

## 複数種類のプリミティブによるパーツレベルの単視点3次元形状再構成 Part-Level Single-View 3D Shape Reconstruction with Multiple Types of Primitives

菊地 真美<sup>†</sup> 伊東 聖矢<sup>†</sup> 金子 直史<sup>†</sup> 鷺見 和彦<sup>†</sup>  
Mami Kikuchi Seiya Ito Naoshi Kaneko Kazuhiko Sumi

### 1. はじめに

単視点画像から物体の3次元形状を再構成することは、コンピュータビジョン分野における重要課題の一つである。従来の単視点3次元形状再構成は、3次元の概形を復元することには成功しているが、複数のパーツから構成される物体ではパーツの接続部が滑らかに表現され、正確さを欠くという問題があった。これに対し、物体を構造的に理解しながら3次元形状を復元することを目的として、パーツレベルで再構成する研究が行われている [1, 2]。Han ら [3] は、Compositional Generalizability [4] という概念を取り入れることで、未知のカテゴリに対してもパーツレベルで頑健に3次元形状を再構成することができる CompNet というニューラルネットワークを提案している。しかし、CompNet ではパーツを表現するプリミティブとして直方体しか用いていないため、曲面をもつ物体の再構成結果が不正確であるという問題がある。

そこで本研究では、パーツレベルの単視点3次元形状再構成において、複数種類のプリミティブを組み合わせることで、より多くの種類の物体を忠実に復元することを目的とする CompNet with Multiple Types of Primitives (CompNet-MTP) を提案する。提案手法では、Han らの手法である CompNet を拡張し、図 1 に示すようにプリミティブとして直方体だけでなく円柱も使用できるようにする。具体的には、CompNet に円柱のパラメータを推定するモジュールと直方体と円柱から適切なプリミティブを選択するモジュールを追加することで、複数種類のプリミティブを用いたパーツレベルの形状復元を可能にする。

### 2. 提案手法

提案する CompNet-MTP の全体像を図 2 に示す。提案手法のベースとなる CompNet [3] では、インスタンスセグメンテーション、回転予測、サイズ予測、接続点予測および位置予測という 5 種類のモジュールでネットワークを構成している。提案手法では、プリミティブとして直方体だけでなく円柱も使用するため、プリミティブごとにサイズ予測モジュールを用意している。また、パーツごとに直方体と円柱のどちらのプリミティブが適切であるかを判定するプリミティブ選択モジュールを追加することにより、各パーツを直方体または円柱で表現した3次元形状を再構成することができる。CompNet からの変更点は、円柱サイズパラメータ予測およびプリミティブ選択モジュールである (図 2 における赤枠)。これらの変更点について、以下で詳細に説明する。

#### 2.1 プリミティブのサイズパラメータ予測

提案手法では、直方体と円柱の 2 種類のプリミティブを用いて、物体パーツの3次元形状を表現する。直方体は、中心位置  $\mathbf{c} = [c_x, c_y, c_z]^T \in \mathbb{R}^3$ 、各辺の長さであるサイズ  $\mathbf{s}_r = [s_x, s_y, s_z]^T \in \mathbb{R}^3$  および回転 (クォータニオン)  $\mathbf{q} \in \mathbb{R}^4$  で表現できる。円柱は、直方体と同様に中心位置、サイズおよび回転で表現できるが、サイズ  $s_c$  は直径  $d$  と



図 1: CompNet-MTP での再構成結果 (上: chair, 下: table)

高さ  $h$  からなり、 $\mathbf{s}_c = [d, h]^T \in \mathbb{R}^2$  となる。

プリミティブのパラメータ予測では、はじめに入力画像  $I$  とインスタンスセグメンテーションにより得られる  $i$  番目のパーツマスク  $M_i$  を回転予測モジュールに入力し、 $i$  番目のパーツの回転パラメータ  $\mathbf{q}_i$  を得る。次に、 $i$  番目のパーツとそれと並進対称性をもつ複数パーツの入力画像、パーツマスクおよび回転パラメータを組みとして各プリミティブのサイズ予測モジュールに入力し、 $i$  番目のパーツの直方体と円柱のサイズパラメータ  $\mathbf{s}_r$ ,  $\mathbf{s}_c$  を獲得する。

#### 2.2 プリミティブ選択モジュール

提案手法では、プリミティブ選択モジュールによって物体パーツに適したプリミティブを選択する。まず、 $i$  番目のパーツとそれと並進対称性をもつパーツに対し、画像とパーツマスクからそれぞれ ResNet18 で特徴抽出する。次に、選択対象となるパーツの特徴を全パーツの特徴に対して Max Pooling を適用した特徴と結合する。最後に、結合した特徴から円柱である確率を求め、このネットワークの出力は、1 に近いほど円柱、0 に近いほど直方体が適していることを意味する。本研究では、閾値を 0.5 に設定し、閾値を超えた場合は円柱、それ以外は直方体とした。また、プリミティブ選択モジュールはバイナリ交差エントロピー損失を用いて学習した。

### 3. 評価実験

#### 3.1 実験設定

本研究では、3次元モデルに対してパーツ情報が付与されている PartNet データセットを用い、ベースライン手法の CompNet [3] と比較して、提案手法の有効性を評価する。ここで、プリミティブで構成された正解3次元モデルはパーツ情報をもとに各パーツに対して直方体または円柱を当てはめることで得られる。本実験では、直方体と円柱の両方を含む3次元モデルを、table カテゴリから 113 個、chair カテゴリから 102 個ランダムに選択した。提案手法の学習には chair カテゴリのみを用い、評価には chair カテゴリと table カテゴリの両方を使用する。

