

# 特別研究報告

題目

三次元点群データを用いた実空間確率場表現に基づく  
物体推定手法の提案と評価

指導教員

村田 正幸 教授

報告者

佐藤 広章

2021年2月9日

大阪大学 基礎工学部 情報科学科

## 内容梗概

近年、画像や動画などの情報から物体の位置やそのカテゴリを識別する物体識別技術が注目されている。さらに最近では、遠隔ショッピングサービスなどの新しいネットワークサービスが期待され、遠隔地の実空間理解や情報処理が求められている。これらの自動化されたサービスを提供するにあたり、計算機を用いた実世界情報の高い識別精度での取得が必要となる。加えて、リアルタイム性や通信量などを考慮すると、高い識別精度に加えて限られたデータ量での識別や低遅延の処理が必要になる。そこで本報告では、実空間の情報を機械学習を用いて確定的に識別するのではなく、実空間を確率的な場として表現することに取り組む。確率的な場とは、実空間にある物体を一意に識別するのではなく、物体カテゴリの確率的な重ね合わせによって表現される場であり、例えば機械学習における識別精度向上に活用したり、遠隔地の実空間情報が欠落した際のデータ補完に活用することができる。本報告では、屋内空間を測定した3次元点群データのデータセットを用いて実空間上に存在する物体の位置関係に基づく物体隣接関係の相関を統計的に求め、実空間を確率的な場として表現する。約1500室の屋内空間データを用いた数値計算によって、例えばデスクが椅子と隣接する確率は38.9%、冷蔵庫がバスタブと隣接する確率が0%となる有意な確率場表現を獲得した。さらに、確率場表現の利用例として、3次元点群データの取得の際に障害物などによって取得できなかった欠落箇所の物体推定、および、機械学習による識別の精度改善を提示し、確率場表現の有効性を評価した。その結果、欠落箇所の物体推定の正解率が平均で71.2%となることがわかった。また、機械学習による識別の精度改善例として、点群データに対するセマンティックセグメンテーションによる識別結果に対して、確率場表現にもとづくベイズ推定による識別精度の改善を図った。その結果、セマンティックセグメンテーションによる識別精度が50%程度となる物体について、識別精度が平均で80%に向上することがわかった。また、ベイズ推定によって識別結果が誤りとなる例を確認し、その発生要因を明らかにした。

## 主な用語

点群データ、確率場表現、物体識別、ディープラーニング、セマンティックセグメンテーション、ベイズ推定

# 目次

<b>1</b>	<b>はじめに</b>	<b>5</b>
<b>2</b>	<b>実空間情報の確率場表現</b>	<b>7</b>
2.1	実空間情報	7
2.2	確率場表現	7
2.3	確率場表現の際の物体検出方法	8
<b>3</b>	<b>3次元点群データセットを用いた確率場表現の例</b>	<b>9</b>
3.1	実空間データ	9
3.2	確率場表現の例	11
<b>4</b>	<b>確率場表現に基づく事前知識を用いた物体推定手法</b>	<b>17</b>
4.1	既存手法	17
4.2	応用例	18
4.3	実行例	19
4.4	評価結果	25
<b>5</b>	<b>おわりに</b>	<b>29</b>
	<b>謝辞</b>	<b>30</b>
	<b>参考文献</b>	<b>31</b>

## 目次

1	屋内3次元点群データ . . . . .	9
2	ScanNet データセットのシーンタイプの出現分布 . . . . .	10
3	物体隣接関係の例 . . . . .	11
4	欠落箇所のカテゴリ推定の実行例 . . . . .	19
5	ドアの隣接関係に対するセグメンテーション予測の実行例 . . . . .	21
6	窓の隣接関係に対するセグメンテーション予測の実行例 . . . . .	23
7	正解ラベルデータに対する予測確率と事後確率の相関 . . . . .	27
8	予測ラベルが変化するパターンに対する予測確率分布 . . . . .	27
9	正解ラベルから間違いラベルに予測が変化する例 . . . . .	28

## 表目次

1	物体カテゴリー	10
2	全シーンに対する存在確率(%表示)	12
3	全シーンに対する隣接確率(%表示)	13
4	シーンタイプ bathroom の物体の存在確率(%表示)	14
5	シーンタイプ bathroom における物体の隣接確率(%表示)	14
6	シーンタイプ bedroom、hotel に対する存在確率(%表示)	15
7	シーンタイプ bedroom、hotel に対する隣接確率(%表示)	15
8	SparseConvNet セマンティックセグメンテーション結果	17
9	実行例に対するカテゴリー推定結果	20
10	物体カテゴリー別のカテゴリー推定の正解率(%表示)	25
11	シーンタイプ別のカテゴリー推定の正解率(%表示)	26

## 1 はじめに

近年、画像や動画などの情報から物体の位置やそのカテゴリを識別する物体識別技術が注目されている [1]。例えば、車や通行人、建物などが混在している道路上のデータに対する物体識別、さまざまな種類の家具が配置されている屋内のデータに対する物体識別など、異なる状況下での物体識別の研究が多く行われている。これらは、自動車の自動運転や自動走行型ロボットの自律制御など、実空間上の情報を高い精度で認識しなければならない環境において、最も必要なタスクの1つに数えられている [2]。さらに最近では、遠隔ショッピングサービスなどの新しいネットワークサービスの登場が期待され、遠隔地で得られる情報から遠隔地の実空間理解や情報処理が求められている。具体的には、遠隔地で取得した実空間情報をネットワーク経由で転送し、アプリケーションを通して実空間情報から得たサービスをユーザに提供することが考えられる。

これらのネットワークサービスを実現するには、計算機を用いた実世界情報の識別が必要となる。また、サービスのリアルタイム性向上やネットワーク通信量削減のためには、高い識別精度に加えて限られたデータ量での識別や低遅延の処理が必要になる。現在は、実世界情報を動画像などの2次元データとして取得し、ディープラーニング等の機械学習技術を適用することで実世界情報をユーザーに提示する研究が進められている [3]。しかし、実世界情報は2次元データに縮約されるため、例えば物体の重なりなどの立体的な情報を扱うことはできない。そこで近年では、3次元データに対する機械学習を用いた物体識別の研究が盛んに行われている [4,5]。3次元データとして、Lidar センサーにより取得できる3次元点群データや、2つのカメラを用いる深度カメラにより取得する深度データなどが用いられる。既存の物体識別手法では、大量の学習データからデータに含まれる特徴を自動的に学習することで実装を行う手法である。しかし、既存の物体識別手法では、どのようなルールにより物体識別結果を導いているかを人間が理解するのは困難であることが指摘されている。また、我々人間が備えている実空間の知識情報、例えば台所とソファが隣接する可能性は低いなど、を無視した物体識別がなされる点に課題がある。

そこで本報告では、実空間の情報を機械学習等を用いて確定的に識別するのではなく、実空間を確率的な場として表現することに取り組む。確率的な場とは、実空間にある物体を一意に識別するのではなく、物体のカテゴリの確率的な重ね合わせによって表現される場であり、例えば機械学習における識別精度向上に活用したり、遠隔地の実空間情報が欠落した際のデータ補完に活用することができる。本報告では、多数のデータから実空間上に存在する物体の位置関係に対して隣接情報を取得し、実空間上の物体隣接にどのような相関関係があるかを統計的に求め、それを実空間の事前知識として与える [6]。さらに、事前知識と、観測等によって得られるデータに対して機械学習を適用して物体識別結果を重ね合わせることで、観測したデータに対する確率的な場としての確度を高める。実空間の統計的情報の取得には、3次元点群データ [7] を使用し、実空間の点群データを確率的な場として表現する。

さらに、確率場表現の利用例として、3次元点群データの取得の際に障害物などによって取得でき

なかった欠落箇所の物体推定、および、機械学習による識別の精度改善を提示する。これにより、遠隔地の実空間情報をネットワークを介して転送し利用するサービスにおいて、データ量が少ない点群データに対して実空間情報の認識向上を目指す。

