

多視点画像からの複合特徴量に基づく三次元形状モデルの分類

奥村 泰明[†] 青野 雅樹[‡]

[†] 豊橋技術科学大学 情報工学系 〒441-8580 愛知県豊橋市天伯町 1-1

E-mail: [†] okumura@kde.cs.tut.ac.jp, aono@tut.jp

あらまし 近年、三次元モデルは幅広い分野で活用されている。これにともない、三次元物体の形状をもとにした分類技術の必要性が高まっている。本稿では、三次元物体を多視点から奥行き陰影法で投影した画像から複合特徴量を抽出し、カテゴリ分類の精度を調査した。提案する複合特徴量は二つの特徴量からなる。第一の特徴量は投影画像から計算した「Bag-of-Features (BoF) 特徴量」である。第二の特徴量は、三次元の深さ方向を考慮した「奥行き距離ヒストグラム特徴量」をベースとし、これにさらに主成分分析で次元削減したものを特徴量とした。これら二つの特徴量を足し合わせたものが、提案する複合特徴量である。比較実験の結果、既存特徴量である球面調和特徴量 (Spherical Harmonic Descriptor: SHD) を識別に用いた場合よりも提案手法が高い識別精度を得た。

キーワード 三次元形状分類, 形状特徴量, Bag-of-Features, SVM

Classification of 3D shape models based on composite features from Multiple-Viewpoints

Hiroaki OKUMURA[†] Masaki AONO[‡]

[†] Department of Information and Computer Science, Toyohashi University of Technology Hibarigaoka 1-1, Tenpaku-cho, Toyohashi-shi, Aichi, 441-8580, Japan

E-mail: [†] okumura@kde.cs.tut.ac.jp, aono@tut.jp

1. はじめに

近年、汎用コンピュータの性能向上などにより、産業・娯楽・医療などさまざまな分野で三次元モデルが使用されている。今後、文書・画像データなどと同様に、三次元モデルデータが、爆発的に増加することが予想される。三次元モデルを効率的に管理・利用するために、三次元モデルを予めカテゴリ分類しておくことが重要である。しかし、現在主に使用されている分類手法は、ユーザが手作業でモデルを分類する手法である。この方法では、ユーザがモデルのデータを分類する手間がかかり、分類も間違っている可能性がある。結果としてそのモデルは管理・利用が困難になってしまう。そのため、三次元モデルの形状を元にした分類技術の必要性が高まっている。

本研究では、三次元モデルの分類精度向上を目的として、三次元モデルを複数の視点から奥行き陰影投影した画像から複合特徴量を抽出し、カテゴリ分類を行った。これを提案手法とする。提案手法の有効性を確かめるため、既存の研究との比較を行った。それらの比較の結果、既存の特徴量を識別に用いた場合より提案する複合特徴量を識別に用いた場合に優れた識別精度を示した。

本論文の構成は以下の通りである。2節では、先行

研究について述べる。3節では、提案手法について述べる。4節では、提案手法と従来研究とを比較した評価実験の結果について述べる。5節では、まとめと今後の課題について述べる。

2. 関連研究

三次元モデルにおける特徴量の抽出アルゴリズムには、三次元物体の面上にランダムな点群を生成し、二点間の距離や角度をヒストグラムで表現するもの[2]、三次元物体をボクセル表現に変換することで立体的な特徴を解析するもの[1][3]、シルエット画像やDepth Buffer 画像など、2次元画像をレンダリングしてその上で画像から特徴量を抽出するもの[1]など様々な手法が提案されている。

Spherical Harmonics Descriptor(SHD)は、Kazhdanらにより提案された特徴量である[2]。SHDは、まず、三次元物体をボクセルに変換する。ボクセルとは、三次元物体を格子空間上に配置した時面と交わる格子を1、交わらない格子を0として、三次元物体の形状を直方体の集合で表したものである。次に、位置・大きさの正規化を行う。最後に、半径が[0,R]の範囲で同心球にボクセルを分解し、各同心球で球面調和関数変換を行う。得られたパワースペクトルの低周波成分を足しあ

わせたものを特徴量とする。

Bag-of-Features(Bag-of-Visual-Words と)は、画像における一般物体認識において広く用いられている手法で、画像を大量の局所特徴量の集合とみなし、局所特徴量のヒストグラムをその画像の特徴量とする手法である。局所特徴量には SIFT 特徴量[9]などが用いられることが多い。Bag-of-Features 表現を作成するには、まず、画像から局所特徴量を抽出する。次に、各画像から得られた局所特徴量の集合から代表点を抽出する代表点をまとめたものをコードブックと呼ぶ。コードブックを作成した後、各画像において各画像における局所特徴量群がどの代表点に近い投票を行い、そのヒストグラムを Bag-of-Features 表現とする。Bag-of-Features 表現の概要を図 1 に示す。

