

機械学習を用いたリバースエンジニアリングのための 形状測定点群データのセグメンテーション

松田優也¹ 小高知宏²

概要: 製造業の現場での 3D スキャナや X 線 CT 等の普及に伴い、形状測定点群データを元にして 3 次元 CAD データのサーフェスを生成するリバースエンジニアリング技術の開発がこれまで行われてきた。リバースエンジニアリングでのサーフェス生成のためには CAD データの面構成に合わせた点群のセグメンテーションが必須となるが、既存手法では手動による作業が必要になる場合が多く、コスト削減のため自動化が求められている。そこで本研究では点群の座標成分、法線成分、法線成分の分散を特徴量とし、変分混合ガウス分布と DBSCAN によるクラスタリングを併せて適用することで、点群のセグメンテーションを自動化する手法を考案したので報告する。

キーワード: 形状測定点群, セグメンテーション, 機械学習, クラスタリング

Segmentation of point clouds for Reverse Engineering based on machine learning

YUYA MATSUDA^{†1} TOMOHIRO ODAKA^{†2}

Abstract: Along with widespread use of 3D scanners and X ray CT scanners on manufacturing scenes, reverse engineering technique creating surfaces for 3D CAD data has been developed so far. For the purpose of creating surfaces on reverse engineering, segmentation of point clouds which suits surface configuration of 3D CAD data is needed, however, it costs a lot with conventional manual operations, and hence, development of automatic operations is required to cut costs. In this manuscript, our new method is presented, which automate segmentation of point clouds with a combination of clustering based on Variational Bayesian Gaussian Mixture model and DBSCAN, where, points, normals, and variances of normals are treated as feature vectors.

Keywords: point cloud, segmentation, machine learning, clustering

1. はじめに

近年の製造業の現場での 3D スキャナや X 線 CT 等の普及に伴い、形状測定点群データを元にして 3 次元 CAD データのサーフェスを生成するリバースエンジニアリング技術の開発がこれまで行われてきた。リバースエンジニアリングにより 3 次元 CAD データを作成することで、機械加工により同一形状のパーツが複製できるだけでなく、部分的に CAD データを編集することにより類似形状のパーツ設計に利用できる。また、構造解析や流体解析などの目的で CAE を利用する際には解析メッシュの生成のため 3 次元 CAD データが必要となるが、試作品や他社製品の現物から形状測定点群データを取得し、リバースエンジニアリングにより 3 次元 CAD データを作成すれば CAE による解析が可能になる。このように、リバースエンジニアリングにより 3 次元 CAD データを作成することで得られるメリットは大きい。

リバースエンジニアリングによるサーフェス生成は一般的に以下の手順により行われる。

- 3D スキャナ等により形状測定点群データを取得
- 形状測定点群データをポリゴン (STL 形式等) に変換
- 形状測定点群データのノイズ除去と穴埋め処理
- CAD データのサーフェス構成に合わせて点群を分割
- セグメンテーション毎にサーフェスを生成

以上のように、リバースエンジニアリングでは点群データを元にしてサーフェスを生成するため、CAD データのサーフェス構成に合わせた点群のセグメンテーションが必須となる。しかし既存手法では手動による作業が必要になる場合が多く、完全な自動処理によりセグメンテーションが完了できるケースは限られている。点群のセグメンテーションはリバースエンジニアリングのプロセスのボトルネックとなっており、リバースエンジニアリングのコスト削減のために自動化が求められている。そこで本研究では機械学習の手法を利用することで点群のセグメンテーションを自動化する手法を考案したので報告する。

2. 従来手法

実用化されているものも含めた従来手法による点群のセグメンテーションにおいては、点群の曲率分布を算出し、立体物のエッジ部に対応する特徴部分の点群の抽出を行うものが多い[1]。それらの手法では抽出したエッジ部を境界線として点群のセグメンテーションを行う。曲率は 2 次微

¹ 福井県工業技術センター
Industrial Technology Center of Fukui Prefecture
² 福井大学工学研究科
Graduate School of Engineering, University of Fukui