本報告の構成は以下の通りである。まず2章では、本報告で示す実空間情報の確率場表現の方法を示す。次に3章では、実際の実空間情報を用いた確率場表現の例を示す。4章では、本報告で取得した確率場表現を事前知識として物体推定手法を提案する。最後に5章では、本報告のまとめと今後の課題を述べる。

## 2 実空間情報の確率場表現

### 2.1 実空間情報

本報告では、実空間に存在する物体の位置やカテゴリの情報に着目する。例えば、屋内データであれば家具などの物体、屋外データであれば車や人などの物体の位置やカテゴリの情報を取り扱うこととする。これらの情報を扱うことで、例えば、自動運転や自律制御に必要な障害物などの検出、遠隔ショッピングサービスなどの新しいネットワークサービスにてユーザの欲しい商品情報などの認識など、様々な応用が可能であると考える。

本報告では、実空間情報として3次元点群データを使用する。点群データとは、空間上の点の情報をまとめたデータであり、点ごとの空間座標情報 (X, Y, Z) と色情報 (R, G, B) の6パラメータが基本の構成要素となる。3次元点群データは、カラー画像と距離画像を同時に取得できるRGB-Dカメラや測定対象物にレーザー光線を照射してレーザーが返ってくるまでの時間と照射角度から対象物の座標値 (X, Y, Z) を取得するレーザー scanner によって測定するデータであり、複雑な形状の実空間を正確に扱うことができる。

### 2.2 確率場表現

確率的な場とは、実空間にある物体を一意に識別するのではなく、物体のカテゴリの確率的な重ね合わせによって表現される場である。本報告では、多数のデータから実空間上に存在する物体の位置関係に対して隣接情報を取得し、実空間上の物体隣接にどのような相関関係があるかを統計的に求め、それを実空間の事前知識とすることを考える。実空間を確率的な場として表現することで、例えばディープラーニング手法に組み入れたり、欠落したデータの補完に活用することができる。

#### 2.2.1 物体の存在確率分布

実空間情報から物体存在の統計的情報を取得する。実空間上に存在する物体情報を用いて、各カテゴリの物体数と物体の総数から各カテゴリの存在確率を算出することで、空間内の存在状況を取得する。

#### 2.2.2 物体の隣接確率分布

実空間情報から物体隣接の統計的情報を取得する。実空間上に存在する物体情報を用いて、物体ごとの重心点間距離から隣接している物体を探索する。注目物体に対してどのカテゴリの物体が隣接しているかの確率を算出し、隣接と判断された物体を集計することで空間内の隣接状況を取得する。隣接確率は、対象の物体カテゴリと隣接の物体カテゴリの隣接数を対象の物体カテゴリの隣接検出数で割ることで算出する。



## 2.3 確率場表現の際の物体検出方法

実空間情報から、物体の位置やカテゴリの情報を取得することを考える。物体検出の方法はいくつかあると考えられるが、本報告では、各カテゴリごとの点群データから密度ベースクラスタリングを実行することでそのカテゴリの物体位置を取得する方法をとる。これは、3次元点群データの取得の際に発生したノイズや鮮明度の問題から、実際の物体とは異なる形状を持つ可能性があるためである。例えば、3次元点群データとして取得した際に、現実空間上での物体と点群データ上での物体の形状に差異が生じる可能性があり、点群データ上で取り扱う際の物体は現実空間上をもとにした情報では不十分であるからである。

加えて、物体の隣接確率分布の取得の際に必要な点群データ内の点の距離計算や最近傍点探索に必要なデータ構造についても記述する。

### 2.3.1 密度ベースクラスタリング

3次元データに対するオープンソースライブラリーである Open3D [8] から、DBSCAN クラスタリング [9] を使用する。ある点から半径  $r$  以内の点に対して探索を行い、到達可能点が閾値  $n$  を超えている場合にクラスターと判断し、閾値  $n$  を超えない点はノイズと判断するアルゴリズムである。

### 2.3.2 近傍点探索

2.3.1 章と同様に、3次元データに対するオープンソースライブラリーである Open3D [8] から、高速最近傍点探索ライブラリ FLANN [10] を用いた KDTree の構築を使用する。KDTree は多次元空間上にある点を分類する空間分割データ構造であり、FLANN は次元数やデータ数によって変化する探索を高速化するライブラリーである。



図 1: 屋内 3 次元点群データ

### 3 3次元点群データセットを用いた確率場表現の例

実空間が反映されたデータセットを使用して、実際に確率場表現を取得する。本報告での確率場表現を得るために使用する実空間が反映されたデータセットとは、屋内空間もしくは屋外空間が反映され、物体情報を取得できるデータセットを意味する。

#### 3.1 実空間データ

実空間が反映されたデータセットとして、本報告では屋内 3 次元点群データセットである ScanNet [11] を用いる。この屋内 3 次元点群データセットは RGB-D カメラによって測量され、オフィス、アパート、バスルームなどシーンタイプにラベル付けされた 1613 部屋分のシーンに対して、座標情報と色情報を持つ図 1 のようなデータの集合である。

全 1613 部屋分のシーンに対して 1513 シーンが学習用データであり、本報告ではこの 1513 部屋分のデータセットを用いて確率場表現を得る。これらのシーンは学習用データセットとして、点群の空間座標と色の情報に加えて、表面の再構成やセマンティックセグメンテーションのアノテーションがつけられている。

図 2 に、ScanNet データセットに含まれる 1513 シーンのシーンタイプの出現分布を示す。縦軸がシーンタイプ名、横軸が各シーンタイプのシーン数を示す。21 種類のシーンタイプから構築される ScanNet データセットには、規模の小さいタイプ（バスルーム、クローゼット、ユニットバスなど）から大きなタイプの（アパート、教室、図書館など）までのさまざまな空間が含まれている。

各シーンには、部屋に設置されている物体に対応した点群オブジェクトが複数含まれている。以降、点群オブジェクトを物体と呼ぶ。物体には、その物体を表す物体カテゴリー識別子（以降、物体カテゴリー）が手作業によりアノテーションされている。表 1 は、ScanNet データセットの物体に対

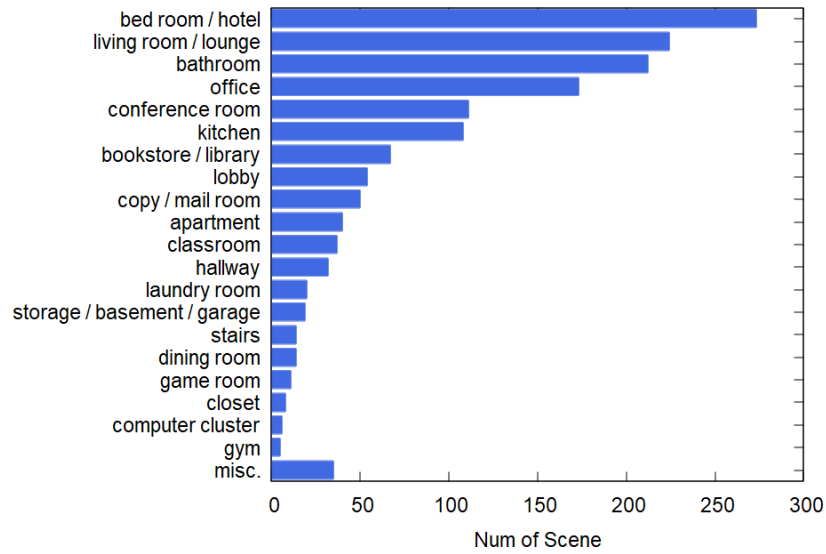


図 2: ScanNet データセットのシーンタイプの出現分布

して付与された物体カテゴリーの分類である。なお、other furniture には piano や display などが分類されている。本報告では、物体を抽出することで実空間情報の取得を試みるため、壁 (wall) と床 (floor) は考慮しないものとする。壁や床との隣接を取得する際、物体と壁や床が接している場合でも壁や床の重心から離れている場合に隣接と判断されない場合があるためである。

表 1: 物体カテゴリー

Label	Category	Label	Category
0	wall	10	picture
1	floor	11	counter
2	cabinet	12	desk
3	bed	13	curtain
4	chair	14	refrigerator
5	sofa	15	shower curtain
6	table	16	toilet
7	door	17	sink
8	window	18	bathtub
9	bookshelf	19	other furniture

