Vol.2018-CVIM-210 No.54
2018/1/19

ポイントクラウドを入力とした物体認識と形状補完ネットワークの提案

伊東 聖矢[1,2,a)]　鷲見 和彦[1]　入江 豪[2]　永野 秀尚[2]

概要：本研究では，ポイントクラウドからクラス識別と形状補完を同時に行うニューラルネットワークを提案する．実世界で観測される物体のポイントクラウドは部分的な形状しか得られない場合が多くあり，物体認識と形状補完を行うことは様々なアプリケーションを実現する上で重要である．昨今主流となっている畳み込みネットワークは，その性質上，3 次元データを扱うために入力をボクセルやデプス画像に変換する必要がある．しかし，これらは元の 3 次元情報を離散化したり，2 次元に写像するため，情報欠落が生じるという問題がある．そこで本研究では，ポイントクラウドから大域的特徴と局所的特徴を抽出し，大域的特徴からクラス識別を行うとともに，その大域的特徴と局所的特徴から形状補完を行う．

1. はじめに

近年，3 次元計測技術の発展により，リアルタイムで手軽に計測できる RGB-D カメラやレーザセンサが普及している．これらのセンサで計測されたデータは 3 次元点の集合であるポイントクラウドとして扱われることが多い．実世界で取得したポイントクラウドデータをコンピュータグラフィクスなどで扱うためには，完全な形状のモデルを再構成する必要があるが，一度の計測で得られるポイントクラウドは，形状の一部である．そのため，対象物体の完全な形状を取得するために，様々な視点から計測し，ポイントクラウドデータを統合する必要がある．しかし，取得データは物体材質によるノイズやオクルージョンなどの影響を受けやすく，データ統合によって完全な形状のモデルを取得することは困難である．

近年，深層学習，特に Convolutional Neural Networks (CNNs) の成功を受け，CNNs を 3 次元データに適用しようとする試みが報告されている．これまでは，2 次元畳み込みを 3 次元畳み込みに置き換えた CNNs の適応が検討の中心であった．ポイントクラウドについても，ボクセルやデプス画像に変換することによって，3 次元畳み込みで扱うことが可能である．しかし，3 次元畳み込みは計算コストが高いため，ポイントクラウドをボクセル化する際，解像度を低くする必要がある．一方，デプス画像に変換する場合は，3 次元の形状を保持するために，複数の視点でデプス画像を生成する必要がある．

そこで本研究では，ポイントクラウドを入力とし，物体クラスと物体形状を補完するポイントクラウドを出力とするネットワーク PointsCompletionNet (PCNet) を提案する．提案する PCNet は，PointNet [1] と生成モデルである Generative Adversarial Networks (GAN) [2] から着想を得ており，GAN の Generator と Discriminator の双方で物体クラス特徴を加味して，物体の形状を補完するポイントクラウド生成を行う．

2. 関連研究

2.1 形状補完

形状補完には様々なアプローチがあり，代表的なものとして，幾何学的なアプローチとテンプレートベースのアプローチがある．

幾何学的なアプローチでは，局所領域に平面，または平滑なボリュームを仮定することで，小さな隙間を補完することができる．欠損が大きい場合は，ある特定の形状カテゴリに対して，経験則に基づいて形状を仮定し，形状補完を行う．Schnabel らは、物体の形状が平面，円柱などのプリミティブな要素の集合であると仮定して，不完全なポイントクラウドから CAD モデルを再構成する手法を提案した [3]．Li らは，局所的にフィッテングされたプリミティブな要素に対し，要素間の向きや配置といった幾何学的関係を考慮し，形状を最適化している [4]．その他に，Speciale らは不完全なポイントクラウドから対称性を検出し，複数

1 青山学院大学
Aoyama Gakuin University
2 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation
a) ito.seiya@vss.it.aoyama.ac.jp

