1/15 (Fri.) 13:30-14:00
14:00-14:30
14:30-15:00

2/22 (Mon.) 10:00-10:30
10:30-11:00
11:00-11:30

金崎研究室の紹介

Introduction of Automation & Knowledge Laboratory

情報理工学院・情報工学系・准教授

金崎 朝子 / Asako Kanezaki

Short bio.

- Mar 2008 Graduated from Department of Mechanical Information Engineering, Faculty of Engineering, The Univ. of Tokyo
- Mar 2010 Completed master's course, Graduate School of Information Science and Technology, The Univ. of Tokyo

  2010-2011 (a half year) visiting research at Technische Universität München

- Mar 2013 Ph.D. (Information Science and Technology), The Univ. of Tokyo,
- Apr 2013 Full-time employee, Research and Development Center, Toshiba Corporation
- Dec 2013 Assistant Professor, Graduate School of Information Science and Technology, The Univ. of Tokyo

  2015.8-9 visiting research at Microsoft Research Redmond

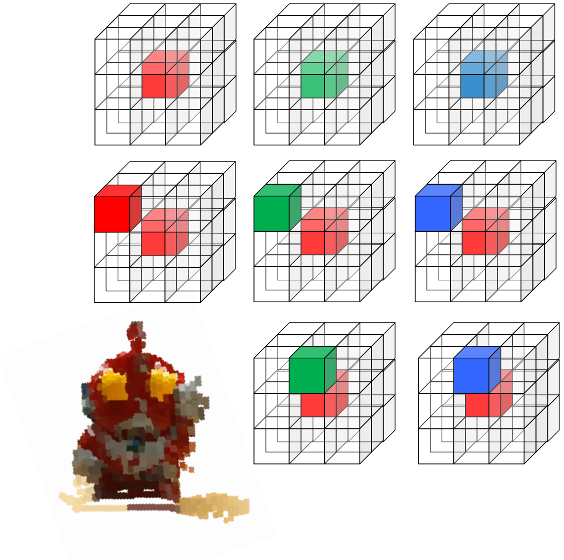
- Apr 2016 Researcher -> Senior Researcher, Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST)
- Apr 2020 Associate Professor, School of Computing, Tokyo Institute of Technology

Research

Computer Vision | Machine Learning | Robotics

Master – Ph.D. student

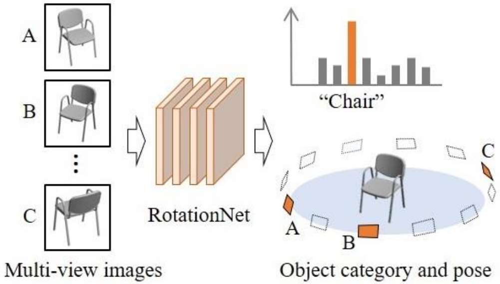
3D features & Object recognition



AIST (last 4 years)

Deep Learning

3D object recognition



Unsupervised Image Segmentation

Robot Navigation

Research

Computer Vision

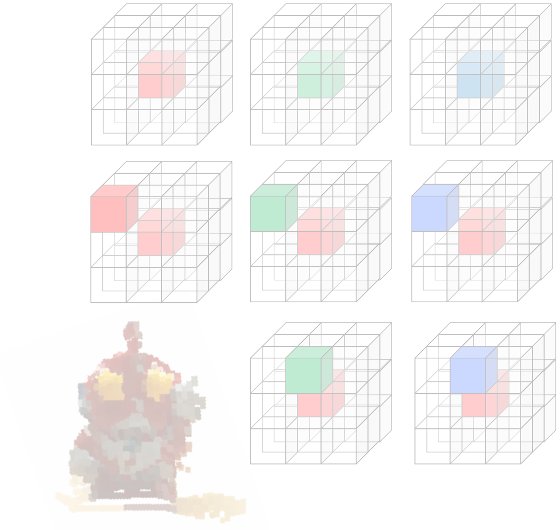
Machine Learning

Robotics

Tokyo Tech

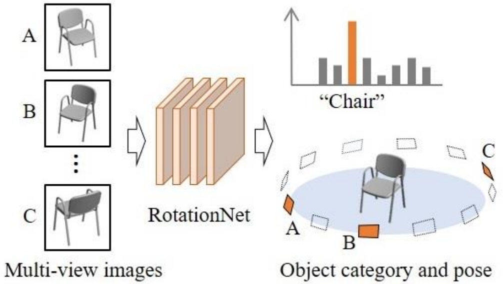
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition



Deep Learning

Unsupervised Image Segmentation

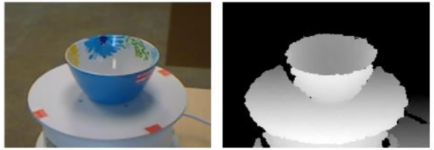
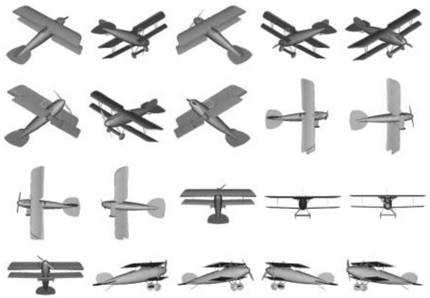
Robot Navigation

RotationNet: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints

Asako Kanezaki, Yasuyuki Matsushita, and Yoshifumi Nishida.
National Institute of Advanced Industrial Science and Technology (AIST)

3D object recognition?

Approaches:

RGBD based	Point Cloud based	Voxel based	Multi-view based
			

Here

RotationNet is the best!

