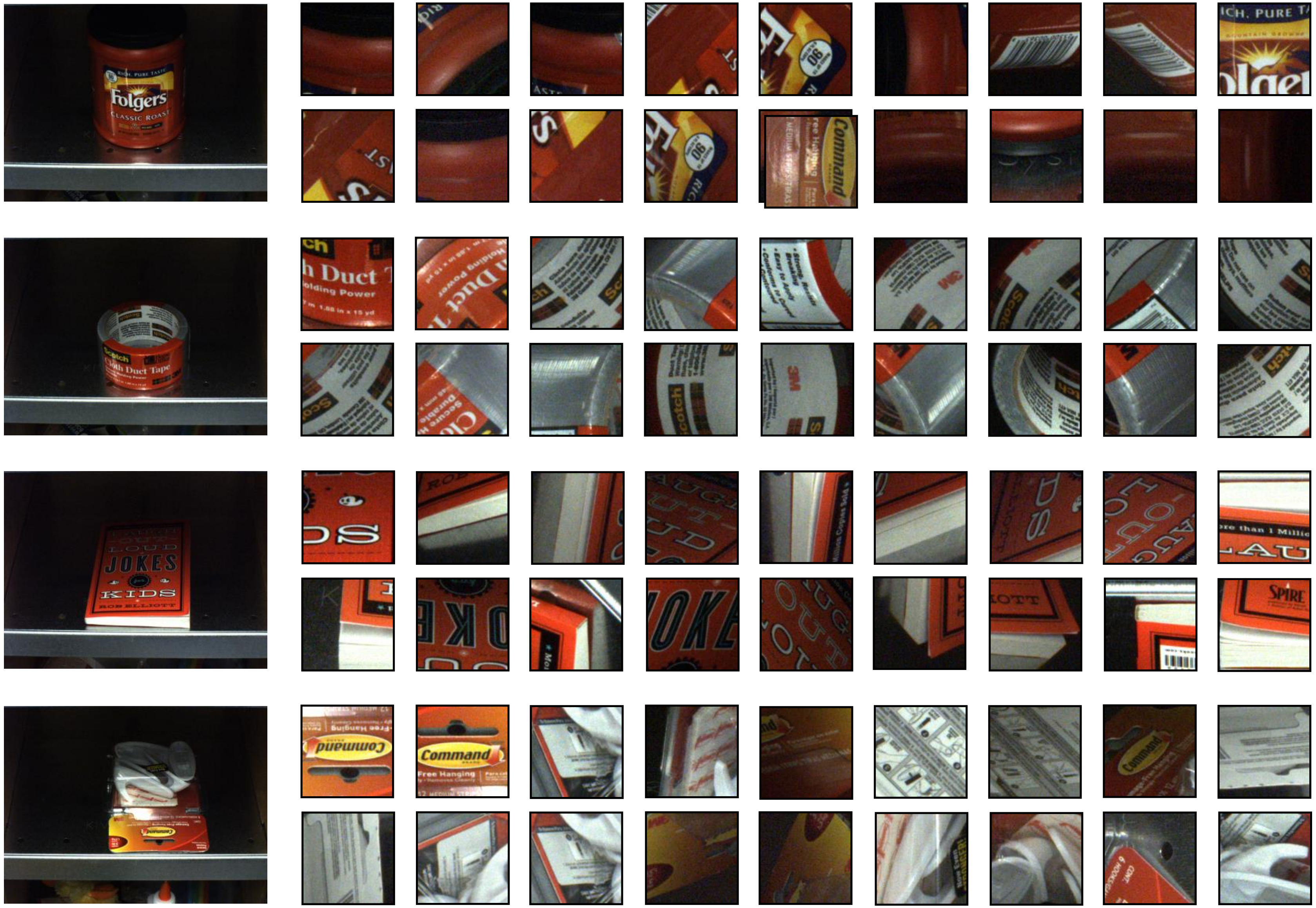
Training image examples

Data augmentation : In-plain rotation and Intensity

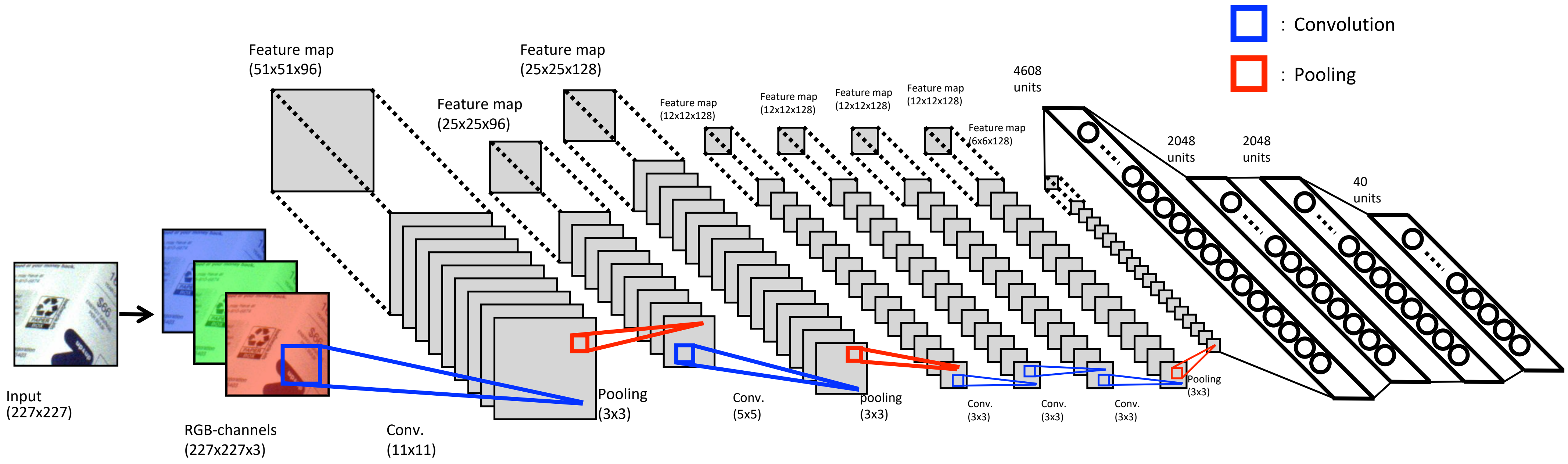


Training image examples

Data augmentation : In-plain rotation and Intensity



CNN Structure

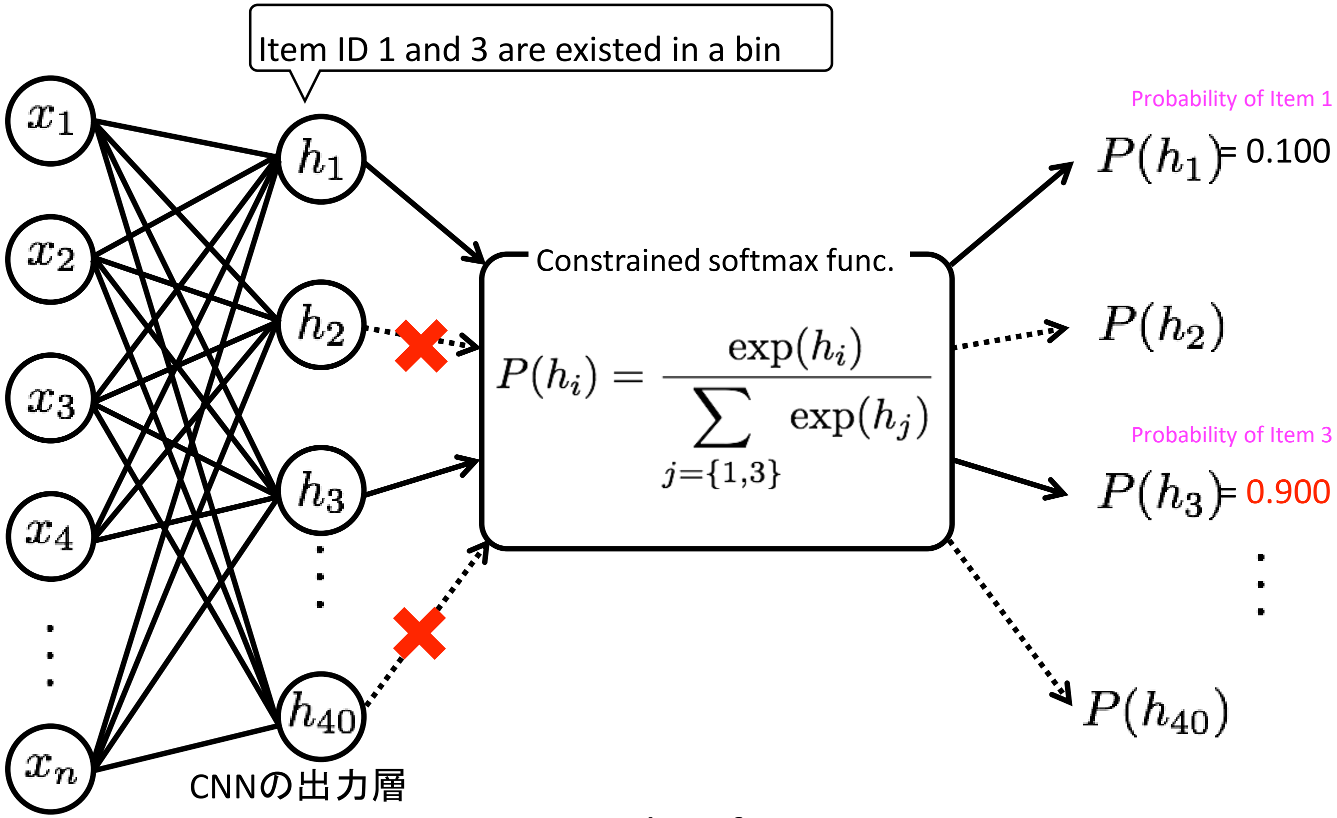


Convolutional layer	Input feature maps	Output feature maps	Filter size	Batch Normalization	Pooling	Activation function
1	3	96	11 × 11	Y	Max pooling	ReLU
2	96	128	5 × 5	Y	Max pooling	ReLU
3	128	128	3 × 3	N	N	ReLU
4	128	128	3 × 3	N	N	ReLU
5	128	128	3 × 3	N	Max pooling	ReLU

Fully connected layer	Input units	Output units	Activation function	Dropout (Learning)
1	4608	2048	ReLU	Y
2	2048	2048	ReLU	Y
3	2048	40	Soft max	N

