

見掛けに基づく特徴量を用いた 3 次元形状の形状類似検索

中澤 昌俊¹, 武井 毅², 南谷 高宏³, 大淵 竜太郎⁴
^{1,3,4} 山梨大学工学部コンピュータ・メディア工学科, ² 山梨大学大学院工学研究科
山梨県甲府市武田 4-3-11

1. はじめに

近年, コンピュータ技術の発達とともに, 3D ゲームや映画、テレビなどの 3 次元 CG, 3 次元形状 CAD が広く普及し, 3 次元モデルの数が増加している. それにともない, 3 次元モデルを対象とした類似検索の必要性が高まっている. 類似検索には, モデルに付加された名前または説明文を対象としたキーワードによる検索もあるが, 今後, モデルの形状そのものの類似性で検索する形状類似検索が重要となると予想される.

本論文では, 人間が視覚に基づいて形の類似性を判断する過程にヒントを得, ある 3 次元形状の「見かけ」, すなわち, レンダリングして得られる 2 次元画像を基に, 3 次元形状の類似検索を行う. 本手法の特長は, z 値のレンダリングが可能でさえあれば, 例えば VRML モデルのようにソリッドでなく, 多数の単連結要素を空なるモデルや, 面の向きが裏返ったモデルでも扱えることである.

本手法を実装して評価実験を行った結果, 同様のモデルを対象とした Osada らの手法¹⁾と比べ, 再現率, 適合率とも上回った.

以下, 2 章で手法を, 3 章で評価実験とその結果を, 述べ, 4 章でまとめと今後の課題を, それぞれ述べる.

2. 形状類似検索手法

2.1 形状類似検索手法の概要

本形状類似検索システムの処理の概要を図 1 に示す. 形状類似検索システムは, 選択された検索モデルに類似しているモデルをデータベースから検索する. まず, 前処理として, データベース内のモデルの特徴量を計算しておく. ユーザが示した検索モデルから特徴量を計算し, データベース内のモデルの特徴量と類似比較する. 結果として,

類似度の高い幾つかのモデルを似ているモデルとしてユーザに提示する.

本手法の特徴量は, モデルをレンダリングして得られる z 値, つまり, カメラからモデルまでの距離の 2 次元画像より計算する. 画像生成の際, 並行投影を仮定すると, カメラの位置で 3 自由度, 向きで 3 自由度ある. まず, カメラ(あるいは, 物体)の位置の 3 自由度は, 形状モデルを原点に置き, 大きさを正規化することで解決する. 物体の向きの 3 自由度は, 多数(数十個)の視点から画像を生成して近似することでその 2 つ(pitch, yaw)を, z 値画像の特徴量に回転不変特徴量を使用することで残りの 1 つ(roll)を, それぞれ解決する.

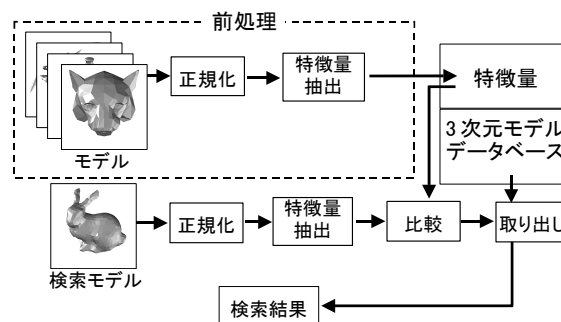


図 1. 検索システムの概要.

2.2 特徴量の抽出

特徴量を計算するのに先立ち, まず, モデルの位置と大きさを正規化する. まず, モデルの中心 G を全頂点の座標の中間値として求め(式(1)), G が原点(0,0,0)にくるようにモデルを平行移動する. ついで, 点 G から最も離れている頂点までの距離が 1 となるようモデルの大きさを正規化する.

$$G_p = \frac{1}{2} \left(\max_{0 \leq i < n} (p_i) + \min_{0 \leq i < n} (p_i) \right) \quad (1)$$

ここで, $p \in \{x, y, z\}$, n は頂点数.

ついで, (pitch, yaw)の自由度を離散的にサンプルした多数の視点から深さ(z 値)画像をレンダリングする(図 2). 今回, 我々は, 点 G からの距離が等しく, 互いにほぼ等間隔の 42 個の視点に原点 G に向けたカメラを置いて 128×128 画素の画像をレンダリングした. 42 個の視点は, 原点を中心とした正

*An appearance-based approach to shape similarity search of 3D models.

