

階層的潜在変数を用いた Mesh VAE による臓器形状アトラスの構築

梅原 隆一[†] 中村 光宏^{††} 中尾 恵^{††}

[†] 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町
^{††} 京都大学大学院医学研究科 〒606-8507 京都市左京区聖護院川原町 53
E-mail: †{u-ryuichi,megumi}@i.kyoto-u.ac.jp

あらまし 生体が有する臓器や骨格の形状と位置を少ないパラメータで表現するモデルは幅広い臨床応用が期待される。しかし、軟臓器の形状と位置は患者間で大きく異なるため、線形モデルでは局所的に変化が大きい形状を再構成することは難しく、従来の非線形モデルでは得られた臓器形状の制御や解釈が難しい。そこで本研究では Mesh Variational AutoEncoder に階層的潜在変数を導入した臓器形状アトラスを提案する。提案手法は潜在変数を階層化することによって、患者間の形状差の潜在変数を階層別に分析することを可能とし、非線形な臓器形状表現の解釈性を向上させることが可能である。124 例からなる肝臓の臓器メッシュデータを対象に構築したモデルの検証を行い、内 19 症例のテストデータに対して、平均 1.4mm の頂点間距離、0.8mm の平均距離で位置と形状の再構成が可能であることを確認した。また、異なる階層の潜在変数を変更することで階層ごとに表現される形状が異なることを確認したので報告する。

キーワード 変分オートエンコーダー, メッシュ, 臓器形状アトラス, グラフ畳み込みネットワーク

Construction of organ shape atlas by Mesh VAE using hierarchical latent variables

Ryuichi UMEHARA[†], Mitsuhiro NAKAMURA^{††}, and Megumi NAKAO^{††}

[†] Graduate School of informatics, Kyoto University
Yoshida Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

^{††} Graduate School of Medicine, Kyoto University
53 Kawahara-chou, Shougoin, Sakyo-ku, Kyoto, 606-8507, Japan
E-mail: †{u-ryuichi,megumi}@i.kyoto-u.ac.jp

Abstract Models that represent the shape and position of the organs and skeleton possessed by a living body with few parameters are expected to have a wide range of clinical applications. However, since the shape and position of soft organs vary greatly from patient to patient, it is difficult for linear models to reconstruct shapes with large local variations, and it is difficult to control and interpret the obtained organ shapes with conventional nonlinear models. Therefore, we propose an organ shape atlas that introduces hierarchical latent variables into Mesh Variational AutoEncoder. The proposed method can improve the interpretability of nonlinear organ shape representation by introducing hierarchical latent variables. We confirmed that position and shape reconstruction was possible with an average vertex-to-vertex distance of 1.4 mm and a mean distance of 0.8 mm for the data. We also report that we confirmed that the shapes represented in each hierarchy were different by morphing.

Key words Variational AutoEncoder, mesh, organ shape atlas, Graph Convolutional Network

1. はじめに

臓器形状アトラスは生体内の臓器の位置や形状などの幾何学的情報とその個人間の変動を少ないパラメータによって表現するモデルとして知られ、医用画像解析や生体医工学分野に

において幅広く利用されている。臓器形状アトラスは統計的形状モデリング (Statistical Shape Model, SSM) [1] の考え方に基いて構築され、主に画像ベースの方法 [2] とメッシュベースの方法 [3] に大別される。画像ベースの方法は医用画像を直接利用して SSM を構築でき、また画像データは格子状のデー

タ構造であるため、特徴量の抽出やフィルタリングなどの計算機上での処理が容易であるが、データ量が画像サイズの3乗に比例して大きくなるため、高解像度なモデルの構築が難しく、複雑な臓器形状や臓器間の非線形かつ不連続な変形を表現することが難しい。一方、メッシュベースの方法では、データの作成に様々な前処理が必要である。またメッシュはグラフデータ構造であるため、統一的な特徴量の計算や変換などの計算機上での処理が画像に比べて難しいが、非常に小さなデータ量で形状を表現でき、変形も扱いやすいという特徴を有する。本研究では後者のメッシュベースのモデルを採用し、その形状表現と汎化性能の向上に着目している。メッシュベースのSSMでは、頂点对応が取られた臓器形状メッシュ群から主成分分析 (Principal Component Analysis, PCA) や独立成分分析 (Independent Component Analysis, ICA) に基づいて特徴量を抽出し、平均形状との荷重和を取ることによって任意の臓器形状を表現する。臓器の形状データベースは多くがデータ数が数百例程度でありかつ、メッシュデータは3次元空間であれば頂点数の3倍の自由度を有するので、高次元少標本データと言える。また、臓器の形状と位置は患者間で非常に多様な形状と位置をとることから、PCAやICAに代表される線形モデルでは局所的に大きな変形を表現することが難しいという課題が存在する。この課題を解決するために近年では非線形モデルを使用した研究[3]も行われているが、非線形モデルは限られたデータ数では過学習に陥りやすい。我々は非線形モデルの一つである変分オートエンコーダー (Variational Autoencoder, VAE) とグラフ畳み込みネットワーク (Graph Convolutional Network, GCN) によるメッシュ形状表現[4]を試みたが、GCNでは入力メッシュと同じ頂点数の潜在変数が必要となり、潜在変数の分析及び生成される臓器形状の制御は困難であった。近年画像分野におけるVAEでは潜在変数を階層化することで大局的な情報と局所的な情報を分離して学習し、VAEモデルの精度を向上させると共に、モデルの解釈性も向上させる研究事例も報告されている[5]。

