

Web 3次元オブジェクトを用いた物体認識に関する検討

Object recognition using multiple 3D Web object

石川 雅浩† 下野 博†† 鈴木 崇之†† 古川 亮††† 川崎 洋† 河合由起子††

Masahiro Ishikawa Hiroshi Shimono Takayuki Suzuki Ryo Furukawa Hiroshi Kawasaki Yukiko Kawai

1. はじめに

近年、プロジェクトやカメラなどを組み合わせることで簡易に物体の3次元形状を計測する手法が数多く提案され、誰もが容易に高精度な3次元形状を取得できるようになりつつある[1]。さらに、3Dコンテンツをweb上に公開する方法も提案されている[2][3]。これに伴い、web上に公開された3次元形状に対する検索要求が高まることが予想される。そこで本論文では、web上にある3次元オブジェクトを高精度に検索する手法を提案する。

3次元オブジェクトを検索するためには、3次元オブジェクトの分類や認識が一般に必要である。例えば、3次元オブジェクトの認識に関しては、これまでCADやモデリングソフトで作成されたCGを対象として盛んに行われており、2006年からはコンテスト[4]が開催され認識精度が競い合われている。しかし、これらの手法は基本的には物体の全周形状が完全に獲得されることを前提としており、実環境において、全周形状を獲得できない状況については十分に考えられていない。また、実物体を計測した3次元形状についてもSpinImage[5]やCCHLAC(カラー立体高次局所自己相関)[6]を用いた認識法が提案されているが、具体的に同一物体を対象としているためwebに不特定多数アップされた3次元オブジェクトの検索にそのまま利用することはできない。

一方で、近年、高精度な画像検索法としてVisual Rankが提案されている[7][8]。これは、Googleの検索エンジンでも採用されているPage Rankの2次元への拡張である。Page Rankは、「多くの良質なページからリンクされているページは、良質なページである」という考えに基づいて設計されているが、Visual Rankも「良い画像は多くの画像と強く結びつき合っている」という仮説に基づいている。最大の特徴は画像特徴を検索基準に採用した点であり、従来のキーワード入力型画像検索エンジンよりも高精度な検索結果が得られる。そこで本研究では、Visual Rankの3Dデータ検索への拡張を目的としてShape Rankを提案する。本手法では、3Dデータの高精度な認識ではなく、3Dデータがweb上に数多く普及した状況においてユーザに高精度な検索結果を提示することを目的とする。具体的には以下の条件下での検索を前提とする。

まず、従来の画像検索同様にキーワード検索によって3次元オブジェクトを絞り込めるとする。これは、3次元オブジェクトが普及しweb上に公開される際にも画像情報同様に文字情報が記載される可能性が高いと考えられる

ことから妥当と考えられる。次に、キーワード検索によって得られる3次元オブジェクトの検索精度は画像検索と同程度の精度とする。これについては、前述の条件を満たしていれば3次元オブジェクトの内容に関係無くテキストデータから求められるため妥当と考えられる。次に、検索で得られた3次元オブジェクトに関しては、形状以外の情報は利用できないとする。これは、テキストチャなどの形状情報以外の情報が必ずしも適切に付与されていないためである。最後に、対象は全周形状に限らない。これは、計測システムが普及しweb上に実環境の3次元オブジェクトが広まった場合、1視点の形状のみで公開するケースも十分に考えられるためである。例えば、レーザーキャナで実物体を計測した距離画像データをオハイオ大学がデータベースとして公開しているが[9]、データベースとしての汎用性の観点から統合された全周データとしては公開されていない。

2. 提案手法

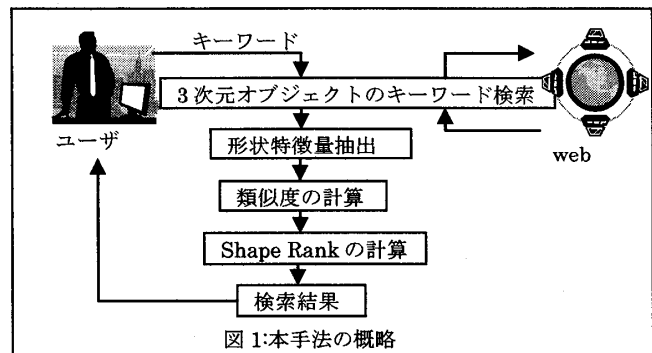


図1: 本手法の概略

本論文では、通常のVisual Rankに形状特徴抽出と類似度計算処理に修正を加えた図1のような手法を提案する。

まず、3次元オブジェクトをキーワード検索によって検索する。次に、3次元オブジェクトから形状特徴を抽出し、各3次元オブジェクト間の類似度計算を行う。最後に、3次元オブジェクトへの対応としてShape Rankを計算する。Shape Rankとは、Visual Rankの3次元オブジェクトへの拡張であり、3次元オブジェクトを扱うため、形状特徴を用いて視点依存性などを解消した手法となっている。以下に詳細を述べる。