- State-of-the-art scores on the ModelNet dataset http://modelnet.cs.princeton.edu/
  - ModelNet40: 40 categories
  - ModelNet10: 10 categories
  - leaderboard ⇒

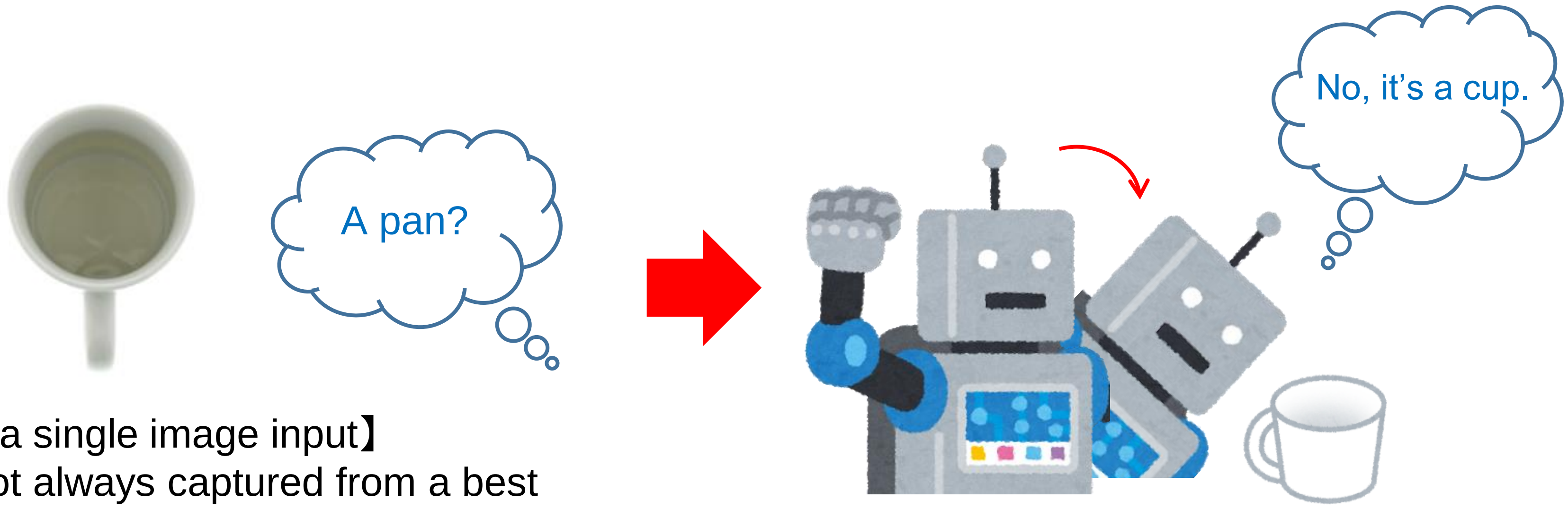
ModelNet40	ModelNet10
First : RotationNet Multi-view based	First : RotationNet Multi-view based
Second : iMHL Multi-view based	Second : SPNet Multi-view based

Algorithm	ModelNet40 Classification (Accuracy)	ModelNet40 Retrieval (mAP)	ModelNet10 Classification (Accuracy)	ModelNet10 Retrieval (mAP)
RS-CNN[63]	93.6%	-	-	-
LP-3DCNN[62]	92.1%	-	94.4%	-
LDGCNN[61]	92.9%	-	-	-
Primitive-GAN[60]	86.4%	-	92.2%	-
3DCapsule [59]	92.7%	-	94.7%	-
3D2SeqViews [58]	93.40%	90.76%	94.71%	92.12%
OrthographicNet [57]	-	-	88.56%	86.85%
Ma et al. [56]	91.05%	84.34%	95.29%	93.19%
MLVCNN [55]	94.16%	92.84%	-	-
iMHL [54]	97.16%	-	-	-
HGNN [53]	96.6%	-	-	-
SPNet [52]	92.63%	85.21%	97.25%	94.20%
MHBN [51]	94.7	-	95.0	-
VIPGAN [50]	91.98	89.23	94.05	90.69
Point2Sequence [49]	92.60	-	95.30	-
Triplet-Center Loss [48]	-	88.0%	-	-
PVNet[47]	93.2%	89.5%	-	-
GVCNN[46]	93.1%	85.7%	-	-
MLH-MV[45]	93.11%		94.80%	
MVCNN-New[44]	95.0%			
SeqViews2SeqLabels[43]	93.40%	89.09%	94.82%	91.43%
G3DNet[42]	91.13%		93.1%	
VSL [41]	84.5%		91.0%	
3D-CapsNets[40]	82.73%	70.1%	93.08%	88.44%
KCNet[39]	91.0%		94.4%	
FoldingNet[38]	88.4%		94.4%	
binVoxNetPlus[37]	85.47%		92.32%	
DeepSets[36]	90.3%			
3D-DescriptorNet[35]			92.4%	
SO-Net[34]	93.4%		95.7%	
Minto et al.[33]	89.3%		93.6%	
RotationNet[32]	97.37%		98.46%	
LonchaNet[31]			94.37	
Achlioptas et al. [30]	84.5%		95.4%	
PANORAMA-ENN [29]	95.56%	86.34%	96.85%	93.28%
3D-A-Nets [28]	90.5%	80.1%		
Soltani et al. [27]	82.10%			
Arvind et al. [26]	86.50%			
LonchaNet [25]			94.37%	
3DmFV-Net [24]	91.6%		95.2%	
Zanuttigh and Minto [23]	87.8%		91.5%	
Wang et al. [22]	93.8%			
ECC [21]	83.2%		90.0%	
PANORAMA-NN [20]	90.7%	83.5%	91.1%	87.4%
MVCNN-MultiRes [19]	91.4%			
FPNN [18]	88.4%			
PointNet[17]	89.2%			
Klokov and Lempitsky[16]	91.8%		94.0%	
LightNet[15]	88.93%		93.94%	
Xu and Todorovic[14]	81.26%		88.00%	
Geometry Image [13]	83.9%	51.3%	88.4%	74.9%
Set-convolution [11]	90%			
PointNet [12]			77.6%	
3D-GAN [10]	83.3%		91.0%	
VRN Ensemble [9]	95.54%		97.14%	
ORION [8]			93.8%	
FusionNet [7]	90.8%		93.11%	
Pairwise [6]	90.7%		92.8%	
MVCNN [3]	90.1%	79.5%		
GIFT [5]	83.10%	81.94%	92.35%	91.12%
VoxNet [2]	83%		92%	
DeepPano [4]	77.63%	76.81%	85.45%	84.18%
3DShapeNets [1]	77%	49.2%	83.5%	68.3%