分量でありノイズの影響を受けやすいため、はじめに点群から測定データのノイズ除去や凹凸のスムージングを行うことが必要となる。それゆえノイズ除去やスムージングが不十分な場合など、何らかの理由で特徴抽出に失敗した場合はセグメンテーションができないという問題がある。また領域成長法を用いた方法も考案されているが、ノイズや点群データの欠損の影響により同一領域の誤抽出が発生する場合がある。実際、リバースエンジニアリングのための商用ソフトウェアとしては *Geomagic Design X*[2]が最も有名であるが、同ソフトウェアでのリバースエンジニアリングにおいても形状測定点群データの品質の影響は大きく、ノイズの除去だけでなく点群データに欠損がある場合には人工的な穴埋めが必要となる[3]。

3. 提案手法

従来手法はミクロな視点で点群をとらえ、逐次的に各々の点を処理するという特徴がある。一方で、マクロな視点で統計的に点群を扱い、機械学習のみでセグメンテーションの自動化を試みた研究はこれまでで見られない。そこで本研究では、従来手法のような点群のエッジ抽出や領域成長法を用いることなく、教師なし機械学習の一手法であるクラスタリングのみを用いることで点群のセグメンテーションを行う手法を考案した。

3.1 クラスタリング

3.1.1 教師なし機械学習

機械学習は教師あり学習と教師なし学習に大別される。両者の違いは正解となる教師データの有無であり、教師なし学習においてはデータのもつ全体的な構造から意味のある情報を抽出する手法の総称である。その中でクラスタリングは教師なし学習の代表的な手法に位置づけられる。本研究で取り扱う点群のセグメンテーションのような問題では当然のことながら教師データは存在せず、データすなわち各測定点のもつ全体的な構造からセグメンテーション結果を推論する必要がある。

3.1.2 クラスタリング手法

クラスタリングには様々な既存手法が存在する。代表的なクラスタリング手法として *k* 平均法[4]が知られているが、*k* 平均法ではあらかじめクラスタ数を指定する必要がある。しかし点群のセグメンテーションでは、元の形状が複雑であるほどクラスタ数を事前に把握することは困難である。

提案手法では変分混合ガウス分布(VBGMM: Variational Bayesian Gaussian Mixture)[5]と DBSCAN(Density-based spatial clustering of applications with noise)[6]を使用するが、これらはいずれもあらかじめクラスタ数を指定する必要があるという長所がある。

また、変分混合ガウス分布ではデータの分布としてガウ

ス分布を仮定しており、外れ値をもつデータをノイズとして省くことが容易である。一方、DBSCAN ではデータの分布として楕円形を仮定しないため、分布の形状が複雑なデータにも適用できる。

3.2 特徴量ベクトル

クラスタリングによりデータをラベル付けするためには各データが何らかのベクトルデータとなっている必要がある。点群のセグメンテーションにおいてははじめ各点が 3 成分の位置座標値を保持しており特徴量とすることができるが、以下で述べるように、さらに周囲の点群との位置関係から法線方向を推定することで特徴量とすることができる。加えて、周囲の点群の法線の分布を参照することによりその分散も特徴量とすることができる。

3.2.1 Kd 木による近傍点探索

まず前節で述べた法線等を推定するためには注目する点の周囲の点群の 3 次元座標値を取得する必要があるが、近傍点を探索するために注目点からすべての点までの距離を計算することは極めて効率が悪い。そこで Kd 木[7]のデータ構造を用いて近傍点探索を高速化する。

3.2.2 法線ベクトル

法線の推定には注目点の近傍の任意の半径内の点群を利用する。その結果、法線ベクトルは点群の外側を向いて並び、例えば立方体点群の場合には図 1 (左) に示すようになる。また法線ベクトルの成分のみに着目し 3 次元空間(法線空間と呼ぶことにする)内にプロットすれば、図 1 (右) のように半径 1 の球面上に点群が分布する。

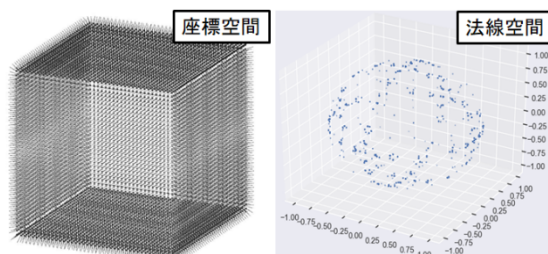


図 1 立方体点群(左)とその法線成分の法線空間での分布(右)

Figure 1 Point cloud of a cube(left) and its distribution of normal components(right).

