

局所特徴量を用いた 3D CNN による 3 次元モデルの部分検索

岩淵 渉[†] 青野 雅樹^{††}

[†] 豊橋技術科学大学 情報・知能工学課程 〒441-8580 愛知県豊橋市天伯町字雲雀ヶ丘 1-1

^{††} 豊橋技術科学大学 情報・知能工学系 〒441-8580 愛知県豊橋市天伯町字雲雀ヶ丘 1-1

E-mail: †iwabuchi@kde.cs.tut.ac.jp, ††aono@tut.jp

あらまし 近年、製造業や医療、建築などさまざまな分野で 3 次元モデルが利用されている。しかし、3 次元モデルが再利用されることは少なく、これらのデータの再利用を目的とした 3 次元モデルの形状類似検索の需要が高まっている。しかし、3 次元モデルを検索質問とした形状類似検索は、利用者が 3 次元モデルのデータを所有していなければ行うことができないという欠点がある。そこでユーザが手軽に検索を行うためには、Kinect などの 3D センサを用いて得られるような不完全な部分モデルデータを検索質問とする部分検索手法が必要である。本研究では局所特徴量を入力とした 3D CNN を用いてこの問題に取り組む。実験は SHREC 2016 Partial データセットを用いて行った。その結果、提案手法が 3D CNN を用いた従来手法である VoxNet より優れた検索性能を示した。

キーワード 3D, 3D 検索, 深層学習, 局所特徴量, ボクセル

1. はじめに

近年、製造業や医療、建築、教育、娯楽などさまざまな分野で 3 次元モデルが利用されている。それに伴い 3 次元モデルのデータ数が増加しており、高精度に検索を行う手法が求められている。しかし、3 次元モデルを検索質問とした形状類似検索は、利用者が 3 次元モデルのデータを所有していなければ行うことができないという欠点がある。そこでユーザが手軽に検索を行うために、比較的安価に入手できる Kinect [1] に代表される 3 次元スキャナを用いて 3 次元形状を入力とすることが簡便である。ただし、3D スキャナから入力された 3 次元データは、一般に不完全な形状である。この不完全な形状は、見方を変えれば、完全な 3 次元形状モデルの部分形状と考えることができる。そこで、3 次元スキャナから入力された 3 次元の部分形状を検索質問と見立てる 3 次元形状の部分検索の研究が行われてきている [4] [6] [8] [15] [18]。

本研究では、3 次元の部分検索としては我々が知る限り、これまで用いられなかった局所特徴量を導入した新しい部分検索手法を提案する。提案手法では、3 次元スキャナから入力された 3 次元メッシュデータを、まず点群に変換し、その後、局所特徴量を抽出し、ボクセルデータで表現する。ボクセルデータを表現とする 3 次元モデルの検索の従来手法例として、VoxNet [12] が有名である。VoxNet では、各ボクセルは 2 値で表現し、メッシュからボクセルに変換する際にボクセル内にデータがあるかどうかで 0 か 1 の値を決めて表現する。提案手法では、ボクセルを 2 値ではなく、局所特徴量で表現する。提案手法の有効性を示すため、VoxNet と検索性能の比較を行い、その結果を報告する。

2. 関連研究

3 次元モデルの部分検索には KAZE+VLAD [15], Randomized Sub-Volumes Partitioning(RSVP) [5] [15], Tran らの手

法 [15], 小林らの手法 [22] などが代表的である。

KAZE+VLAD は画像ベースの手法であり、KAZE 特徴量と Vector of Locally Aggregated Descriptor(VLAD) を組み合わせた手法である。まず、3 次元物体を複数視点からレンダリングを行い、生成した 2 次元画像から KAZE 特徴量を抽出する。抽出した KAZE 特徴量に VLAD を組み合わせ、特徴ベクトルを作成する。

RSVP は 3 次元モデルから部分形状を抽出する手法である。3 次元モデルをグリッド状に区切り、それらを Sub-Volumes とする。Sub-Volumes は部分形状に対応しており、Sub-Volumes を特徴ベクトル化しそれらを比較することで部分形状の比較を行うことができる。

Tran らは RootRoPS と呼ばれる局所特徴量を用いた BoW による手法と ICP による 3 次元モデルの Registration による手法を組み合わせることによって 3 次元モデルの部分検索を行っている。