Motivation

- Object recognition by robots

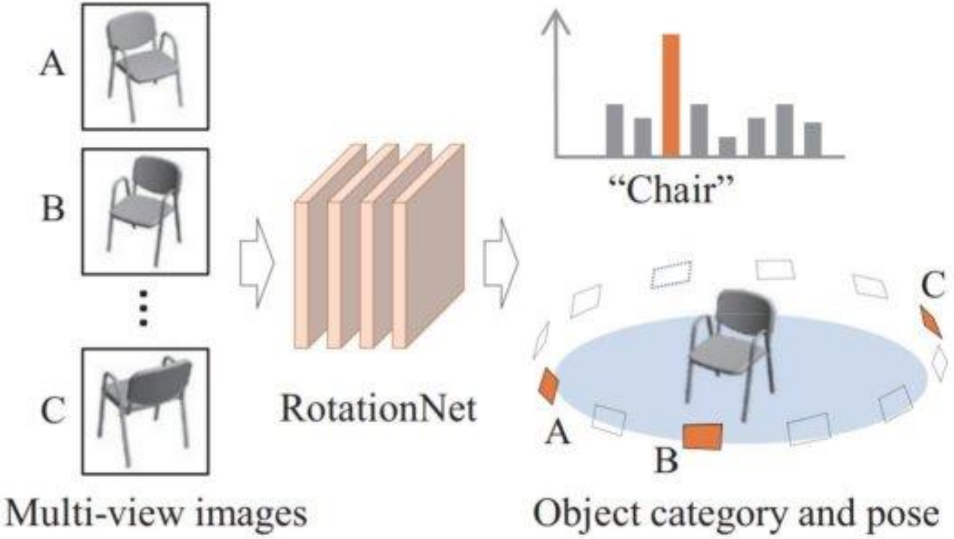
"Move and see" to achieve better performance



【a single image input】
Not always captured from a best view to recognize an object.

RotationNet (Overview)

- Takes multi-view images as input and predicts object category and pose
  - Applicable to real-time applications
  - Improve accuracy by rotating in a direction that is easy to recognize



RotationNet (Overview)

Video
https://kanezaki.github.io/rotationnet/

Research

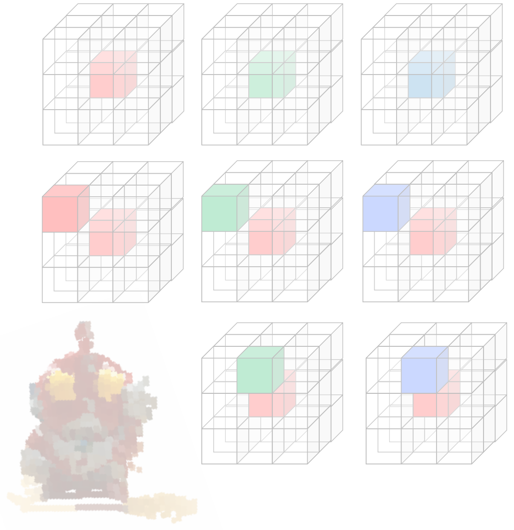
Computer Vision

Machine Learning

Robotics

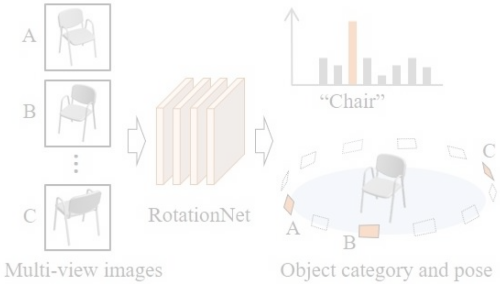
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition

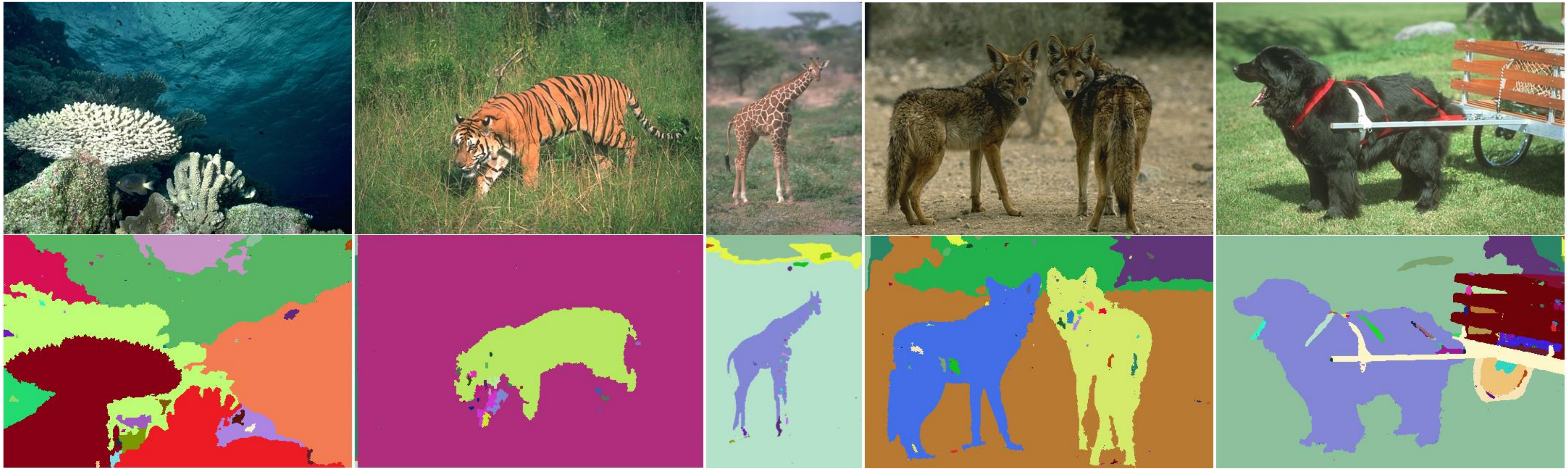


Deep Learning

Unsupervised Image Segmentation

Robot Navigation

Deep Unsupervised Image Segmentation



Each segment is shown in a random identical color.