点群の密度にもよるが、図 2 に示すように近傍点探索の半径が過小な場合はノイズの影響が大きくなり法線がばらつきやすくなるので注意が必要である。

2 平面で形成されるエッジ部分のように、位置連続で接続する曲面に対応する点群についての法線ベクトルを法線空間内にプロットすると、異なる位置にデータが集中する傾向がある。たとえば平面に対応する点群では法線空間内においては 1 点にデータが集中する傾向があるため、2 平面が交わりエッジを形成するような点群の場合は法線空間

内においてはデータが2点に集中する. その性質を利用し, 次節で述べるように法線ベクトルは位置連続で接続した曲面に対応する点群を分割するために使用する.

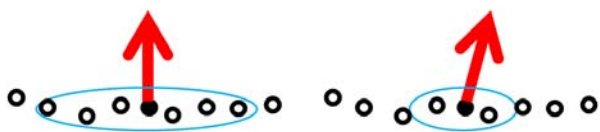


図 2 適切な近傍点半径 (左) と過小な近傍点半径 (右)

Figure 2 An appropriate neighborhood radius(left) and a too small neighborhood radius(right).

3.2.3 法線ベクトル成分の分散

注目点の近傍の点のもつ法線ベクトルの分散を計算することで特徴量とすることができる. 図 1 (右) のように法線ベクトルの成分は半径 1 の球面上に分布するが, 法線空間内において注目点の近傍の点を探索し, その分散を調べることで特徴量とする (図 3).

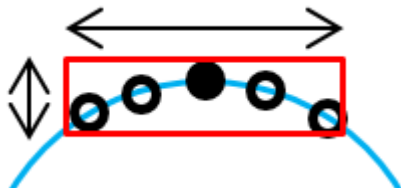


図 3 法線空間内の注目点近傍における分散

Figure 3 Variance of normal components.

法線ベクトルは 3 成分をもつので, 近傍点の法線成分について主成分分析を行うことで第一主成分から第三主成分までの 3 個の分散が得られ, それを新たな特徴量ベクトルとし 3 次元空間 (分散空間と呼ぶことにする) 内にプロットする. すると, 平面に対応する点群については分散空間内の原点付近にデータが集中する傾向がある. また, 曲率が近い曲面に対応する点群のデータは分散空間内で同じ位置に集中する傾向がある. そこで次節で述べるように, 法線ベクトルの分散は接線連続で接続した曲面に対応する点群を分割するために使用する.

3.3 位置連続曲面に対応する点群のセグメンテーション

リバースエンジニアリングで一般的に対象となるのは機械部品などであり, それらの多くは位置連続で接続する面と接線連続で接続した面から構成される. 曲率連続で接続する曲面が含まれることは稀であるため, 本研究でのセグメンテーションの対象は接線連続までとし, 意匠面のように曲率連続の曲面群で構成されるようなセグメンテーション境界が判然としない形状は取り扱わない. 本節では位

置連続で接続する曲面に対応する点群のセグメンテーション手法について述べる.

3.3.1 法線成分の VBGMM によるクラスタリング

まず点群の法線成分のみに着目しクラスタリングを行う. クラスタリングアルゴリズムには変分混合ガウス分布を用いる.

簡単な例として立方体点群 (図 1) について生成したクラスタをウェイト (要素数) 順にソートすると, 図 4 のようにノイズおよびエッジ部と, 立方体面上に対応する 6 クラスタの点群が判然とする.

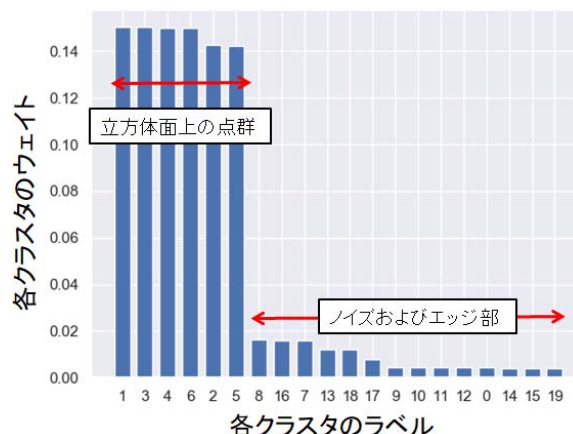


図 4 立方体点群の法線空間でのクラスタリング結果の各クラスタのウェイト

Figure 4 Weights of each clusters, the result of clustering on normal components of a point cloud of a cube.