小林らは可視領域に着目した部分モデルの抽出と法線による特徴量などを用いている。

ボクセルと 3D CNN を用いた形状類似検索手法には Voxnet [12] が代表的である。これはボクセルを 0 または 1 の 2 値で表現する。3 次元モデルをボクセルで区切り、そのボクセル内に点が存在すれば 1, 存在しなければ 0 という値をそのボクセルに設定する。

3 次元の局所特徴量に関する研究には Fast Point Feature Histogram(FPFH) [16] や 3DMatch [21] などがある。FPFH は法線やヒストグラムを用いている。代表点とその近傍点との相対的な位置関係や各点の法線によって局所特徴量が定まる。3DMatch は 3D CNN を用いて局所特徴量を求める手法である。3D CNN の学習に Siamese Network [9] を組み合わせることで、3D CNN によって局所特徴量を学習することを可能としている。

3. 提案手法

提案手法の大まかな流れを図1に示す。図内の点線部分が、今回の主たる提案部分である。これを詳細に記述した図が図2である。

提案手法について3.1節では3次元モデルの正規化，3.2節では部分モデルの抽出，3.3節では局所特徴量を持ったボクセルの生成，3.4節では部分モデルの大域特徴量，3.5節では検索を説明する。

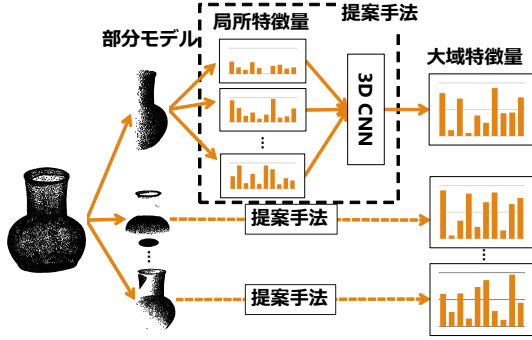


図1 提案手法の概要

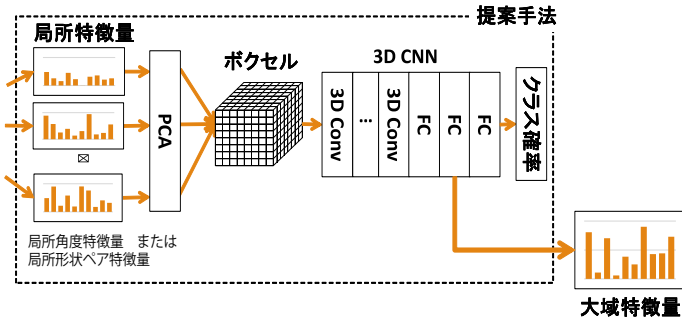


図2 提案手法の詳細

3.1 3次元モデルの正規化

最初に3次元モデルの正規化を行う。これは3次元モデルの位置や大きさ、向きといった情報が異なれば、本来同じ3次元モデル同士でも異なる3次元モデルと判定されてしまうためである。これらの影響を抑制するために正規化を行う。

まず位置の正規化を行う。モデルの重心の座標が原点に重なるように3次元モデルを平行移動させる。次に大きさの正規化を行う。大きさは3次元モデルが単位円に納まるように正規化を行う。重心と各点とのユークリッド距離を計算し、その最大値を求める。得られた最大値を用いて各点の座標を除算することで3次元モデルが単位円に納まるように変換される。次に向きの正規化を行う。向きの正規化は点群発生と特異値分解からなるPointSVD [23]によって行う。

具体的には、以下の式に基づいて正規化を行った。まず、一般性を失うことなく、3次元モデルは三角形メッシュからなると仮定できる。そこで、3次元モデルの面上にランダムな点を m 個の点群として生成する [13]。この際、3次元モデルの面上

に一様に分布する点の座標 \mathbf{p} は、頂点と接続性により構成される3角形の頂点座標 \mathbf{a} , \mathbf{b} , \mathbf{c} から以下の式で求める。

$$\mathbf{p} = (1 - \sqrt{r_1})\mathbf{a} + \sqrt{r_1}(1 - r_2)\mathbf{b} + \sqrt{r_1}r_2\mathbf{c}$$