¹Masatoshi Nakazawa, ²Tuyoshi Takei, ³Takahiro Minamitani,

⁴Ryutarou Ohbuchi

^{1,3,4}Department of Computer Science and Media Engineering, University of Yamanashi

²Graduate School of Engineering, University of Yamanashi

20 面体を囲む球上にあり, その位置は 12 個の頂点, および, 各辺の中点の 30 個と原点 G を結ぶ直線がその球と交差する点とした. レンダリング結果の z 値は $[0,1]$ を取り, 距離が遠い程小さい. また背景の z 値は 0 とした.

最後の roll の自由度は, 得られた z 値の 2 次元画像の特徴量を求めて比較する際, 画像の原点廻りの回転に不変な特徴量を用いることで無視できる. 2 次元画像の類似比較を行うための特徴量として, 本論文では, Zhang らの手法²⁾を用いた. Zhang らの手法では, まず, z 値画像を極座標(r, θ)で表現し, r を x - y 直交座標系の y 軸に, θ を x 軸に, それぞれ写す変換(ここでは極座標変換と呼ぶ)を行う(図 3). これにより, 回転の自由度が, 平行移動の自由度に変換され, Fourier 変換の平行移動不変性を利用すれば, 全体として, 回転不変な特徴量を得ることができる. 後述の評価実験では, 128×128 の z 値画像を, r 軸を 64, θ 軸を 256 に離散化し, 極座標変換して 256×64 の z 値画像 $f(r, \theta)$ を生成した. 図 4 はウサギの z 値画像を極座標変換した例で, 色が黒いほど z 値が小さく, 視点に近い.

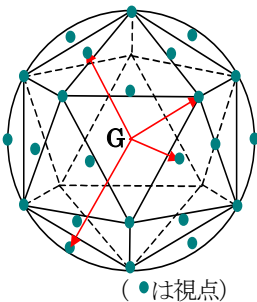


図 2. 視点の位置.

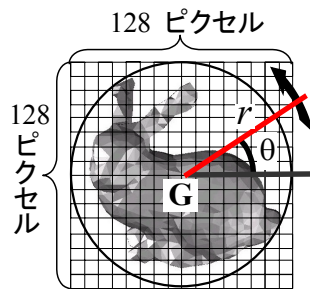


図 3. 極座標変換.



図 4. 極座標変換後の z 値の画像 $f(r, \theta)$.

次いで, この画像 $f(r, \theta)$ に対し式(2)のように Fourier 変換を施して得られた $PF(\rho, \phi)$ のスペクトル係数のうち低周波域の係数を比較したい画像の特徴量とする.

$$PF(\rho, \phi) = \sum_r \sum_\theta f(r, \theta) \exp \left[j 2 \pi \left(\frac{r}{R} \rho + \frac{\theta}{T} \phi \right) \right] \quad (2)$$

ここで, ρ は縦方向の空間周波数, ϕ は横方向の空間周波数である. $f(r, \theta)$ の画像が 256×64 画素なので, $0 \leq r < R = 64$, $0 \leq \theta < T = 256$, (r, θ は整数値) $0 \leq \rho < R$, $0 \leq \phi < T$ となる.

図 5 の(a)と(b)は, 同じ視点から見たときのモデルであり, (b)は(a)の視点の roll を 60 度回転させたモデルである. (c), (d)はそれぞれ極座標変換した z 値画像の Fourier 変換後のパワーを表わしている. 大きな値ほど高い輝度を割り当てている. 回転していても同じパワーを取り出すことができる. また, 低域にエネルギーが集中していることがわかる.

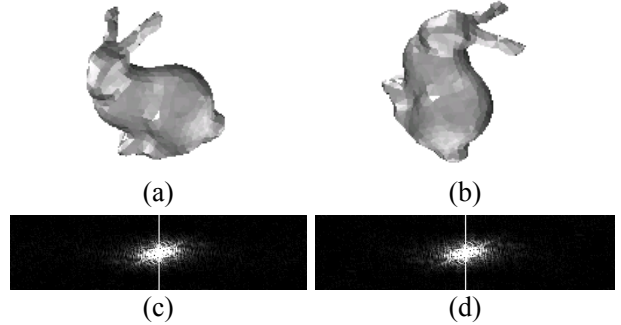


図 5. (a) ウサギのモデル, (b) ウサギのモデル(a)を 60 度回転; (c)ウサギのモデル(a)の Fourier スペクトル; (d) 回転したモデル(b)の Fourier スペクトル.