© 2018 Information Processing Society of Japan

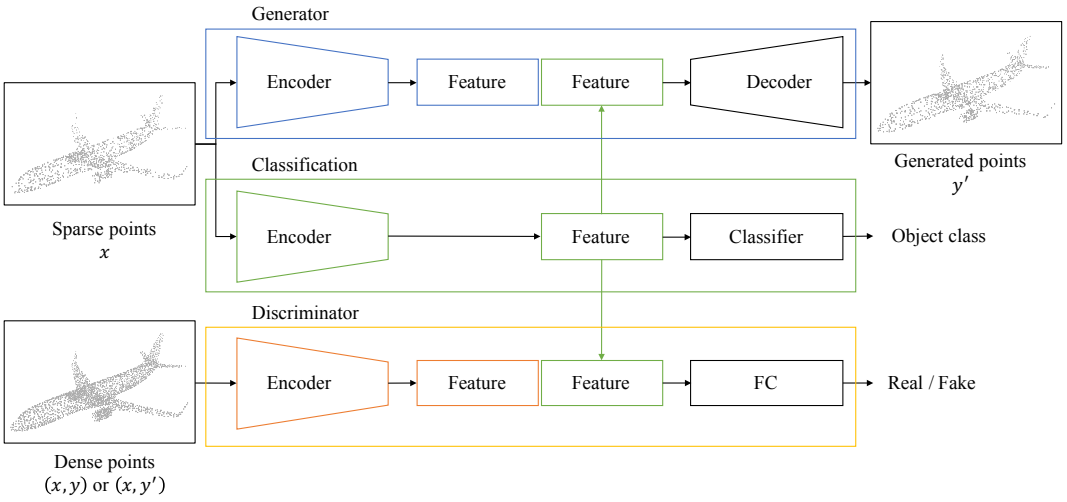


図 1: PointsCompletionNet のネットワーク構造

の対称性から形状を補完した [5].

テンプレートベースのアプローチは，変形可能なテンプレートまたは最近傍によって入力形状の再構成を行う．Pauly らは，データベースから最も類似した形状を取り出し，それをテンプレートとして使用して，入力データに合わせて変形した [6]．しかし，これらの手法は，複雑な入力に対してはユーザーへの対話が必要であり，入力に近いテンプレートが存在しないときは，形状補完に失敗することがある．この問題に対して，Sung らは構造と部品の関係に着目し，データベースから最良の部分候補を見つけ出し，その組み合わせの最適化問題を解くことによって，部分スキャンデータから形状の再構成を行った [7].

近年では，深層学習による形状補完手法が提案されてきている．Wu らは，3 次元グリッドにおけるボクセル占有率を表すバイナリ変数に対する確率分布を Convolutional Deep Belief Networks (CDBN) で学習し，形状を補完している [8]．Nguyen らは，CDBN とマルコフ確率場を組み合わせて，最大事後確率推論問題として形状補完を定式化した [9]．Sharma らは，Encoder-Decoder 型のネットワークを End-to-End で学習し，形状補完を行なっている [10]．これらの手法は，3 次元畳み込みの計算コストが高いため，入力を粗いグリッドでボクセル化している．Dai らは，粗いグリッド上で完全な形状を推定するための 3 次元 Encoder-Predictor Network (EPN) を提案し，最近傍探索に基づくボリュームのパッチ合成をすることで，詳細な形状を再構成している [11]．Han らは，さらに大域的な形状構造と，パッチレベルのサーフェスを推定する局所的な EPN を学習している [12].

本研究では，粗いグリッドでのボクセル化や，詳細な形状を得るためのパッチ合成をせず，ポイントクラウドを直接入力とし，形状補完を行うネットワークを構築する．

2.2 ポイントクラウドを入力とした深層学習

ポイントクラウドは，画素やボクセルと異なり，順序性がなく，データの順序が変わっても同一のデータとして扱わなくてはならない．そのため，ボクセルやデプス画像などに変換した上で処理されることが多かった．これに対し，Qi らは，ポイントクラウドを入力として，物体認識やセマンティックセグメンテーションを行うネットワーク構造 PointNet を提案している [1]．ポイントクラウドの各点に対して畳み込みを行い，得られた特徴に対して Max pooling を行うことで，データ順序に不変な大域的特徴を抽出している．ポイントクラウドから特徴抽出する際は，T-Net という回転を考慮したネットワークを中間に挟むことで，ポイントクラウドの回転不変性に対処している．