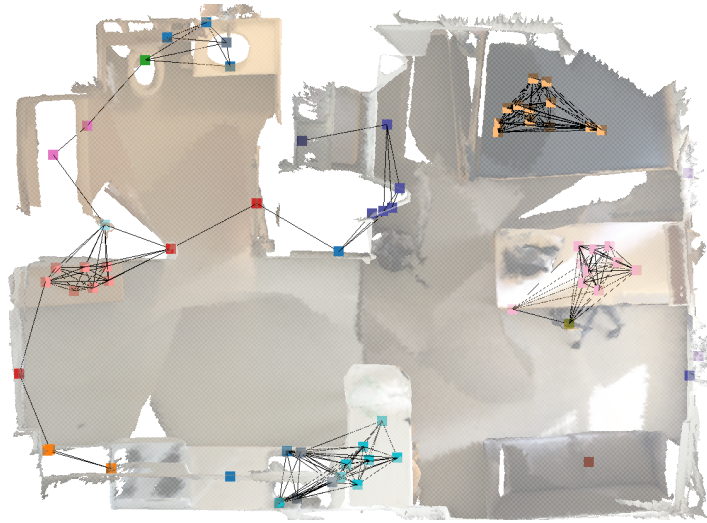


図 3: 物体隣接関係の例

## 3.2 確率場表現の例

本報告では、2章で述べた手法を適用して、2.3.1章の閾値を10、半径を5 [cm]、物体隣接の重心点間距離を1 [m]として取得する。

図3に確率場表現を取得した際の1つのシーンでの物体隣接関係を示す。点が物体の重心点、色が物体カテゴリー、直線が隣接を表している。この図から、2.3.1章で示す方法で物体を取得する場合、点群データの取得時に起こったノイズや鮮明度の違いから、現実空間上では同一である物体が複数の物体に分割される（図左中央部や図右上部、図下中央部など）ことが読み取れる。また、隣接状況において、deskとchairの隣接（図右中央部）、sinkとcounterとcabinetの隣接（図中央下部）、ユニットバスでのcabinetとtoiletとbathtubとshower curtainの隣接（図左上部）など、現実空間で現れる隣接状況が確認できるため、物体隣接情報を取得できていることが分かる。

### 3.2.1 データセット全体

学習用データの全1513シーンに対する、物体存在確率分布と物体隣接確率分布の情報を表2,3に示す。

表3から得た特徴を以下に記述する。まず、表の対角の値が総じて大きいことがわかる。例えば、bookshelf (Label 9) は92%であり、本棚の隣に本棚が置かれる性質が反映されている。また、shower curtain, toilet, sink, bathtub (Label 15-18) は低いパーセンテージとなっているが、機能上1つで十分なものであるにも関わらず、複数の物体が出現している。これは、2.3.1章で示す方法で物体を取得する場合、点群データの取得時に起こったノイズや鮮明度の違いから、現実空間上では同一である物体が複数の物体に分割されるためである。また、ユニットバス関係 shower curtain, toilet, bathtub

表 2: 全シーンに対する存在確率 (%表示)

Label	Category	Prob	Label	Category	Prob
0	wall	-	10	picture	2.35
1	floor	-	11	counter	2.11
2	cabinet	7.54	12	desk	6.26
3	bed	4.98	13	curtain	1.27
4	chair	19.61	14	refrigerator	1.28
5	sofa	3.44	15	shower curtain	0.41
6	table	22.52	16	toilet	0.81
7	door	8.56	17	sink	1.45
8	window	5.55	18	bathtub	0.57
9	bookshelf	3.55	19	other furniture	7.76

(Label 15, 16, 18) は相互的に隣接確率が高く、chair (Label 4) に着目すると、その隣接は、対角を除けば table (Label 6) 38% であり、続いて desk (Label 12) 8% である。加えて、toilet (Label 16) の隣に refrigerator (Label 14) は存在しない。これは、我々が理解している実空間上の特徴を反映していると考えられ、室内空間の特徴が確率的に表現できる可能性を示唆している。なお、bed (Label 3) の隣に toilet (Label 16) が 1 件 (0.005%) 発生しているが、これは、シーンの中で壁を挟んだ物体の隣接が反映されているものであり、重心点距離の観点から隣接を取得する方法から認識された隣接関係が抽出されたものであった。

表 3: 全シーンに対する隣接確率 (%表示)

隣接 対象	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	71.9	0.1	3.4	0.3	1.2	1.4	0.8	0.4	0.5	7.7	3.2	0.3	1.6	0.0	0.8	3.9	0.1	2.7
3	0.1	95.3	1.0	0.0	0.2	0.2	0.4	0.2	0.3	0.0	1.3	0.4	0.0	0.0	0.0	0.0	0.0	0.7
4	1.1	0.3	48.1	0.2	38.9	0.5	0.9	0.2	0.1	0.0	8.0	0.3	0.1	0.0	0.0	0.0	0.0	1.5
5	0.7	0.1	1.3	80.5	11.4	0.2	0.6	0.0	0.4	1.4	0.5	0.3	0.0	0.0	0.0	0.0	0.0	2.5
6	0.2	0.0	17.7	0.7	79.8	0.2	0.3	0.0	0.1	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.7
7	2.1	0.2	2.1	0.1	1.6	81.9	1.1	1.1	0.6	0.5	1.1	0.6	0.5	0.6	0.8	0.7	0.3	4.1
8	1.5	0.7	5.0	0.4	4.2	1.4	75.2	0.5	0.5	0.9	4.0	1.7	0.1	0.0	0.1	0.5	0.0	3.3
9	0.9	0.3	1.6	0.0	0.1	1.6	0.6	92.6	0.2	0.0	0.7	0.1	0.0	0.0	0.0	0.0	0.0	1.2
10	3.1	1.6	1.7	0.9	4.2	2.6	1.8	0.6	74.9	0.5	3.2	0.3	0.2	0.6	0.4	0.5	0.0	2.8
11	27.1	0.0	0.2	2.0	0.3	1.1	1.6	0.0	0.3	50.0	0.0	0.2	3.4	0.0	0.1	10.8	0.0	2.9
12	3.1	1.2	23.2	0.2	1.2	0.7	2.1	0.3	0.5	0.0	64.6	0.3	0.1	0.0	0.0	0.0	0.0	2.6
13	3.8	4.3	9.5	1.6	6.6	4.7	11.8	0.7	0.6	0.6	4.1	45.4	0.5	0.0	0.1	0.5	0.0	5.0
14	11.7	0.0	1.2	0.1	0.3	2.6	0.3	0.0	0.3	7.3	0.7	0.3	67.8	0.0	0.0	2.7	0.0	4.7
15	1.9	0.0	0.0	0.0	1.3	17.7	0.2	0.0	4.2	0.0	0.0	0.0	0.0	23.0	19.0	2.7	25.0	5.0
16	15.1	0.1	0.1	0.0	0.6	10.7	0.7	0.0	1.3	0.4	0.0	0.2	0.0	8.3	25.1	8.4	11.0	18.1
17	30.4	0.0	0.3	0.0	1.3	3.4	2.1	0.0	0.7	24.3	0.0	0.3	2.8	0.5	3.3	22.0	0.8	7.8
18	2.4	0.0	0.0	0.0	1.1	6.9	0.4	0.0	0.0	0.0	0.0	0.0	0.0	17.6	17.8	3.1	45.9	4.7
19	3.8	1.0	6.3	1.4	6.1	3.8	2.5	0.7	0.6	1.2	3.7	0.6	0.9	0.2	1.3	1.4	0.2	64.1

### 3.2.2 シーンタイプ別

3.2.1 章と同様に、シーンタイプ別の存在確率と隣接確率を表 4~7 に示す。シーンタイプが bathroom であるシーンに対する存在確率と隣接確率を表 4 と表 5、シーンタイプが bedroom や hotel であるシーンに対する存在確率と隣接確率を表 6 と表 7 に示す。

