

# 実空間の知識情報を利用した 物体推定手法の提案

大阪大学 大学院情報科学研究科  
村田研究室 佐藤広章

- **画像や動画から、物体の位置やカテゴリーを識別**
  - 実世界情報に対して高い識別精度での取得が必要
    - 自動車の自動運転
    - 自動走行型ロボットの自律制御 など
- **現在：ディープラーニングなどの機械学習を使用**
  - 物体識別結果のプロセスを人間が理解するのは困難
  - 人間が持つ実空間の知識情報が考慮されていない
    - 台所とソファが隣接する可能性は低いなど



深層学習 ↓



カテゴリーによって色分け  
茶色：ソファ、黄緑：床、など

- **実空間上の事前知識の取得**

- 物体の隣接関係に着目し、実空間上での相関関係を取得
  - 物体カテゴリーの存在確率分布
  - 物体カテゴリー間の隣接確率分布

- **事前知識を利用した物体推定手法の提案と評価**

- 物体の識別推定に事前知識を組み込み、物体の確からしさを修正
  - 物体カテゴリー間の隣接確率分布を使用

## ● ScanNet データセット[1]

- 屋内 3 次元点群データのデータセット
- 1613 部屋分のシーンに対して、座標情報と色情報を持つデータの集合
- 21 種類のシーンタイプ、20 種類の物体カテゴリーから構成
  - シーンタイプ：規模の小さいもの（浴室、押入、倉庫など）から大きなもの（アパート、教室、図書館など）まで
  - 物体カテゴリー：右表の 20 種類

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



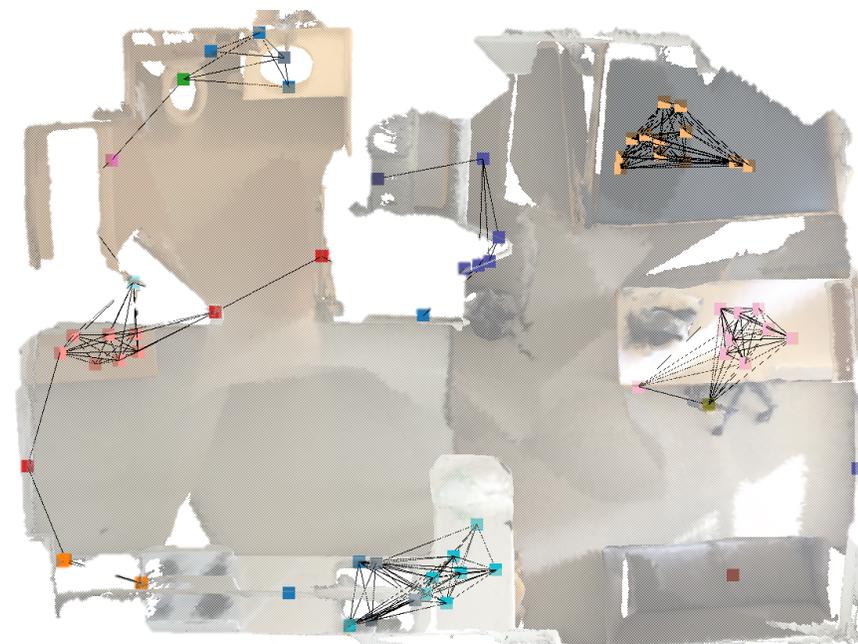
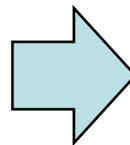
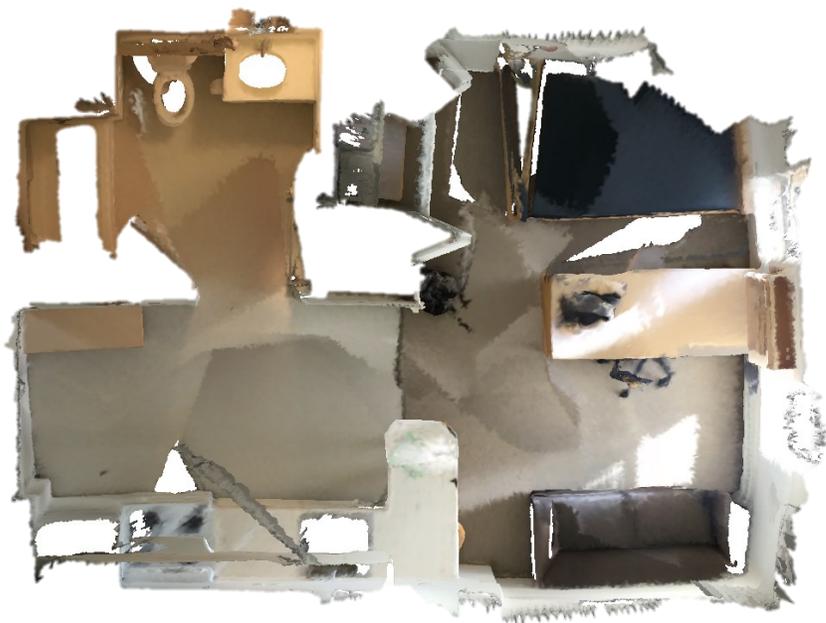
屋内3次元点群データ