ウェイトが数%以下のクラスタはエッジ部またはノイズとみなして除外する. これによりクラスタリングとノイズの除去を同時に実行できる.

3.3.2 法線成分の DBSCAN によるクラスタリング

変分混合ガウス分布ではクラスタ形状を球状と仮定するため, 法線成分の分布領域が大きな点群については前節での法線空間での変分混合ガウス分布によるクラスタリングの結果, 同一曲面に対応するはずの点群が別クラスタに分割されてしまうことがある (図 5).

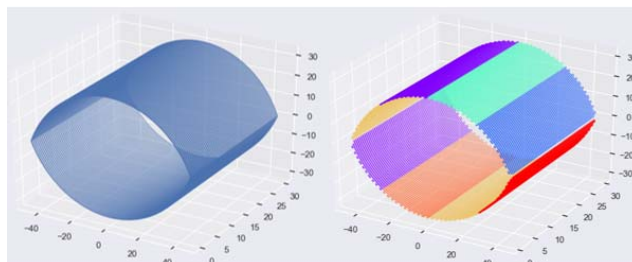


図 5 曲面に対応する点群 (左) と法線空間の変分混合ガウス分布によるクラスタリング結果 (右)

Figure 5 Point cloud of a curved surface(left) and the result of clustering on normal components by VBGMM.

そこで次に法線空間で DBSCAN を実行することにより、別クラスタに分割されてしまったクラスタを統合する (図 6)。

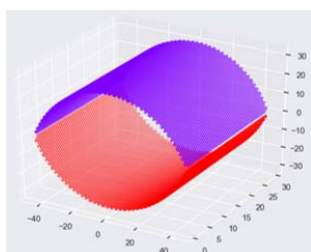


図 6 図 5 (右) で分割されてしまった点群に対する法線空間での DBSCAN によるクラスタリング結果

Figure 6 The result of clustering on normal components by DBSCAN

3.3.3 位置座標成分の DBSCAN によるクラスタリング

点群の形状が非凸包の場合は、法線空間でのクラスタリングの結果、位置座標空間では明らかに別のクラスタに属する点群が、同一クラスタに併合されてしまうことがある (図 7)。

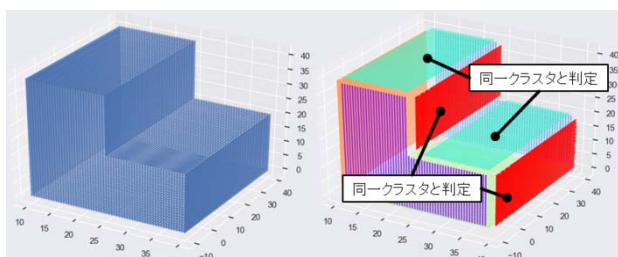


図 7 非凸包の点群 (左) と法線空間でのクラスタリング結果 (右)

Figure 7 Point cloud of a concave hull(left) and the result of clustering(right) on normal components.

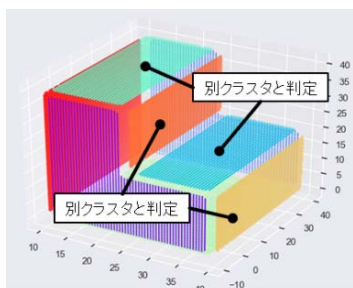


図 8 座標空間での DBSCAN によるクラスタリング結果

Figure 8 The result of clustering with DBSCAN in position coordinates space.