3. 形状特徴量抽出

類似度計算を行うためには、各3次元オブジェクトの形状特徴量を抽出する必要がある。形状特徴量は、シンプルな方法としては距離値を用いた手法[10]やpixel depthを用いる手法などが提案されている。本研究では、視点依存性を避けるため不変特徴量として良く使われる曲率を用いることとした。

†埼玉大学

††京都産業大学

†††広島市立大学

3. 1 主曲率

3次元形状の曲率を計算する場合、主曲率を計算する。主曲率は、いくつかの計算方法が考えられるが本研究では2次元面当てはめによって求めることとした。当てはめる2次元面を、式(1)のような2次多項式で表す。

$$h(x, y) = a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6y^2 \quad (1)$$

この時、 \mathbf{a} は2次元面を求めるパラメータである。最適な \mathbf{a} は式(2)によって求まる。

$$\mathbf{a} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{h} \quad (2)$$

この時、 \mathbf{a} , \mathbf{h} , \mathbf{X} はそれぞれ式(3), 式(4), 式(5)のようになる。

$$\mathbf{a} = (a_1, \dots, a_6)^T \quad (3)$$

$$\mathbf{h} = (h_1, \dots, h_6)^T \quad (4)$$

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_N & y_N & x_N^2 & x_Ny_N & y_N^2 \end{pmatrix} \quad (5)$$

係数ベクトル \mathbf{a} から式(6)を用いて1次と2次の偏微分を求めることができる。

$$h_x = a_2, h_y = a_3, h_{xx} = 2a_4, h_{xy} = a_5, h_{yy} = 2a_6 \quad (6)$$

以上を用いて、最大曲率と最小曲率を求めることができる。

3. 2 shape index

主曲率を形状情報として扱う場合、ガウス曲率と平均曲率を用いて領域分割する手法や平均曲率が0であるような曲面を用いる手法など様々な扱い方が考えられる。本研究では、曲面変化の凹凸情報を強調した画像を用いることを目的として shape index 値を用いることとした。Shape index 値は式(7)を用いて求めることができる。

$$S_I = \frac{1}{2} - \frac{1}{\pi} * \arctan \frac{k_{max}(i) + k_{min}(i)}{k_{max}(i) - k_{min}(i)} \quad (7)$$

ここで、 k_{max} は最大曲率、 k_{min} は最小曲率である。この時、Shape index 値は[0, 1]の範囲で求まる。本研究では、Shape Rank として扱いやすいよう[0, 255]に正規化し、グレースケール画像に変換する。

4. Shape Rank

Shape Rank は基本的には Page Rank と同様の構造をしており、ハイパーリンクの代わりに各形状間の類似度を定義する。これによって、キーワード検索によって求めた形状ランキングの修正を行う。形状間の類似度がより高いものが上位にランキングされるため、キーワード検索によって集まった形状群の代表的な形状が上位に表示されることになる。これは、基本的には Visual Rank と同様のアプローチである。

4. 1 3D データ特徴量の記述方法

Shape Rank の計算において最も問題となるのが、3次元オブジェクトどうしの類似度を計算するための特徴量の記述方法である。本論文では、Shape index が不変3次元特徴量であることから、これを2次元的に配置した Shape index 画像を作成し、この Shape Index 画像に対して SIFT を適用し特徴抽出を行うものとした。ただし、Shape Index 自体は不変特徴量ではあるものの、Shape index 画像は、視点変化の影響を受けてしまうため、Shape Rank の計算においてはこれを解決する必要がある。以下にその手法を述べる。