A. Kanezaki*, W. Kim*, and M. Tanaka. “Unsupervised Learning of Image Segmentation Based on Differentiable Feature Clustering.” IEEE Transactions on Image Processing, vol. 29, pp. 8055-8068, 2020.
*equal contribution

Deep Unsupervised Image Segmentation

Tokyo Tech

- In this paper, good conditions for image segmentation are defined as follows:

  (a) Pixels of similar features should be assigned the same label.

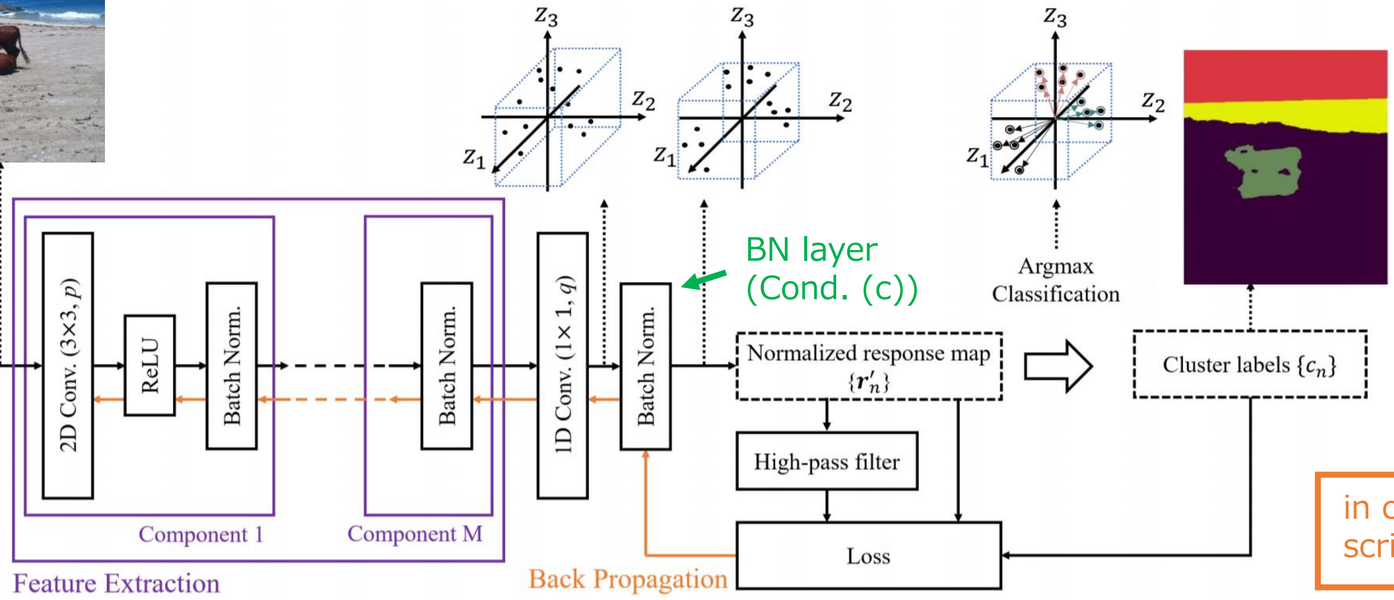
  (b) Spatially continuous pixels should be assigned the same label.

  (c) The number of unique cluster labels should be large.

The three conditions will never be met at the same time, but they will settle at a moderate balance.

Deep Unsupervised Image Segmentation

Differentiable Clustering: image features are trained in an end-to-end manner.



in case with scribbles

$$\sum_{n=1}^{N}\sum_{i=1}^{q} -\delta(i-c_n)\ln r'_{n,i} + \sum_{\xi=1}^{W-1}\sum_{\eta=1}^{H-1} \| \boldsymbol{r}'_{\xi+1,\eta} - \boldsymbol{r}'_{\xi,\eta} \|_1 + \| \boldsymbol{r}'_{\xi,\eta+1} - \boldsymbol{r}'_{\xi,\eta} \|_1 + \sum_{n=1}^{N}\sum_{i=1}^{q} -u_n\delta(i-s_n)\ln r'_{n,i}$$