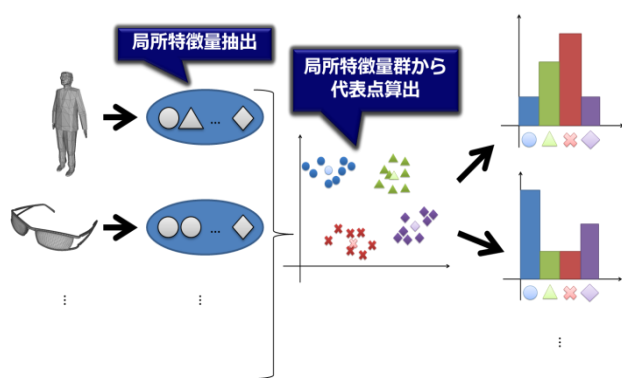


図 1: Bag-of-Features

Furuya らは、三次元モデルの周囲の複数視点から Depth Buffer 画像を生成し、その画像から局所特徴量としてグリッド上の特徴点から SIFT 特徴量を抽出し、その SIFT 特徴量の集合から Bag-of-Features 表現の特徴量を生成する手法 BF-DSIFT を提案した[4]。また、Lian らもほぼ同様の手法を提案した[5]。両者の主な違いは局所特徴量の集合から Bag-of-Features 表現をどのように構成するかである。Furuya らは Bag-of-Features 表現における出現頻度の投票を行うとき、全ての視点の画像から抽出した特徴点を用いる。Lian らの手法では視点毎に Bag-of-Features 表現を持つ。

奥行き距離ヒストグラム特徴量とは、画像の局所特徴から得られる距離ヒストグラムから得られる距離ヒストグラムの類似度を特徴量とした大域特徴量である[6]。

奥行き距離ヒストグラムは、画像を任意個のセルに分割し、セルごとに m 個のビンからなる輝度のヒストグラムを取る。このときヒストグラムは $L1$ 距離を用い正規化する。その後注目セルの輝度ヒストグラムと 8 近傍のセルの輝度ヒストグラムとの Bhattacharyya 距離による類似度を算出し、各セルとの距離を特徴量とする。Bhattacharyya 距離の式は以下の通りである。

$$d(p, q) = -\ln(BC(p, q))$$

$$BC = \sum_{i=1}^n \sqrt{p_i q_i}$$

これを奥行き距離ヒストグラム特徴量とする。この特徴量は注目セルに対する近傍セルの相対的な距離関係を表現した特徴量である。奥行き距離ヒストグラムの概要を図 2 に示す。

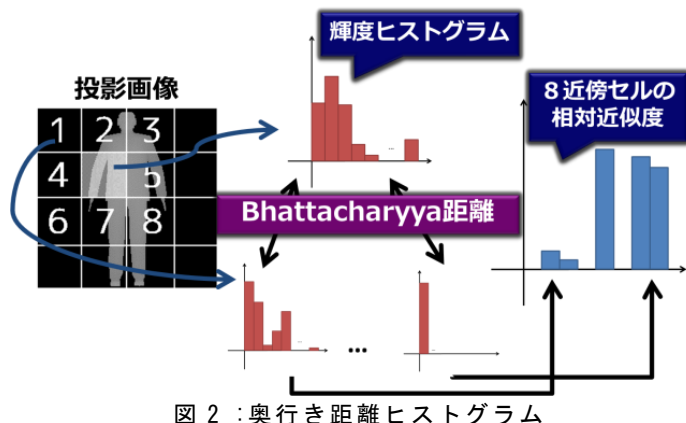


図 2 : 奥行き距離ヒストグラム

3. 複合特徴量

本節では、三次元物体を奥行き陰影法で投影した画像内で、局所特徴量のベクトル量子化特徴量である Bag-of-Features と、大域特徴量である PCA-距離ヒストグラムを特徴量とした複合特徴量の提案を行う。

3.1. 三次元モデルの正規化

三次元モデルは、位置・大きさ・向きに関して任意性が存在する。これらの任意性に対して立間らが提案した Point SVD[1]により正規化を行う。Point SVD は、三次元物体の面上にランダムに生成した点をサンプル点として、三次元モデルの姿勢正規化を行う。まず、サンプル点の平均を、三次元モデルの重心として、重心が三次元空間の原点となるように平行移動する。次に、サンプル値の特異値分解することで、三次元物体の主軸を求め、主軸が三次元空間の x 軸、 y 軸、 z 軸に沿うように回転する。最後に三次元物体の頂点と重心との最大距離で各頂点の座標値を割ることで三次元物体を単位球体に収める。