4. 1. 1 距離画像

対象の3次元オブジェクトが1視点の場合(距離画像)でも、3次元の場合、見る方向により見え方が変化するため、Shape Index 画像1枚だけを用いた SIFT 特徴量は本来、望ましくない。しかし、1視点で計測される形状は、全体の一部分でしかなく、視点を少し変化させるだけで、その形状のほとんどが観測されなくなり、特徴点も非常に少ししか抽出されなくなる。そこで、本論文では、処理の簡便さを優先し、このような場合には、1枚の Shape Index 画像のみを用いることとした。

4. 1. 2 全周形状

全周形状の場合、視点変化により、観測可能な部分が減少することなく、その見え方のみが大きく変化する。これに対する単純な解決法として、複数視点からの Shape Index 画像を作成して用いることが考えられる。例えば、過去に、x, y, z 方向の正負、計6視点の距離画像から、SIFT 特徴量を算出する手法が提案されている[10]。しかし、6視点のみでは視点依存性の解消に十分とは言えないため、本論文では、SIFT 特徴量が30度程度の視点変化に対しては頑健であることを考慮して、図2のように正二十面体の12個頂点を用いるものとした。具体的には、正二十面体の12個の各頂点から12枚の shape index 画像を生成し、これに対して SIFT 特徴量を計算し、それらを1まとめにして3D オブジェクトの特徴量として用いた。

4. 2 類似度の計算

Shape Rank を計算するには、類似度を要素とする隣接行列 \mathbf{M} が必要である。隣接行列の類似度 C_{ij} を SIFT(Scale Invariant Feature Transform)を用いて画像間の対応点数を計算し、式(8)のように定義する[11]。

$$C_{ij} = \frac{M_{ij}}{(T_i + T_j)/2} \quad (8)$$

ここで、 M_{ij} は各画像間の対応点数、 T_i , T_j は各画像の SIFT キーポイントの数である。

ジオデシックドームの頂点を用いることで、均一な視点からの SIFT 記述子を獲得することができる。しかし、これらを Bags of Keypoints として用いる場合、1視点の3D オブジェクトと比較する際にキーポイントの数を考慮する必要がある。本研究では、1視点の Shape index 画像と全周形状を比較する場合、全周形状のある一面の SIFT 記述子と1視点の SIFT 記述子が一致すると考え、類似度計算の式(8)において、全周形状のキーポイント数を頂点数で割ることとした。

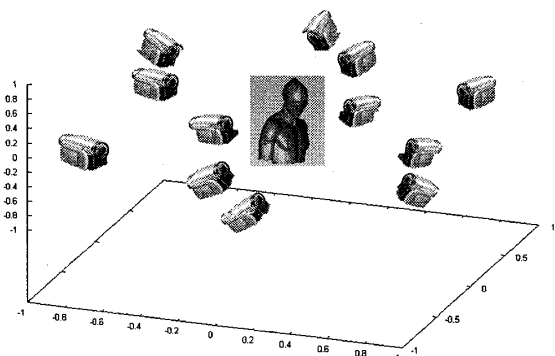


図2:ジオデシックドームの頂点

4.3 ランキング計算

隣接行列 M が求まれば、式(9)を用いてランキングが収束するまで計算を繰り返す。

$$r = (1 - \alpha)Mr' + \alpha P \quad (9)$$

これは、Page Rank と同じ方法で、 P は直接のつながりが無いケースを反映するためである。従って、 α はランキングベクトル r と P の重み付けである。この時、基本的には r は M の最大固有値の固有ベクトルである。

5. 実験と結果

5.1 実験環境

実験には、図3のようなプロジェクタ・カメラシステムによるレンジセンサを用いることとした。全周形状は、複数方向から対象オブジェクトを計測し、位置合わせ、形状統合を行い、3次元形状を獲得する。

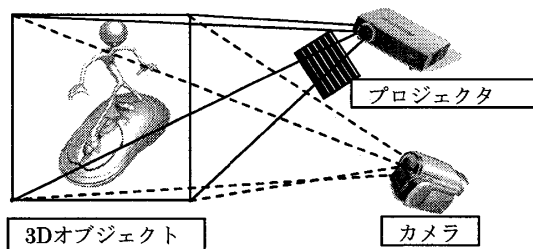


図3:プロジェクタ・カメラシステム

5.2 全周形状と複数視点オブジェクトのランキング

Shape Rank の動作確認として、同一オブジェクトの全周形状と様々な視点から撮影した距離画像群を用いてランキング計算を行った。用いた3Dオブジェクトは図4のような石膏像を用いた。全周形状としては、図5のような形状統合した3Dオブジェクトを用いている。距離画像群には、計測時のカメラの視線ベクトルからの形状(全14方向)を用いることとし、全周画像の際に計算した shape index 画像とは異なる視点の画像群を与えたことになる。用いた3次元オブジェクトの No. 1-14 が距離画像であり、No. 15 が全周形状からジオデシックドームによって SIFT 特徴量を算出したものである。表1にランキング結果を、図6に全周形状について上位にランキングされた距

