

形状マップによる三次元形状データカテゴリ分類のための形状特徴抽出

金井 廉[†] 藤田 悟[‡]

法政大学大学院 情報科学研究科^{†‡}

1 まえがき

近年, Kinect や 3D スキャナ等のデバイスの普及等により, 大量の三次元形状データ (以下形状データ) が蓄積されるようになり, それらを整理・検索するための高精度な分類手法への需要が高まりつつある. 本研究では, 形状データ分類のための特徴量として形状マップと呼ばれる二次元画像を提案する. 形状マップとは, 形状データの凹凸情報を含む画像であり, 形状データの重心から線分を放射し, 重心からサーフェスまでの距離を測定することで生成される. さらに別の手法として, 従来より見かけの特徴量として形状データ分類に用いられた深度マップ [1] を, 三次元空間中の形状データを複数視点から撮影することで複数取得し, それらを複数チャンネルの画像とする方法も提案する. 本研究ではこれら二つの特徴量を入力可能な分類器として, 画像認識に特化したニューラルネットワークである Covolutional Neural Network (CNN) [2] を用いる. これらの手法により, 高精度な形状データの分類を可能にする. 提案手法である特徴量を用いた形状データの分類実験を三次元形状データセットである SHREC[3] を使って行い, 高い精度で分類を行うことができた.

2 提案手法

2.1 形状マップ

形状マップは, 形状データの形状情報を画素値として持つ二次元画像である. 図 1 に, 形状マップの生成方法を示す. 形状マップの生成工程は次の四つからなる. はじめに, グリッドと呼ばれる多角形を三次元空間上に定義する. グリッドの面は再帰的に分割されており, 多数の等間隔な頂点を面上に含む. また, グリッドは形状データと重心が一致するように配置され, 形状データの全ての頂点とサーフェスがグリッドのサーフェスよりも内側に来るようにスケールされる. 次に, 形状データの重心からグリッド上の各頂点へ線分を定義する. その後, 各線分と形状データのサーフェスの交点を求め, 各

交点と重心間の距離を求める. 最後に, グリッドの各面の頂点に対応する解像度を持つマップを生成し, 各画素値に対応する線分の距離を格納する. このマップの形状は, グリッドの面の形状と等しい. 形状データの重心点が形状データの内部に存在すれば, 形状マップのすべての画素は距離値を持つ. 一方, 形状マップの距離値は重心から放射状に平面へマップされるため, マップの中心に近いほど凹凸の分解能は小さくなり, 遠いほど大きくなる. この減少はグリッドの面数が少ないほど顕著になる. 本研究では, 立方体状のグリッドを用いて形状マップの生成を行った. 立方体状のグリッドを使うことで, 生成される形状マップは正方形となり, 二次元画像として CNN に入力可能な特徴量となる.

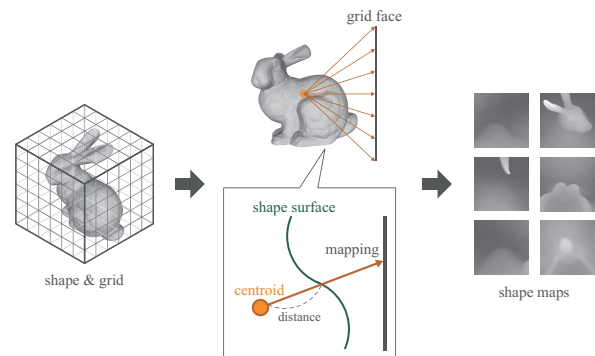


図 1 形状マップの生成

2.2 深度マップ

深度マップは, 各画素に深度値が埋め込まれた画像である. 三次元空間上のオブジェクトをレンダリングしてグラフィックボード上の深度バッファを参照したり, 深度カメラで現実世界を撮影することで生成できる. 形状マップと異なり, 三次元空間中の形状データのある視点からレンダリングして深度マップを生成した場合, 深度マップ内には形状データへの深度を持つ領域とそれ以外の領域ができるため, 深度マップは任意の一視点から見た形状データのエッジ情報を保持することができる. Chen ら [1] は三次元空間上の形状データを複数の視点から撮影し, 各視点からの深度マップを生成している.

