原著論文

多視点画像の局所的な特徴の相関に基づく 三次元物体の形状類似検索†

立間 淳司*1・青野 雅樹*2

近年、スマートフォンの普及やWebブラウザの機能向上により、三次元物体による豊かな表現を利用したアプリケーションやサービスが増加している。今後、文書・画像データのように、Web上に三次元物体データが爆発的に増加することが予想される。これにともない、大量にある三次元物体から、形状的特徴を手がかりに検索を行う、三次元物体の形状類似検索技術が必要とされている。本論文では、三次元物体から生成したDepth Buffer画像の局所特徴量の相関による、新しい三次元物体の特徴量Local Feature Correlation Descriptor (LCoD)を提案する。LCoDでは、まず、複数視点よりレンダリングしたDepth Buffer画像から局所特徴量を抽出する。そして、各Depth Buffer画像ごとに、局所特徴量の相関行列を求め、相関行列の要素を並べたものを、三次元物体の特徴量とする。LCoDは、Princeton Shape Benchmarkを用いた比較実験において、検索結果上位での適合率を表すFirst Tierが0.4708となり、Light-Field Descriptorなどの従来手法よりも優れた検索性能を示した。

キーワード：情報検索、特徴抽出、三次元物体、相関、データベース

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・映画・医療など様々な分野で、三次元物体が多く用いられている。Webの分野においても、三次元物体による豊かな表現を利用した、アプリケーションやサービスが増加しつつある。スマートフォンには、三次元コンピュータグラフィックスAPIであるOpenGL for Embedded Systems[11]が搭載され、拡張現実(Augmented Reality; AR)アプリケーションをはじめ、三次元物体が多く利用されている。また、Webブラウザ上で、三次元コンピュータグラフィックスを扱うための標準仕様であるWebGL[12]が策定され、多くのWebブラウザで、特別なプラグインをインストールすることなく、三次元物体を表示することができるようになった。このような状況から、今後、文書・画像データなどと同様に、三次元物体データが、Web上に爆発的に増加することが予想される。

爆発的に増加する三次元物体を、効率的に管理・利

用するために、三次元物体の形状そのものを手がかりに検索を行う、三次元物体の形状類似検索技術が注目されている[8, 29, 36]。大量に存在する三次元物体から、目的とする形状のものを検索できれば、既存の三次元物体を組み合わせ、新たな三次元物体を作成することも可能となり、制作コストの削減や効率化につながる[5]。

三次元物体の形状類似検索のプロセスは、文書や画像などの内容に基づく類似検索と同様である。図1は、一般的な三次元物体の形状類似検索システムの例である。システムは、利用者から検索質問として三次元物体が与えられると、まず、三次元物体の形状的特徴を捉えた特徴量を抽出する。次に、あらかじめ抽出しておいた、検索対象データベース内の特徴量との相違度を計算する。最後に、検索対象の三次元物体を、

図1 三次元物体の形状類似検索システムの概要

† 3D Object Retrieval based on Correlation of Multi-view Image Local Feature

Atsushi TATSUMA and Masaki AONO

*1 豊橋技術科学大学大学院電子・情報工学専攻
Department of Electronic and Information Engineering,
Toyohashi University of Technology

*2 豊橋技術科学大学情報・知能工学系
Department of Computer Sciences and Engineering, Toyohashi
University of Technology

相違度をもとに並び替え、検索結果として利用者に提示する。三次元物体の形状類似検索システムでは、三次元物体の形状的特徴を捉えた特徴量の抽出アルゴリズムが、重要なものとなる。

特徴量の抽出アルゴリズムには、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[10, 19, 33]、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[20, 21, 34]、シルエット画像や Depth Buffer 画像など、二次元画像をレンダリングして解析するもの[3, 6, 15]、三次元物体の接続情報からグラフ解析を行うもの[7, 24, 25]など様々な手法が提案されている。なかでも、複数視点でレンダリングした Depth Buffer 画像より抽出した局所特徴量から、画像の物体認識の研究で用いられる、Bag-of-Visual-Wordsと呼ばれる特徴量を抽出する手法が、優れた検索性能を得ている[6, 15, 23]。