本研究では階層的潜在変数を用いたメッシュ変分オートエンコーダ (Mesh VAE) による臓器形状アトラスを提案する。提案手法では、生体臓器の複雑な形状と個人間の非線形な形状差を表現するために Mesh VAE [4] による形状表現に新たに二つの技術的要素を加える。一つ目は潜在変数の階層化である。潜在変数が階層化されることで臓器形状の再構成性能を維持したまま、全体としてより少ない次元数での形状再構成を見込める。二つ目にメッシュ再分割による異なる解像度間の頂点对応を有した臓器形状テンプレートの構築である。メッシュ再分割を用いることでメッシュデータにおける階層的な表現を可能とし、臓器形状の大局的特徴と局所的特徴を制御可能にしている。これら二つの要素をモデルに組み込むことで高い再構成性能と解釈性を両立する臓器形状アトラスの実現を目指す。実験ではこれらの技術要素の有効性を検証するために、124名の患者の3次元CT画像から生成した肝臓形状のメッシュ群を用いてモデルの学習を行い、要素の有無による再構成性能の差を比較する。また、異なる二つの臓器メッシュ

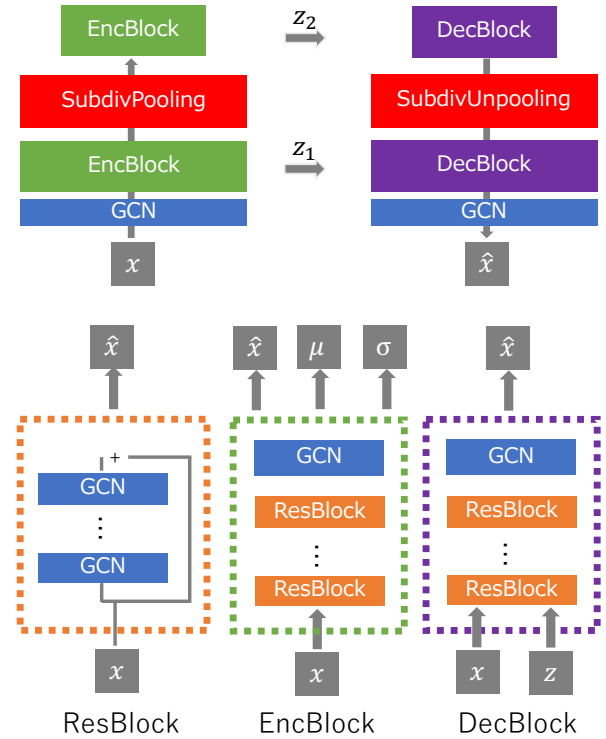


図1 提案モデル。モデルにはGCN層を重ねて作成された3種類のBlockから構成され、Pooling/Unpooling層を用いることで潜在変数を階層化している。

から抽出された2階層の潜在変数の線形補間によって得られる新たな潜在変数から臓器形状を生成し、階層ごとに実現される臓器形状間の補間の違いを比較する。

2. 提案方法

2.1 提案モデルの概要

図1に設計した階層的メッシュ変分オートエンコーダーのモデルを示す。本モデルはEncoder Block (EncBlock) と Subdivision Neighbor Pooling (SubdivPooling) からなるエンコーダーと Decoder Block (DecBlock) と Subdivision Neighbor Unpooling (SubdivUnpooling) からなるデコーダーで構成される。EncBlockとDecBlockでは、GCNで構成した残差ブロック (Residual connection block, ResBlock) を利用している。本モデルにおける重みの学習はモデルの入出力から定義される損失関数の最小化によって達成される。