3. PointsCompletionNet

本研究では，ポイントクラウドを入力とし，物体認識と形状補完を同時に行うネットワーク PointsCompletionNet (PCNet) を提案する．提案する PCNet のネットワーク構造を図 1 に示す．入力となる疎なポイントクラウドを x，入力の形状を補完するポイントクラウドを y'，形状補完の教師となるポイントクラウドを y とする．PCNet は，入力ポイントクラウド x に対して物体認識を行う Classification Network，形状補完ポイントクラウド y' を生成する Generator，入力ポイントクラウドと補完ポイントクラウドの集合 (x, y')，または入力ポイントクラウドと教師ポイントクラウドの集合 (x, y) の Real / Fake を識別する Discriminator で構成される．本研究では，Classification Network に PointNet [1] を使用する．

3.1 ネットワーク構造

PCNet の Generator 及び Discriminator のネットワー

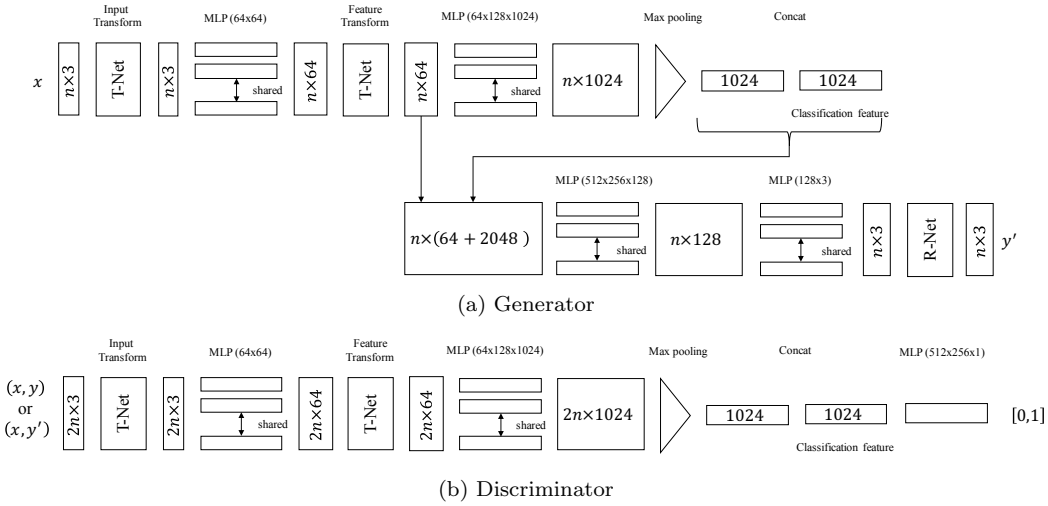


図 2: Generator と Discriminator のネットワーク構造の詳細

ク構造の詳細を図 2 に示す．Generator と Discriminator は，まず入力ポイントクラウドの各点から局所的特徴を抽出した後，全特徴の各次元に対して Max pooling を行い，大域的特徴を得る．PCNet では，入力ポイントクラウドから抽出した大域的特徴と，Classfication Network で得られるクラス識別特徴を結合する．クラス識別特徴を結合することで，入力ポイントクラウドに対して，明示的に物体形状に関する特徴を与えることができる．Generator では，各点から抽出した中間の局所的特徴に対して，大域的特徴を結合し，畳み込みを行って 3 次元のベクトルを得る．最終層では，入力ポイントクラウドと生成されたポイントクラウドの回転と平行移動による位置ずれを R-Net によって軽減する．R-Net は，PointNet [1] における T-Net を拡張し，回転行列だけでなく，平行移動ベクトルを学習する．Discriminator では，結合した大域的特徴に対して畳み込みを行い，最終層の活性化関数にシグモイド関数を用いて，区間 $[0, 1]$ の値を出力する．

3.2 損失関数

PCNet における Generator を $G(x;\theta_g)$，Discriminator を $D(x, y;\theta_g)$ とする．$G(x;\theta_g)$ は，観測ポイントクラウド x に対して，形状補完ポイントクラウド生成の確率分布 $p_{(x)}$ を近似するようにパラメータ θ_g を学習する．一方 $D(x, y;\theta_d)$ は，形状補完ポイントクラウド y が，Generator によって生成されたポイントクラウドである確率を算出するようにパラメータ θ_d を学習する．D は $logD(x)$ が最大となり，G は $log(1 - D(x, G(x))$ が最小となるように，それぞれ学習する．これらをまとめた損失関数 $V(D, G)$ とすれば，

$$\begin{aligned} V(D,G) =& \mathbb{E}_{(x,y)\sim p_{(x,y)}}[logD(x,y)] \\ &+ \mathbb{E}_{x\sim p_{(x)}}[log(1 - D(x, G(x))] \end{aligned} \tag{1}$$