以下に、表 5 から得た特徴を以下に記述する。まず、表 3 と同様に、同一カテゴリーの隣接 (対角) の確率が高く、実空間情報を正しく認識していると考えられる。また、Label 15-18 (shower curtain, toilet, sink, bathtub) の物体数が他のシーンタイプに比べて多いため、関連する隣接関係が強く現れている。例えば、ユニットバスなどで見られるような相互的な隣接である shower curtain と toilet、bathtub との隣接や、同様に水回り周辺での counter と sink の隣接など、現実空間上で現れる隣接状況が正しく認識できていると考える。加えて、部屋の規模が小さいことや表 4 で物体数が多いことから、物体と cabinet、door の隣接が多くあらわれている。

以下に、表 7 から得た特徴をまとめる。まず、表 3 と同様に、同一カテゴリーの隣接 (対角) の確率が高く、また、chair と table、desk での隣接、shower curtain、toilet、bathtub での隣接も多くみられるため、実空間情報を正しく認識していると考えられる。加えて、window と curtain の隣接、desk と bookshelf の隣接など、実空間上でも考えられる隣接が表 3 と比べて現れている。さらに、表 3 と同様の隣接関係を持つが、chair に対する隣接が desk に比べて table が高いという特徴的な隣接が含まれていることがわかる。

表 4: シーンタイプ bathroom の物体の存在確率 (%表示)

Label	Category	Prob	Label	Category	Prob
0	wall	-	10	picture	3.47
1	floor	-	11	counter	1.91
2	cabinet	12.28	12	desk	0.00
3	bed	0.00	13	curtain	0.36
4	chair	0.28	14	refrigerator	0.12
5	sofa	0.40	15	shower curtain	6.02
6	table	0.48	16	toilet	12.75
7	door	26.54	17	sink	13.35
8	window	2.27	18	bathtub	8.57
9	bookshelf	0.00	19	other furniture	11.20

表 5: シーンタイプ bathroom における物体の隣接確率 (%表示)

隣接 対象	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	54.8	0.0	0.0	0.0	0.4	4.2	0.5	0.0	0.8	0.7	0.0	0.0	0.1	0.3	8.7	21.0	1.1	7.4
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	73.7	0.0	0.0	21.1	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	12.8	0.0	0.0	0.0	29.8	0.0	0.0	0.0	6.4	0.0	0.0	0.0	0.0	0.0	14.9	4.3	14.9	17.0
7	2.4	0.0	0.1	0.0	0.0	81.7	0.0	0.0	0.3	0.2	0.0	0.5	0.0	2.9	4.0	3.1	1.4	3.3
8	12.5	0.0	0.0	0.0	0.0	1.6	25.0	0.0	15.6	0.0	0.0	0.0	0.0	1.6	12.5	7.8	4.7	18.8
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	10.8	0.0	0.0	0.0	2.5	5.8	8.3	0.0	38.3	0.0	0.0	0.0	0.0	13.3	12.5	6.7	0.0	1.7
11	5.9	0.0	0.5	0.0	0.0	2.7	0.0	0.0	0.0	26.7	0.0	0.0	1.6	0.0	2.1	40.6	0.0	19.8
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	77.8	0.0	0.0	0.0	0.0	0.0	11.1	0.0	0.0	11.1	0.0	0.0	0.0
14	40.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	60.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	1.0	0.0	0.0	0.0	0.0	19.8	0.2	0.0	4.0	0.0	0.0	0.0	0.0	23.2	20.2	2.0	24.7	4.9
16	13.8	0.0	0.0	0.0	0.7	11.1	0.8	0.0	1.5	0.4	0.0	0.2	0.0	8.1	26.4	7.5	11.0	18.4
17	30.4	0.0	0.0	0.0	0.2	7.7	0.5	0.0	0.7	6.9	0.0	0.0	0.0	0.7	6.9	32.3	1.6	12.1
18	3.0	0.0	0.0	0.0	1.2	6.6	0.5	0.0	0.0	0.0	0.0	0.0	0.0	16.6	18.4	3.0	46.8	4.0
19	16.6	0.0	0.0	0.0	1.1	12.8	1.7	0.0	0.3	5.2	0.0	0.0	0.0	2.8	25.8	18.7	3.3	11.7

表 6: シーンタイプ bedroom、hotel に対する存在確率 (%表示)

Label	Category	Prob	Label	Category	Prob
0	wall	-	10	picture	4.46
1	floor	-	11	counter	1.05
2	cabinet	3.81	12	desk	11.92
3	bed	28.09	13	curtain	4.30
4	chair	10.73	14	refrigerator	1.45
5	sofa	0.99	15	shower curtain	0.25
6	table	2.61	16	toilet	0.33
7	door	12.76	17	sink	0.59
8	window	5.23	18	bathtub	0.29
9	bookshelf	1.27	19	other furniture	9.87

表 7: シーンタイプ bedroom、hotel に対する隣接確率 (%表示)

隣接 対象	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	47.8	1.6	4.1	0.8	2.3	4.6	0.4	0.6	1.5	16.2	3.5	0.6	3.5	0.2	0.7	5.5	0.0	5.9
3	0.1	95.6	0.8	0.0	0.1	0.2	0.4	0.2	0.3	0.0	1.4	0.4	0.0	0.0	0.0	0.0	0.0	0.8
4	1.1	3.5	35.0	0.1	7.2	1.7	0.8	0.4	0.6	0.1	39.5	2.7	0.7	0.0	0.0	0.0	0.0	6.4
5	4.5	1.8	2.3	67.0	9.1	0.0	0.5	0.0	2.7	0.0	2.7	1.8	1.4	0.0	0.0	0.0	0.0	6.3
6	2.8	1.1	31.4	1.9	46.2	1.4	1.4	0.0	0.9	1.1	0.3	6.7	0.0	0.0	0.0	0.0	0.0	5.0
7	1.3	0.8	1.7	0.0	0.3	84.8	0.3	0.4	0.5	0.2	2.9	0.2	0.1	0.2	0.1	0.1	0.1	5.9
8	0.5	7.4	3.5	0.1	1.3	1.4	50.8	0.7	1.9	0.0	8.8	15.9	0.0	0.0	0.0	0.0	0.0	7.6
9	2.4	11.8	6.9	0.0	0.0	6.6	2.4	42.9	2.1	0.0	15.2	4.5	0.0	0.0	0.0	0.0	0.0	5.2
10	2.2	6.7	3.2	0.7	1.0	2.8	2.3	0.7	62.6	0.0	9.3	0.8	0.6	0.7	0.1	0.2	0.0	6.0
11	20.4	0.0	0.5	0.0	1.1	0.8	0.0	0.0	68.5	0.0	0.0	0.0	1.5	0.0	0.0	4.6	0.0	2.5
12	0.6	4.0	24.7	0.1	0.0	1.8	1.3	0.6	1.1	0.0	61.2	1.0	0.3	0.0	0.0	0.0	0.0	3.2
13	0.9	8.6	13.8	0.5	7.8	0.8	18.8	1.5	0.8	0.0	8.5	32.4	0.6	0.0	0.0	0.0	0.0	5.1
14	10.9	0.2	8.2	0.7	0.0	0.7	0.0	0.0	1.2	3.7	5.0	1.2	57.9	0.0	0.0	0.0	0.0	10.1
15	2.4	0.0	0.0	0.0	0.0	11.9	0.0	0.0	7.1	0.0	0.0	0.0	0.0	23.8	20.2	1.2	26.2	7.1
16	11.0	1.2	1.2	0.0	0.0	6.1	0.0	0.0	1.2	0.0	0.0	0.0	0.0	20.7	17.1	2.4	19.5	19.5
17	40.6	0.0	0.0	0.0	0.0	3.5	0.0	0.0	1.2	27.1	0.0	0.0	0.0	0.6	1.2	11.8	0.0	14.1
18	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	27.8	20.3	0.0	38.0	10.1
19	2.5	5.2	9.6	0.5	1.7	8.8	2.6	0.5	1.7	0.8	7.7	1.5	1.4	0.2	0.5	0.8	0.3	53.8