離画像を示す。実験の結果、全周画像が1位にランキングされた。これは、別視点からの距離画像を用いたが、今回構築したジオデシックドームの頂点から計測した全周形状と類似形状が見つかったことを示している。2つの視点は一致していないものの正二十面体ジオデシックドームを用いたため視点のずれは高々30度程度であり、多くの SIFT に関する論文において、高い精度でのマッチングを保障する角度変化と一致しており妥当と言える。

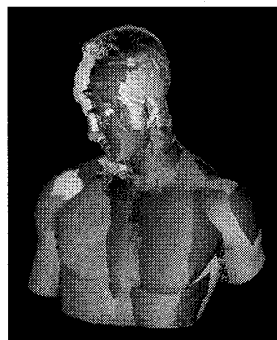


図4:石膏像

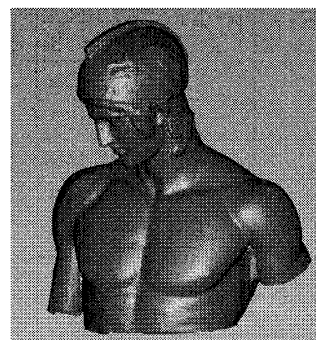


図5:統合結果

No	Rank
15	0.361162
5	0.067785
3	0.066091
2	0.056591
13	0.054775
12	0.05313
6	0.051408
14	0.047496
4	0.046444
1	0.042264
8	0.038725
11	0.03826
9	0.031863
10	0.028803
7	0.015201

表1:全周形状と複数距離画像

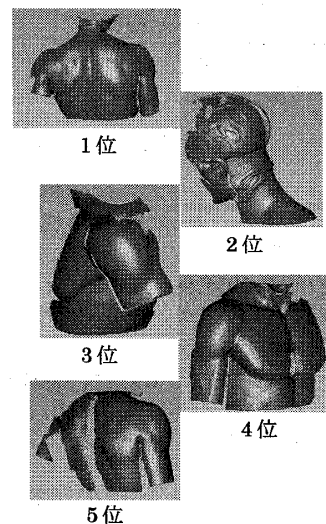


図6:ランキング結果(5位以上)

5.3 キーワード検索によるノイズの影響

画像のキーワード検索では、画像情報を加味していないためまったく関係ない画像が含まれている可能性がある。こういった状況下での動作を確認するため5.2節の実験に「No. 16:土偶」「No. 17:ウサギの人形」「No. 18:埴輪」「No. 19:カエルの玩具」の距離画像を加えて実験を行った。ランキング結果を表2に示す。また、ノイズとして与えたオブジェクトを図7に示す。表2から、石膏像の全周形状がノイズに影響されずに1位にランクインしていることが分かる。また、ノイズとして加えたオブジェクトは比較的下位にランクインされ期待通りの結果が得られた。

No	Rank
15	0.357152
5	0.067856
3	0.064232
2	0.060214
12	0.055671
13	0.055214
14	0.055133
6	0.050704
4	0.040552
8	0.035666
11	0.028942
16	0.023576
9	0.023458
18	0.022379
1	0.022282
10	0.020981
7	0.008823
19	0.00524
17	0.001923

表 2:検索結果にノイズを加えた場合

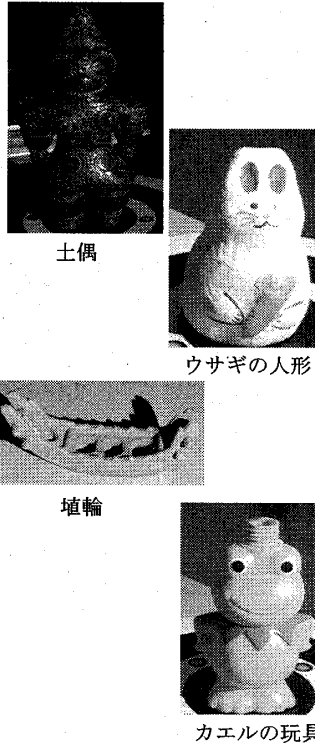


図 7:ノイズとして与えたオブジェクト

No	Rank
15	0.30334
20	0.134474
5	0.059998
3	0.056431
2	0.052133
13	0.047627
14	0.047554
12	0.046898
6	0.043207
4	0.036386
8	0.031182
11	0.025023
16	0.021589
9	0.020653
18	0.020277
1	0.019807
10	0.017668
7	0.008192
19	0.005505
17	0.002057