Bag-of-Visual-Words(BoVW)は、統計的言語処理におけるBag-of-Wordsモデルのアナログであり、画像を局所特徴量の集合として考える[4, 28]。BoVWでは、画像データベースから抽出した大量の局所特徴量に対して、ベクトル量子化を行うことで得られる代表ベクトルを、文書の単語になぞらえて Visual Word と呼び、画像の特徴量を、この Visual Word の出現頻度によるヒストグラムで表す。BoVWは、局所特徴量の分布を、頻度という統計量で表したものであると言える。画像の物体認識の研究では、頻度のみならず、平均や分散などの高次統計量を考慮した手法も提案されている[22]

Furuyaら、Lianらは、BoVWを、三次元物体から生成した Depth Buffer 画像を用いて計算し、三次元物体の特徴量とする手法を提案した[6, 15]。BoVWによる特徴量は、優れた検索性能を得られるが、訓練データセットに含まれる三次元物体の形状により、検索性能が変化する問題がある[15]。そのため、訓練データセットに含まれる三次元物体が、検索対象となる三次元物体と、形状が類似したものでなければ、十分な検索性能を得ることができない。そこで、我々は、Visual Wordを使用せず、直接、局所特徴量のみから計算できる統計量を、三次元物体の特徴量とすることを考えた。

本論文では、三次元物体から生成した Depth Buffer 画像の局所特徴量の相関による、新しい三次元物体の特徴量 Local Feature Correlation Descriptor(LCoD)を提案する。LCoDでは、まず、BoVWによる手法と同様に、複数視点よりレンダリングした Depth Buffer 画像から局所特徴量を抽出する。そして、各 Depth

Buffer 画像ごとに、局所特徴量の相関行列を求め、相関行列の上三角要素を並べたものを、三次元物体の特徴量とする。LCoDは、BoVWによる手法と同様に、Depth Buffer 画像から抽出した局所特徴量の統計量による特徴量であるが、Visual Wordを必要としない。LCoDは、Princeton Shape Benchmark Test Sets[26]を用いた比較実験において、検索結果上位での適合率を表す First Tier で0.4708、検索結果全体の適合度を表す Document Cumulative Gainで0.7119となり、Light-Field Descriptor[3]やDESIRE Descriptor[32]といった従来手法よりも、優れた検索性能を示した。

以下に、本論文の構成を示す。第2節では、三次元物体の特徴量の関連研究について具体的に述べる。第3節では、提案手法のアルゴリズムについて詳細を述べる。第4節では、提案手法の有効性を確認するために行った比較実験について述べる。最後に、第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

三次元物体の形状類似検索では、三次元物体の特徴抽出アルゴリズムが、重要なものとなる。特徴量が、どのような三次元物体の形状的特徴を捉えているかは、検索結果や検索性能に大きな影響を与える。

三次元物体の特徴抽出において、問題となるのは、ISOの国際標準モデルとして普及している VRML (Virtual Reality Modeling Language) など、三次元物体データの多くが、ポリゴンスープモデルと呼ばれる、非連結な多角形の集合で定義されることである。そのため、三次元物体データに対して、直接、立体を前提とした解析手法を適用することは難しい。そこで、多くの三次元物体の特徴抽出アルゴリズムは、ポ

図2 三次元物体の特徴量抽出で一般的に用いられる形状表現 (a)ボクセル表現, (b)点群表現, (c) Depth Buffer 画像, (d)シルエット画像

リゴンスープモデルを、図2の様な、ボクセル表現やDepth Buffer画像など、別の形状表現に変換して解析を行う。

Osadaらが提案したShape Distribution D2(D2)は、三次元物体の面上にランダムな点を生成し、二点間の距離によるヒストグラムを、三次元物体の特徴量とする[21]。D2は、そのシンプルなアルゴリズムから、実装が容易であり、高速に計算することができるが、検索性能は低い。このD2に対して、Ohbuchiらは、法線で表される面の向き情報を加えた、二次元ヒストグラムによる特徴量Absolute Angle Distance (AAD)を提案した[20]。さらに、Wahlらは、AADと同様にして、曲率により表される幾何学的な情報を加えた、四次元ヒストグラムによる特徴量Surflet-Pair Relation Histograms (SPRH)を提案した[34]。