3.2. Depth Buffer 画像の生成

姿勢正規化を行った三次元モデルから、複数視点で Depth Buffer 画像を生成し、そこから局所特徴量を抽出する。ここでの Depth Buffer 画像とは、視点と三次元モデルの面との距離を輝度で表現したものであり、三次元物体の見た目の形状を表現する事ができる。

今回の提案手法では、姿勢正規化を行った三次元モデルに対して、 $+x$ 軸、 $-x$ 軸、 y 軸、 $-y$ 軸、 z 軸、 $-z$ 軸

の 6 方向から 256×256 の大きさで Depth Buffer 画像を生成する．Depth Buffer 画像の距離は[0,255]の値を取り，距離が小さいほど値は小さくなる．背景の値は 0 とした．

モデルの周囲から生成される Depth Buffer 画像の例を図 3 に示す．

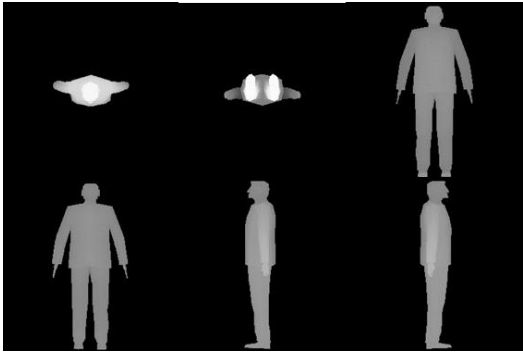


図 3:Depth Buffer 画像例

3.3. Bag of Features

まず，6 方向から投影した Depth Buffer 画像から局所特徴量である Dense-SIFT 特徴量を抽出する．次に Dense-SIFT 特徴量を用いて Bag-of-Features 表現を作成し，複合特徴量の一つとする．

通常の SIFT 特徴量では，画像の輝度が顕著に変化する点を検出し，その点を中心にパッチと呼ばれる小領域を切り出すが，Dense-SIFT ではこの点は一定の間隔で抽出する．そののちこのパッチから SIFT 記述子を計算したものが SIFT 特徴量となる．SIFT 特徴量は 1 つの点ごとに 128 次元の特徴量を計算する．今回の提案手法では，5 ピクセルごとに 32 ピクセルの大きさのパッチを切り出し，SIFT 記述子を計算した Dense-SIFT 特徴量を Bag-of-Features の計算に用いる．次に Dense-SIFT 特徴量を用い Bag-of-Features 表現を作成する．全ての特徴量を k 個の代表点にクラスタリングしコードブックを作成する．その後，モデルを対象とし，モデルを投影した画像 1 枚ごとに得られた局所特徴に対して，コードブック内の k 個のベクトルと最も近いものを選び投票を行う．最終的に各モデルに対し k 次元のヒストグラムができ，このヒストグラムを正規化したものをそのモデルの Bag-of-Features の特徴量となる．提案手法では k = 1000 とし，クラスタリングのアルゴリズムには k-means++[10]を用いた．

3.4.PCA 奥行き距離ヒストグラム特徴量の抽出

6 方向から投影した Depth Buffer 画像を用い，大域特徴量である奥行き距離ヒストグラム特徴量を抽出した後，主成分分析を用いて次元圧縮を行ったものを複合特徴量の一つとする．

奥行き距離ヒストグラム特徴量は，セルごとに 8 次元の特徴量を持つが，このセルごとの次元を n 次元に削減することで次元数の削減を行った．

本提案手法では，奥行き距離ヒストグラムのパラメータとして，輝度ヒストグラムのビン数を 16 とした．

提案する特徴量は，モデル周辺から投影した画像から (1)Bag-of-Features を用いた特徴量，(2)PCA-距離ヒストグラム特徴量の複合特徴量とする．

4. 実験と評価

評価実験では，既存特徴量として SHD を用い，提案手法との識別精度を比較する．なお，提案手法，SHD 共に識別器に SVM を用いている．SVM の実装には LibSVM[7]を用いた．

比較実験のベンチマークデータには NIST Shape Benchmark(NSB)[8]と呼ばれる三次元モデルデータベースを使用する．このデータベースには 40 のカテゴリが存在し，カテゴリごとに 20 のモデルが存在する．総モデル数は 800 である．

評価には識別精度を用いる．識別精度とは，正しいカテゴリに分類されたモデルの数をすべてのモデルの数で割ったものである．

識別の検証方法には 10-fold-cross-validation を用いた．

4.1. 単体特徴量での比較実験

まず，提案手法の複合特徴量を構成する Bag-of-Features 特徴量，奥行き距離ヒストグラム，奥行き距離ヒストグラム特徴量を主成分分析で次元削減した特徴量をそれぞれ単独で用いた場合での識別精度を調べる．