上式の2つの乱数 r_1 , r_2 には、Sobol あるいはNiederreiterなどの擬似乱数を用いる [2]。次に、3次元モデルの重心 \mathbf{g} と向きを決める回転行列 Q を求める。生成した点群の平均値を求め、それを重心とする。

$$\mathbf{g} = \frac{1}{m} \sum_{i=1}^m \mathbf{p}^{(i)}$$

重心が原点になるよう点群を平行移動し、それを点群行列 P とする。

$$P = \begin{pmatrix} p_x^{(1)} - g_x & p_x^{(2)} - g_x & \dots & p_x^{(m)} - g_x \\ p_y^{(1)} - g_y & p_y^{(2)} - g_y & \dots & p_y^{(m)} - g_y \\ p_z^{(1)} - g_z & p_z^{(2)} - g_z & \dots & p_z^{(m)} - g_z \end{pmatrix}$$

点群行列 P を以下のように特異値分解する。

$$P = U\Sigma W^T$$

ここで、 U は 3×3 の直交行列、 Σ は 3×3 で対角に特異値が降順に並ぶ対角行列、 W^T は $3 \times m$ の直交行列である。求めた左特異ベクトル行列 U を、各特異ベクトルで正規化する。正規化された左特異ベクトル行列 \hat{U} を転置したものを回転行列 Q とする。

$$Q = \hat{U}^T$$

最後に、回転した点群 $P' = QP$ より、3次元モデルの前後を決める鏡像行列 F を求める。

$$F = \begin{pmatrix} \text{sign}(f_x) & 0 & 0 \\ 0 & \text{sign}(f_y) & 0 \\ 0 & 0 & \text{sign}(f_z) \end{pmatrix}$$

ただし、

$$f_x = \sum_{i=0}^m \text{sign}(p_x^{(i)})(p_x^{(i)})^2 \quad (f_y, f_z \text{ も同様})$$

である。3次元モデルの k 個の頂点情報を行列 V 、回転後の頂点情報を行列 V' で表し、位置と向きの任意性を以下のように解決する。

$$V = \begin{pmatrix} v_x^{(1)} - g_x & v_x^{(2)} - g_x & \dots & v_x^{(k)} - g_x \\ v_y^{(1)} - g_y & v_y^{(2)} - g_y & \dots & v_y^{(k)} - g_y \\ v_z^{(1)} - g_z & v_z^{(2)} - g_z & \dots & v_z^{(k)} - g_z \end{pmatrix}$$

$$V' = FQV$$

回転後の各頂点の重心からの距離を求め、その最大距離で各頂点の座標値を割ることで、半径1の球 Bounding Sphere に3次元モデルを納める。

3.2 部分形状モデルの抽出

部分形状モデルの抽出は、我々の先行特許を用いて行う [22]。この手法ではある視点から 3 次元モデルを見た場合をシミュレートし、可視部分を抽出する。特許の実施例では、3 次元の部分形状抽出を 66 視点で行っている例を示しているが、ここでは高速化のため、38 視点から部分モデルを抽出する。

3.3 局所特徴量を持ったボクセルの生成

本研究では、3 次元モデルの局所特徴量として、モバイルロボティクス分野で導入された FPFH (Fast Point Feature Histogram) と、3 次元の物体認識・再構築で導入された 3DMatch の 2 種類を、3 次元形状モデルの部分検索のための局所特徴量として採用する。いずれの特徴量も、3 次元形状モデル上の「代表点」をピックアップし、その周辺の局所形状を求めていくアプローチである。

3.3.1 局所角度特徴量

局所角度特徴量として、FPFH に関して、ここでの実装手順を簡単に述べる。FPFH では、3 次元の局所的な形状を以下の 3 つの角度に関連する量を計算することで求める。すなわち、以下の式を用いて α , ϕ , θ を計算する。

$$\begin{aligned}\alpha &= \mathbf{v} \cdot \mathbf{n}_t \\ \phi &= \mathbf{u} \cdot \frac{(\mathbf{p}_t - \mathbf{p}_s)}{d} \\ \theta &= \arctan\left(\frac{\mathbf{w} \cdot \mathbf{n}_t}{\mathbf{u} \cdot \mathbf{n}_t}\right)\end{aligned}$$

ここで \mathbf{p}_s は代表点、 \mathbf{p}_t はその近傍点、 \mathbf{n}_s は \mathbf{p}_s の法線、 \mathbf{n}_t は \mathbf{p}_t の法線、 d は \mathbf{p}_s と \mathbf{p}_t のユークリッド距離を表している。ただし、局所的な座標系を与える $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ は互いに直交する単位ベクトルで以下の式で求める。

$$\begin{aligned}\mathbf{u} &= \mathbf{n}_s \\ \mathbf{v} &= \mathbf{u} \times \frac{(\mathbf{p}_t - \mathbf{p}_s)}{\|\mathbf{p}_t - \mathbf{p}_s\|_2} \\ \mathbf{w} &= \mathbf{u} \times \mathbf{v}\end{aligned}$$