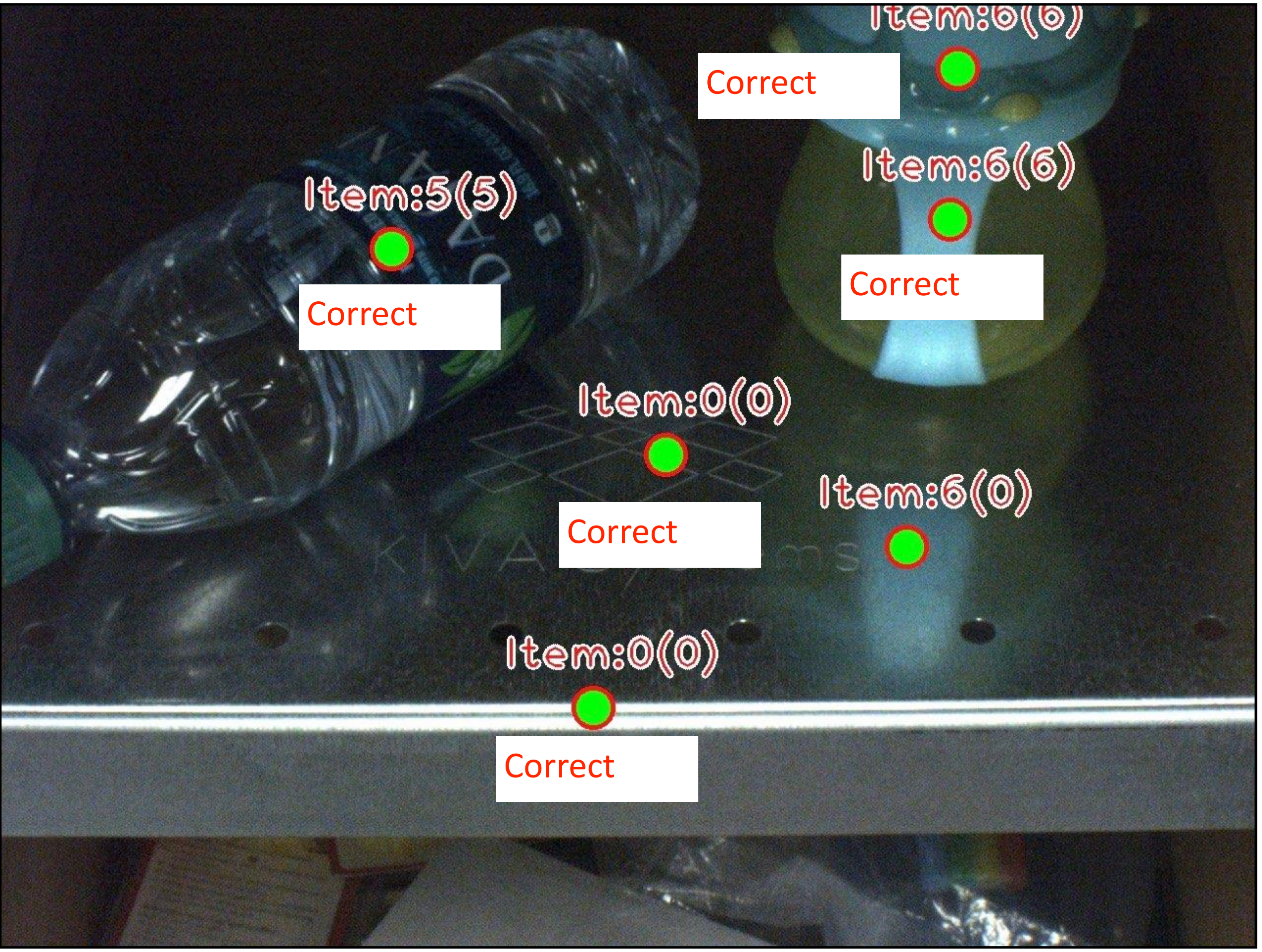
Constrained softmax function

If items are constrained and thus doesn't exist in the bin, we didin't compute corresponded softmax functions.

