

《様式B》

| | | | |
|-------|----------------------------|-----------------|------------------------------|
| 研究テーマ | 「 機械部品に特化した三次元モデルの形状類似検索 」 | | |
| 研究責任者 | 所属機関名 | 国立大学法人 豊橋技術科学大学 | |
| | 官職又は役職 | 助教 | |
| | 氏名 | 立間 淳司 | メールアドレス tatsuma@cs.tut.ac.jp |
| 共同研究者 | 所属機関名 | アイコクアルファ株式会社 | |
| | 官職又は役職 | シニアマネージャー | |
| | 氏名 | 東野 千春 | |

(平成 27 年度募集) 第 28 回 助成研究 完了報告書

1. 実施内容および成果ならびに今後予想される効果の概要 (1, 000 字程度)

※産業技術として実用化の可能性や特許出願 (予定も含む) の有無についてもご記載ください。

本研究課題の目的は、三次元機械部品モデルデータを対象とした、形状類似検索システムの開発であった。結果として、本研究課題で開発した三次元モデルの形状類似検索技術は、共同研究者のアイコクアルファ株式会社が開発するモデリングソフト **Simple Modeler** の加工実績検索機能として製品化された。

三次元モデルの形状類似検索では、形状を数値で表現した特徴ベクトルを計算する特徴抽出処理が重要となる。本研究課題では、まず、本研究グループが以前に考案した **Multi-Resolution Representation Local Binary Pattern (MRLBP)** 法をもとに三次元モデルの形状類似検索エンジンを開発し、実システムでの検索性能・速度の評価を行った。**MRLBP** 法は、機械部品形状の三次元モデルの検索に関して検索性能が優れており、特徴ベクトルのファイルサイズも小さいという利点をもつ。**MRLBP** 法は、実際の機械部品データに対しても、十分な検索性能・速度が得られ、**Simple Modeler** の検索機能として組み込まれることとなった。

次に、より直感的な検索システムの実現のため、スケッチ画像による検索を進めた。本研究グループが以前に提案した **Overlapped Pyramid Histogram of Oriented Gradients (OPHOG)** 法は、優れた検索性能が得られる一方で、特徴ベクトルのファイルサイズが大きくなってしまいう問題があった。そこで、**OPHOG** 法のファイルサイズに関連するパラメータである、エッジ画像生成数・重複分割レベル・ヒストグラムビン数の調整を試みた。結果として、実装上の工夫と組み合わせて、元の **OPHOG** 法のファイルサイズからの約 $1/23$ のサイズへの削減に成功した。しかし一方で、わずかではあるが、検索性能の低下が発生した。そこで、類似度計算処理を改善することでこれを補うことを考え、スパースコーディングで用いられる再構成誤差の概念を導入した類似度計算法を開発した。これは、検索クエリとして与えられたスケッチ画像の特徴ベクトルを、三次元モデルのエッジ画像の特徴ベクトルの重み付き線形和で表現 (再構成) し、その重みを類似度とする手法である。これにより、大きな検索性能の改善が実現でき、特許出願を行った。

この他、今後の需要を見込み、色付きの三次元モデルの類似検索を目的とした、特徴抽出手法

Volumetric Spiral Scan Order 法の開発、三次元モデルから作成した深さバッファ画像に最適化した畳み込みニューラルネットワークである Reduced VGG-M ネットワークの開発を進めた。

2. 実施内容および成果の説明（A 4 で、5 ページ以内）

本研究課題の目的は、三次元機械部品モデルデータを対象とした、形状類似検索システムの開発である。三次元モデルの形状類似検索において、検索性能に大きな影響を与える重要な処理は、形状を数値で表現した特徴ベクトルを計算する特徴抽出処理と、特徴ベクトル間の違いを表す類似度の計算処理である。また、実用化においては、検索性能だけでなく、特徴ベクトルファイルのサイズや実行速度も重要となる。本研究課題では、実用化も視野に入れながら、三次元モデルの形状類似検索に関する特徴抽出処理・類似度計算処理の研究を行った。