実際に、欲する 3 つの角度特徴量 α , ϕ , θ に関しては、代表点の半径 r の範囲内にあるすべての近傍点に対して計算し、それらを適当なビン数 (例 11) のヒストグラムに変換し連結する。3 つの角度の意味としては、 α が、代表点での法線ベクトルと \mathbf{v} 軸とのなす角度の余弦、 ϕ が、代表点から近傍点と接線ベクトル \mathbf{u} とのなす角度の余弦、 θ が、代表点での \mathbf{u} 軸から \mathbf{w} 軸を見上げた仰角を表す。この連結したものをその代表点の SPFH とする。

最後に SPFH から FPFH を以下の式で求める。点 \mathbf{p}_q の FPFH を求める式は次のようになる。

$$\text{FPFH}(\mathbf{p}_q) = \text{SPFH}(\mathbf{p}_q) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_i} \cdot \text{SPFH}(\mathbf{p}_i)$$

ここで ω_i は \mathbf{p}_q と \mathbf{p}_i の距離を表す。 k は、探索する近傍点を決める半径 r に依存して決定され、FPFH は、 k 個の SPFH の重みつき平均として定義される量である。なお今回の実験では近傍点を探索する半径 r は 0.3 とした。

3.3.2 局所形状ペア特徴量

もうひとつの局所特徴量として、3DMatch を改良した局所形状ペア特徴量を説明する。この局所特徴量を求める手順は以下のようなものである。まず、ある代表点を中心とし一辺の長さを l とする立方体を生成する。生成した立方体の内部に存在する点群を取り出す。次に取り出した点群を局所的なボクセルデータに変換する。ここでは生成した局所的なボクセルデータをパッチと呼ぶ。このパッチを入力として、2 つのパッチのペアを入力とする Siamese Network [3] を訓練する。訓練後はペアとなるパッチを入力とし局所特徴量が出力される特徴抽出器として使用することができる。

今回の実験では、この局所形状ペア特徴量に向きの影響を減らすための処理を加えた。つまり、図 3 に示すように、3DMatch に基づく元来の局所形状ペア特徴量では、パッチを取り出す向きによって結果として得られる特徴量が変化してしまうという問題がある。そこで、図 3 に示すような 3 次元モデル上の曲率が類似する代表点のペアの局所形状ペア特徴量が類似する値となるように、我々の実装では、点群を取り出した後に 3.1 で述べた PointSVD を用いて向きの正規化を行うという処理を加えた。なお、前節で述べた FPFH では図 3 のような場合でも似たような特徴量が得られるためこの処理は行っていない。

今回の実験では局所形状ペア特徴量は Analysis-by-Synthesis [19], 7Scenes [17], SUN3D [20], RGB-D Scenes v.2 [11], Halber et al [7] といったデータを用いてすでに学習済みのニューラルネットワークを使用した。特徴量の次元数は 512 とし、パッチの一片の長さ l は $l = 0.3$ として実験を行った。

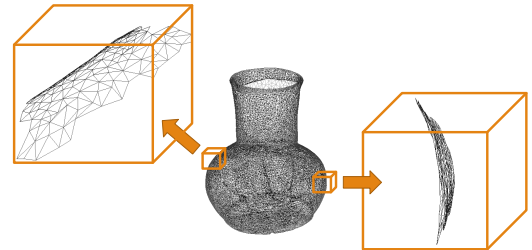


図 3 局所形状ペア特徴量で、曲率が類似するとしていたペア例。本来の 3DMatch では向きが異なると判定される

3.3.3 ボクセル生成

局所特徴量を持ったボクセルの生成方法に関して説明する。まず、与えられた任意の 3 次元モデル上にランダムに K 個の代表点 (実験では $K = 500$) を選び、各代表点に対して局所特徴量を求める。次に、 $L \times L \times L$ (実験では、 $L = 8$) のボクセルを生成する。この時点では、ボクセルによって内部に複数の局所特徴量が含まれているものや、まったく含まないものが存在する。続いて、ボクセルごとに局所特徴量の平均を計算し、その結果をボクセルごとの局所特徴量とする。ボクセルに局所特徴量が含まれていないものは 0 ベクトルとした。

