

研究終了報告書

「高精度 3次元物体認識を実現する Denoising & Mesher DNN の構築」

研究期間：2019年10月～2023年3月

研究者：宮内 翔子

1. 研究のねらい

医療支援システムや農業支援システムなどのヒト支援システムにおいては、専門知識を有する人間が認識できる情報を、いかに的確にシステムに認識させるか、ということが重要な課題の一つである。この課題を解決する手法として、画像情報を入力とするDNNが多数提案されており、DNNを用いることで、高精度な画像認識が可能となっている(W. Liu et al., Neurocomputing 2017)。しかし、実際にヒトが認識している空間は3次元であり、支援システムにおける識別・分類精度や安全性を高めるためには、画像から得られる2次元情報のみでなく、実環境に即した3次元情報も考慮できることが望ましい。

この要求に対し、RGB-Dカメラのような安価な3Dセンサから得られる注目物体の3次元形状情報を入力とした、3次元物体認識のためのDNNの需要が高まりつつある。DNNの入力情報は、固定長のベクトルで表され、この入力ベクトルの要素の並びには規則性がなければならない。また、大量のデータを効率的に学習し、高精度に物体を認識するためには、特徴的な形状を捉えた短いベクトルで、3次元形状の細部まで記述する必要がある。

これらの条件を満たす手法として、メッシュモデルを用いる手法がある(N. Umetani, Siggraph Asia 2017)。しかし、3Dセンサから固定長ベクトルで表されるメッシュモデルを直接得ることはできず、ノイズも発生しやすいため、形状の細部まで捉えた学習データを作成することは依然として難しい。質の高い学習データを大量に取得することが困難であるため、画像認識精度に比べ、DNNを用いた3次元物体認識精度は未だ低い。

センサから取得されたノイズや欠損を含む物体の点群から、同一クラス内の物体間の意味的対応を考慮しつつ、共通のメッシュ構造で記述された「意味的構造化モデル」を生成する、Denoising & Mesher DNN (DMDNN) (図1)を構築する。ここでの意味的対応とは、同一クラス内で物体を構成する共通のパーツ間が対応することを指す。



図1. 提案する Denoising & Mesher DNN

2. 研究成果

(1) 概要

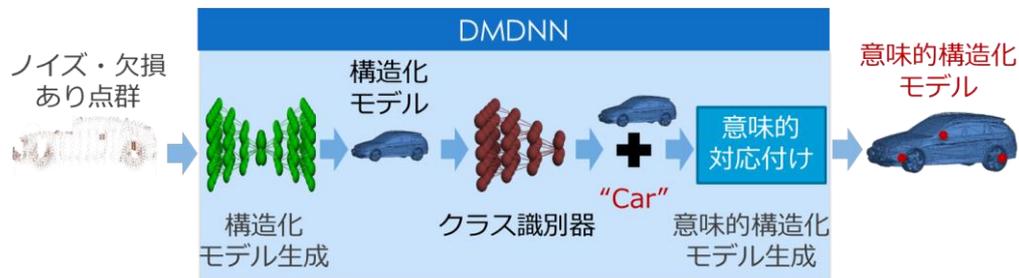


図 2. クラスが未知の点群を入力とする DMDNN

図 2 に示すように、DMDNN は、「構造化モデル生成」「クラス識別器」「意味的構造化モデル生成」の 3 つの部分からなる。ここで、構造化モデルとは、異なる物体間で共通のメッシュ構造を有するメッシュモデルを指す。ACT-X 研究期間中には、特徴の異なる 2 種類の構造化モデル生成法の構築を行い、その研究成果を発表した (5. (1)-1, 5. (3)-(ア))。また、構造化モデルに対して有効なクラス識別器の検討を行い、その結果を発表した (5. (3)-(ウ))。さらに、意味的構造化モデル生成法の構築に着手し、現在も研究を進めている。構造化モデルに対するクラス識別器の研究成果については、大阪大学歯学部との共同研究に繋がり、3 次元歯列咬合評価支援システムの構築およびその研究成果の発表を行った (5. (3)-(イ)および(エ))。

(2) 詳細

研究テーマ1「構造化モデル生成法の構築」

構造化モデルとは、異なる物体間で共通のメッシュ構造を有するメッシュモデルを指す。これまでに構築した 2 つの構造化モデル生成法では、球形の基準メッシュを対象物体の入力点群にあわせて変形することで、構造化モデルを生成した。各手法の概要を以下に示す。