本研究課題では、まず、研究責任者の研究グループが以前に考案した **Multi-Resolution Representation Local Binary Pattern (MRLBP)** 法（特開 2015-176484）[1] による三次元モデルの形状類似検索エンジンを開発し、実際の機械部品モデルデータでの検索性能・速度の評価を、共同研究者の協力のもとで行った。**MRLBP** 法（図 1）では、まず、三次元モデルから複数視点で深さバッファ画像をレンダリングし、それらから多重解像度 **LBP** 画像を作成する。多重解像度 **LBP** 画像は、三次元モデル表面の凹凸を異なる解像度でとらえた画像となる。そして、多重解像度 **LBP** 画像の輝度値によるヒストグラムを作成し、それを特徴ベクトルとする。**MRLBP** 法は、機械部品形状の三次元モデルに関して優れた検索性能を得られ、特徴ベクトルのファイルサイズも小さいという利点をもつ。評価の結果、**MRLBP** 法は、実際の機械部品データに対しても、十分な検索性能・速度が得られ、共同研究者であるアイコクアルファ株式会社が開発する **Simple Modeler** の検索機能として組み込まれた（図 2）。

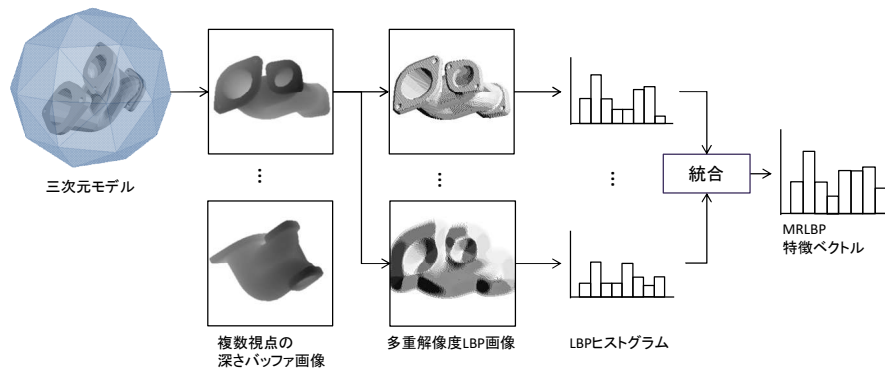


図 1 MRLBP 法による特徴抽出処理の概要

三次元モデルの形状類似検索システムは、三次元モデルそのものを検索クエリとするため、すでに三次元モデルを所有するデザイナーなど、専門家が使用するツールとなっていた。これに対して、より直感的なインターフェースでの検索を実現するため、手書きのスケッチ画像を検索クエリとした三次元モデルの検索が注目されている。研究責任者の研究グループが、以前に、スケッチによる三次元モデル検索のために開発した特徴抽出手法 Overlapped Pyramid Histogram of Oriented Gradients (OPHOG) 法 (特開 2015-201151) [2] は、優れた検索性能が得られる一方で、特徴ベクトルのファイルサイズが大きくなる問題があった。これは、大規模データを検索対象とする実際の検索システムでは、ストレージ容量と検索速度に悪影響をおよぼす。OPHOG 法は、三次元モデルから複



図 2 Simple Modeler に実装された形状類似検索機能

数視点で作成したエッジ画像およびスケッチ画像に対して、画像を小領域に重複分割し、小領域ごとに輝度勾配ヒストグラム (Histogram of Oriented Gradients) を計算する。よって、OPHOG 法では、三次元モデルから生成するエッジ画像数、重複分割レベル、ヒストグラムビン数がパラメータとなる。これらを、大きく検索性能を低下させることなく、特徴ベクトルのファイルサイズが小さくなる様に調整した。結果として、実装上の工夫とあわせて、元の OPHOG 法のファイルサイズからの約 1/23 のファイルサイズへの削減に成功した。しかし、わずかではあるが、検索性能の低下が発生した。そこで、特徴抽出処理ではなく、類似度計算処理により検索性能を改善することを考えた。

これまでの OPHOG 法の類似度計算では、ユークリッド距離を用いていた。特徴ベクトルの次元

数が大きくなると、次元の呪いと呼ばれる現象により、ユークリッド距離では特徴ベクトル間の差異を正しく表現できなくなることが知られている。これに対して、本研究では、距離ではなく、スパースコーディングで用いられる再構成誤差（Reconstruction Error）の重みを類似度に用いる手法を考案した（特願 2017-042156）。この手法は、検索クエリとして与えられたスケッチ画像の特徴ベクトルを、三次元モデルのエッジ画像の特徴ベクトルの重み付き線形和で表現（再構成）し、その重みを類似度とする。このとき、MRLBP 法で評価した、三次元モデル間の類似度も反映させる。