奥行き距離ヒストグラム，PCA 奥行き距離ヒストグラムはそれぞれ最も精度がよくなるようセル数を調整した結果である．識別精度の結果を表 1 に示す．奥行き距離ヒストグラム，PCA 奥行き距離ヒストグラムのセル数を変化させた時の詳しい結果をそれぞれ表 2, 3 に示す．表 1 より，奥行き距離ヒストグラムをそのまま用いるよりも次元削減したものを特徴量とし識別に用いると識別精度が良いことがわかる．

表 1:単独特徴量での識別精度

	特徴量		
	Bag-of-Features	奥行き距離ヒストグラム	PCA 奥行き距離ヒストグラム
識別精度 [%]	83.25	63.75	65.88

表 2:奥行き距離ヒストグラムのための識別精度

	セル数			
	4	16	64	256
識別精度 [%]	3.38	14.88	56.63	63.75

表 3:PCA 奥行き距離ヒストグラムのための識別精度

		セルあたりの削減後次元			
		1	2	3	4
セル数	4	4.88	4.88	5.00	5.13
	16	18.88	18.88	18.88	19.13
	64	60.38	58.25	57.63	58.25
	256	64.38	64.13	65.13	65.88

4.2. 複合特徴量での比較実験

先行研究である SHD を分類に用いた場合の識別精度と提案手法の識別精度の比較を行う。

両者の識別精度の結果を表 3 に示す。提案手法のパラメータはセル数 8, セルあたりの削減後次元を 1 とした。複合特徴量のパラメータによる識別精度の違いについて表 4 に示す。表 3 より, 既存特徴量を識別に用いるより提案する手法を識別に用いたほうが識別精度が良いことがわかる。表 4 より, セル数を大きくしすぎると識別精度が下がっていき, 削減後の次元も大きすぎると識別精度が下がってしまうことがわかる。これは, セル数, セルあたりの削減後次元を上げると特徴量の次元も上がり, 識別に悪影響を及ぼすためであると考えられる。

表 3:複合特徴量と SHD 特徴量の識別精度比較

	提案手法	SHD
識別精度 [%]	84.75	72.38

表 4:複合特徴量のパラメータを操作した場合の識別精度

		セルあたりの削減後次元			
		1	2	3	4
セル数	4	84.25	84.50	84.63	84.50
	16	84.50	84.50	84.50	84.50
	64	84.75	84.50	84.50	84.10
	256	84.12	83.25	82.25	82.00

5. まとめと今後の課題

本研究では三次元モデルを投影した画像内で, Bag-of-Features 特徴量と PCA 奥行き距離ヒストグラムを複合した特徴量を提案した。この複合特徴量を用いることにより, 既存の特徴量を用いて分類するより精

度の良い分類ができた。

今後の課題として, 異なるデータベースでの実験を行うことや, 複合特徴量の特徴量の改良, さらに他の特徴量との複合などを行うことでさらなる識別精度の向上を図ってゆく事が挙げられる。

参 考 文 献

- [1] 立間 淳司, 関 洋平, 青野 雅樹, 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, Vol. J91-D, No. 1, pp. 23-36, 2008.
- [2] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Matching 3d models with shape distributions. In Shape Modeling International, pp. 154-166. IEEE Computer Society, 2001.
- [3] Michael Kazhdan and Thomas Funkhouser. Harmonic 3d shape matching. In SIG-GRAPH 2002 Technical Sketches, p. 191, San Antonio, TX, 2002.
- [4] Takahiko Furuya and Ryutarou Ohbuchi. Dense sampling and fast encoding for 3d model retrieval using bag-of-visual features. In Proceedings of the ACM International Conference on Image and Video Retrieval, CIVR '09, pp. 26:1-26:8. ACM, 2009.
- [5] Zhouhui Lian, Afzal Godil, and Xianfang Sun. Visual similarity based 3d shape retrieval using bag-of-features. In Proceedings of the 2010 Shape Modeling International Conference, pp. 25-36, Washington, DC, USA, 2010. IEEE Computer Society.
- [6] 藤吉 弘亘, 局所特徴量の関連性に着目した Joint 特徴による物体検出, 情報処理学会 研究報告 CVIM 166, 43-54, 2009.
- [7] Chih-Chung Chang and Chih-Jen Lin, LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1-27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [8] Rui Fang, Afzal Godil, Xiaolan Li and Asim Wagan. A New Shape Benchmark for 3D Object Retrieval. 4th International Symposium, ISVC 2008, Las Vegas, NV, USA, December 1-3, 2008. Proceedings, Part I, pp 381-392, 2008
- [9] David G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal Computer Vision, Vol. 60, No. 2, pp. 91-110, November 2004.
- [10] David Arthur, Sergei Vassilvitskii. k-means++: the advantages of careful seeding. Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. pp. 1027-1035, 2009.