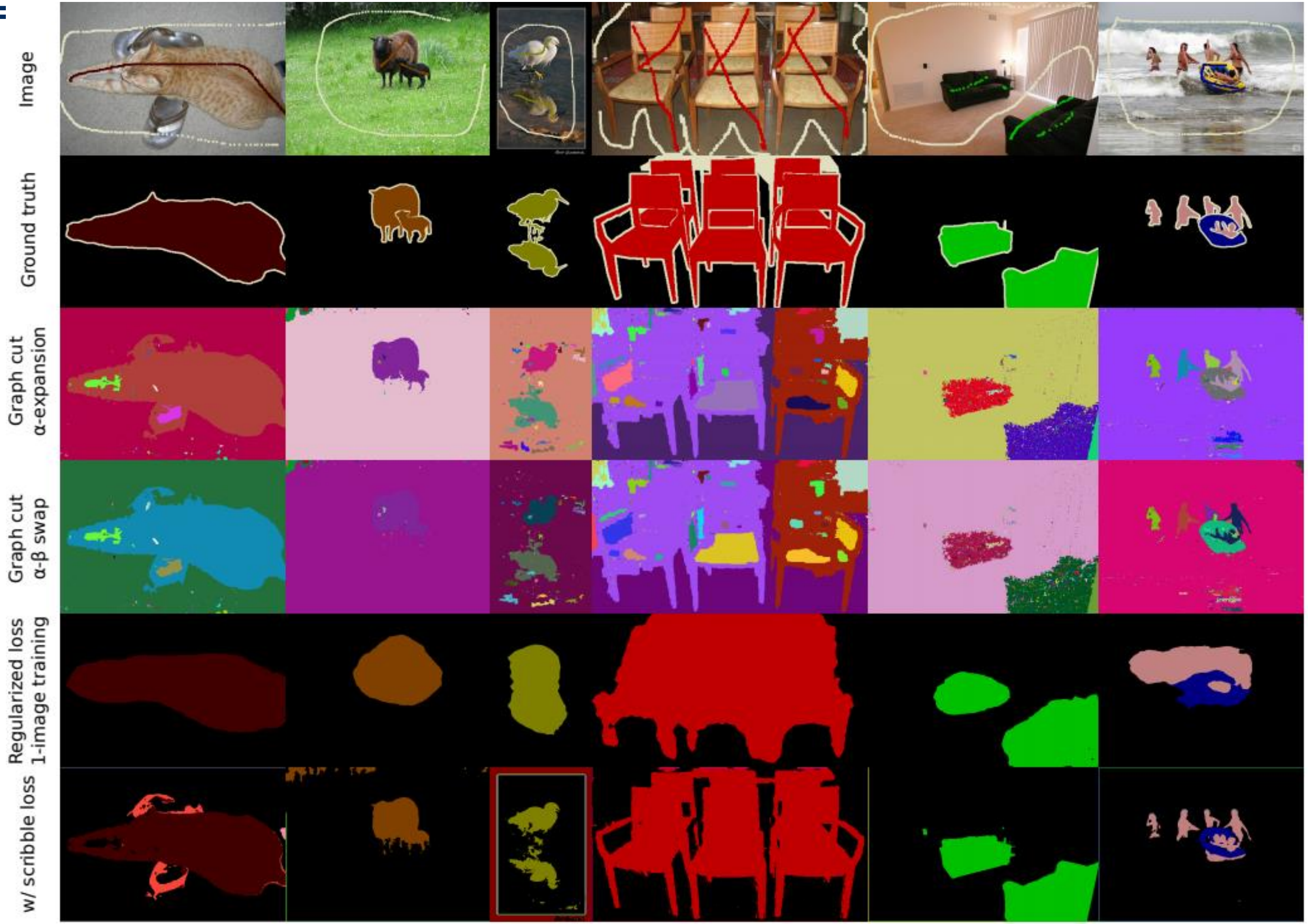
Cross-Entropy Loss（Cond. (a)）

Differential Filter（Cond. (b)）

Partial Cross-Entropy Loss

Deep Unsupervised Image Segmentation

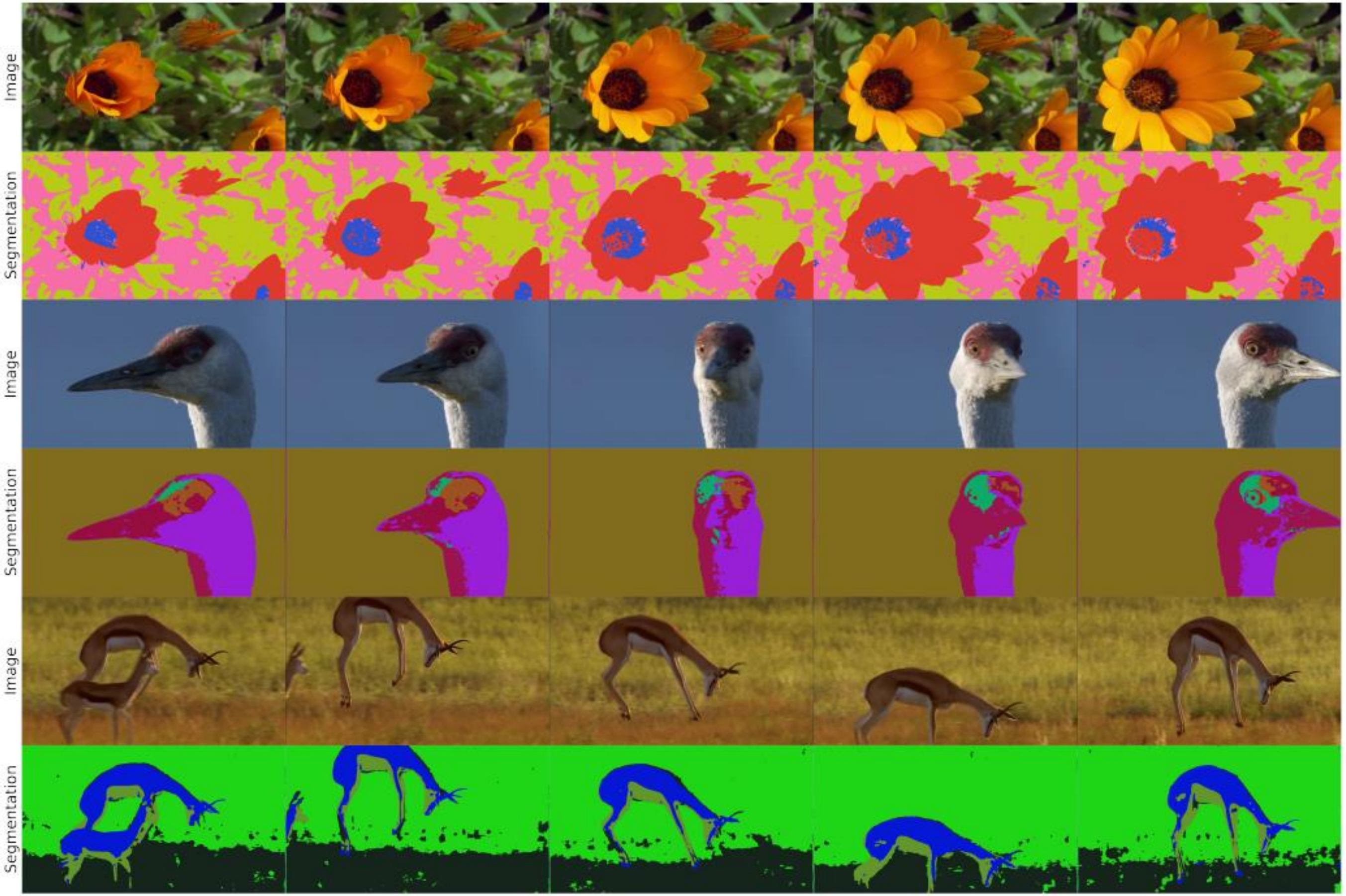
- Results with scribbles as input



Deep Unsupervised Image Segmentation

- Results with reference image(s)
  - video

Network is unsupervisedly trained with a single first frame



Tokyo Tech

Research

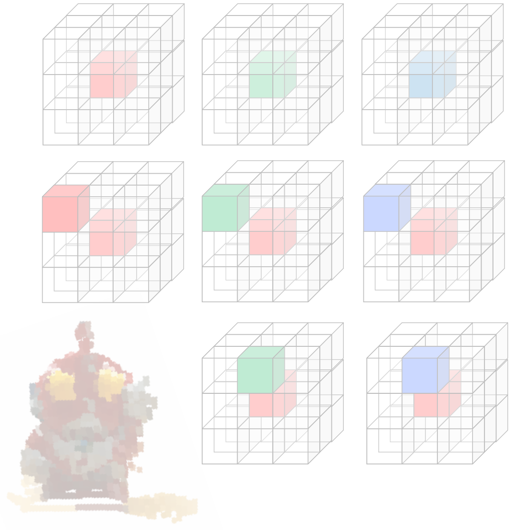
Computer Vision

Machine Learning

Robotics

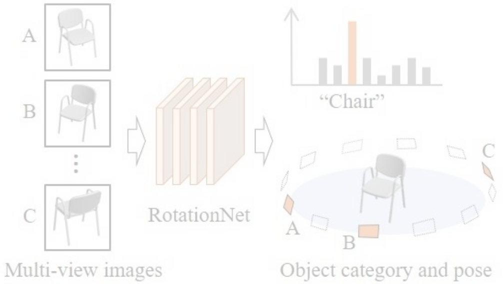
Master – Ph.D. student

3D features & Object recognition



AIST (last 4 years)

3D object recognition



Deep Learning

Unsupervised Image Segmentation

Robot Navigation

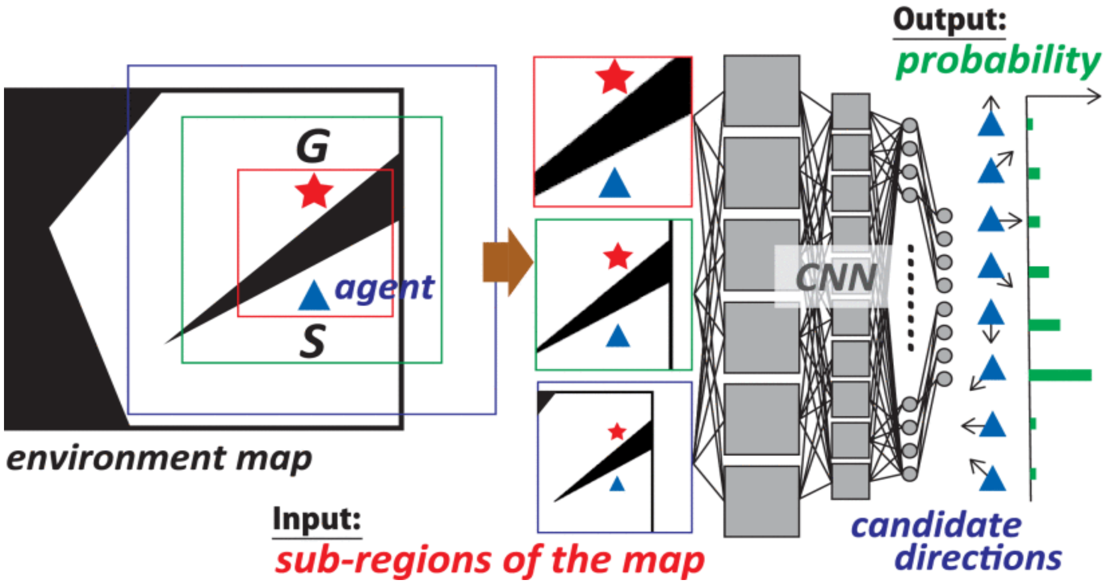
Robot navigation using deep learning

Video
https://kanezaki.github.io/goselo/