局所特徴量は主成分分析 (PCA) を適用することで次元削減を行う。これは局所特徴量の次元が大きすぎる場合、3D CNN の学習が進まなかったためである。PCA の各係数は 3d pottery

dataset [10] を用いて計算する。このデータセットの各モデルに対しても同様に局所特徴量を求める。それらに対して主成分分析を行い各係数を計算する。実験ではこの係数を用いて次元削減を行う。すなわち、教師無し学習としての PCA の訓練データに 3D pottery dataset を用いることとした。局所特徴量の次元数はきりのよい数字であり、加えて累積寄与率が 0.95 以上である 8 とした。このときの累積寄与率のグラフを図 4 に示す。図 4 において、赤色のグラフが局所角度特徴量に対応し、実装では、33 次元 (α, ϕ, θ それぞれ 11 ビン) から PCA で 8 次元に次元削減し、一方、青色のグラフが局所形状ベア特徴量に対応し、512 次元から PCA で 8 次元に次元削減することになる。

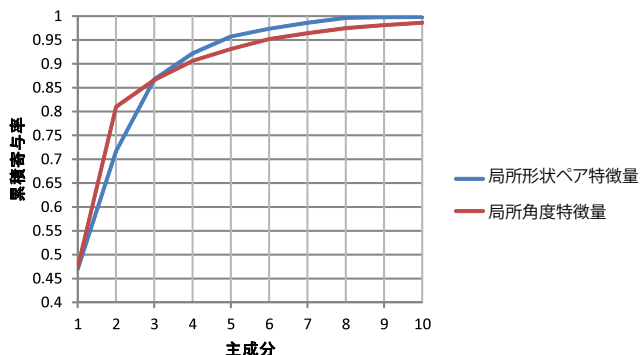


図 4 累積寄与率

3.4 部分モデルの大域特徴量

部分モデルの大域特徴量は 3D CNN を用いて求める。3D CNN とは画像に対して用いられる畳込みニューラルネットワーク (Convolutional Neural Network (CNN)) を 3 次元に拡張したものである。3 次元の畳込み処理は図 5 のようになる。3D CNN の学習はクラス分類問題として行い、学習後の最後から 2 番目の全結合層 (Fully Connected (FC)) の出力を取り出し大域特徴量とする。

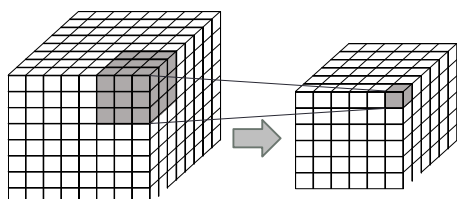


図 5 3 次元の畳込み

3.5 検索

検索は大域特徴量のユークリッド距離を比較することで行う。大域特徴量は 1 つの検索質問からは 1 つ得られ、1 つの検索対象からは抽出した部分モデルの数だけ得られる。検索質問の大域特徴量と検索対象のすべての部分モデルの大域特徴量のユークリッド距離を計算し、もっともユークリッド距離が小さいものをその検索質問と検索対象のユークリッド距離とする。これをすべての検索対象に対して行い、ユークリッド距離が小さいものから順に並べたものを検索結果とする。

4. 実験

実験について 4.1 節ではデータセット、4.2 節では評価尺度、4.3 節では比較手法、4.4 節では 3D CNN の学習方法、4.5 節では実験結果と考察を説明する。

4.1 データセット

データセットは SHREC16 Partial [15] を用いて行った。このデータセットに含まれるデータは 6 クラスに分けられており、検索対象として完全な 3 次元モデルを 383 個含んでいる。検索質問には大きく分けて Artificial (図 6), Breuckmann (図 7), Kinect (図 8) の 3 種類があり、これらはそれぞれ部分の抽出方法が異なる。Artificial は完全なモデルを人の手によって境界を定め部分の抽出を行ったものである。Artificial は Q25 と Q40 という 2 つのグループに分かれており、Q25 は全体形状の 25%, Q40 は 40% を部分形状としている。Breuckmann は高精度なレンジスキャナである SmartScan Breuckmann Scanner によって得られたデータである。Kinect は低精度なレンジスキャナである Microsoft Kinect V2 sensor によって得られたデータである。Breuckmann と Kinect データは View1, View2, View3 の 3 つの視点が異なるグループに分かれている。