表 8: SparseConvNet セマンティックセグメンテーション結果

Label	Category	IoU	Label	Category	IoU
0	wall	0.6584	10	picture	0.0011
1	floor	0.9222	11	counter	0.2883
2	cabinet	0.3690	12	desk	0.1460
3	bed	0.5673	13	curtain	0.2806
4	chair	0.7786	14	refrigerator	0.3087
5	sofa	0.6725	15	shower curtain	0.3683
6	table	0.4942	16	toilet	0.5172
7	door	0.2026	17	sink	0.3958
8	window	0.1047	18	bathtub	0.4552
9	bookshelf	0.1336	19	other furniture	0.2937
				avg IoU	0.4903

## 4 確率場表現に基づく事前知識を用いた物体推定手法

3.2.1 章で得た全 1513 シーンに対する確率場に対して、事前知識にもとづく物体推定手法への利用例を提示する。

### 4.1 既存手法

ScanNet データセットのセマンティックセグメンテーションタスクのベンチマーク [12] として上位に挙げられている SparseConvNet [13] モデルを使用する。3次元点群データに対するセマンティックセグメンテーションは、各点に対して各物体カテゴリーのラベル付けを行うタスクであり、ラベル付けは物体カテゴリーそれぞれの確率を算出して最大のもの取得する。

Scannet データセットを用いた SparseConvNet の識別精度を表 8 に示す。セマンティックセグメンテーションの識別精度は IoU (Intersection over Union) と呼ばれる指標を用い、物体カテゴリーを正しく認識できた領域の割合を示す。既存手法では、50%を超える正解率で物体識別ができていない物体カテゴリーがいくつかあることに対して、10%台の予測確率を持つ物体カテゴリーも存在する。

SparseConvNet などのディープラーニング手法は、大量の学習データからデータに含まれる特徴を自動的に学習することで実装を行う手法であり、どのようなルールにより物体識別結果を導いているかを人間が理解するのは困難であることが指摘されている。また、我々人間が備えている実空間の知識情報、例えば台所とソファが隣接する可能性は低いなど、を無視した物体識別がなされる点に課題がある。

---

**Algorithm 1** Algorithm for Object Category Estimation

---

**Require:**  $Next : N, P_{next} : 20 \times 20$

```
1:  $P_{label}[0..19] \leftarrow 0$   
2: for all  $l \in Label$  do  
3:    $P_{label}[l] \leftarrow prob(l, Next, P_{next})$   
4: end for  
5: return  $argmax(P_{label})$ 
```

---

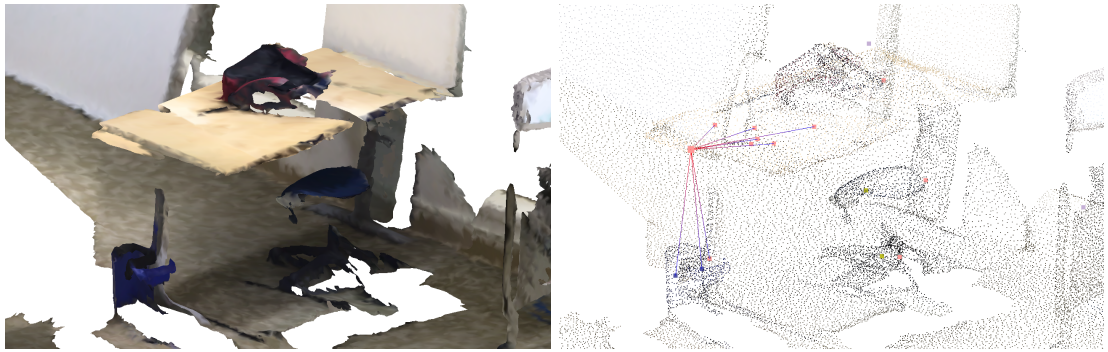
## 4.2 応用例

### 4.2.1 欠落箇所のカテゴリー推定

本節では、点群データが疎である箇所や空洞である箇所などの欠落箇所に対して、確率場情報にもとづいて点群データを補完することを想定し、欠落箇所のカテゴリー推定への応用例を示す。欠落箇所は、RGB-D カメラやレーザースキャナーなどで実空間の情報を取得する際に、ノイズの発生や障害物などによって取得できなかった箇所が該当する。ScanNet データセットにおいても欠落が発生している箇所が確認されている。このような欠落箇所を補完することで、実空間情報の情報量を増やすことが可能になったり、遠隔地で取得した実空間情報をネットワーク経由で転送する際に少ないデータ量でも良くなるなどが考えられる。

なお、欠落箇所の正解ラベルは付与されていないことから、本報告では、シーンの点群データからランダムな点を1つ選び取り除き、その点に隣接する物体のラベルと事前知識を用いて、その点の物体カテゴリーを推定することとした。ランダムに選んだ点に対して、物体のカテゴリーを推定するアルゴリズムを Algorithm 1 に示す。このアルゴリズムへの入力には、選んだ点に隣接する物体  $N$  個の物体カテゴリー  $Next$  と、事前知識である、物体カテゴリー全 20 種類間の隣接確率  $P_{next}$  を与える。点に対して全 20 物体カテゴリーの各ラベルを仮定し、推定確率として全隣接物体に対する隣接確率の重み付き平均を得る (line 2-4)。ここで隣接確率は、対象の物体カテゴリーと隣接の物体カテゴリーの隣接数を対象の物体カテゴリーの隣接検出数で割ることで算出しており、重み付き平均は全隣接物体に対する各カテゴリーの物体数を重みとしている。その後、最も推定確率の高いラベルを推定カテゴリーとする (line 5)。

シーンはアノテーション済みであり、取り除いた点の正解ラベルと推定結果の誤差を算出する。各シーンからランダムな1点に対して、アノテーションラベルと予測カテゴリーが一致するかを確認し、これを100回繰り返すことで全アノテーションシーンに対するの正解率を示す。なお、wall や floor とアノテーションされた点は、今回は確率場表現を取得していないため選択されないようにしている。



(a) 元データ

(b) 隣接点の探索結果

図 4: 欠落箇所のカテゴリ推定の実行例

#### 4.2.2 事前知識を考慮したセグメンテーション予測

取得した確率場表現を用いて、セマンティックセグメンテーション予測に対する物体推定を行う。セマンティックセグメンテーションの予測から、ベイズ推定により事前知識を含めた事後確率を算出する。

点群データ内のある物体  $Z$  に着目する。今、セマンティックセグメンテーションによる物体  $Z$  の予測ラベル  $P(A)$  ( $A$  はカテゴリ) を、物体  $Z$  の予測確率とする。さらに、物体  $Z$  の隣接物体集合とその予測ラベルを  $X$  とするとき、 $X$  を観察下での物体  $Z$  の事後確率  $P(A|X)$  は、ベイズ推定を用いて、

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)} \quad (1)$$

$$= \frac{P(X|A)P(A)}{\sum_A P(X|A)P(A)} \quad (2)$$

と求めることができる。ここで、 $P(X|A)$  は事前知識であり、4.2.1 章で述べたものと同様の、カテゴリ  $A$  の物体が各隣接物体に対する隣接確率の重み付き平均を用いている。

### 4.3 実行例

推定手法を適用して、物体が正しく認識できている例を示す。

#### 4.3.1 欠落箇所のカテゴリ推定

図 4 に実行例を示す。図 4a はシーンの点群データであり、図 4b は、その点群データに対して選んだランダムな 1 点と、抽出した物体の重心点との隣接を表している。図 4b において、大きい点が今回ランダムに選んだ点、一回り小さい点が物体の重心点、直線が隣接とみなされた関係である。

表 9: 実行例に対するカテゴリ推定結果

Label	Category	Prob	Label	Category	Prob
0	wall	0.0000	10	picture	0.0393
1	floor	0.0000	11	counter	0.0089
2	cabinet	0.0150	12	desk	0.0147
3	bed	0.0031	13	curtain	0.0623
4	chair	0.3057	14	refrigerator	0.0131
5	sofa	0.0944	15	shower curtain	0.0215
6	table	0.6222	16	toilet	0.0448
7	door	0.0216	17	sink	0.0276
8	window	0.0401	18	bathtub	0.0190
9	bookshelf	0.0037	19	other furniture	0.1903