手法 1(5-(1)-1): Multilayer Perceptron(MLP)を基準メッシュから入力点群への写像関数と見なし、入力点群ごとに MLP を最適化することで、基準メッシュの変形を行っている。また、基準メッシュと対象物体の形状が大きく異なる場合でも安定して構造化モデルを生成するため、基準メッシュの大局的な変形を行った後、基準メッシュと入力点群を複数の局所領域に分割し、局所領域ごとに変形している。これにより、様々な対象物体の構造化モデルを高精度に生成できる。しかし、局所領域ごとの変形では、局所領域の数だけ MLP の最適化が必要であり、計算コストが非常に高い(入力点群が 2500 点の場合、計算時間 3060 秒) (図 3)。



図 3. 手法 1 による構造化モデル生成結果

手法 2(5-(3)-(ア)): クラスごとに、それに属する物体間で共通する形状特徴を有するテンプレート形状を、陰関数表現を用いて学習する。次に、学習した全てのクラスのテンプレート形状を、事前に用意した基準メッシュモデルのメッシュ構造で記述する。推論時には、入力物体が属するクラスのテンプレート形状に対する構造化モデルから入力物体への逆写像を、学習済みのネットワークを用いて行うことで、手法 1 よりも 143 倍高速に入力物体に対する構造化モデルを生成できる（入力点群が 2500 点の場合、計算時間 21.4 秒）（図 4）。しかし、入力された点群のクラスが既知である必要があり、学習済みのクラス以外に属する物体の形状復元精度は保証されない。

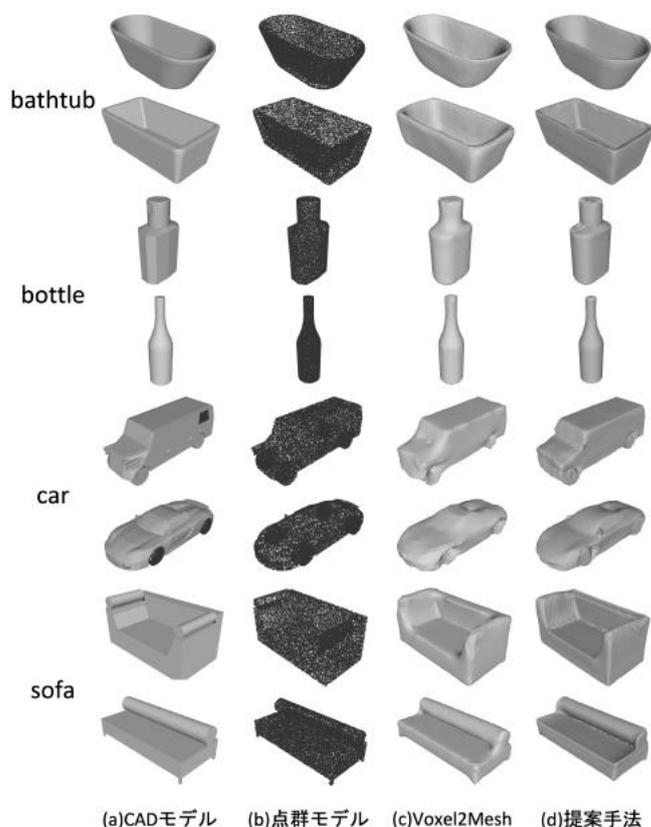


図 4. 手法 2 による構造化モデル生成結果

以上のように、各手法にはそれぞれ異なる利点と問題点がある。そこで今後は、これらの構造化モデル生成法を組み合わせることで、それぞれの手法の問題点を解決し、クラスが未知の点群に対しても高速かつ高精度な構造化モデル生成が可能な、実用性の高い構造化モデル生成法の構築を行う予定である。

研究テーマ2「クラス識別器の構築」

メッシュモデルに対して有効なニューラルネットワークとして、Graph Convolutional Network (GCN)がある。構造化モデルでは、全ての物体のメッシュ構造が同じであるため、GCNで効率的にメッシュに対する畳み込みやプーリングを行える。このような構造化モデ

ルに対して有効なネットワーク構造を明らかにするため、3種類のグラフ畳み込み手法と2種類のプーリング手法を組み合わせることで、6種類のネットワークを構築し、分類精度を比較した。その結果、グラフ畳み込み手法としてChebyconv(Kipfら, 2017)を、プーリング手法としてGraclus(Dhillonら, 2017)を用いたネットワークの物体識別精度が最も高く、畳み込む範囲が広い場合に、識別精度が向上することを確認した(表1, 5.(3)-(ウ))。