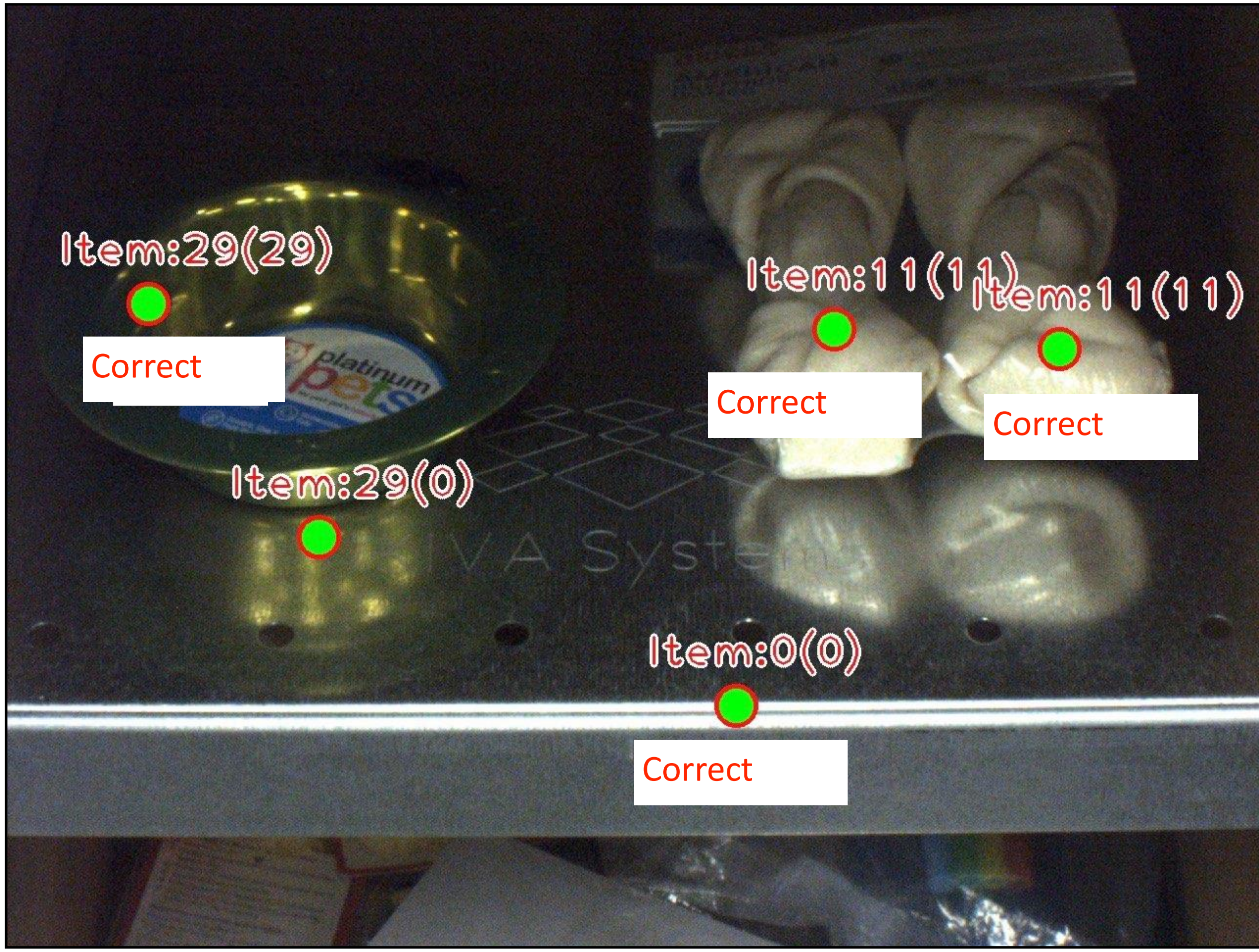


- Less classification error
- Less computational time

Recognition results



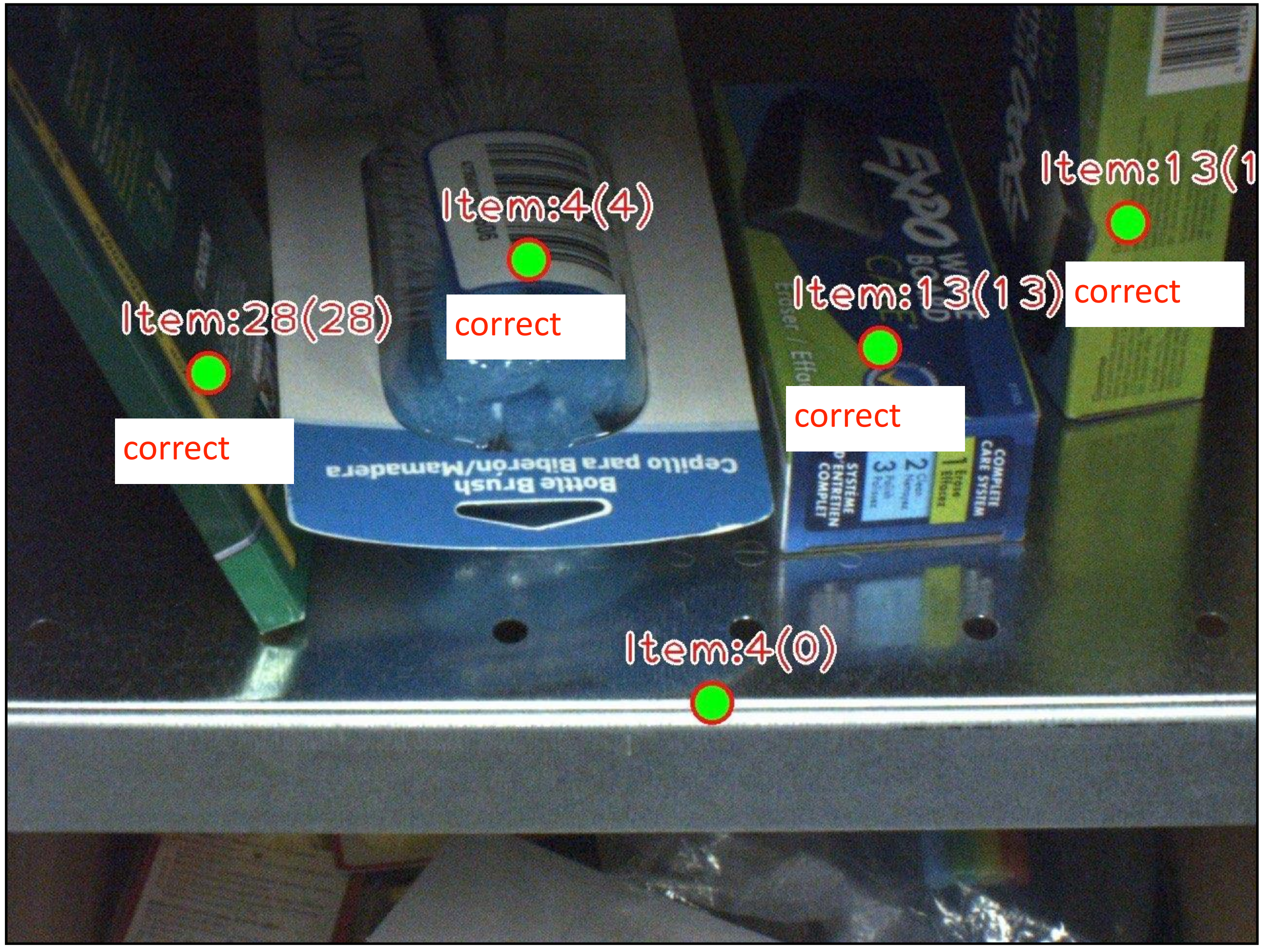
Recognition rate : 5/6



Recognition rate : 4/5