点がカテゴリ  $a$  であると仮定して、各  $a \in A$  に対してカテゴリ  $a$  の推定確率は、この点が table 7つと other furniture 2つと隣接していることから、

$$\text{カテゴリ } a \text{ が table と隣接する確率} \times \frac{7}{9} + \text{カテゴリ } a \text{ が other furniture と隣接する確率} \times \frac{2}{9}$$

と計算する。隣接とみなされた関係に対して事前知識を用いたカテゴリ推定を行った結果を表 9 にまとめる。結果、最も確率の高いカテゴリである table と推定する。実際に取り除いた点のラベルは table であり、カテゴリ推定が正しく行われていることがわかる。

#### 4.3.2 事前知識を考慮したセグメンテーション予測

例として、図 5, 6 に実行例を示す。点がセグメンテーション予測に対する物体の重心点、色が家具カテゴリ、直線が隣接と判断された箇所を表す。

図 5 の左部の door と判断された赤色で表された物体の重心 3 点に着目する。以下の表は、予測確率とベイズ推定による事後確率を求めた結果である。予測確率は、既存手法であるセマンティックセグメンテーションタスクである SparseConvNet での推定結果を使用する。ドア中央に着目すると、予測確率（セマンティックセグメンテーション手法による推定結果）では、wall である確率が 48.71%、door である確率が 51.25%であったが、隣接物体が door 2つである情報を加えた事後確率では door である確率が 99.99%となっていることがわかる。ドア上部、ドア下部でも同様の傾向を示している。

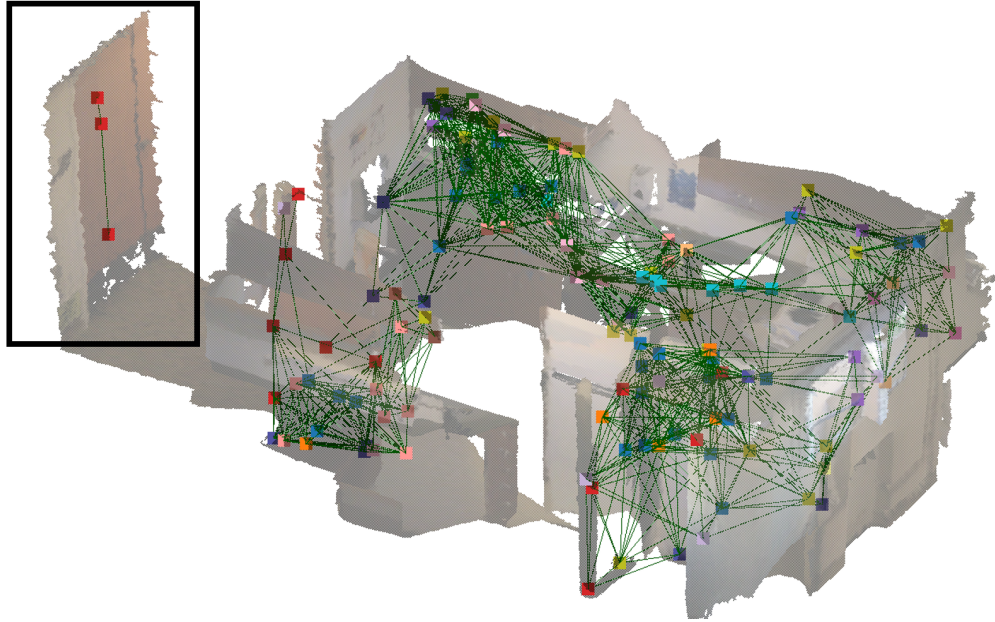


図 5: ドアの隣接関係に対するセグメンテーション予測の実行例

ドア上部 (座標 0.3298 1.7071 0.5403)

Label	Category	予測確率	事後確率	Label	Category	予測確率	事後確率
0	wall	0.2944	-	10	picture	0.0000	0.0000
1	floor	0.0238	-	11	counter	0.0000	0.0000
2	cabinet	0.0000	0.0000	12	desk	0.0000	0.0000
3	bed	0.0001	0.0000	13	curtain	0.0203	0.0018
4	chair	0.0000	0.0000	14	refrigerator	0.0000	0.0000
5	sofa	0.0003	0.0000	15	shower curtain	0.0000	0.0000
6	table	0.0000	0.0000	16	toilet	0.0000	0.0000
7	door	0.6597	0.9982	17	sink	0.0000	0.0000
8	window	0.0010	0.0000	18	bathtub	0.0000	0.0000
9	bookshelf	0.0001	0.0000	19	other furniture	0.0003	0.0000

ドア中央（座標 0.2541 1.6949 1.3984）

Label	Category	予測確率	事後確率	Label	Category	予測確率	事後確率
0	wall	0.4871	-	10	picture	0.0001	0.0000
1	floor	0.0000	-	11	counter	0.0000	0.0000
2	cabinet	0.0000	0.0000	12	desk	0.0000	0.0000
3	bed	0.0000	0.0000	13	curtain	0.0000	0.0000
4	chair	0.0000	0.0000	14	refrigerator	0.0000	0.0000
5	sofa	0.0000	0.0000	15	shower curtain	0.0000	0.0000
6	table	0.0000	0.0000	16	toilet	0.0000	0.0000
7	door	0.5125	0.9999	17	sink	0.0000	0.0000
8	window	0.0001	0.0000	18	bathtub	0.0000	0.0000
9	bookshelf	0.0001	0.0000	19	other furniture	0.0000	0.0000

ドア下部（座標 0.2221 1.6854 1.6016）

Label	Category	予測確率	事後確率	Label	Category	予測確率	事後確率
0	wall	0.4855	-	10	picture	0.0006	0.0000
1	floor	0.0000	-	11	counter	0.0000	0.0000
2	cabinet	0.0000	0.0000	12	desk	0.0000	0.0000
3	bed	0.0000	0.0000	13	curtain	0.0000	0.0000
4	chair	0.0000	0.0000	14	refrigerator	0.0000	0.0000
5	sofa	0.0000	0.0000	15	shower curtain	0.0000	0.0000
6	table	0.0000	0.0000	16	toilet	0.0000	0.0000
7	door	0.5137	0.9999	17	sink	0.0000	0.0000
8	window	0.0001	0.0000	18	bathtub	0.0000	0.0000
9	bookshelf	0.0001	0.0000	19	other furniture	0.0000	0.0000

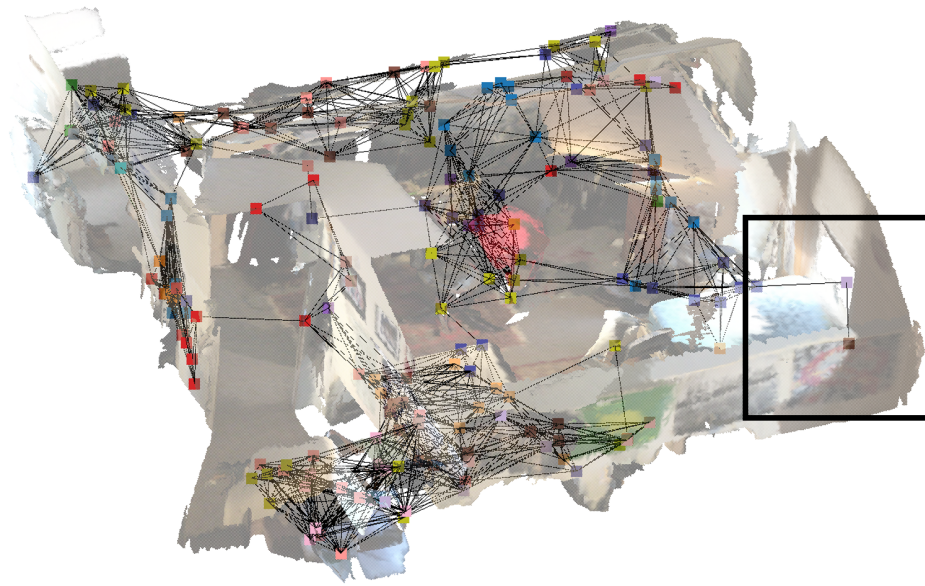


図 6: 窓の隣接関係に対するセグメンテーション予測の実行例

図 6 の右下部の窓周辺の重心 2 点に着目する。以下の表は、予測確率とベイズ推定による事後確率を求めた結果である。window (薄紫色) に着目すると、予測確率 (セマンティックセグメンテーション手法による推定結果) では、wall である確率が 23.88%、window である確率が 72.25%であったが、隣接物体が bed と sofa である情報を加えた事後確率では window である確率が 91.39%となっていることがわかる。同様に、sofa (茶色) に着目すると、予測確率 (セマンティックセグメンテーション手法による推定結果) では、floor である確率が 34.51%、sofa である確率が 36.40%であったが、隣接物体が window である情報を加えた事後確率では sofa である確率が 41.53%となっていることがわかる。floor は物体抽出の観点から確率場表現を取得していないため、事後確率を計算していない。この例では実ラベルである sofa の事後確率の上昇は見られず、物体の高さなど他の情報を加味することで改善されると考える。