なお、本研究において再構成対象とする臓器メッシュは頂点数と辺の数が同じで、かつ、頂点对応が取れたメッシュ群を想定している。患者固有の3次元CT画像から得られるメッシュはそのままでは患者間で頂点や辺が統一されていない。そこで、前処理として平均的な形状を有するメッシュ M_f を初期テンプレートとし、それを患者固有の臓器メッシュに対して事前にメッシュ変形位置合わせ[3]を行う。この位置合わせ処理によって、患者固有の形状を持ちつつ、 M_f と同一の構造を持つ頂点对応が取られたメッシュが得られる。

2.2 階層的 Mesh VAE

本モデルでは潜在変数を階層化し、階層ごとに学習される形状特徴を分割することによってモデルの再構成性能の向上

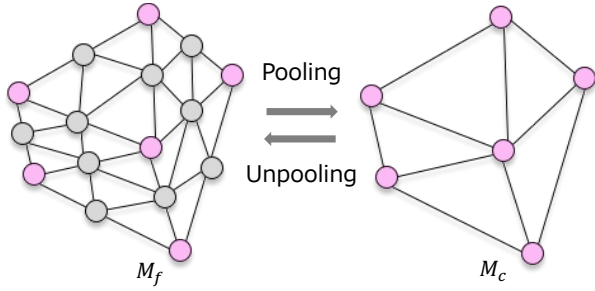


図2 Subdivision Neighbor Pooling と Subdivision Neighbor Unpooling の概念図。

よび潜在変数の分析を容易にする。そのためには入出力メッシュの頂点群と同次元の高解像度潜在変数とは別により低次元の低解像度潜在変数を用意する必要がある。加えて、それらは双方向に一意に変換される必要がある。

これら二つの要件を満たすために、本研究では SubdivPooling / Unpooling を提案する。Pooling / Unpooling は Convolution Neural Network (CNN) において、データの次元を削減し、大局的な情報を集約する演算子として広く利用されている。Pooling には隣接する値に対して様々な集約関数を用いられるが、メッシュ頂点によって臓器形状は表現されるため、Pooling の前後でなるべく形状が一致するような集約関数として隣接頂点の平均を用いた Pooling / Unpooling とする。具体的には、図2に示すように、SubdivPooling では出力される頂点(ピンク)が持つ情報は入力において自身と自身に隣接する頂点(グレー)の平均とし、SubdivUnpooling ではメッシュの辺の中点に新しい頂点を挿入する。

SubdivPooling では高解像度のメッシュにおける頂点のサンプリング方法を定める必要がある。本研究では、SubdivPooling 前後における頂点の対応関係を Loop 再分割法 [6] に基づいて定義する。まず、与えられたテンプレートメッシュに対して簡約化を行い、低解像度テンプレートメッシュ M_c を得る。 M_c に対して Loop 再分割を適用し、三角形要素数が M_c の4倍になった高解像度テンプレートメッシュ M_f を生成する。この M_f と M_c の頂点対応からサンプリング対象の頂点が一意に定めることが可能である。

2.3 損失関数

本研究では汎化性能を達成しながら、高精度の再構成を行うために、潜在変数の分布に対する損失と入出力の再構成誤差を組み合わせた式1の損失関数を導入する。

$$\mathcal{L} = \mathcal{L}_{KL} + \alpha \mathcal{L}_P \quad (1)$$

ここで、 \mathcal{L}_{KL} は正規分布と潜在変数のカルバック・ライブラー情報量 (Kullback Leibler divergence, KLD) であり、 \mathcal{L}_P は入力頂点群と出力頂点群の対応する頂点座標から得られる Mean Squared Error (MSE) である。

3. 実験

提案モデルの実装には Python3.9, Pytorch, Pytorch-Geometric を用いた。学習におけるバッチサイズを32, トレーニング epoch は最大を1000として、検証用データの \mathcal{L}_P を監

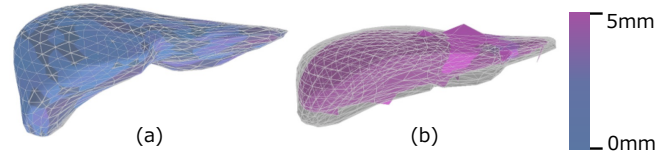


図3 提案手法によるテストデータ2例の再構成結果。(a) 平均的な再構成精度の症例, (b) 最も精度が悪かった例。メッシュ: 目的形状, サーフェス: 再構成結果。