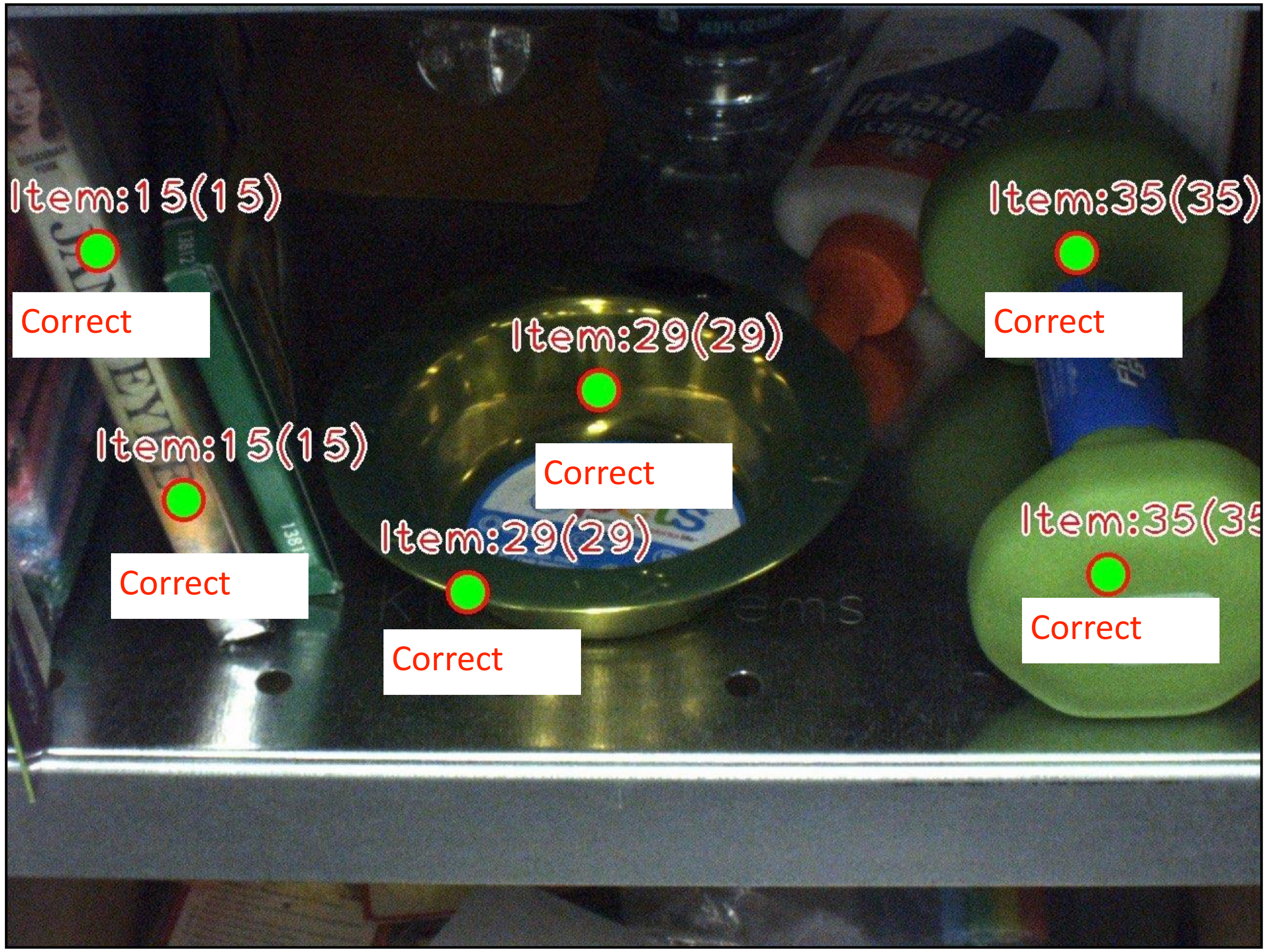
Recognition results

Recognition rate : 4/4



Recognition rate : 4/5

Recognition results



Recognition rate ： 5/5

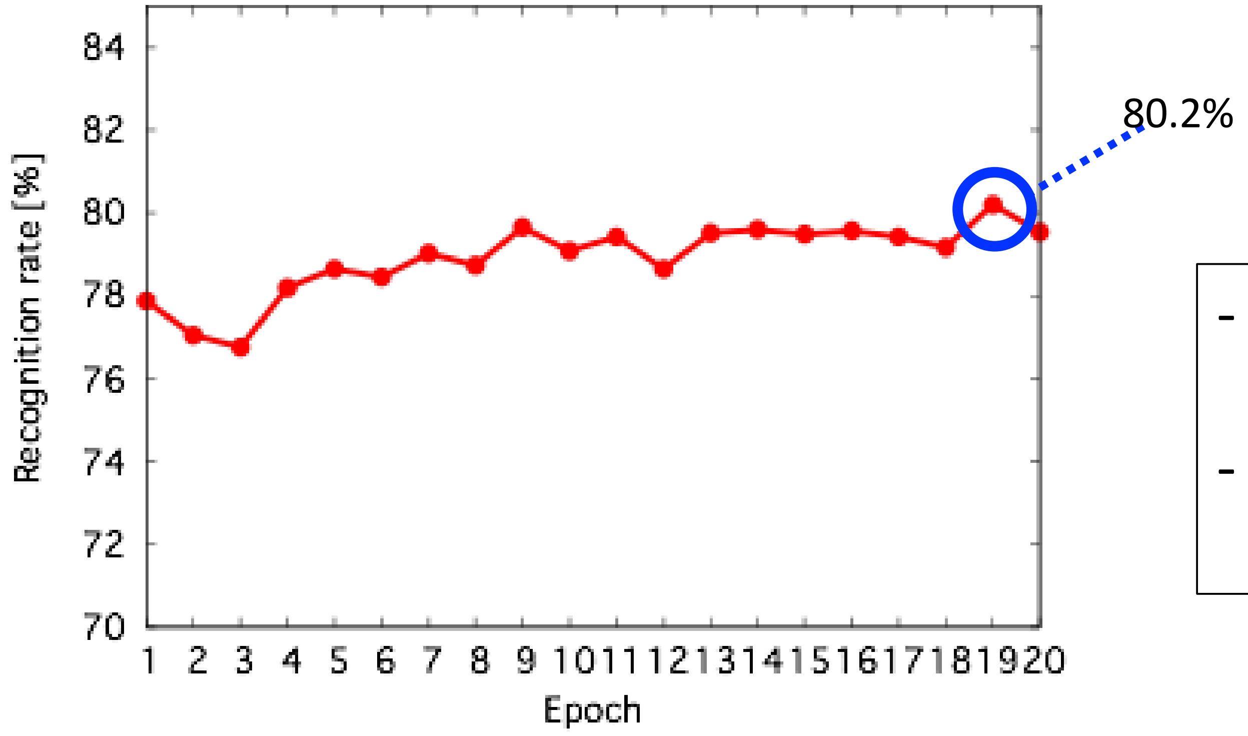
Recognition rate ： 6/6

Recognition successful rate