そこで次に、前節でエッジ部とノイズを除外した点群に

対して、位置座標空間において DBSCAN によるクラスタリングを行う。DBSCAN はクラスタ数の指定が不要であるのに加えてクラスタ形状が球状でなくとも適用可能である。これにより図 7 で法線空間では併合されてしまった点群が正しく別クラスタに分割される (図 8)。

3.4 接線連続曲面に対応する点群のセグメンテーション

前節までの処理により位置連続で接続する曲面に対応する点群のセグメンテーションは可能であるが、接線連続で接続する曲面に対応する点群のセグメンテーションはできない。そこで本節では接線連続で接続する曲面に対応する点群のセグメンテーション手法について述べる。

3.4.1 法線の分散成分の VBGM によるクラスタリング

図 9 (左) のような、接線連続で接続する曲面に対応する点群について 3.2.3 で述べた法線成分の分散空間内でクラスタリングを行う。図 10 のように、分散空間内ではデータは曲率が同一な曲面ごとに塊状にクラスタを形成していると考えられるので、クラスタリングアルゴリズムには変分混合ガウス分布を用いる。クラスタリングの結果として、図 9 (右) のように曲率の異なる点群が別クラスタに分割される。

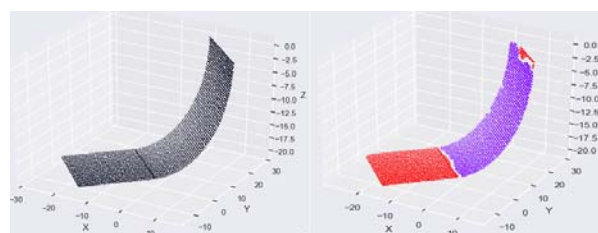


図 9 接線連続で接続した点群 (左) と分散空間での変分混合ガウス分布によるクラスタリング結果 (右)

Figure 9 A point cloud of a tangent continuous surface(left) and the result of clustering with VBGM on variance components(right).

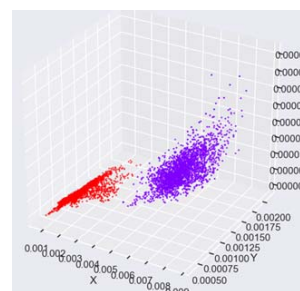


図 10 図 9 の点群の分散空間での分布

Figure 10 The distribution of variance components for data shown in Figure 9.

3.4.2 位置座標成分の DBSCAN によるクラスタリング

法線の分散成分の変分混合ガウス分布によるクラスタリングの結果、接線連続で接続する曲面に対応する点群は分

割られるが、複数の接線連続曲面がひと続きになっているような場合に、位置連続でない点群が同じクラスタに属したままとなる場合がある。そこでさらに位置座標空間内で DBSCAN によるクラスタリングを行うことで分割する。

3.5 エッジ部の点群の扱い

点群のセグメンテーションに関する既存の研究では、エッジ部分を含めたすべての点をセグメンテーションの対象としている。しかし、実用的にはリバースエンジニアリングにおいてはエッジやフィレット部分などの細かな部分については点群から忠実に再現する必要は無い場合が多い。むしろサーフェスを生成した段階ではサーフェス間にはフィレットが無いほうがデータを扱い易い。必要であればフィレットをリバース後に追加することは一般的には容易である。以上のことから、リバースエンジニアリングという目的に特化すれば、サーフェスを生成する段階でサーフェス同士の連続性が保証されるのであれば、点群のセグメンテーションの段階ではエッジ部分の点群は除外してもさほど問題ではない。前節までの手法によるセグメンテーションの結果、ノイズ除去と同時にエッジ部分の点群は自然と除外されるが、上述の理由によりリバースエンジニアリングの場合は問題ではない。

4. 実験および結果

提案手法を用いていくつかの点群についてセグメンテーションの実験を行った。使用した点群について表 1 に点群の点群数とセグメント数およびセグメンテーションの処理に要した時間を示した。

表 1 実験で使った点群データ

Table 1 Point clouds used in experiments.

データ名	点群数	セグメント数	処理時間(min)
Fandisk	134299	22	20
非凸立方体	88779	9	3
葉形	71283	4	2

4.1 実験の環境

前節で紹介した手法を、ウェブブラウザ上で実行可能なデータ分析ツールである Jupyter Notebook[8]上で Python により実装した。実装するにあたり使用したライブラリは以下の通りである。まず、多次元配列を用いた効率的な演算を実行するために NumPy[9]を、変分混合ガウス分布や DBSCAN や主成分分析などの機械学習アルゴリズムを実行するために scikit-learn[10]を利用した。さらに、入力する点群データのファイル形式は一般的に STL 形式なので、STL ファイルのインポートのために numpy-stl[11]を利用した。加えて Kd 木を用いた近傍点探索アルゴリズムなどの点群処理のために Open3D[12]を利用した。なお、実験に使