予測カテゴリー：window 薄紫色（座標 6.1749 2.3014 0.8043）

Label	Category	予測確率	事後確率	Label	Category	予測確率	事後確率
0	wall	0.2388	-	10	picture	0.0003	0.0004
1	floor	0.0006	-	11	counter	0.0001	0.0002
2	cabinet	0.0001	0.0001	12	desk	0.0023	0.0022
3	bed	0.0118	0.0032	13	curtain	0.0207	0.0470
4	chair	0.0000	0.0000	14	refrigerator	0.0000	0.0000
5	sofa	0.0011	0.0313	15	shower curtain	0.0000	0.0000
6	table	0.0007	0.0003	16	toilet	0.0000	0.0000
7	door	0.0010	0.0014	17	sink	0.0000	0.0000
8	window	0.7225	0.9139	18	bathtub	0.0000	0.0000
9	bookshelf	0.0000	0.0000	19	other furniture	0.0000	0.0000

予測カテゴリー：sofa 茶色（座標 6.0302 1.8774 0.5305）

Label	Category	予測確率	事後確率	Label	Category	予測確率	事後確率
0	wall	0.0901	-	10	picture	0.0002	0.0007
1	floor	0.3451	-	11	counter	0.0000	0.0000
2	cabinet	0.0000	0.0000	12	desk	0.0014	0.0059
3	bed	0.1807	0.1435	13	curtain	0.0087	0.2062
4	chair	0.0016	0.0029	14	refrigerator	0.0000	0.0000
5	sofa	0.3640	0.4153	15	shower curtain	0.0000	0.0000
6	table	0.0016	0.0011	16	toilet	0.0000	0.0000
7	door	0.0019	0.0044	17	sink	0.0000	0.0000
8	window	0.0014	0.2122	18	bathtub	0.0020	0.0016
9	bookshelf	0.0001	0.0001	19	other furniture	0.0012	0.0060

表 10: 物体カテゴリー別のカテゴリー推定の正解率 (%表示)

Label	Category	Prob	Label	Category	Prob
0	wall	-	10	picture	86.11
1	floor	-	11	counter	42.05
2	cabinet	68.74	12	desk	83.08
3	bed	87.88	13	curtain	51.86
4	chair	47.68	14	refrigerator	79.27
5	sofa	80.80	15	shower curtain	31.07
6	table	82.70	16	toilet	48.38
7	door	92.18	17	sink	37.18
8	window	76.03	18	bathtub	89.17
9	bookshelf	88.51	19	other furniture	68.67

## 4.4 評価結果

### 4.4.1 欠落箇所のカテゴリー推定結果

全 1513 シーンに対して、シーンから 1 点を欠落させてカテゴリー推定を行う操作を 100 回行い、71.22%の結果を得た。加えて、表 10 は物体カテゴリー毎の正解率、表 11 はシーンタイプ別の正解率をまとめたものである。表 10 の物体カテゴリー別の正解率を見ると、bed, door, bookshelf など 90%近くのカテゴリーもあれば、shower curtain, sink など 40%を下回るカテゴリーもある。正解率の低い物体カテゴリーは、表 3 にて対角の確率が低いカテゴリーである。これは、欠落点が位置する物体に対して隣接確率が低くなるため、正解率が下がると考える。表 11 のシーンタイプ別の推定結果から、突出した推定結果を示すシーンタイプはないが、game room, stairs など比較的家具数の少ないシーンタイプに対して高い推定結果を得ている。

### 4.4.2 事前知識を考慮したセグメンテーション予測

全ての物体に対して、セマンティックセグメンテーションに対する予測確率と事後確率の相関を正解ラベルデータを元に調べ、有効性を確認する。図 7 に、予測ラベルと事後ラベルが物体カテゴリーと等しい物体予測に対して、予測確率と事後確率の相関を箱ひげ図で示す。補助線として破線が予測確率と事後確率が等しい位置を表している。予測確率を 0.1 で分割したものを示しており、全区分で事後確率が予測確率を上回っていることが分かる。したがって、予測ラベルと事後ラベルが物体カテゴリーと等しい物体予測に対しては確率場表現を用いることは有効であると考えられる。

確率場表現を用いた事後確率によって認識結果が変わる場合を考える。図 8 は、予測ラベルが変化したケースのヒストグラムを示している。横軸は物体の予測確率であり、区間を 0.05 としている。

表 11: シーンタイプ別のカテゴリ推定の正解率 (%表示)

Label	Type	Prob	Label	Type	Prob
0	apartment	67.43	10	game room	88.47
1	bathroom	69.12	11	gym	82.67
2	bedroom / hotel	73.96	12	hallway	85.76
3	bookstore / library	76.43	13	kitchen	62.99
4	classroom	72.00	14	laundry room	84.66
5	closet	74.87	15	living room / lounge	70.40
6	computer cluster	67.12	16	lobby	74.24
7	conference room	62.94	17	office	69.65
8	copy rooms / mail room	73.06	18	stairs	94.67
9	dining room	64.97	19	storage / basement / garage	81.11
			20	misc.	75.44

縦軸は区間内に現れるパターン数を示す。予測ラベルの変化は、予測確率最大のラベルと正解ラベルが不一致の状態から正解ラベルを導いたパターンと、一致の状態から間違いラベルを導いたパターンがあり、図中ではそれぞれピンク色の領域と緑色の領域で示している。図を見ると、予測確率が0.3以下では、予測ラベルが正解ラベルに変化したパターンが多いが、0.3~0.6では予測ラベルが間違いラベルに変化したパターンも現れていることがわかる。

図9に、予測ラベルが正解ラベルが間違いラベルに変化したパターンの一例を示す。toiletの一部がchairと予測され、toiletの予測確率が0.8751、tableの予測確率が0.080となっている。従って、セマンティックセグメンテーションでは高い確度でtoiletと識別していると言える。しかし、本手法を適用すると、隣接物体としてchairが存在することから、toiletの事後確率が0.0187、tableの事後確率が0.9786と変化し、結果としてtoiletではなくtableと識別される。このような例の発生を回避するには、予測確率が十分に高い物体に対しては手法を適用せず、図8に示すように予測確率が0.3以下となる確度の低い物体に対して手法を適用することが考えられる。また、シーンタイプを特定したのちにシーンタイプ別の確率場表現を用いることも考えられる。