was **80.2%** by

inputting **3485 patches**

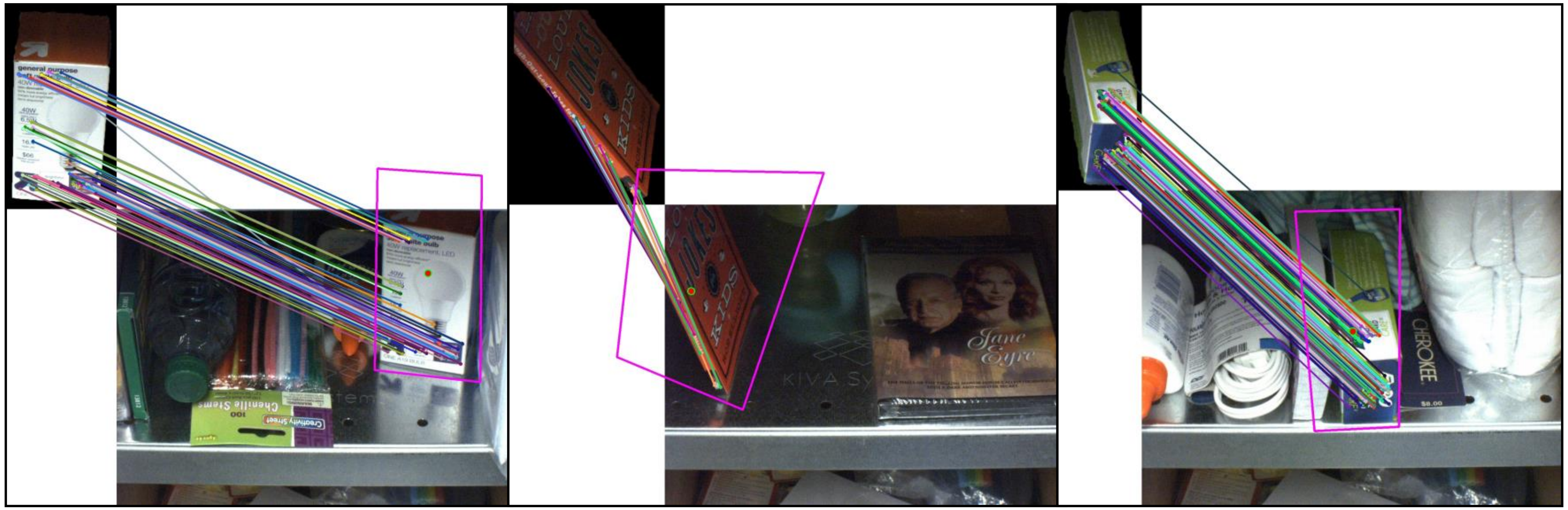
being extracted from **400 images.**



- Applicable to various items
- Not applicable to items which has no RGBD features