https://kanezaki.github.io/goselo/goselodemo

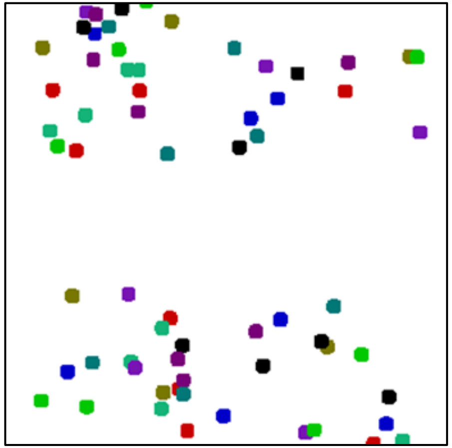
Robot navigation using deep learning

Key Idea: By creating a goal-directed map representation, we can learn the relationship between visual patterns of the surrounding environment and movement patterns without being constrained by the shape of a particular environment!



Follow-up studies（1/2）

- Yoko Sasaki, Syusuke Matsuo, Asako Kanezaki, and Hiroshi Takemura. "A3C Based Motion Learning for an Autonomous Mobile Robot in Crowds." *IEEE International Conference on System Man and Cybernetics (SMC2019)*, pp.1046-1052, 2019.
- 佐々木洋子，松尾修佑，金崎朝子，竹村裕. 歩行者観測履歴を用いた深層強化学習による車輪ロボットの雑踏切り抜け動作生成. 日本機械学会ロボティクス・メカトロニクス講演会, 2019.
- 渋谷 薫, 金崎 朝子, 大西 正輝. 深層学習による画像識別問題に帰着した人の流れのシミュレーション. Meeting on Image Recognition and Understandings (MIRU), 2018.