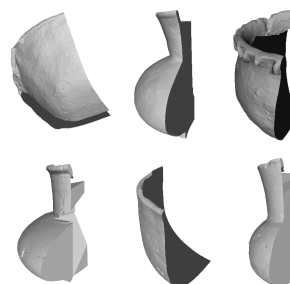


図 6 Artificial データの一例。断面が人工的にカットされている

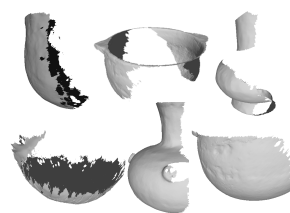


図 7 高精細な Breuckmann データの一例

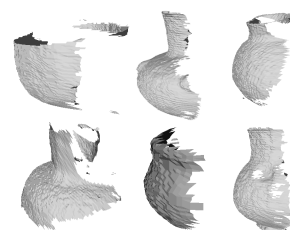


図 8 Kinect の一例。Breuckmann に比べてメッシュが粗い

4.2 評価尺度

評価尺度には Nearest Neighbor(NN), First Tier(FT), Second Tier(ST), Discounted Cumulative Gain(DCG) を用いた。NN は P@1 と呼ばれ、

$$\text{Nearest Neighbor (NN)} = \text{rel}(1)$$

で求められる。ここで $\text{rel}(k)$ は検索上位 k 件のうちに含まれる検索質問と同じクラスに属する 3 次元物体の数である。NN は最上位の検索結果の適合率を表す評価尺度である。

FT は R-precision と呼ばれる。FT, ST の計算式は以下のとおりである。

$$\text{First Tier (FT)} = \frac{\text{rel}(c-1)}{c-1}$$

$$\text{Second Tier (ST)} = \frac{\text{rel}\{2(c-1)\}}{c-1}$$

ここで c は検索質問と同じクラスに属する 3 次元モデルの数である。

DCG は順位を含めて正解データのランキングをどれだけ再現できるかをあらわした評価尺度である。値が大きくなるほど、正解データが検索の上位に出力されていることがわかる。計算式は以下のようになる。

$$\text{DCG}(i) = \begin{cases} G(1) & (i = 1) \\ \text{DCG}(i-1) + \frac{G(i)}{\log_2(i)} & (\text{otherwise}) \end{cases}$$

$$\text{DCG@N} = \frac{\text{DCG}(N)}{1 + \sum_{j=2}^N \frac{1}{\log_2(j)}}$$

ここで、 i は順位、 G は正解リスト、 N は 3 次元モデルの総数を表す。

4.3 比較手法

比較手法には提案手法のボクセルの生成方法を Voxnet に置き換えた手法を用いる。Voxnet では 2 値で表現された $32 \times 32 \times 32$ のボクセルを使用する。

4.4 3D CNN の学習方法

3D CNN の学習は検索対象となるデータを各クラスに分類する問題として行った。5-分割交差検証を行い、学習に用いたデータは検索性能のテスト時には使用しないようにした。5-分割交差検証では検索対象となるデータを 5 つのグループに分割し、1 つのグループを検証用、残りの 4 つのグループを学習用として用い、実験を行う。これを検証に用いるグループを変えながら 5 回繰り返し、それらの結果の平均を最終的な実験結果として用いる。

4.5 実験結果と考察

Artificial を用いた実験結果を表 1, 表 2 に示す。Breuckmann を用いた実験結果を表 3, 表 4, 表 5 に示す。Kinect を用いた実験結果を表 6, 表 7, 表 8 に示す。

実験結果を見ると条件によっては精度が下回ってしまうこともあったが多くの条件では Voxnet の精度を上回っていた。これは提案手法によってボクセルの持つ情報量が増えたためであると考えられる。

表 1 Artificial(Q25) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.15 | 0.19 | 0.40 | 0.56 |
| 提案手法 (局所角度特徴量) | 0.37 | 0.33 | 0.56 | 0.64 |
| 提案手法 (局所形状ペア特徴量) | 0.23 | 0.24 | 0.47 | 0.58 |

表 2 Artificial(Q40) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.50 | 0.37 | 0.59 | 0.68 |
| 提案手法 (局所角度特徴量) | 0.54 | 0.49 | 0.73 | 0.76 |
| 提案手法 (局所形状ペア特徴量) | 0.24 | 0.26 | 0.51 | 0.59 |

表 3 Breuckmann(View1) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.32 | 0.40 | 0.66 | 0.70 |
| 提案手法 (局所角度特徴量) | 0.58 | 0.50 | 0.76 | 0.77 |
| 提案手法 (局所形状ペア特徴量) | 0.54 | 0.48 | 0.72 | 0.74 |