Kazhdanらが提案したSpherical Harmonic Descriptor (SHD)は、三次元物体を、立方体の集合で表したボクセル表現に変換し、球面調和関数変換から得られるスペクトルを特徴量とする[10]。SHDは、ボクセル表現に基づくアルゴリズムにおいて、優れた検索性能を得られることが知られている[31]。ボクセル表現に基づく特徴量は、その他に、Vranicらによる三次元フーリエ変換から得られるスペクトルを特徴量とするもの[33]、Novotniらによる三次元ツェルニケモーメントを特徴量とするものなどがある[19]。

Chenらが提案したLight-Field Descriptor (LFD)は、三次元物体から複数視点でシルエット画像を生成し、シルエット画像のフーリエスペクトルとツェルニケモーメントを特徴量とする[3]。Furuyaらは、複数視点からDepth Buffer画像を生成し、画像の物体認識の研究で用いられるBag-of-Visual-Words (BoVW) [4, 28]を抽出する手法BF-DSIFTを提案した[6]。同様にして、Lianらも、Depth Buffer画像からBoVWを抽出する手法CM-BOFを提案した[15]。BF-DSIFTが、三次元物体から複数視点でレンダリングしたDepth Buffer画像全てを用いてBoVWを計算するのに対して、CM-BOFでは、Depth Buffer画像ごとにBoVWを計算する

三次元物体の特徴量が、ボクセル表現やシルエット画像など、どの形状表現に基づくかにより、捉えられる形状的特徴が異なる[2]。これに対して、複数の特徴量を組み合わせる手法が、いくつか提案されている[13, 32, 37]。Vranicは、Depth Buffer画像・シルエット画像から抽出したフーリエスペクトルと、三次元物体の重心から面に向かって放射状に照射したベクトルからなるDESIRE Descriptorを提案した[32]。また、立間らは、Depth Buffer画像・シルエット画像・輪郭画

像・ボクセル表現より抽出したフーリエスペクトルからなるMulti-Fourier Spectra Descriptorを提案した[37]。

その他、三次元物体を、ボクセル表現やシルエット画像といった、別の形状表現に変換せず、面や頂点の接続情報を用いて、グラフ解析から特徴量を抽出する手法もある[7, 24, 25]。しかし、これらの手法は、正確な面や頂点の接続情報を前提とするため、本論文が対象とするポリゴンスープモデルには、適応することが難しい。

様々な特徴量のなかでも、近年、複数視点で生成したDepth Buffer画像から、BoVWを計算する手法が、優れた検索性能を得ている[23]。しかし、BoVWによる特徴量は、訓練データセットに含まれる三次元物体の形状により、検索性能が変化する問題がある[15]。そのため、訓練データセットに含まれる三次元物体が、検索対象となる三次元物体と、形状が類似したものでなければ、十分な検索性能を得ることができない。

本論文では、三次元物体から生成したDepth Buffer画像に基づく、新しい特徴量Local Feature Correlation Descriptor (LCoD)を提案する。LCoDは、複数視点で生成したDepth Buffer画像から抽出した局所特徴量の相関を、三次元物体の特徴量とする。BoVWによる手法との違いは、Visual Wordを使用せず、直接、局所特徴量から三次元物体の特徴量を計算することにある。

3. 局所的な特徴の相関に基づく三次元物体の形状特徴量

本論文では、Depth Buffer画像に基づく新しい特徴量Local Feature Correlation Descriptor (LCoD)を提案する。LCoDでは、まず、複数視点より生成したDepth Buffer画像から局所特徴量を抽出する。そして、各Depth Buffer画像ごとに、局所特徴量の相関行列を求め、それら行列の要素を並べたものを、三次元物体の特徴量とする。