用したノート PC の CPU は Intel Core i7-4510U(2.0GHz)である。

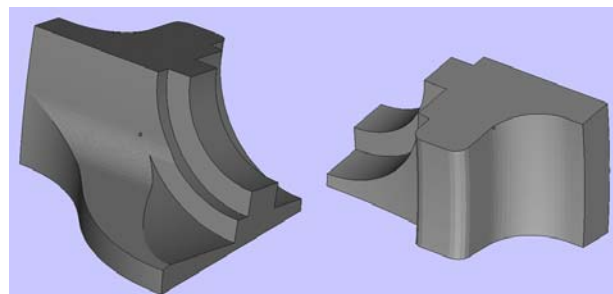


図 11 Fandisk の STL データの前面 (左) と背面 (右)
 Figure 11 Front(left) and back(right) of Fandisk in STL.

4.2 Fandisk の点群データのセグメンテーション

点群のセグメンテーション手法に関する論文でよく用いられる代表的なデータである、Fandisk (図 11) の形状測定点群データに提案手法を適用した。

図 12 は(a)~(e)の順に、それぞれ 3.3.1, 3.3.2, 3.3.3, 3.4.1, 3.4.2 で述べたクラスタリング手法を順に適用したセグメンテーション結果を示す。図 12(a)の段階では同一曲面に属する点群が分割されてしまっているが、次の(b)の段階では同一クラスタに併合されている。また(c)の段階では同一クラスタに併合されている接線連続の曲面に属する点群が、(d)の段階で曲率ごとに分割されている。最終的に(e)では位置連続でないクラスタが別クラスタに分割されている。

4.3 実際の形状測定点群データのセグメンテーション

実物の形状を 3D スキャナで測定して得られた形状測定点群データに提案手法を適用し、セグメンテーションの実験を行った。使用したのは図 13 (上) の非凸立方体と葉形の模型の形状測定点群データである。実際の形状測定データにはかなりのノイズとデータの欠損が含まれるため、たとえ平面对応する点群であっても同一平面上には存在しない。にもかかわらず、ノイズのフィルタリング処理と穴埋め処理を施すことなく適切なセグメンテーション結果となった (図 13 (下))。

5. おわりに

リバースエンジニアリングという目的に特化し、点群の位置座標成分、法線成分、法線成分の分散を特徴量とし、変分混合ガウス分布によるクラスタリングと DBSCAN によるクラスタリングを併せて適用することで、位置連続および接線連続で接続した曲面に対応する点群のセグメンテーションを自動化する手法を考案した。

有名な Fandisk の点群データおよび非凸立方体と葉形の模型の形状測定点群データに対して提案手法によりセグメンテーションの実験を行ったところ、適切なセグメンテ

ション結果が得られた。

今後は今回使用した Fandisk のような、より複雑な形状のサンプルについての、実際の形状測定点群データに対して提案手法を適用し、その有効性を確認していきたい。