<sup>†</sup> 青山学院大学 Aoyama Gakuin University

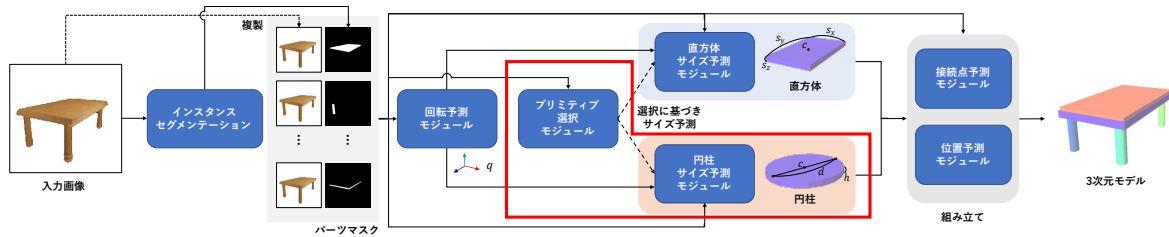


図 2: 提案手法の全体像 (赤枠はベース手法からの変更点の図示)

表 2: 提案手法のプリミティブ選択の精度

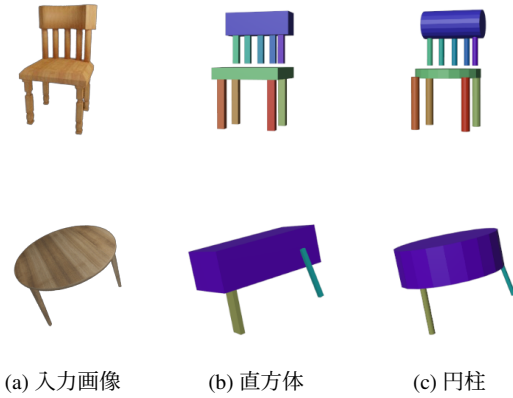


図 3: 各プリミティブ単体での再構成結果

	chair	table
	0.52	0.65

### 3.3 プリミティブ選択の精度

プリミティブ選択モジュールの性能を評価するために、直方体と円柱の2値分類における精度 (Accuracy) を算出した。結果を表2に示す。table カテゴリにおいて、精度が高いことが確認できる。これは学習データに平らな円柱パーツを持つモデルを多く使用したためだと考えられる。

図1は、提案手法のプリミティブ選択モジュールで選択したプリミティブを用いて生成した3次元モデルである。学習に使用した chair カテゴリでは、概形は正解モデルに近いものの、座面など直方体が適しているパーツで円柱が選択されてしまっている。このような誤ったプリミティブの選択は、特に脚部分で多く見られた。細長い円柱は直方体との形状の差が小さいことから、プリミティブの選択が難しいと考えられる。学習に使用していない table カテゴリにおいても、脚部分の選択は chair カテゴリと同様に誤っているが、円柱が適切なプリミティブである天板では正しいプリミティブが選択されている。このことから、未知カテゴリに対してもプリミティブ選択が可能であると言える。

### 4. おわりに

本研究では、直方体と円柱という2種類のプリミティブを組み合わせた、パーツレベルの単視点3次元形状再構成手法である CompNet-MTP を提案した。提案手法では、プリミティブとして直方体のみを使用する先行研究と同程度の精度を維持しながら、別のプリミティブを加えることができた。今後の課題として、プリミティブを高精度に選択することやプリミティブの種類を増やすことが挙げられる。

### 謝辞

データセット構築にご協力頂いた Songfang Han 氏に感謝の意を表す。本研究は JSPS 科研費 JP22K17978 の助成を受けたものである。また、本研究の一部は青山学院大学先端情報技術研究センター (CAIR) のプロジェクトの一環として行われた。

### 参考文献

- [1] G. Zhan et al., "Generative 3D Part Assembly via Dynamic Graph Learning", In NeurIPS (2019).
- [2] S. Tulsiani et al., "Learning Shape Abstractions by Assembling Volumetric Primitives", In CVPR (2017).
- [3] S. Han et al., "Compositionally Generalizable 3D Structure Prediction", arXiv:2012.02493 (2020).
- [4] T. Mu et al., "Refactoring Policy for Compositional Generalizability using Self-Supervised Object Proposals", In NeurIPS (2020).

### 3.2 ベースライン手法との再構成結果の比較

提案手法では、各パーツに対して直方体と円柱というそれぞれのプリミティブの形状を予測し、適切なプリミティブを選択する。提案手法の CompNet-MTP とベースライン手法の CompNet を Earth Mover's Distance (EMD) で比較した結果を表1に示す。ここで、EMD は3次元モデルからパーツ単位で3次元点を一様にサンプリングし、すべてのパーツに対する EMD の総和を算出したものである。表1の CompNet-MTP w/ GT はプリミティブの選択に正解のパーツ情報を使用した場合の結果である。CompNet-MTP は、学習に使用した chair カテゴリでは CompNet と同等の精度であり、未知の table カテゴリでは精度が低下する。一方、CompNet-MTP w/ GT は両カテゴリで CompNet より高精度であることから、プリミティブを適切に選択できれば、CompNet-MTP は CompNet よりも精度が向上することがわかった。

直方体と円柱それぞれのサイズパラメータを用いて生成した3次元モデルを図3に示す。円柱のみをプリミティブとした場合であっても、直方体のみとサイズに大きく差はないことが確認できる。なお、サイズパラメータ予測モジュールはモジュール単体で学習しているため、ベースライン手法と提案手法で結果に差はない。また、図3に示した机の天板は円形であるため、直方体よりも円柱のほうがプリミティブとして適切だと考えられる。

表 1: 各パーツからサンプリングした点群における EMD

Method	chair	table
CompNet [3]	0.056	0.124
CompNet-MTP	0.056	0.127
CompNet-MTP w/ GT	<b>0.055</b>	<b>0.123</b>