と表せる．

本研究では，$V(D, G)$ に加え，生成点群が物体の形状を再現するような損失関数 R を設計する．まず，入力ポイントクラウドと形状補完の教師ポイントクラウドに対し，安定マッチングによって，一対一の点対応ペアを作成する．安定マッチングには，Gale-Shapley のアルゴリズム [13] を用いる．安定マッチングは選好リストに基づいており，得られるペアが，現在の対応よりも選好順位が高くなるペアが存在しないように，最適化する．本研究では，選好リストをユークリッドの距離の昇順で作成し，入力ポイントクラウドと教師ポイントクラウドの各点の距離がより小さくなるようにペアを作成する．Generator は，入力ポイントクラウドの各点から生成される形状補完点が，教師点を近似するように学習する．ここで，損失関数 R は，

$$R(G) = \sum_i ||y'(i) - y(i)||_2 \tag{2}$$

とし，最終的に最適化する損失関数は，

$$\mathcal{L} = \min_G \max_D V(D, G) + \lambda R(G) \tag{3}$$

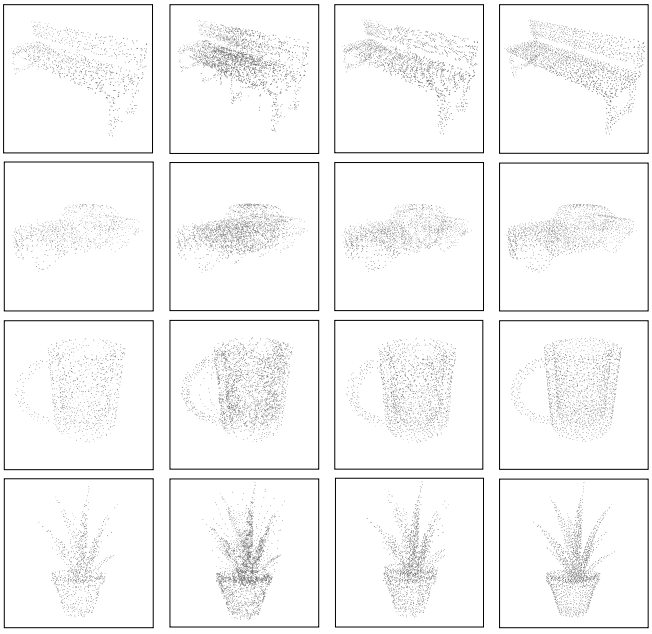
であり，λ は重み係数である．

4. 実験

提案手法の性能を評価するために，ModelNet40 [8] を用いて実験を行う．ModelNet40 は，40 クラスの単一オブジェクトの CAD モデルであり，PointNet[1] では頂点数を 2048 点に統一している．本実験では，2048 点からなるポイントクラウドを，入力ポイントクラウドと形状補完の教師ポイントクラウドに分割し，それぞれ 1024 点となるように整形した．データの分割にはランダムサンプリングを用い，1 つのモデルから 10 回のサンプリングを行なって，データ拡張を行った．PCNet は，まず Classfication Network を学習し，パラメータを固定する．次に，R-Net

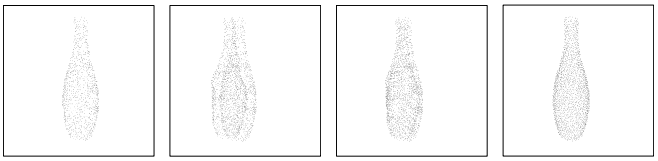
を除いた Generator と，Discriminator を交互に最適化し，最後に R-Net 単体を最適化する．なお，損失関数の最適化に確率的勾配降下法を用い，学習率は 10^{-10}，損失関数の重み係数は $\lambda = 0.1$ とした．