3.1 三次元物体の姿勢正規化

三次元物体の特徴量抽出において問題となるのは、形状は同じであっても、三次元物体の位置・向き・大きさが、制作者やツールなどによって、異なる場合があることである。この問題に対して、立間らが提案したPoint SVDとNormal SVDを用いて、姿勢正規化を行う[37]。

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、三次元物体の姿勢正規化

を行う．まず，サンプル点の平均を，三次元物体の重心として，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，サンプル点を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内に納める．Normal SVDは，三次元物体の主軸を，特異値分解により求める際，サンプル点に法線ベクトルを用いる．以上から，三次元物体の位置・向き・大きさを一意に決定し，三次元物体の姿勢正規化を行う．

立間らは，Point SVDとNormal SVDで姿勢正規化した三次元物体それぞれで特徴量を抽出し，相違度が小さい方を，最適な姿勢正規化手法であるとした[37]．LCoDでも同様にして，姿勢正規化手法ごとに特徴量を計算し，相違度計算において，最適な姿勢正規化手法を選択する．

3.2 局所特徴量の抽出

姿勢正規化を行った三次元物体から，複数視点でDepth Buffer画像を生成し，局所特徴量を抽出する．Depth Buffer画像は，視点と三次元物体の面との距離を濃淡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ことができる．

LCoDでは，まず，姿勢正規化を行った三次元物体で，図3の様に，正八面体の頂点と辺の中点を視点として，計18方向から， 256×256 の大きさで，Depth Buffer画像を生成する．そして，各Depth Buffer画像から，局所特徴量としてDIFTを抽出する．DIFT(Dense SIFTやDense HoGとも呼ばれる)[14, 30, 35]は，一定の間隔と大きさで切り出したパッチから，SIFT記述子[16]を計算したものである．

SIFT記述子では，まず，切り出したパッチを， 4×4 の計16ブロックにグリッド分割し，ついで，ブロックごとに8方向の勾配方向ヒストグラムを作成する．勾配方向ヒストグラムを計算するために，切り出したパッチ $P(x, y)$ の勾配強度 $m(x, y)$ と勾配方向 $\theta(x, y)$ を，以下の式から求める．

$$m(x, y) = w(x, y) \cdot (L_x(x, y)^2 + L_y(x, y)^2)^{1/2}$$

$$\theta(x, y) = \tan^{-1} \frac{L_y(x, y)}{L_x(x, y)}$$

$$L_x(x, y) = P(x+1, y) - P(x-1, y)$$

$$L_y(x, y) = P(x, y+1) - P(x, y-1)$$

ここで， $w(x, y)$ はガウス窓である．勾配方向 $\theta(x, y)$ を， $\pi/4$ ずつ8方向に量子化し，輝度強度 $m(x, y)$ を

図3 正八面体の頂点と辺の中点による18視点 (a) 三次元物体と正八面体による18視点，(b) 正八面体の展開図と18視点の位置

重みとするヒストグラムを作成する．そして，ブロックごとに求めたヒストグラムを並べて，一つのベクトルとし，L2ノルムで正規化したものが，SIFT記述子である．

LCoDでは，Depth Buffer画像の局所特徴量として，4ピクセル間隔で， 58×58 の大きさのパッチを切り出し，SIFT記述子を計算したDIFTを用いる．

3.3 Local Feature Correlation Descriptor

各Depth Buffer画像から抽出した局所特徴量の相関行列を計算し，相関行列の上三角部分の要素を並べたベクトルを，Local Feature Correlation Descriptor (LCoD) とする．

正八面体の頂点と辺の中点を視点として，三次元物体からレンダリングした18枚のDepth Buffer画像を $I_1(x, y), \dots, I_{18}(x, y)$ とする．その内のある一枚のDepth Buffer画像 $I_i(x, y)$ から， n 個の D 次元の局所

特徴量 $\mathbf{v}_1^{(i)}, \dots, \mathbf{v}_n^{(i)}$ を抽出したとすると, Depth Buffer 画像 $I_i(x, y)$ における, 局所特徴量の相関行列 $R^{(i)}$ は, 以下の様になる.