次いで, 式(3)のようにして振幅スペクトル(フーリエ係数の絶対値)の低周波成分のみを取り出す. これが物体をある方向(yaw と pitch)から見た場合の特徴量で, この特徴量を以後 *Depth Fourier Descriptor* (DFD)と呼ぶ.

$$DFD = \left\{ \frac{|PF(0,0)|}{area}, \frac{|PF(0,1)|}{|PF(0,0)|}, \dots, \frac{|PF(0,n)|}{|PF(0,0)|}, \dots, \frac{|PF(m,0)|}{|PF(0,0)|}, \dots, \frac{|PF(m,n)|}{|PF(0,0)|} \right\} \quad (3)$$

ただし, 面積 $area = 64 \times 256$, また m は, 縦方向の空間周波数のサンプル数, n は, 横方向の空間周波数のサンプル数である.

DFD において物体の形状特徴を表わすエネルギーは低周波に集中しており, 低周波域の係数のみで形状特徴の大部分を捉えることができる. 実験の結果, どの形状モデルについても, これを各種の方向からレンダリングした z 値画像において, 式(3)において m を 3, n を 15 とした計 4×16 個(全係数の $1/256$)の低周波成分係数で全パワーの約 80%を表現している.

ある 1 つの形状の特徴量は, 42 の向きから見た z 値画像に基づき, 従って 42 個の DFD の組からなる. 以後, これを *Multiple Orientation DFD* (MODFD)と呼ぶ.

2.3 モデル間の類似度計算と取り出し

2 つのモデル間の距離(相違度)は, MODFD より計算する. 距離が小さいほど似ているモデルであ

る. モデル X とモデル Y の MODFD をそれぞれ, $\mathbf{x}=(X_{ij}), \mathbf{y}=(Y_{ij})$ ($i=1,2,\dots,p, j=1,2,\dots,q$) とする. p は MODFD あたりの DFD の個数 (= 42), q は各 DFD 特徴ベクトルの次元 (= 64 = 4×16) である. X からモデル Y への距離 $D(\mathbf{x}, \mathbf{y})$ を求めるには, まず, $\mathbf{x}$ の i 番目の DFD と $\mathbf{y}$ の p 個の DFD の間で Manhattan 距離を求め (式(4)), さらにその最小値をとる. これを $i=1,2,\dots,p$ の p 組について行い, 得られた p 組の距離の最小値を平均した結果を X と Y 間の距離 $D(\mathbf{x}, \mathbf{y})$ とする (式(5)).

$$d(i, k) = \sum_{j=1}^q |X_{i,j} - Y_{k,j}| \quad (4)$$

$$D(\mathbf{x}, \mathbf{y}) = \frac{1}{p} \sum_{i=1}^p \left(\min_{1 \leq k \leq q} (d(i, k)) \right) \quad (5)$$

検索システムは, 提示したモデルとデータベース中の全モデルの距離を求め, 距離が小さい t 個をデータベースから取り出してユーザに提示する.

3. 評価実験

本論文で提案する手法を C++ を用いて実装し, 評価実験を行った. また, 比較のため, Osada らの手法¹⁾も実装した. ただし, Osada らの手法の実装の詳細には不明な点があり, 我々の実装とは異なる可能性がある. 実験には, VRML (Virtual Reality Modeling Language) モデルを用いた.

データベース中の 216 個のモデルは 32 個のカテゴリに分類し, 類似検索を行った. 例えば, Car のカテゴリには, 異なる形状の Car のモデルが 7 つ含まれていて, Chair のカテゴリには, 異なる形状の Chair のモデルが 9 個含まれている. 検索結果の例を図 6, 7, 8, 9, 10, 11 に示す. 各図の左上のモデルが検索モデルである. 検索モデルは, それぞれ “Car”, “Plane”, “Human”, “Chair”, “Ball” の各カテゴリ中のモデルの一つである. 検索の結果, 全モデル (216 個) の中で検索モデルにもっとも類似していると判断された上位 5 個のモデルを表示している. 図 6 の 2, 3 番目のモデルや, 図 7 の検索モデルや 2 番目のモデルなどでは, モデルの一部または全部が黒くなっているが, これは面が裏返っているためである. このように「整っていない」モデルでも, 本手法の検索には影響がない.

本手法が苦手なのは, 人間の (意味的) 判断では似ていても, 形やレンダリング結果の画像が異なる場合である. 例えば図 11 では窓ガラスの部分の無い自動車の外板のモデルで検索しており, 本手法ではいわゆる自動車のモデルが検索結果に現れない.