形状補完の比較結果を図 3 に示す．図 3 はそれぞれ，(a) 入力ポイントクラウド， (b) ドロネー分割で得られる三角形メッシュの重心に点を補完する手法， (c) 提案手法，(d) 形状補完の教師ポイントクラウドであり， (b)-(d) は入力ポイントクラウドを加えている．ドロネー分割を利用した形状補完では，全体の形状を保ったままポイントクラウドを密にできているが，全体的にノイズがかかったような結果となった．一方で，提案手法による形状補完は，Generator のポイントクラウド生成に物体クラス識別特徴を加えたことで，ノイズがかかることなく，入力ポイントクラウドに沿って補完ポイントクラウドを生成している．また，図 4 に示すように，R-Net は形状補完ポイントクラウドが教師ポイントクラウドに近づくような変換行列を学習し，R-Net なしの場合と比べて，位置ずれを軽減している．



(a) input (b) Delaunay (c) PCNet (d) GT

図 3: 形状補完の結果の比較



(a) input (b) w/o R-Net (c) w/ R-Net (d) GT

図 4: R-Net の効果

5. おわりに

本研究では，疎なポイントクラウドに対して物体認識を行い，物体クラス特徴を加味してポイントクラウドの形状補完を行うネットワーク PointsCompletionNet を提案した．今後，様々な欠損パターンで提案手法の有効性を検証する必要がある．また，単一オブジェクトの形状補完だけでなく，シーンを対象としたポイントクラウド補完ができるように手法を改良していく．

参考文献

[1] C. R. Qi, H. Su, K. Mo and L. J. Guibas: PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation, *Proc. Computer Vision and Pattern Recognition* (2017).

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio: Generative Adversarial Nets, *Advances in Neural Information Processing Systems 27*, pp. 2672–2680 (2014).

[3] R. Schnabel, P. Degener and R. Klein: Completion and Reconstruction with Primitive Shapes, *Computer Graphics Forum (Proc. of Eurographics)*, Vol. 28, No. 2, pp. 503–512 (2009).

[4] Y. Li, X. Wu, Y. Chrysathou, A. Sharf, D. Cohen-Or and N. J. Mitra: GlobFit: Consistently Fitting Primitives by Discovering Global Relations, *ACM SIGGRAPH 2011 Papers*, ACM, pp. 52:1–52:12 (2011).

[5] A. C. Pablo Speciale, Martin R. Oswald and M. Pollefeys: A Symmetry Prior for Convex Variational 3D Reconstruction, *European Conference on Computer Vision* (2016).

[6] M. Pauly, N. J. Mitra, J. Giesen, M. Gross and L. J. Guibas: Example-based 3D Scan Completion, *Proc. the Third Eurographics Symposium on Geometry Processing* (2005).

[7] M. Sung, V. G. Kim, R. Angst and L. Guibas: Data-driven Structural Priors for Shape Completion, *ACM Trans. Graphics*, Vol. 34, No. 6 (2015).

[8] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang and J. Xiao: 3D ShapeNets: A Deep Representation for Volumetric Shapes, *Proc. Computer Vision and Pattern Recognition* (2015).

[9] D. T. Nguyen, B. S. Hua, M. K. Tran, Q. H. Pham and S. K. Yeung: A Field Model for Repairing 3D Shapes, *Proc. Computer Vision and Pattern Recognition* (2016).

[10] A. Sharma, O. Grau and M. Fritz: *VConv-DAE: Deep Volumetric Shape Learning Without Object Labels*, pp. 236–250 (2016).

[11] A. Dai, C. R. Qi and M. Nießner: Shape Completion using 3D-Encoder-Predictor CNNs and Shape Synthesis, *Proc. Computer Vision and Pattern Recognition* (2017).

[12] X. Han, Z. Li, H. Huang, E. Kalogerakis and Y. Yu: High-Resolution Shape Completion Using Deep Neural Networks for Global Structure and Local Geometry Inference, *International Conference on Computer Vision* (2017).

[13] D. Gale and L. S. Shapley: College Admissions and the Stability of Marriage, *The American Mathematical Monthly*, Vol. 69, No. 1, pp. 9–15 (1962).

© 2018 Information Processing Society of Japan