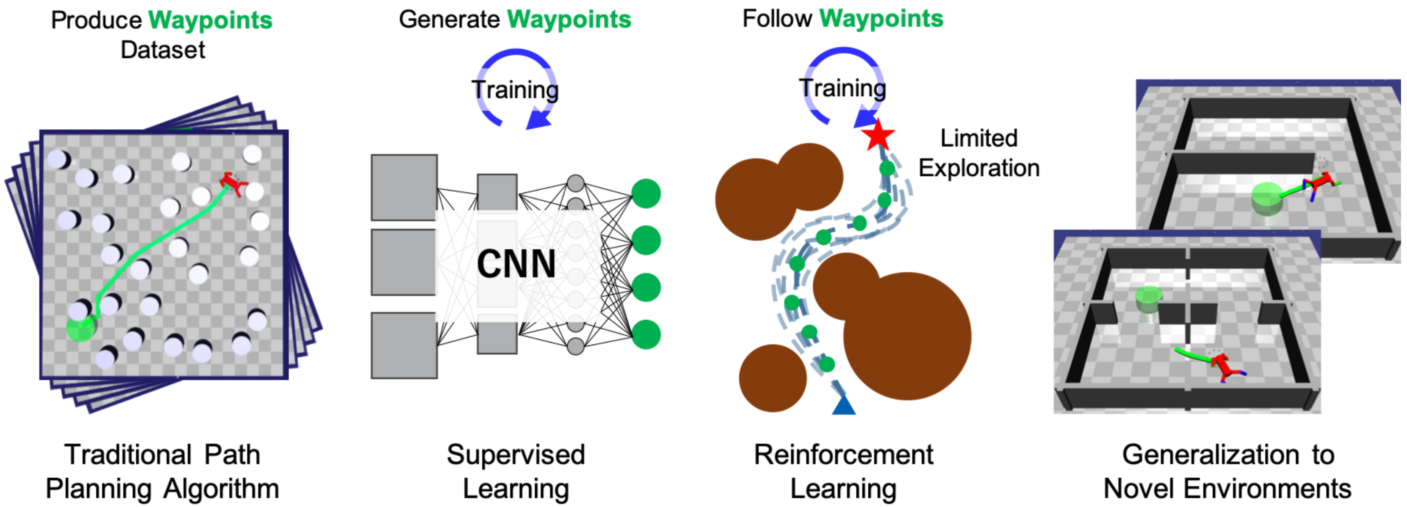


Learned policy simulation

Follow-up studies (2/2)

- Imitation learning of waypoints estimation using path planning as an expert
- Reinforcement learning using the estimated Waypoints as a guide

Improve learning efficiency and generalizability

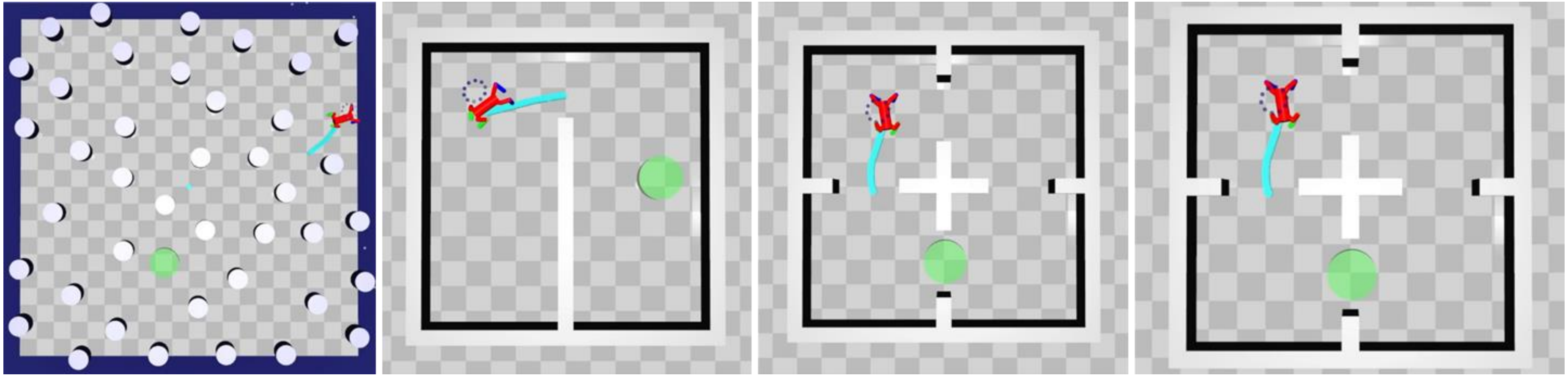


K. Ota, Y. Sasaki, D. K. Jha, Y. Yoshiyasu, and A. Kanezaki. "Efficient Exploration in Constrained Environments with Goal-Oriented Reference Path." IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.

Follow-up studies（2/2）

- Imitation learning of waypoints estimation using path planning as an expert
- Reinforcement learning using the estimated Waypoints as a guide

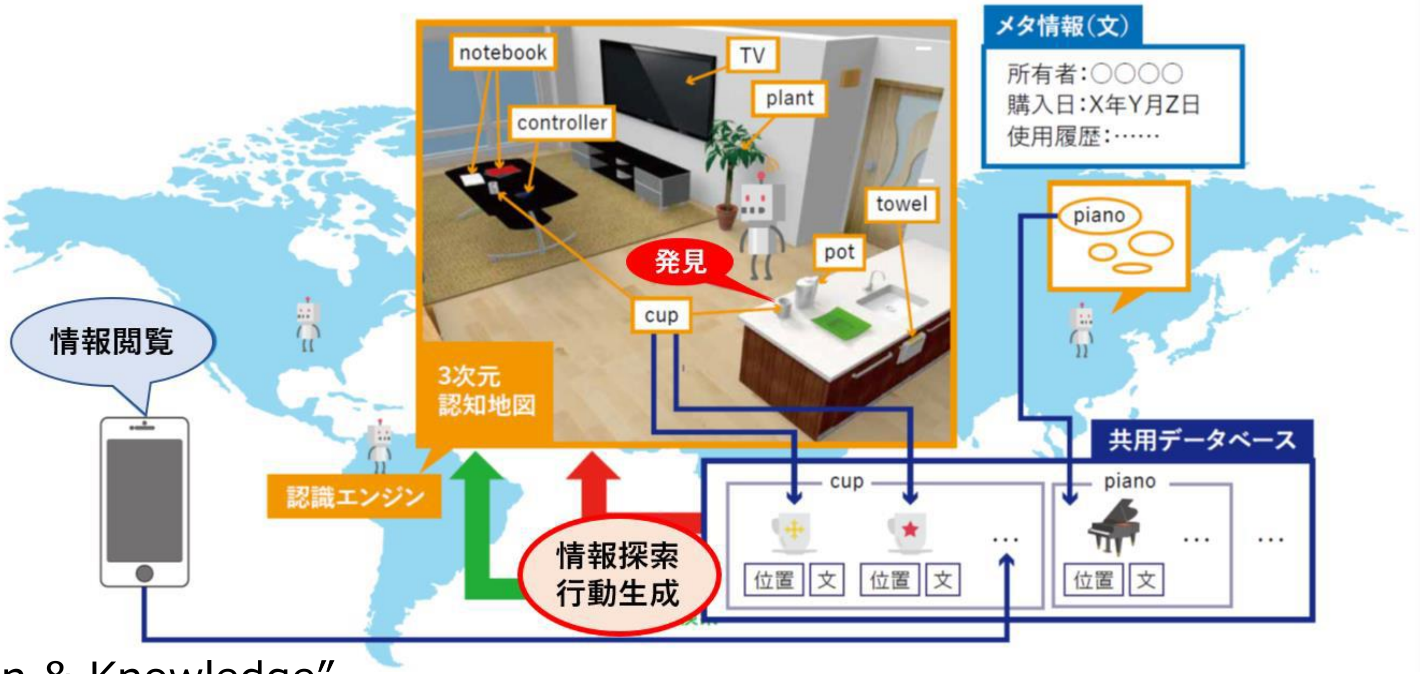
Improve learning efficiency and generalizability