視する EarlyStopping を用いて学習を行った。EarlyStopping の patience は 50epoch に設定した。またネットワークの最適化には学習率を 1×10^{-2} に設定した Adam を用いた。損失関数のハイパーパラメーターは幾つかの組み合わせについてモデルの性能を確認し、最も性能が高かった $\alpha = 1.0 \times 10^{12}$ を用いることとした。

3.1 データセット及び前処理

京都大学医学部附属病院放射線治療科において、膵がんの放射線治療を受けた124名の患者の腹部3D-CTデータから得られた肝臓の臓器メッシュを対象に実験を行った。放射線腫瘍専門医によって定義された輪郭情報から各臓器領域のサーフェスメッシュを得た後、テンプレートベースのメッシュ変形位置合わせによって位置合わせ済みの肝臓メッシュを得た。これら前処理は過去の研究 [3] において実施された手続きと同じである。なお、一連の実験は京都大学における医の倫理委員会によって承認された(承認番号: R1446)。メッシュ頂点は放射線治療における治療用ビームの照射対象である膵癌の位置を原点として、左右: x軸, 背腹: y軸, 頭尾: z軸のそれぞれにおいて $[-256, 256]$ (単位: mm) の範囲に存在している。モデルの入出力として頂点の座標の範囲が $[-1, 1]$ になるように正規化した。124名のメッシュに対して、86を学習用に、19を検証用に、19をテスト用に分割して実験を行った。

3.2 評価指標

結果の定量評価にはメッシュの頂点間の誤差を定量するために式(2)に基づいて Mean Absolute Error (MAE) を使い、表面間距離を定量するために Mean Distance (MD) [7] を用いた。

$$E_{MAE} = \frac{1}{NM} \sum_i^M \sum_j^N \|v_i^{(j)} - v_i^{\prime(j)}\|_2 \quad (2)$$

ここで N はメッシュの頂点数を表し、 M は評価対象の症例数を表す。

3.3 メッシュ再構成性能の比較

最初に表1に示す複数の従来モデルによるメッシュ再構成結果と提案方法のメッシュ再構成性能を確認する実験を行った。FCは全結合層のみによるモデル、GCNは図1内の ResBlockのみを用いて構築されたモデル、Poolingは ResBlock と SubdivPooling を組み合わせ、潜在変数の次元を圧縮したモデルである。表1から、提案手法が最も精度良くメッシュを再構成できており、GCNとPoolingの行の比較から SubdivPooling の有効性が、Poolingと提案手法の比較から潜在変数の階層化の有効性が確認できる。提案手法における再構成結果を2例可視化した図を図3に示す。提案手法は肝臓の臓器形状を精度よく再構成でき、最も精度が悪かった例でも大まかな形状