[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

- 実空間の知識情報を確率を用いて取得

- 物体の存在確率
  - カテゴリ別の物体の数を取得
- 物体間の隣接確率
  - カテゴリ間の隣接の相関関係を取得

点：物体 (カテゴリを色分け)  
直線：隣接を取得 (1m)



# 取得した確率場表現 (全シーンを対象)

## ● 存在確率分布

- table, chair, door など、一般的な部屋に置かれる家具の存在確率が高い
- other furniture は、piano, display, trash can などが分類され、様々な家具が含まれたため存在確率が高い

## ● 隣接確率分布

- 同一カテゴリーの隣接確率が総じて高い
  - bookshelf は 92% → 本棚の隣に本棚が置かれる性質が反映
- 取得した隣接情報
  - shower curtain, toilet, bathtub は相互的に隣接確率が高い
  - chair の隣接は、table: 38%, desk: 8% (chair: 48%)
  - toilet の隣に refrigerator は存在しない

Category	Prob	Category	Prob
cabinet	7.54	counter	2.11
bed	4.98	desk	6.26
chair	19.61	curtain	1.27
sofa	3.44	refrigerator	1.28
table	22.52	shower curtain	0.41
door	8.56	toilet	0.81
window	5.55	sink	1.45
bookshelf	3.55	bathtub	0.57
picture	2.35	other furniture	7.76

	shower curtain	toilet	bathtub
shower curtain	23.0	19.0	25.0
toilet	8.3	25.1	11.0
bathtub	17.6	17.8	45.9

	chair	table	desk
chair	48.1	38.9	8.0
table	17.7	79.8	0.2
desk	23.2	1.2	64.6

## ● シーンタイプ別（隣接確率分布）

### ● Bathroom

- 全データと同様に、同一カテゴリーの隣接確率が総じて高い
- shower curtain, toilet, sink, bathtub に関連する隣接関係が強く出現
  - shower curtain, toilet, sink, bathtub の存在確率も他のシーンタイプに比べて高いため
- 物体と cabinet, door との隣接が多く検出
  - 部屋の規模が小さい、物体数が多いため

### ● Bedroom / hotel

- 全データと同様に、同一カテゴリーの隣接確率が総じて高い
- chair と table, desk の隣接、shower curtain, toilet, bathtub の隣接が多く検出
- 全データと比べて、window と curtain の隣接、desk と bookshelf の隣接が多く検出

## 確率場表現を事前知識として利用したベイズ推定

- セグメンテーション予測 (SparseConvNet<sup>[2]</sup> 手法)
  - 予測確率： 機械学習を用いて、物体が各カテゴリーであると予測した確率
  - 予測ラベル： 機械学習で取得した予測確率に対して、確率が最も高いカテゴリー
- 事前知識を用いた物体の確度修正
  - セグメンテーション予測の予測確率  $P(x)$  からベイズ推定により事前知識  $g(z_i|x)$  を含めた事後確率  $P(x|z_{1:k})$  を算出

$$P(x|z_{1:k}) = \frac{G(z_k|x) \times P(x|z_{1:k-1})}{\sum_{x' \in X} G(z_k|x') \times P(x'|z_{1:k-1})}$$

$$G(z_k|x) = p_{z_k}(z_k) \times g(z_k|x) + (1 - p_{z_k}(x_k)) \times (1 - g(z_k|x))$$

$$P(x|z_{1:0}) = pA(x)$$

$P_A(x)$  : セグメンテーション予測による物体 A のカテゴリー x の予測確率

$P(x|z_{1:k})$  : 物体の k 個の隣接物体  $Z_{1:k}$  とその予測ラベルを  $z_{1:k}$  とするとき、 $Z_{1:k}$  を観察下での物体の予測確率

$g(z_i|x)$  : カテゴリー x である物体の隣接物体  $z_i$  に対する隣接確率

## ● 多クラス分類の性能指標数値による評価

- Accuracy (正解率) : すべてのデータのうち正解であった割合
- Precision (適合率) : クラス  $L_i$  と予測したもののうちの正解の割合
- Recall (再現率) : クラス  $L_i$  データのうちの正しく予測された割合
- F1-Score (F-尺度) : Precision と Recall の調和平均

	Accuracy (正解率)	Precision (適合率)	Recall (再現率)	F1-score (F尺度)
ベイズ推定 適用前	65.3%	60.1%	59.9%	60.0%
ベイズ推定 適用後	67.5%	70.8%	62.4%	66.3%

適合率は 10% 上昇、再現率は 3% 上昇したが依然として低いクラスが多い

- ⇒
- ・ ベイズ推定を適用することで一定の物体予測の補正を達成
  - ・ ベイズ推定では補正できなかった物体予測も多数存在

- **まとめ**

- 実空間情報を確率的な場として表現
- 取得した確率場を用いた事前知識にもとづく物体推定手法への応用例を提示
  - 事前知識を用いた物体の確度修正として、予測確率の変化を確認

- **今後の展望**

- 複数の実空間情報を合わせた確率場表現の獲得
- 時系列を含めた実空間情報の確率場表現の獲得