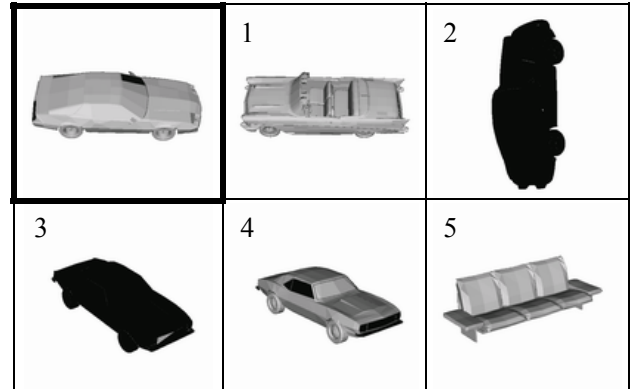


図 6. 検索モデルが“Car”の類似検索結果.

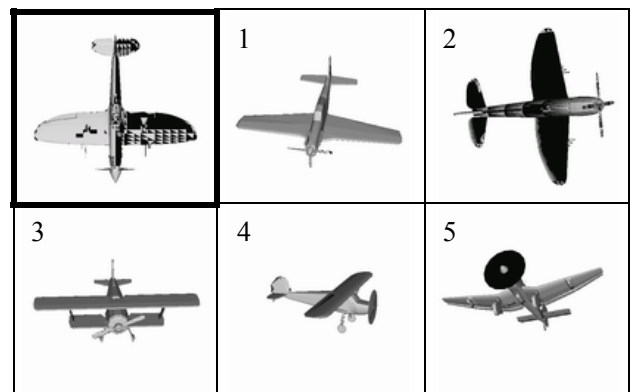


図 7. 検索モデルが“Plane”の類似検索結果.

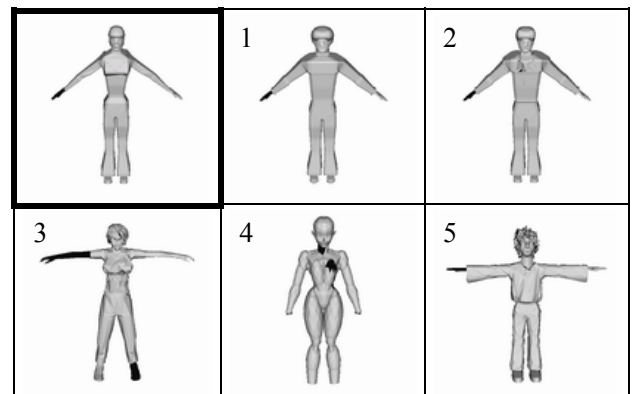


図 8. 検索モデルが“Human”の類似検索結果.

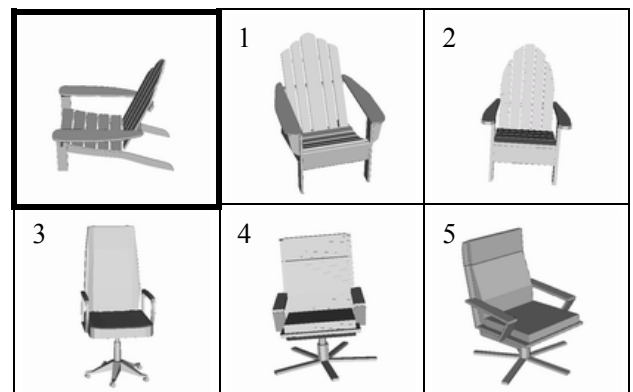


図 9. 検索モデルが“Chair”の類似検索結果.

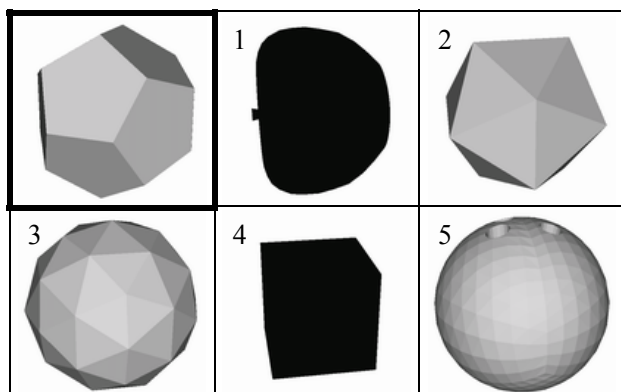


図 10. 検索モデルが“Ball”の類似検索結果.