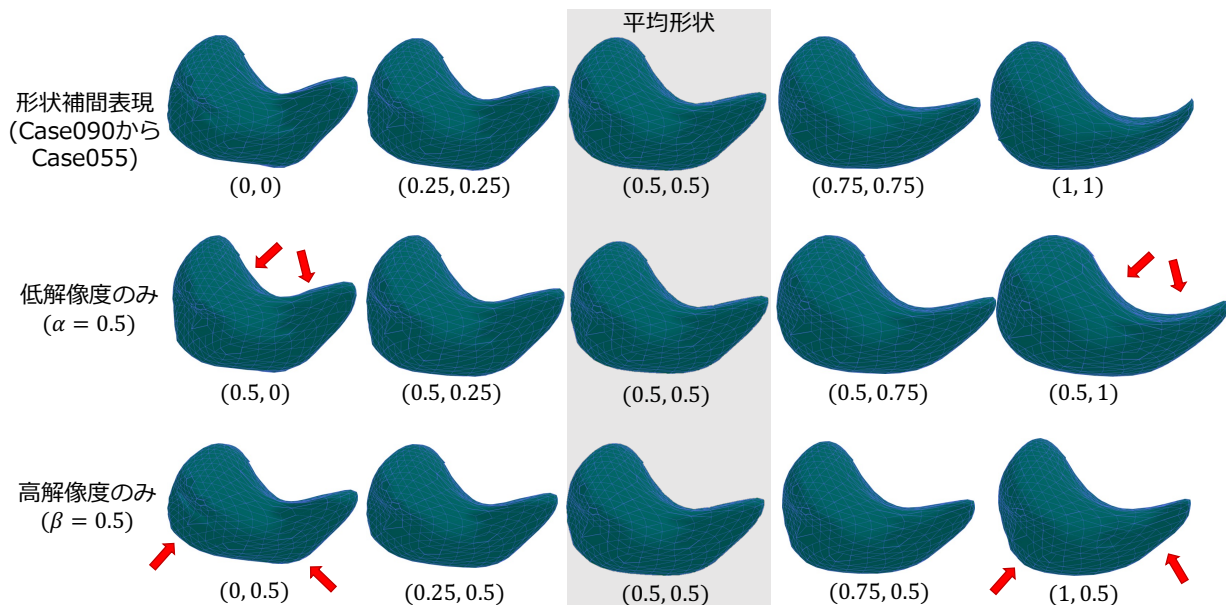


図4 階層的潜在変数の変更によるテストデータ2例の形状補間表現. 上段: (α, β) の同時変更による Case090 から Case055 への形状補間. 中段: $\alpha = 0.5$ で固定し, β のみを変更した場合. 下段: $\beta = 0.5$ で固定し, α のみを変更した場合.

Model	MAE [mm]			MD [mm]		
	Train	Val	Test	Train	Val	Test
FC	7.697	14.52	13.60	3.774	7.190	6.386
GCN	10.39	10.18	10.11	5.708	5.599	5.330
Pooling	2.278	2.505	2.495	1.112	1.282	1.253
Proposed	1.331	1.417	1.472	0.7214	0.7780	0.8070

表1 各モデルにおける胃メッシュの再構成性能の比較.

は生成可能であることを確認した.

3.4 階層的潜在変数による形状補間表現

テストデータから形状が大きく異なる2症例として Case055 と Case090 を選択し, 3パターンの形状補間表現を試みた. 生成された形状を図4に示す. 上段では式3において, α, β を同時に0から1へ連続的に変化させた結果を示しており, 潜在変数の変更によって2例間の中間形状を様々に生成できることを確認した.

$$\begin{cases} z_t^1 = (1 - \alpha)z_{090}^1 + \alpha z_{055}^1 \\ z_t^2 = (1 - \beta)z_{090}^2 + \beta z_{055}^2 \end{cases} \quad (3)$$

図4中段と下段にそれぞれ $\alpha = 0.5, \beta = 0.5$ に固定して他方を0から1へ連続的に変化させた結果を示す. 中段では赤矢印で示すように全体に大きく形状が変わっている一方で, 下段では主に一部の形状の曲率が変化するなどの局所的な変化が生じていることが確認できる. 以上から低解像度の潜在変数では形状の大局的な情報が, 高解像度の潜在変数では形状の局所的な情報が学習されたことが示唆される.

4. おわりに

本研究では Mesh VAE に階層的潜在変数を導入することによって, 患者間の形状差の潜在変数を階層別に分析すること

を可能とし, 非線形な臓器形状表現の解釈性を向上させた臓器形状アトラスを構築した. 124例からなる肝臓の臓器メッシュデータを対象に構築したモデルの検証を行い, 内19症例のテストデータに対して, 平均1.4mmの頂点間距離, 0.8mmの平均距離で位置と形状の再構成が可能であることを確認した. また, 異なる階層の潜在変数を変更することで階層ごとに表現される形状が異なることを確認した. 今後はさらに階層を増やしたモデルを他臓器にも適用し, 潜在変数の分析を行う予定である.

謝辞

本研究は日本学術振興会 科研費補助金 基盤研究 (B) 22H03021 及び 19H04484 の支援による.

文献

- [1] T. Heimann, H. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543 - 563, 2009.
- [2] S. Oh, S. Kim, "Deformable image registration in radiation therapy," *Radiation Oncology Journal*, vol. 35, no. 2, pp. 101 - 111, 2017.
- [3] M. Nakao, M. Nakamura, T. Mizowaki, T. Matsuda, "Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization," *Medical Image Analysis*, vol. 67, p. 101829, 2021.
- [4] R. Umehara, M. Nakamura, M. Nakao, "Construction of an organ shape and position atlas using 3D Mesh Variational Autoencoder," *IEICE Technical Report (MI)*, vol. 122, no. 417, pp. 205-209, 2023.
- [5] R. Child, "Very deep vaes generalize autoregressive models and can outperform them on images," *arXiv preprint*, art no. 2011.10650, 2020.
- [6] L. Charles, "Smooth subdivision surfaces based on triangles," M.S. Mathematics thesis, University of Utah, 1987.
- [7] B. Rigaud et al., "Statistical Shape Model to Generate a Planning Library for Cervical Adaptive Radiotherapy," *IEEE Trans. on Medical Imaging*, vol. 38, no. 2, pp. 406-416, 2019.