Osaka University 1

Bayesian Estimation for 3D-Point Object Identification Based on Probabilistic Field Representation

Graduate School of Information Science and Technology
Osaka University
Hiroaki Sato

1

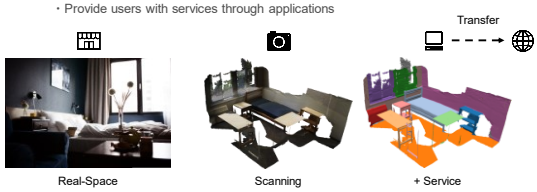
Outline 2

- **Introduction**
- **Represent as probabilistic field of objects**
  - Example of probabilistic field representation
- **Examine application of obtained probabilistic field**
  - Segmentation prediction with prior knowledge by Bayesian estimation
  - Evaluate
- **Conclusion**

2

Introduction 3

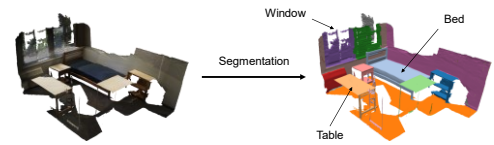
- **New network services using real-space information**
  - Require understanding real-space & processing information
    e.g. Remote shopping services
    - Transfer real-space information acquired remotely via a network
    - Provide users with services through applications



3

Introduction 4

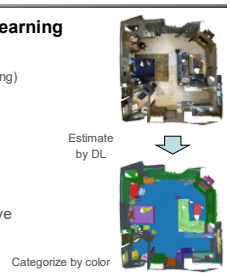
- **Object Identification**
  - **Estimate** object location & category
    Toward understanding real-space where object exist & what object is
  - Task must be recognized with **High accuracy**
    e.g. automatic driving of automobiles & control of self-driving robots



4

Introduction 5

- **Research machine learning with deep learning**
  - 2D data such as image & video
    Problem: Cannot handle 3D information (e.g. Overlapping)
  - 3D spatial information
    Scan by Lidar sensors & depth camera
- **Problem of deep learning**
  - Hard to understand rules of result
  - Ignores real-space knowledge we humans have
    e.g. kitchen and a sofa are unlikely to be adjacent



5

Aim 6

- **Represent as probabilistic field about objects**
  - Represented by probabilistic superposition of object categories
    - Quantify empirical knowledge that we potentially acquired during our lives
  - Instead of directly identifying real-space information using DL
- **Examine application of probabilistic field**
  - Modification of object identification results using probabilistic field
  - Improve recognition of real-space information with empirical knowledge

6

3D Point Cloud Dataset

7

- **ScanNet[1]**
  - Indoor 3D point cloud dataset for 1613 room scenes
    Coordinate and color by RGB-D camera
  - 21 scene type
    e.g. office, apartment, bathroom
  - 20 object category
    "other furniture":
    piano, display, bicycle, etc.

wall	picture
floor	counter
cabinet	desk
bed	curtain
chair	refrigerator
sofa	shower curtain
table	toilet
door	sink
window	bathtub
bookshelf	other furniture

Indoor 3D point cloud data

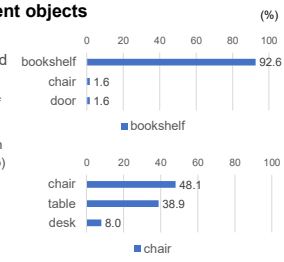
[1] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser and M. Nießner, "ScanNet: Richly-Annotated 3D Reconstructions of Indoor Scenes," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2432-2443

7

Example of Probabilistic Field Representation

8

- **Probability distribution of adjacent objects**
  - Characteristic Information
    → reflect real-space as we understand
    - Prob in same category is High
      Show bookshelf placed next to bookshelf
    - Chair: chair 48%, table 38%, desk 8%
    - High prob of mutual adjacency in Bathroom
      (shower curtain, toilets, bathtub)
  - Exceptional Information
    - One case: toilet occurs next to bed
      adjacency along opposite wall

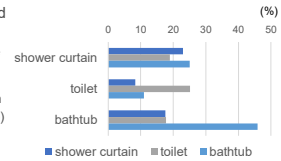


8

Example of Probabilistic Field Representation

9

- **Probability distribution of adjacent objects**
  - Characteristic Information
    → reflect real-space as we understand
    - Prob in same category is High
      Show bookshelf placed next to bookshelf
    - Chair: chair 48%, table 38%, desk 8%
    - High prob of mutual adjacency in Bathroom
      (shower curtain, toilets, bathtub)
  - Exceptional Information
    - One case: toilet occurs next to bed
      adjacency along opposite wall



9

Object Estimation Method Using Prior Knowledge

10

- **Segmentation prediction considering prior knowledge**
  - Use probabilistic field representation
    to perform object identification for semantic segmentation predictions
    - With DL ignores real-space knowledge humans have
      e.g. kitchen and sofa are unlikely to be adjacent to each other
  - Semantic segmentation predictions (SparseConvNet[2] method)
    - Prediction Probability: Probability an object is predicted in each category
    - Predicted Class: Category with the highest prediction probability

[2] B. Graham, M. Engelcke and L. v. d. Maaten, "3D Semantic Segmentation with Submanifold Sparse Convolutional Networks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 9224-9232

10

Evaluation

11

- **Performance indicator before / after Bayesian estimation**
  - Evaluation of Multi-Class Classification

	Accuracy	Precision	Recall	F1-score
Before Bayes	65.3%	60.1%	59.9%	60.0%
After Bayes	67.5%	70.8%	62.4%	66.3%

  - Precision increased 10% & Recall increased 3%
  - Many classes still have low reproducibility
    → ・Achieved correction of some predictions
    ・Many predictions could not be corrected

11

Evaluation

12

- **Distribution of posterior labels with predicted labels**

	posterior label = annotated	posterior label ≠ annotated
predicted label = annotated	6678 (88.7%)	850 (11.3%)
predicted label ≠ annotated	1099 (27.5%)	2896 (72.5%)

  - Maintain 88.7% of "predicted = annotated to posterior label"
    → Prediction probability would be higher using prior knowledge
  - Correctly modified 27.5% of "predicted ≠ annotated to posterior label"
  - Mistakenly modified 11.3% of "predicted = annotated to posterior label"
    → Affect object identification indicator with prior knowledge

12

Conclusion & Future Work

- **Conclusion**
  - Represent of real-space information as a probabilistic field
  - Present an example of object estimation method based on prior knowledge using the acquired probabilistic field
    - Confirm of changes in probability as object accuracy modification using prior knowledge
- **Future Work**
  - Obtain probability field that combines multiple real-space information
  - Obtain probability field of real-space information including time series
  - Update by segmentation prediction results