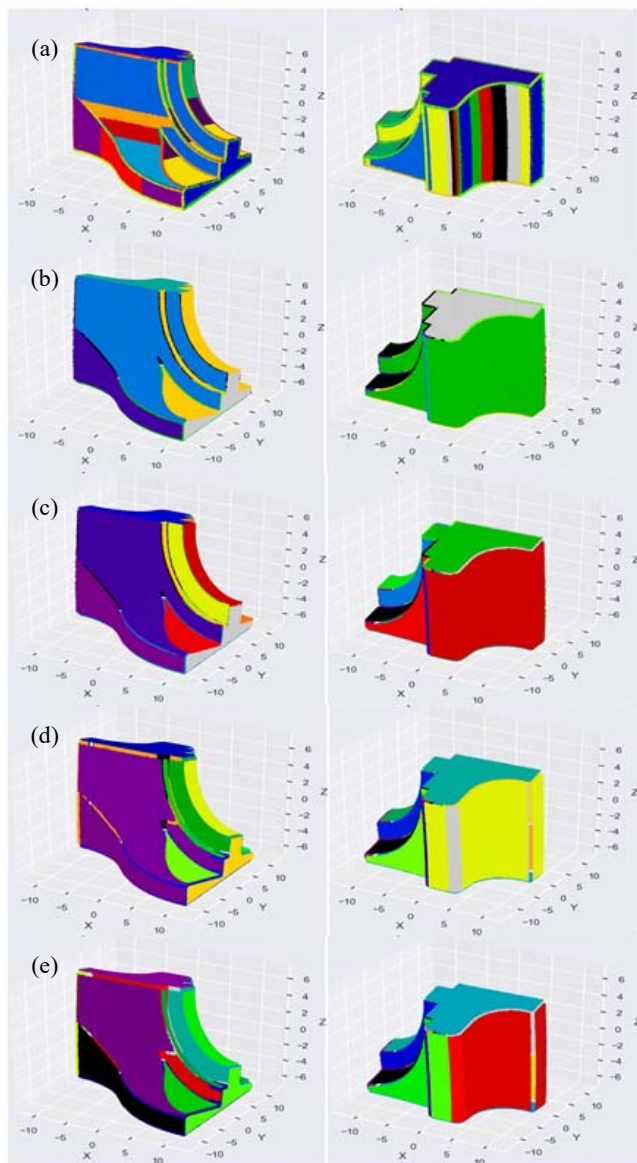


図 12 Fandisk の点群の各処理ステップにおけるセグメンテーション結果の前面（左側）と背面（右側）

Figure 12 Front(left side) and back(right side) of segmentation results of each steps for Fandisk.

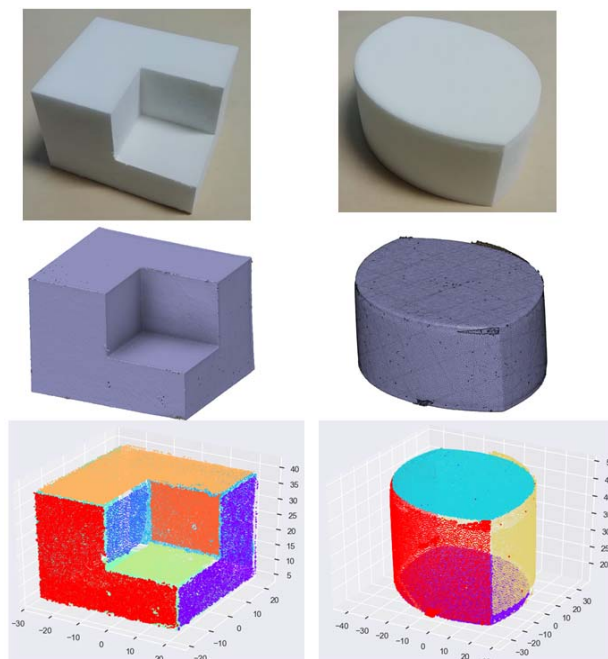


図 13 非凸立方体(左側)と葉形(右側) (模型の実物(上), 形状測定データ(中), セグメンテーション結果(下))

Figure 13 Concave cube(left side) and leaf shape(right side). Actual object (upper), acquired point cloud (middle), and segmentation result (lower).

参考文献

- [1] Gumhold, S. et al.. Feature Extraction from Point Clouds. Proc. of 10th IMR, 2001, p.293-305.
- [2] “Geomagic Design X” . <https://ja.3dsystems.com/software/geomagic-design-x/features/>, (参照 2020-01-27).
- [3] Varady, T.. Automatic procedures to create cad models from measured data. Computer Aided Design and Application, 2008, vol. 5, no. 5, p. 577-588.
- [4] MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations, Proc. of the 5th Berkeley Symposium on Math, 1967, vol. 1, p. 281-296.
- [5] Bishop, C. M.. Pattern Recognition and Machine Learning. Springer, 2006.
- [6] Ester, M. et al.. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Proc. of KDD-96, 1996, p. 226-231
- [7] Bentley, J. L.. Multidimensional binary search trees used for associative searching. Commun. ACM, 1975, vol. 18, no. 9, p. 509-517.
- [8] Jupyter Notebook. <https://jupyter.org/>, (参照 2020-01-27).
- [9] NumPy. <https://numpy.org/>, (参照 2020-01-27).
- [10] scikit-learn. <https://scikit-learn.org/stable/>, (参照 2020-01-27).
- [11] numpy-stl. <https://pypi.org/project/numpy-stl/>, (参照 2020-01-27).
- [12] Open3D. <http://www.open3d.org/>, (参照 2020-01-27).