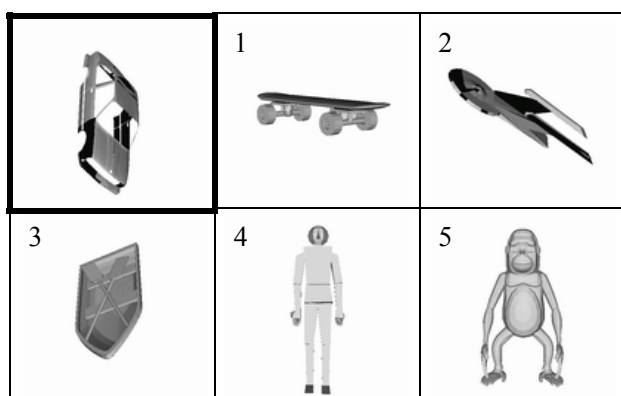


図 11. 検索モデルが“Car”の類似検索結果.

本手法の性能を再現率-適合率のグラフで表したものが図 12 である. プロットしたのはデータベース中の, 全カテゴリ(32)に分類したすべてのモデル(216)の再現率と適合率の平均である. 比較の対象としたのは Osada らの D2 手法¹⁾である.

再現率-適合率のプロットでは, カーブが右上に近づくほど性能が良い. 再現率と適合率は以下のように計算される. 検索モデルが含まれるカテゴリの数を C, 検索結果の上位 A 個 (A=1, 2, ..., 全てのモデルの数) のモデルの数が検索モデルと同じカテゴリに含まれている数を N とする.

$$\text{再現率} = \frac{N}{C}, \quad \text{適合率} = \frac{N}{A}$$

このグラフにおいて再現率が約 0.2 以下値をとらないのは, 一部のカテゴリに含まれるモデルの数が少ないためである. 例えば, あるカテゴリの中の全モデル数が 2 個の場合, 検索結果には自分(検索モデル)も含み, 1 位のモデルが自分となるため, 再現率 0.5, 適合率 1.0 となる. つまり 0.5 未満の再現率は現れない.

図 12 より, どの再現率-適合率の組み合わせにおいても本 MODFD 法の方が Osada らの D2 よりも類似検索の性能が高いことが読み取れる.

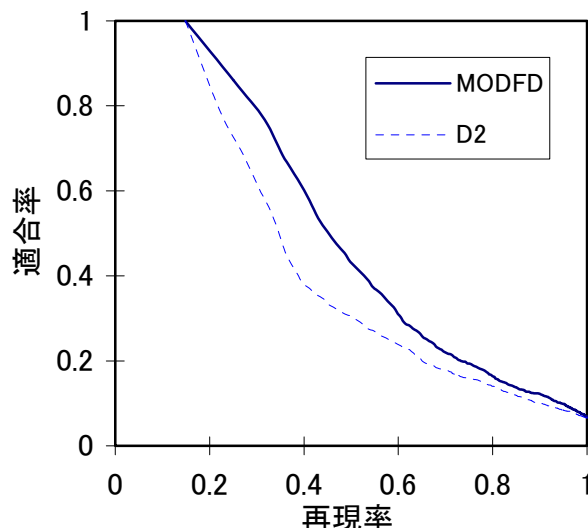


図 12. 再現率-適合率.

4. まとめと今後の課題

本論文では, 3 次元モデルの形状類似検索手法を提案し, その検索性能を評価した. 本手法では, 多方向から見たときの z 値画像を Fourier 変換し, 振幅スペクトルを特徴量とした. 特徴量同士の計算には, Manhattan 距離を用いて類似度を求めた. 評価実験より, 本手法の方が Osada らの D2 手法¹⁾よりも類似検索の性能が高いことがわかった.

今後の課題として, まず, 使用した種々のパラメタ(視点の数など)の最適値を求める必要がある. 同時に, 実験で使用するモデル数とカテゴリ数を増やし, より詳細な評価実験を行う必要もある. その上で, Osada らの D2 よりもさらに性能が高い Funkhouser らの新たな手法³⁾との性能比較も行いたい. また, 検索提示方式を拡張し, 2 次元の画像を元に検索できるようにする予定である. 本論文の手法では検索要求として 3 次元形状を用いたが, そのような 3 次元形状が無く, 手書きの 2 次元画像などを用いざるを得ない場合もある. そのような場合に対応できるよう, 本手法の特徴量が 2 次元画像に基づくことを利用した拡張を行う予定である.

参考文献

- [1] R. Osada, T. Funkhouser, Bernard Chazelle, and David Dobkin Shape Distributions, *ACM TOGS*, **21**(4), pp. 807-832, (October 2002).
- [2] D. S. Zhang, G. Lu, Shape-based image retrieval using generic Fourier descriptor, *Signal Processing: Image Communication*, **17**(10), pp. 825-848, (November, 2002).
- [3] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, D. Jacobs, A search engine for 3D models, *ACM TOGS*, **22**(1), pp. 83-105, (January, 2003).