K. Ota, Y. Sasaki, D. K. Jha, Y. Yoshiyasu, and A. Kanezaki. "Efficient Exploration in Constrained Environments with Goal-Oriented Reference Path." IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.

Research plan　*IoP (Internet of Perception)?*

- A robot system for active information search and database construction



"Automation & Knowledge"

Actively collect data and create knowledge ⇔ Collect knowledge and act wisely

Research topics (examples)

- Action planning for autonomous mobile robots using reinforcement learning
- Automatic generation of daily life search engines
- Information collection planning for interactive robots
- Robot intelligence with superhuman recognition technology
- Global and local action planning for robots

Machine Learning

Computer Vision

Robotics

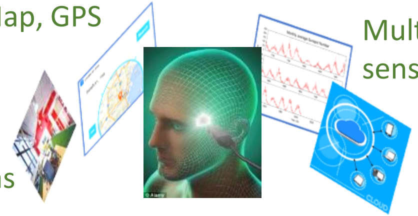
Human-Computer Interaction

Map, GPS

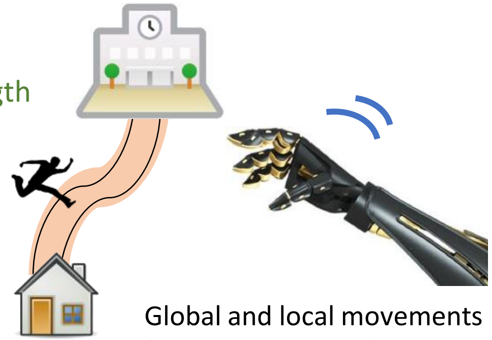
Multi-wavelength sensing

Multi-cameras

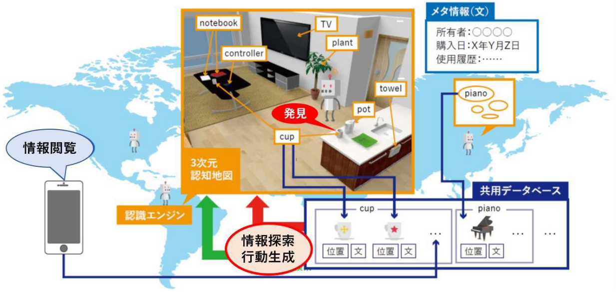
Database



Construction of robot intelligence with superhuman recognition technology



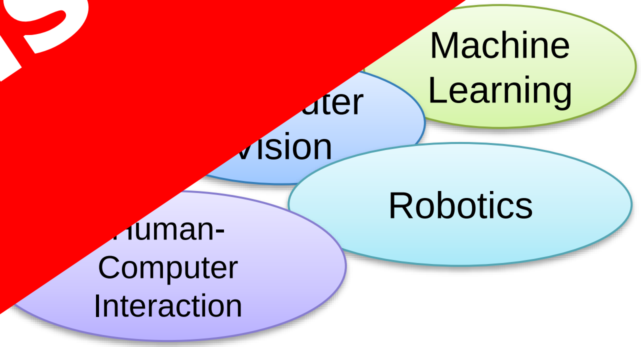
Global and local movements (Manipulation, etc.)



Research topics (examples)

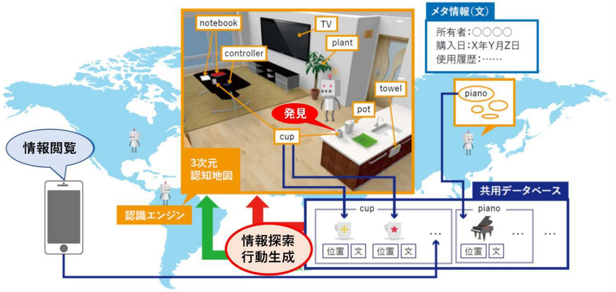
- Action planning for autonomous m
- Automatic generation of daily
- Information collection pl
- Robot intelligence
- Global and lo

Anything else is fine.



Map, GPS

Global and local movements
(Manipulation, etc.)



Lab information

- Location: West Building 8, Ookayama Campus
- Core time: Weekly lab meeting
- Members:
  - 2020.9～ 2 Master
  - 2021.4～ 5 Master and 3 Ph.D cand.
- Infrastructure :
  - Use of computer clusters such as TSUBAME, AIST-ABCI
  - Managing source code and tips on GitHub
  - Sharing of library know-how such as deep learning and robot simulation

Lab information

- Location: West Building 8, Ookayama Campus
- Core time: Weekly lab meeting

Sofa and coffee machine

(now vacant) desks