Key-point matching

- Key-point detector : Cascaded FAST [T. Hasegawa, ICIP 2014]
- Local features : ORB features [E. Rublee, ICCV 2011]



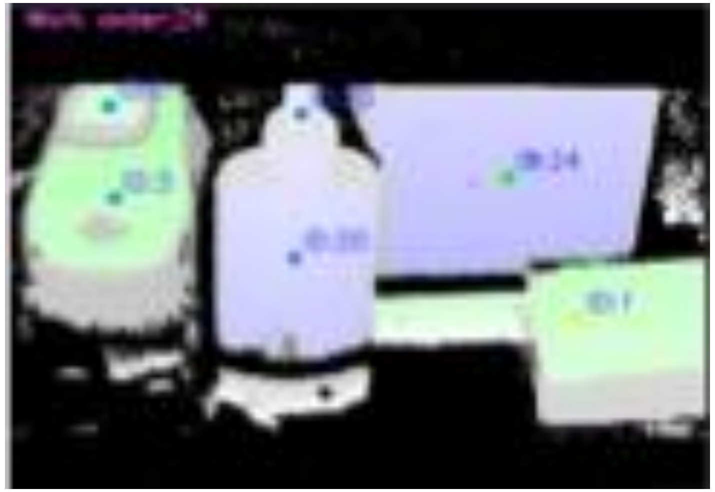
Key-point matching results

Combination of Key-point matching and CNN

Classification
results by CNN

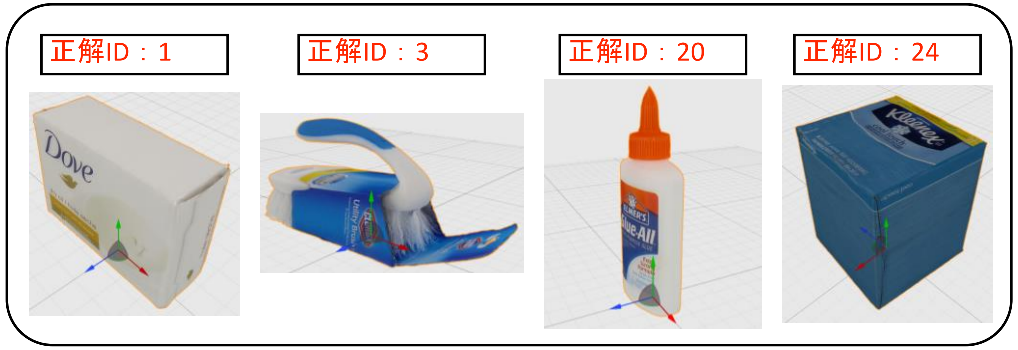
Classification results by Key-point
matching

Classification error estimation and
rejection for each segment which
are extracted from depth image