42 種のクラスに分けられた機械部品形状の三次元モデル 867 個、検索クエリとなるスケッチ画像 420 枚からなるデータセットを作成し、検索性能を評価した。評価尺度は、三次元モデルの形状類似検索で一般的に用いられる、Nearest Neighbor (NN)、First Tier (FT)、Second Tier (ST) を選択した。いずれの評価尺度も、値が大きいほど検索性能が優れていると判断できる。表 1 は、OPHOG 法、OPHOG 法と類似度に再構成誤差の重みを用いたものの検索性能をまとめたものである。類似度計算を改善することで、大きく検索性能が向上していることがわかる。

表 1 OPHOG 法と再構成誤差の重みによる類似度計算法の検索性能比較

| 手法 | FT | ST | NN |
|------------------|-------|-------|-------|
| OPHOG 法 | 0.471 | 0.240 | 0.318 |
| OPHOG 法 + 類似度計算法 | 0.538 | 0.286 | 0.373 |

今後の需要を見込み、色付き三次元モデルの類似検索に取り組んだ。機械部品設計では、加工方法などのメタ情報を色によって表す場合がある（図 3）。よりの確な機械部品の検索のため、三次元モデルの形状だけでなく色も考慮した検索技術が必要となる。これに対して、本研究では、形状と色の分布の両方を捉えた特徴抽出手法 Volumetric Spiral Scan Order (VSSO) 法 [3] を考案した。

VSSO 法では、まず、三次元モデルを、形状を立方体の集合で表現するボクセル表現に変換する。次に、ボクセル表現を、一定の大きさと間隔で、小領域にブロック分割する。そして、各小領域において、図 4 の様に、ボクセルをらせん状に走査しながら、現在走査しているボクセルと、次に走査するボクセルとの色の距離を計算する。得られた色の距離の重みとして、色ヒストグラムを作成することで、三次元モデルの色の分布を捉える。さらに、各小領域で、空ボクセルではない RGB 値を持つボクセルの数をカウントし特徴量とすることで、三次元モデルの形状を捉える。

色付きの機械部品モデルのデータセットは公開されておらず、研究グループでも作成することが難しかった。そこで、The Virtual Hampson Museum で公開されている考古遺物を三次元スキャンしたモデル 302 個を 16 種のクラス分けすることで、評価用のデータセットを用意した。比較実験から、色付き三次元モデルや点群を考慮した従来手法よりも優れた検索性能を得ることができた（表 2）。

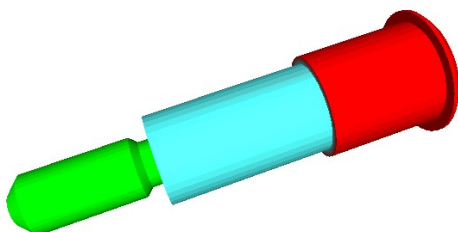


図 3 色によって加工方法を示した機械部品モデルの例

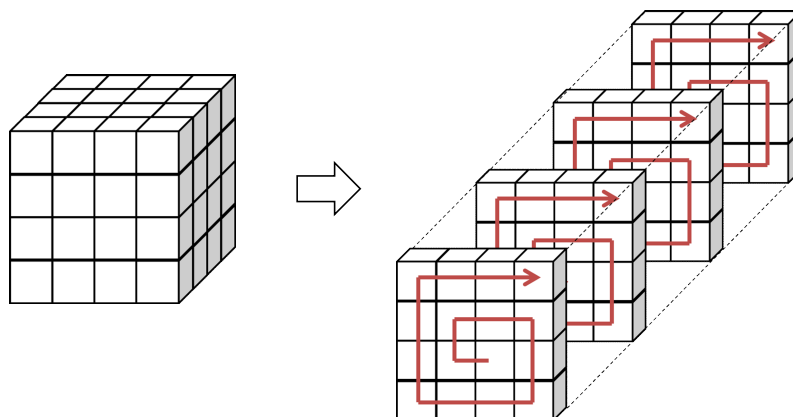


図 4 VSSO 法におけるボクセル表現のらせん状走査の例

表 2 色付き三次元モデルにおける検索性能比較

| 手法 | NN | FT | ST |
|------------------------------------|-------|-------|-------|
| SHOT (S. Salti et al., 2014) | 0.500 | 0.285 | 0.449 |
| SH-Lab (A. Giachetti et al., 2015) | 0.527 | 0.402 | 0.632 |
| C-CHLAC (A. Kanazaki et al., 2010) | 0.530 | 0.478 | 0.660 |
| VSSO (提案手法) | 0.676 | 0.494 | 0.678 |