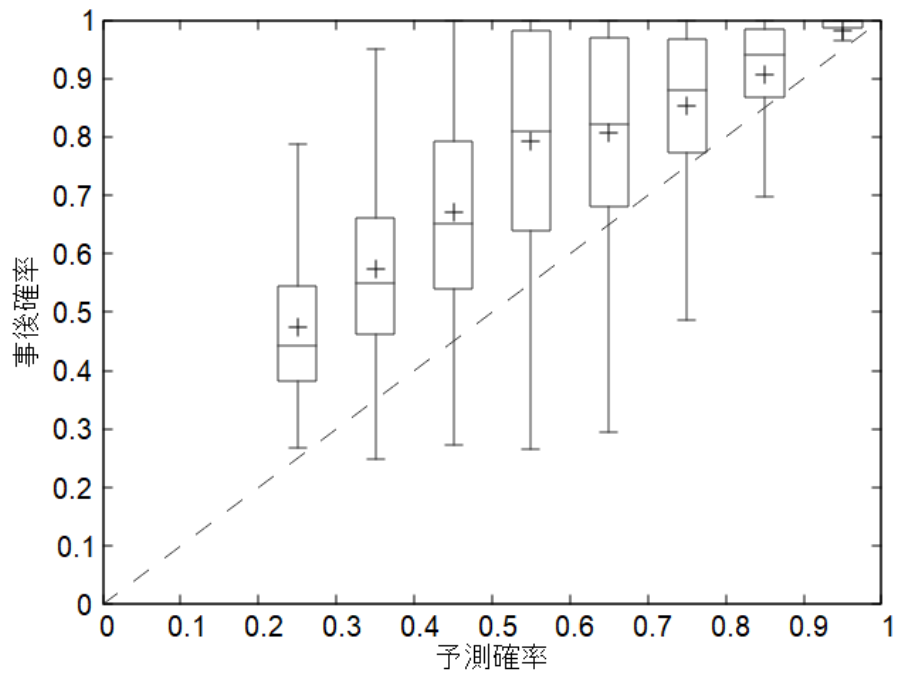


図 7: 正解ラベルデータに対する予測確率と事後確率の相関

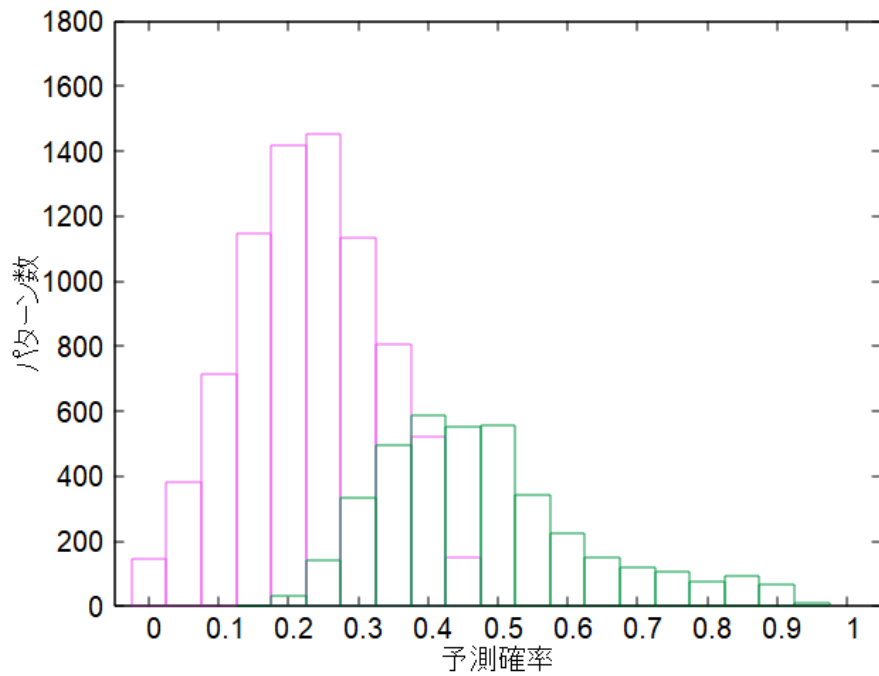


図 8: 予測ラベルが変化するパターンに対する予測確率分布

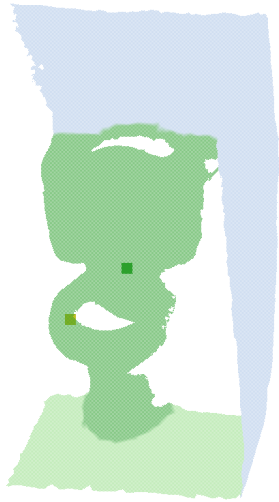


図 9: 正解ラベルから間違いラベルに予測が変化する例

## 5 おわりに

本報告では、実空間の情報を確率的な場として表現することに取り組んだ。実空間の情報のデータとして、屋内3次元点群データセットである ScanNet [11] から、実空間上に存在する物体の位置関係に対して隣接情報を取得し、実空間上の物体隣接にどのような相関関係があるかを統計的に求めた。その結果、実空間情報を正しく認識していると考えられる確率場を取得した。次に、取得した確率場を用いて、事前知識にもとづく物体推定手法への利用例を提示した。物体の一部の面に穴があり、その場所が推定された際にその推定箇所位置する点の物体カテゴリーを、事前知識を用いて推定することを考え、全体として71.22%の正解率を得た。加えて、取得した確率場表現を用いてセマンティックセグメンテーション予測に対する物体推定を行い、予測確率の変化を確認した。

本報告の測定では、実空間情報の確率場表現を物体推定手法に組み込めることを確認したが、実空間上で利用する物体推定手法としてはまだまだ改善の余地があると考えられる。加えて、例えば、他の屋内データセットや屋外データセットなど、複数のデータセットから実空間情報の確率場を取得することで、より正確な実空間情報を取得できると考えている。また、セマンティックセグメンテーションを実行することで、その予測結果を事前知識としてアップデートをしていくことも考えられる。

## 謝辞

本報告を終えるにあたり、大阪大学大学院情報科学研究科村田正幸教授にはご多忙の中貴重なご指導とご助言を賜りましたこと深く感謝いたします。大阪大学大学院情報科学研究科荒川伸一准教授には、進捗と方針の確認や論文執筆指導、加えて細かい質問や私事のことにお付き合いいただくなど手厚くご指導していただきました。心より厚く感謝申し上げます。また、平素よりご指導いただきました大阪大学先導的学際研究機構大下裕一准教授並びに大阪大学大学院情報科学研究科小南大智助教に心より感謝申し上げます。最後に、日々の学生生活を支えてくださった研究室の皆様に感謝の意を表し、謝辞とさせていただきます。

## 参考文献

- [1] T. de Vries, I. Misra, C. Wang, and L. van der Maaten, “Does object recognition work for everyone?,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 52–59, June 2019.
- [2] L. Wang, X. Fan, J. Chen, J. Cheng, J. Tan, and X. Ma, “3D object detection based on sparse convolution neural network and feature fusion for autonomous driving in smart cities,” *Sustainable Cities and Society*, vol. 54, p. 102002, Mar. 2020.
- [3] S. Takagi, J. Kaneda, S. Arakawa, and M. Murata, “An improvement of service qualities by edge computing in network-oriented mixed reality application,” in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 773–778, Apr. 2019.
- [4] J. Zhang, X. Zhao, Z. Chen, and Z. Lu, “A review of deep learning-based semantic segmentation for point cloud,” *IEEE Access*, vol. 7, pp. 179118–179133, Dec. 2019.
- [5] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, “Deep learning for generic object detection: A survey,” *International Journal of Computer Vision*, vol. 128, pp. 261–318, Oct. 2019.
- [6] J. Lafferty, A. McCallum, and F. C. Pereira, “Conditional random fields: Probabilistic models for segmenting and labeling sequence data,” in *Proceedings of the Eighteenth International Conference on Machine Learning*, pp. 282–289, June 2001.
- [7] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, “Deep learning for 3D point clouds: A survey,” *arXiv:1912.12033*, June 2020.
- [8] “Open3D - A Modern Library for 3D Data Processing.” <http://www.open3d.org/>. (Accessed: 2020-11-18).
- [9] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 226–231, Aug. 1996.
- [10] M. Muja and D. G. Lowe, “Fast approximate nearest neighbors with automatic algorithm configuration,” in *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pp. 331–340, Feb. 2009.



- [11] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, and M. Nießner, “ScanNet: Richly-annotated 3D reconstructions of indoor scenes,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2432–2443, July 2017.
- [12] “ScanNet Benchmark.” [http://kaldir.vc.in.tum.de/scannet\\_benchmark/](http://kaldir.vc.in.tum.de/scannet_benchmark/). (Accessed: 2020-10-05).
- [13] B. Graham, M. Engelcke, and L. van der Maaten, “3D semantic segmentation with submanifold sparse convolutional networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9224–9232, June 2018.