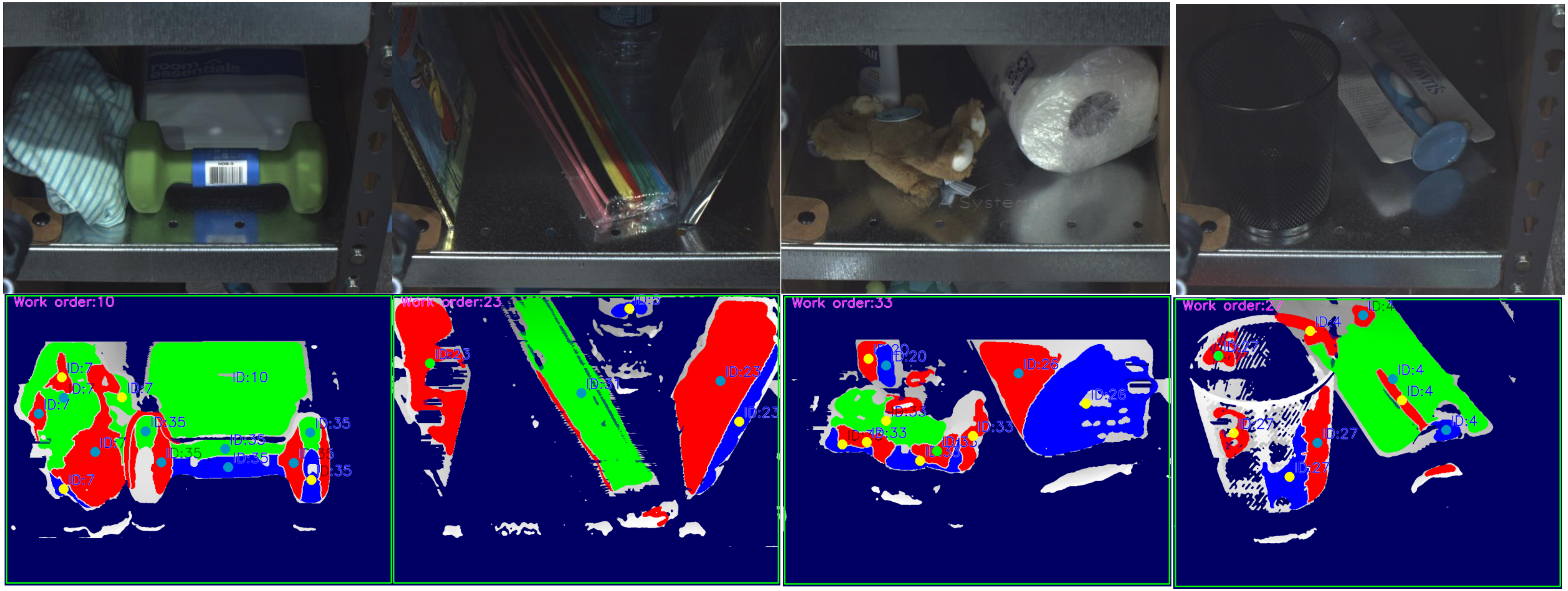


Synthesized result

- Matching Score
- Distance from CoG
- One item ID for each segment



Recognition result by using combination method



Recognition successful rate becomes **92.5%.**

What's the next? : Amazon Robotics Challenge 2017

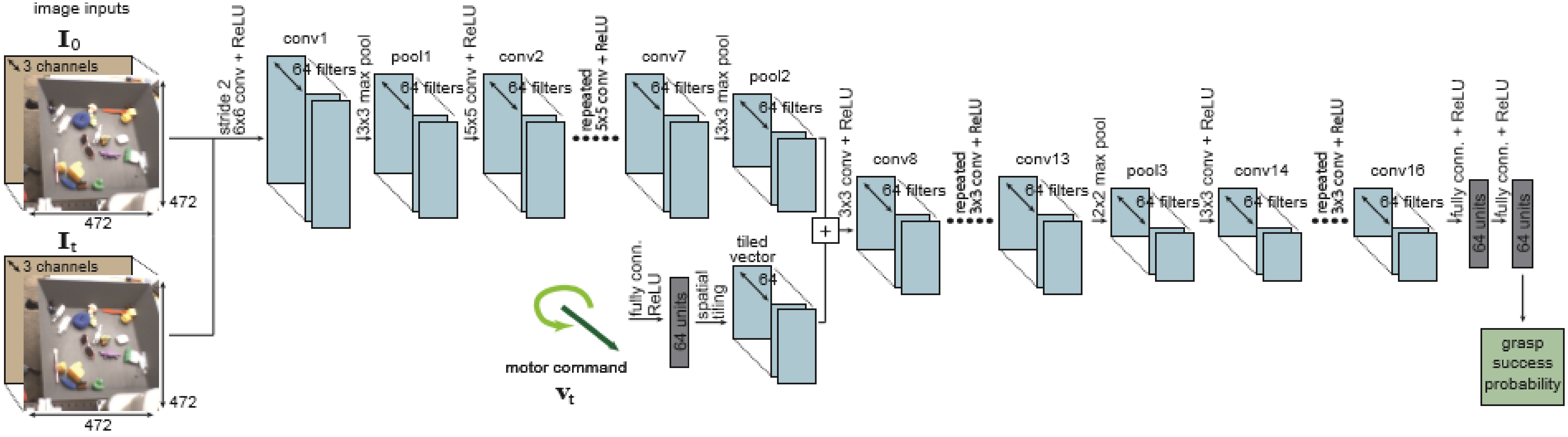
1． "Unknown" items
- Half of items are supplied "just before" the challenge
- Hard to use Deep Learning (because of learning time)
- Category classification(general object recognition) becomes important

2． Additional task : Stow and pick
- Robot must stow items in order to pick

Other approach: End-to-End learning

Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection
http://arxiv.org/abs/1603.02199

Grasp predictor by CNN



Input : RGB images and Movement vector of robot picking approach
Output : Grasp success probability

When we acquire an RGB image, we can estimate picking success probability for each picking approach.

- No requirement of hand-eye calibration
- Need many trials for learning

Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection
http://arxiv.org/abs/1603.02199

Conclusions

1. Factory and Warehouse
   - Various parts are supplied by hand
   - Stowing and Picking is by hand
   - Picking and item classification is important

2. algorithms
   - voting-based pose estimation: point pair features
   - gripper-pose-based : fast graspability evaluation
   - Regrasping : graph searching
   - Item classification : Deep learning,
     Appearance-based method

3. What's the Next?
   - End-to-End learning
   - Combination of general/specific object resognition

Thank you !

Self-introuction

Yukiyasu Domae, Ph.D.
Principal researcher
Advanced R&D center, Mitsubishi Electric Corporation

Research Interests: Machine vision, robot vision, manipulation

Awards:
R&D100 award(2014), IPSJ Kiyasu special industrial achievement award(2014),
Best paper awards of JRM(2012) and RSJ(2016), Japanese robot award(2012),
Good Design Best100(2016) and so on,
with the contribution of machine vision research and development.