近年、主に画像認識の分野において、深層学習による特徴抽出技術が大きく発展している。三次元モデルの形状類似検索においても、三次元モデルから複数視点で深さバッファ画像を作成し、それらを畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) に入力することで特徴ベクトルを得る手法が、一定の成果を得ている [4]。CNN は、様々なネットワーク構成が提案されているが、三次元モデルの深さバッファ画像を用いた特徴抽出では、VGG-M ネットワーク [5] が多く用いられている。VGG-M は、物体や風景などをデジタルカメラで撮影した画像による画像認識を目的としている。一方、三次元モデルから作成した深さバッファ画像は、グレースケールであり、背景部分の多くの輝度値がゼロとなる画像であり、一般的な写真画像とは大きく異なる。そこで、VGG-M は、深さバッファ画像からの特徴抽出には、ユニット数が多く冗長な構成であると考え、VGG-M におけるユニット数を削減した Reduced VGG-M ネットワーク [6] を考案した。

Reduced VGG-M は、VGG-M とネットワーク構造は同様であるが、各層のユニット数が異なる (表 3)。学習段階では、ラベルが付与された三次元モデルデータセットを用意し、それらモデルから複数視点から作成した深さバッファ画像を入力として、画像認識における CNN の学習と同様にして、深さバッファ画像が正しく分類されるようにネットワークを学習する。テスト段階 (特徴抽出) では、三次元モデルから複数視点から深さバッファ画像を作成し、学習済みのネットワークに入力し、ネットワークの全結合層 (分類結果を出力する層の一つ手前の層) の出力を得る。そして、それら全結合層の出力のベクトル和を L2 正規化したものを、三次元モデルの特徴ベクトルとする (図 5)。

深層学習による三次元モデルの形状類似検索では、評価のために、Princeton 大が提供する ModelNet40 データセットが一般に用いられる。ModelNet40 には、40 種のクラスに分けられた、学習用の三次元モデルが 9,843 個、テスト用の三次元モデルが 2,468 個ふくまれている。表 3 は、従来

の VGG-M で特徴抽出したものと、Reduced VGG-M で特徴抽出したものの検索性能をまとめたものである。Reduced VGG-M が、優れた検索性能を得ていることがわかる。また、Reduced VGG-M は、VGG-M と比較して、ユニット数が約 1/4 であり、メモリ消費量は少なく、実行速度もより高速となる。これは、実際のシステムに組み込む場合に、大きな利点となる。ModelNet40 では、Accuracy（これは三次元モデルの分類性能を表す）と Mean Average Precision（MAP）で評価するのが一般的である。参考までに、Reduced VGG-M と従来手法の評価尺度の値をまとめたものを、表 5 に示す。Reduced VGG-M は、コンパクトなネットワーク構成であるにも関わらず、従来手法よりも優れた性能もしくは同等の性能を得ていることがわかる。ここで、参考までに、GIFT などの従来手法の一部は、深層学習による特徴抽出処理だけでなく、リランキング処理も含まれていることを記す。

表 3 VGG-M と Reduced VGG-M の各層のユニット数

| 構成 | 畳み込み 1 | 畳み込み 2 | 畳み込み 3 | 畳み込み 4 | 畳み込み 5 | 全結合 6 | 全結合 7 |
|---------------|--------|---------|---------|---------|---------|-------|-------|
| VGG-M | 96×7×7 | 256×5×5 | 512×3×3 | 512×3×3 | 512×3×3 | 4096 | 4096 |
| Reduced VGG-M | 48×7×7 | 64×5×5 | 128×3×3 | 128×3×3 | 128×3×3 | 1024 | 1024 |