表 4 Breuckmann(View2) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.36 | 0.40 | 0.68 | 0.71 |
| 提案手法 (局所角度特徴量) | 0.48 | 0.47 | 0.69 | 0.73 |
| 提案手法 (局所形状ペア特徴量) | 0.58 | 0.54 | 0.76 | 0.77 |

表 5 Breuckmann(View3) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.30 | 0.37 | 0.63 | 0.69 |
| 提案手法 (局所角度特徴量) | 0.28 | 0.24 | 0.48 | 0.60 |
| 提案手法 (局所形状ペア特徴量) | 0.34 | 0.28 | 0.54 | 0.63 |

表 6 Kinect(View1) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.42 | 0.41 | 0.69 | 0.73 |
| 提案手法 (局所角度特徴量) | 0.59 | 0.48 | 0.73 | 0.76 |
| 提案手法 (局所形状ペア特徴量) | 0.50 | 0.46 | 0.71 | 0.73 |

表 7 Kinect(View2) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.25 | 0.30 | 0.59 | 0.67 |
| 提案手法 (局所角度特徴量) | 0.44 | 0.38 | 0.63 | 0.70 |
| 提案手法 (局所形状ペア特徴量) | 0.43 | 0.41 | 0.69 | 0.72 |

表 8 Kinect(View3) の実験結果

| 手法 | NN | FT | ST | DCG |
|------------------|-------------|-------------|-------------|-------------|
| VoxNet | 0.34 | 0.37 | 0.65 | 0.69 |
| 提案手法 (局所角度特徴量) | 0.59 | 0.48 | 0.73 | 0.76 |
| 提案手法 (局所形状ペア特徴量) | 0.45 | 0.43 | 0.70 | 0.72 |

5. おわりに

本研究では局所特徴量を入力とした 3D CNN による形状類似検索手法を提案した。実験の結果、提案手法が 3D CNN を用いた手法である VoxNet より優れていることを示した。

今後の課題としては、ネットワーク構造の改良やほかのデータセットに対する実験、及びボクセル以外のデータ表現での 3 次元モデルの部分検索手法 [14] との比較などがあげられる。