表 3:検索結果に全周形状のノイズを加えた場合

5. 4 他の全周形状がノイズとして含まれた場合

キーワード検索では、物体の形状を考慮しないため他のオブジェクトの全周形状が含まれることも十分に考えられる。そこで、5. 3 節の実験で用いた土偶の全周囲形状(No. 20:土偶全周)をノイズとして加えて実験を行った。ランキング結果を表3に示す。結果として、石膏像がやはり1位にランクインしたが、土偶の全周形状が2位にランクインされた。これは、全周形状の場合、部分形状よりも多くのキーポイントが検出されるため類似度が高くなる傾向が表れているものと考えられる。

6. まとめと今後の課題

本研究では、web上に公開された多数の3次元オブジェクトの高精度な検索法に関する検討として、形状情報を用いた Shape Rank の提案し、実験によりその有効性を確認した。Shape Rank とは、Visual Rank を3次元形状へと拡張したものであり、正二十面体の各頂点から見た Shape Index 画像を用いることで、視点依存性のない特徴抽出を実現する手法となっている。

今後の課題として、全周囲形状がノイズとして含まれた場合、上位にランクインしてしまう問題の解決が必要と考えられる。また、キーワードを用いた3次元オブジェクトの検索方法の実装も必要である。また、今回計算時間について考慮しなかったが、SIFT キーポイントから対応点を探索する際に全周形状の場合、キーポイントが多く隣接行列の計算に時間がかかってしまった。これについては PCA-SIFT などを用いて次元削減を行い高速化が必要と考えている。

謝辞

本研究の一部は、総務省戦略的情報通信研究開発制度(SCOPE)若手 ICT(072103013)の一環として実施されたものである。ここに記して謝意を表す。

参考文献

- [1]川崎 洋,大澤 裕,古川 亮,中村 泰明“空間コード化法を用いた未校正ステレオシステムによる密な3次元形状復元手法”,情報処理学会論文誌,Vo47,2006.
- [2]加納 利英,川嶋 紀弘,田澤 翔吾,成田 直樹,河合 由起子,川崎 洋“ユーザの選好に基づく3Dコンテンツの情報推薦と視点に依存した提示手法”,インタラクティオン2008,3,2008.
- [3]google O3D API “http://code.google.com/intl/ja/apis/o3d/”,2009.
- [4]SHREC “http://www.aimatshape.net/event/SHREC”,2009.
- [5]A.E.Johnson and M.Hebert, “Using spin images for efficient object recognition in cluttered 3D scenes”, IEEE Trans. Pattern Anal. And Mach. Intell.,21,pp.433-449,1999.
- [6]金崎 朝子,原田 達也,國吉 康夫“三次元環境地図からの物体探索タスク応用を目指したカラー立体高次局所自己相関特徴の開発”,第26回日本ロボット学会学術講演会,1L3-06,2008.
- [7]Y. Jing, and S. Baluja, “Visual Rank: Applying Page Rank to large-scale image search” IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.30, no.11, pp.1877-1890, Nov.2008.
- [8]Y. Jing, and S. Baluja, “Page Rank for product image search” Proceeding of the 17th international conference on World Wide Web, pp.307-316, Beijing, China, Apr.2008.
- [9]Ohaio data base, “http://sampl.ece.ohio-state.edu/”, 2009.
- [10]長田 邦男,坂野 智久,大淵 竜太郎“顕著度を考慮した多視点画像特徴量を用いた三次元形状類似検索”,情報科学技術レターズ(Information Technology Letters), Vol.6, pp.223-226, 2007.
- [11]安部 満,吉田 悠一“Visual Rank の多クラスへの拡張画像特徴量を用いた類似画像の自動分類とランキング付け手法” PRMU2008-178,2008.