ShapeMap: The Shape Feature Extraction for 3D Object Classification

[†]Ren KANAI [‡]Satoru FUJITA

^{†‡} Graduate School of Computer and Information Sciences

本研究においても、上記の手法と同様に三次元空間中の形状データを複数視点からレンダリングし、一つの形状データから複数の深度マップを取得する。しかし、それらの深度マップを複数チャンネルの画像としてCNNに入力し、カテゴリ分類を行うという点で従来手法と異なる。図2に、本研究における深度マップの生成の様子を示す。今回は形状マップとの分類精度の比較を行うため、モデルの上下左右前後の6方向から深度マップの生成を行った。

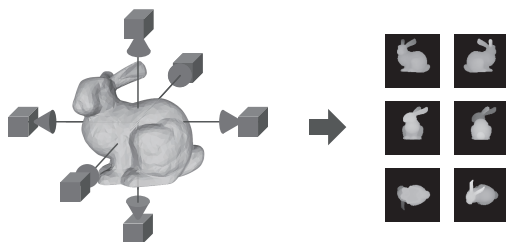


図2 三次元空間上における形状データからの深度マップの生成

3 実験

提案手法である特徴量に対する分類精度を検証するため、三次元形状データセット SHREC[3]を用いた実験を行った。SHRECは1200個の形状データからなり、クラス数は50、1クラスあたり24個のデータが存在する。形状マップは 128×128 の頂点を持つ面からなる立方体グリッドによって一つの形状データにつき6枚生成され、6チャンネルの画像としてCNNの入力となる。深度マップも同様に解像度を 128×128 とし、一つの形状データにつき6視点からの深度マップを生成して、6チャンネルの画像とする。また深度マップを生成する際、カメラ平面への投影図法を平行投影とし、形状データはカメラ平面内に収まるようにスケールした。各マップの分類器の構成にはCNNであるAlexNet[2]を用い、バッチサイズ10のミニバッチ学習によって最適化を行った。但し、AlexNetの畳み込み層におけるストライドを1とし、全結合層部分のユニット数の構成を1024-512-50としている。図3に、6-fold交差検証による、形状マップと深度マップに対するテストデータ誤分類率の推移をプロットしたグラフを示す。また、解像度 128×128 の立方体形状マップを生成するのに要した時間は、頂点数34835の形状データに対し約4.6sであった。なお、形状マップ生成時にはCPUにIntel Core i7 3.60GHz、GPUにGForce GTX780を用いており、放射線とサーフェスの交差判定には並列化したTomas-Moller[4]の交差判定アルゴリズムを用いている。

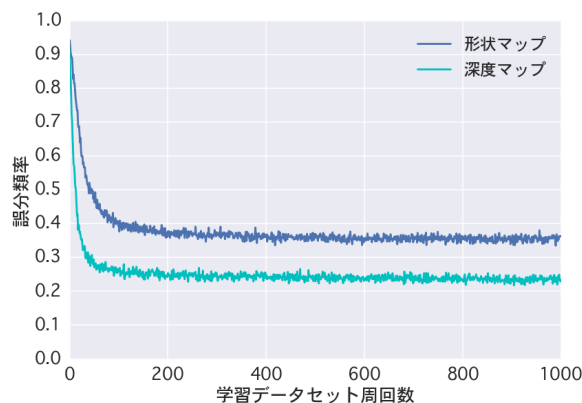


図3 6-fold交差検証によるテストデータ誤分類率の推移

4 まとめ

形状マップと深度マップによるSHRECの誤分類率の推移を比較すると、深度マップによる分類がより低い値を示した。このことから、すべての画素に重心からサーフェスまでの距離値を持つ形状マップよりも、深度情報に加えエッジ情報を持つ深度マップの方が、CNNを用いた分類に有用な特徴を多く含むと考えられる。一方で、形状マップの生成時間の結果から、大規模なデータセットに対して形状マップを生成する場合、多くの処理時間を要してしまうと考えられる。今後の課題として、分類精度をより高めるための分類器のパラメータ選定や、異なる分類器による分類の検証が挙げられるほか、より短時間で生成可能かつ十分な精度で分類可能な二次元特徴量の提案が望まれる。

参考文献

- [1] Ding-Yun Chen et al. On Visual Similarity Based 3D Model Retrieval. *Computer Graphics Forum*.
- [2] Alex Krizhevsky et al. Imagenet classification with deep convolutional neural networks. In P. Bartlett et al., editors, *Advances in Neural Information Processing Systems 25*, pages 1106–1114. 2012.
- [3] Z. Lian et al. Non-rigid 3D Shape Retrieval. In I. Pratikakis, M. Spagnuolo, T. Theoharis, L. Van Gool, and R. Veltkamp, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2015.
- [4] Tomas Moller. Practical analysis of optimized ray-triangle intersection. http://fileadmin.cs.lth.se/cs/Personal/Tomas_Akenine-Moller/raytri/.