謝 辞

本研究の一部は、科研費基盤 (B) (課題番号 17H01746) の支援を受けて遂行した。

文 献

- [1] Kinect. <https://en.wikipedia.org/wiki/Kinect>.
- [2] Paul Bratley and Bennett L. Fox. Algorithm 659: Implementing sobol's quasirandom sequence generator. *ACM Trans. Math. Softw.*, 14(1):88–100, March 1988.
- [3] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a "siamese" time delay neural network. In *Proceedings of the 6th International Conference on Neural Information Processing Systems, NIPS'93*, pages 737–744, San Francisco, CA, USA, 1993. Morgan Kaufmann Publishers Inc.
- [4] H. Dutagaci, A. Godil, C. P. Cheung, T. Furuya, U. Hillenbrand, and R. Ohbuchi. SHREC'10 Track: Range Scan Retrieval. In Mohamed Daoudi and Tobias Schreck, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2010.
- [5] T. Furuya, S. Kurabe, and R. Ohbuchi. Randomized sub-volume partitioning for part-based 3d model retrieval. In *Proceedings of the 2015 Eurographics Workshop on 3D Object Retrieval*, 3DOR, pages 15–22, Aire-la-Ville, Switzerland, Switzerland, 2015. Eurographics Association.
- [6] A. Godil, H. Dutagaci, B. Bustos, S. Choi, S. Dong, T. Furuya, H. Li, N. Link, A. Moriyama, R. Meruane, R. Ohbuchi, D. Paulus, T. Schreck, V. Seib, I. Sipiran, H. Yin, and C. Zhang. Range scans based 3d shape retrieval. In *Proceedings of the 2015 Eurographics Workshop on 3D Object Retrieval*, 3DOR, pages 153–160, Aire-la-Ville, Switzerland, Switzerland, 2015. Eurographics Association.
- [7] Maciej Halber and Thomas A. Funkhouser. Structured global registration of RGB-D scans in indoor environments. *CoRR*, abs/1607.08539, 2016.
- [8] Binh-Son Hua, Quang-Trung Truong, Minh-Khoi Tran, Quang-Hieu Pham, Asako Kanezaki, Tang Lee, HungYueh Chiang, Winston Hsu, Bo Li, Yijuan Lu, Henry Johan, Shoki Tashiro, Masaki Aono, Minh-Triet Tran, Viet-Khoi Pham, Hai-Dang Nguyen, Vinh-Tiep Nguyen, Quang-Thang Tran, Thuyen V. Phan, Bao Truong, Minh N. Do, Anh-Duc Duong, Lap-Fai Yu, Duc Thanh Nguyen, and Sai-Kit Yeung. RGB-D to CAD Retrieval with ObjectNN Dataset. In Ioannis Pratikakis, Florent Dupont, and Maks Ovsjanikov, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2017.
- [9] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. 2015.
- [10] Anestis Koutsoudis, George Pavlidis, Vassiliki Liami, Despoina Tsiafakis, and Christodoulos Chamzas. 3d pottery content-based retrieval based on pose normalisation and segmentation. *Journal of Cultural Heritage*, 11(3):329 – 338, 2010.
- [11] K. Lai, L. Bo, and D. Fox. Unsupervised feature learning for 3d scene labeling. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3050–3057, May 2014.
- [12] Daniel Maturana and Sebastian Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Pittsburgh, PA, September 2015.
- [13] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Trans. Graph.*, 21(4):807–832, October 2002.
- [14] I. Pratikakis, M. A. Savelonas, F. Arnaoutoglou, G. Ioannakis, A. Koutsoudis, T. Theoharis, M.-T. Tran, V.-T. Nguyen, V.-K. Pham, H.-D. Nguyen, H.-A. Le, B.-H. Tran, Q. H. To, M.-B. Truong, T. V. Phan, M.-D. Nguyen, T.-A. Than, K.-N. C. Mac, M. N. Do, A.-D. Duong, T. Furuya, R. Ohbuchi, M. Aono, S. Tashiro, D. Pickup, X. Sun, P. L. Rosin, and R. R. Martin. Partial shape queries for 3d object retrieval. In *Proceedings of the Eurographics 2016 Workshop on 3D Object Retrieval*, 3DOR '16, pages 79–88, Goslar Germany, Germany, 2016. Eurographics Association.
- [15] I Pratikakis, MA Savelonas, Fotis Arnaoutoglou, G Ioannakis, Anestis Koutsoudis, T Theoharis, MT Tran, VT Nguyen, VK Pham, HD Nguyen, et al. Shrec '16 track: Partial shape queries for 3d object retrieval. *Proc. 3DOR*, 1(8), 2016.
- [16] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *Proceedings of the 2009 IEEE International Conference on Robotics and Automation, ICRA'09*, pages 1848–1853, Piscataway, NJ, USA, 2009. IEEE Press.
- [17] J. Shotton, B. Glocker, C. Zach, S. Izadi, A. Criminisi, and A. Fitzgibbon. Scene coordinate regression forests for camera relocation in rgb-d images. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2930–2937, June 2013.
- [18] I. Sipiran, R. Meruane, B. Bustos, Tobias Schreck, H. Johan, B. Li, and Y. Lu. SHREC'13 Track: Large-Scale Partial Shape Retrieval Using Simulated Range Images. In Umberto Castellani, Tobias Schreck, Silvia Biasotti, Ioannis Pratikakis, Afzal Godil, and Remco Veltkamp, editors, *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2013.
- [19] Julien P. C. Valentin, Angela Dai, Matthias Nießner, Pushmeet Kohli, Philip H. S. Torr, Shahram Izadi, and Cem Keskin. Learning to navigate the energy landscape. *CoRR*, abs/1603.05772, 2016.
- [20] J. Xiao, A. Owens, and A. Torralba. Sun3d: A database of big spaces reconstructed using sfm and object labels. In *2013 IEEE International Conference on Computer Vision*, pages 1625–1632, Dec 2013.
- [21] Andy Zeng, Shuran Song, Matthias Nießner, Matthew Fisher, Jianxiong Xiao, and Thomas Funkhouser. 3dmatch: Learning local geometric descriptors from rgb-d reconstructions. In *CVPR*, 2017.
- [22] 青野雅樹 小林祐輝. 三次元モデル検索方法及び三次元モデル検索システム, March 2017. 特願 2017-029425, March 2017.
- [23] 立間淳司, 関洋平, 青野雅樹, and 大淵竜太郎. 多重フーリエスペクトル表現に基づく三次元モデルの形状類似検索. *電子情報通信学会論文誌. D, 情報・システム = The IEICE transactions on information and systems (Japanese edition)*, 91(1):23–36, jan 2008.