表 1. 10 クラスの構造化モデルを用いた3次元物体分類精度

ネットワーク	平均識別精度
Chebyconv+Graclus	95.1 (±1.55)
Graphconv+Graclus	94.0 (±1.61)
Spatialconv+Graclus	93.9 (±1.44)
Chebyconv+Diffpool	77.6 (±3.62)
Graphconv+Diffpool	80.2 (±4.46)
Spatialconv+Diffpool	80.5 (±3.34)

また、クラス識別を行う際には、Topological Data Analysis(TDA)の枠組みを導入することを検討し、ACT-X 一期生である池 祐一先生と共に、検証実験を行った。識別精度を向上させることはできなかったが、TDAの考え方を習得でき、今後の研究の幅を広げることができた。

さらに、構造化モデルに対するクラス識別器の研究成果については、大阪大学歯学部との共同研究に繋がり、3次元歯列咬合評価支援システムの構築およびその研究成果の発表を行った(5.(3)-(イ)および(エ))。現在も共同研究を続けており、さらなる研究成果が見込まれる。

3. 今後の展開

2023年度中に、実用性の高い構造化モデル生成法の構築およびクラスが既知の場合の意味的構造化モデル生成法の構築を行う。2024年度中に、DMDNNの構築と精度検証を行う。その後、完成したDMDNNを公開することで、3次元形状を研究対象とする研究者に広く使用してもらうことを目指す。

4. 自己評価

研究目的の達成状況としては、本来予定していた最終目標までは期間中に達成できなかったものの、期間中に新たに判明した問題点を解決する手法の構築を追加で行い、その研究成果を期間中にまとめて発表することができた。期間後の研究の方向性も確立できたことから、達成状況に対する総合的な自己評価は、5段階中の4相当であると考えられる。

研究の進め方としては、ライフイベントによる中断により、研究期間や研究費執行時期等の変更が発生したものの、期間全体として大きな問題は発生しなかった。そのため、自己評価は5である。

研究成果の科学技術及び社会・経済への波及効果については、現在継続して手法を構築中のため、現時点で見極めることは難しい。しかし、ACT-Xを通して、異分野の研究者と議論を行ったり、異分野の研究会に参加したりと、異分野間で積極的に交流する機会が増え、今までにない知見や人脈を獲得できた。これにより、今後の研究の方向性に幅ができた実感している。将来的に科学技術及び社会・経済への波及効果

の高い研究を行える可能性が高まったという点で、自己評価は4である。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数:1件

1. Shoko Miyauchi, Ken'ichi Morooka, Ryo Kurazume, "Isomorphic mesh generation from point clouds with multilayer perceptrons", arXiv preprint arXiv:2210.14157, pp.1-18, 2022.10

ノイズや欠損を含む点群を、同一のメッシュ構造で記述する手法を提案した。シミュレーションデータと、センサを用いて計測した実データに対して提案手法を適用した結果、どちらのデータに対しても、従来手法よりも高い精度で形状復元が可能であることを確認した。

(2) 特許出願

該当なし

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

(ア) 板谷響, 宮内翔子, 諸岡健一, "陰関数表現を用いた同一構造を持つ3次元物体メッシュモデル生成法の構築", コンピュータビジョンとイメージメディア研究会, pp.1-8, 2022.11

(イ) Towako Tanimura, Chihiro Tanikawa, Shoko Miyauchi, Ken'ichi Morooka, Hajime Nagahara, Takashi Yamashiro, "Artificial intelligence system that estimates the GOSLON Yardstick index in patients with unilateral cleft lip and palate", Proc. 14th International CLEFT Congress, 2022.7

(ウ) 板谷響, 宮内翔子, 諸岡健一, "同一構造を持つ3次元物体メッシュモデルの識別に適した Graph Convolutional Network 構造の検討", 画像の認識・理解シンポジウム MIRU2021, 2021.7

(エ) 宮内翔子・渡邊匠吾・板谷 響・谷川千尋・谷村百和子・山城 隆・長原 一・諸岡健一, "Graph Convolutional Network による口唇口蓋裂患者の咬合評価", 電子情報通信学会医用画像研究会, 2021.3