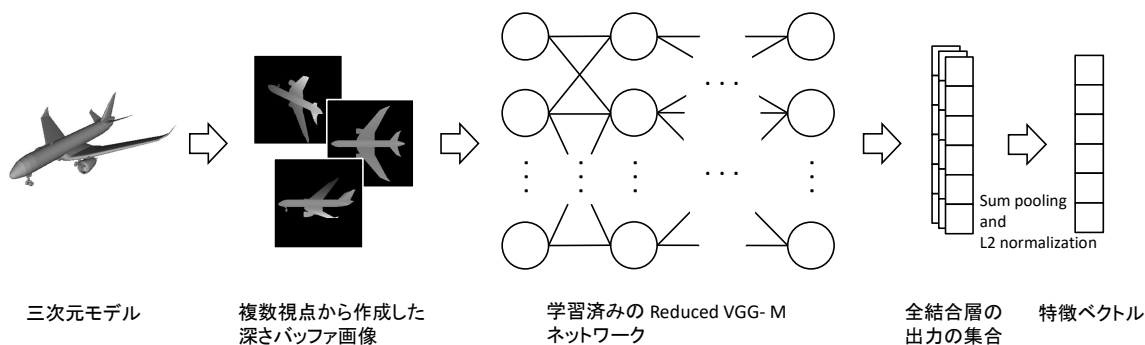


図 5 Reduced VGG-M による特徴抽出処理の概要

表 4 ModelNet40 における各ネットワーク構成の検索性能比較

| 手法 | NN | FT | ST |
|----------------------|-------|-------|-------|
| VGG-M | 0.878 | 0.700 | 0.824 |
| Reduced VGG-M (提案手法) | 0.874 | 0.747 | 0.867 |

表 5 ModelNet40 における Reduced VGG-M と従来手法の性能比較

| 手法 | Accuracy | MAP |
|----------------------------------|----------|--------|
| 3DShapeNets (Z. Wu et al., 2015) | 77.0 % | 49.2 % |
| DeepPano (B. Shi et al., 2015) | 77.6 % | 76.8 % |
| MVCNN (H. Su et al., 2015) | 90.1 % | 79.5 % |
| Geometry Image (A. Sinha, 2016) | 83.9 % | 51.3 % |
| GIFT (S, Bai et al., 2016) | 83.1% | 81.9% |
| Reduced VGG-M (提案手法) | 87.4 % | 77.0 % |

【文献】

- [1] S. Biasotti, A. Cerri, M. Aono, A. Ben Hamza, V. Garro, A. Giachetti, D. Giorgi, A. Godil, C. Li, C. Sanada, M. Spagnuolo, A. Tatsuma, S. Velasco-Forero, “Retrieval and classification methods for textured 3D models: a comparative study”, *The Visual Computer*, Vol. 32, No. 2, Springer, pp. 217-241, 2016.
- [2] B. Li, Y. Lu, C. Li, A. Godil, T. Schreck, M. Aono, M. Burtscher, Q. Chen, N. Chowdhury, B. Fang, H. Fu, T. Furuya, H. Li, J. Liu, H. Johan, R. Kosaka, H. Koyanagi, R. Ohbuchi, A. Tatsuma, Y. Wan, C. Zhang, C. Zou, “A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries”, *Computer Vision and Image Understanding*, Vol. 131, Elsevier, pp. 1-27, 2015.
- [3] 濱田和真, 立間淳司, 青野雅樹, “Volumetric Spiral Scan Order による三次元考古遺物モデルの類似検索”, *情報処理学会論文誌データベース*, Vol.74, 2017.
- [4] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller, “Multi-view Convolutional Neural Networks for 3D Shape Recognition”, *Proc. of the 2015 IEEE International Conference on Computer Vision*, pp. 945-953, 2015.
- [5] K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman, “Return of the Devil in the Details: Delving Deep into Convolutional Networks”, *Proc. of British Machine Vision Conference*, 2014.
- [6] M. Sarva, F. Yu, H. Su, A. Kanazaki, T. Furuya, R. Ohbuchi, Z. Zhou, R. Yu, S. Bai, X. Bai, M. Aono, A. Tatsuma, S. Thermos, A. Axenopoulos, G. Papadopoulos, P. Daras, X. Deng, Z. Lian, B. Li, H. Johan, Y. Lu, S. Mk, “SHREC’17 Track: Large-Scale 3D Shape Retrieval from ShapeNet Core55”, *Proc. of Eurographics Workshop on 3D Object Retrieval*, 2017.