$$R^{(i)} = \frac{1}{n} \sum_{j,k} \mathbf{v}_j^{(i)} \mathbf{v}_k^{(i)T}$$

相関行列 $R^{(i)}$ は, 対象行列であることから, Depth Buffer 画像 $I_i(x, y)$ における特徴量 $\mathbf{r}^{(i)}$ には, 上三角部分の要素のみを使用する.

$$\begin{aligned} \mathbf{r}^{(i)} &= [r_1^{(i)}, \dots, r_p^{(i)}] \\ &= [R_{1,1}^{(i)}, \dots, R_{1,D}^{(i)}, R_{2,2}^{(i)}, \dots, R_{2,D}^{(i)}, \dots, R_{D,D}^{(i)}] \end{aligned}$$

ここで, $p = D(D+1)/2$ である. 18枚の Depth Buffer 画像ごとに計算した, 相関行列による特徴量 $\mathbf{r}^{(i)}$ を並べたベクトル $\mathbf{f}$ を作成する.

$$\begin{aligned} \mathbf{f} &= [f_1, \dots, f_{18p}]^T \\ &= [\mathbf{r}^{(1)}, \dots, \mathbf{r}^{(18)}]^T \\ &= [r_1^{(1)}, \dots, r_p^{(1)}, \dots, r_1^{(18)}, \dots, r_p^{(18)}]^T \end{aligned}$$

相関行列によるベクトル $\mathbf{f}$ に対して, Power Normalization[22]を行う. 家具や機械部品など, 直線的な三次元物体の Depth Buffer 画像では, 勾配方向に偏りが生じるため, Depth Buffer 画像から抽出した SIFT 記述子は, スパースなものも多くある. そのため, 相関行列によるベクトルもスパースなものとなる. スパースなベクトル同士のユークリッド距離では, 相違が表れないことが知られている[18, 22]. そこで, ベクトルのスパースネスを緩和する働きがある Power Normalization を行う. 具体的には, ベクトルの各要素に対して, 以下の計算を行う.

$$\bar{f} = \text{sign}(f)|f|^\alpha$$

ここで, $\text{sign}(\cdot)$ は符号関数であり, $|\cdot|$ は絶対値を表す. α は任意の実数値をとるパラメータであるが, Perronnin らの知見から $\alpha=0.5$ とした[22]. 最後に, ベクトル全体を, L2 ノルムで正規化したものが, LCoD である.

3.4 特徴量間の相違度計算

LCoD の相違度には, ユークリッド距離を用いる. 二つの三次元物体から抽出した LCoD を, それぞれ, $\mathbf{a}, \mathbf{b}$ とすると, LCoD は L2 ノルムで正規化されたベクトルであることから, LCoD の相違度は以下の様になる.

$$\begin{aligned} d(\mathbf{a}, \mathbf{b}) &= \|\mathbf{a} - \mathbf{b}\|_2 \\ &= \left(\mathbf{a}^T \mathbf{a} + \mathbf{b}^T \mathbf{b} - 2\mathbf{a}^T \mathbf{b} \right)^{1/2} \\ &= \left(2 - 2\mathbf{a}^T \mathbf{b} \right)^{1/2} \end{aligned}$$

さらに, 前述したように, Point SVD と Normal SVD の姿勢正規化手法の選択を, 相違度計算にて行う. Point SVD と Normal SVD で姿勢正規化した三次元物体それぞれで LCoD を計算し, 相違度が小さい方を, その三次元物体にとって最適な姿勢正規化手法であるとする. 姿勢正規化に Point SVD を用いた LCoD を $\mathbf{a}^{(P)}, \mathbf{b}^{(P)}$, Normal SVD を用いた LCoD を $\mathbf{a}^{(N)}, \mathbf{b}^{(N)}$ とすると, 相違度は, 以下の様になる.

$$d(\mathbf{a}, \mathbf{b}) = \min \left(d(\mathbf{a}^{(P)}, \mathbf{b}^{(P)}), d(\mathbf{a}^{(N)}, \mathbf{b}^{(N)}) \right)$$

4. 比較実験

提案手法である Local Feature Correlation Descriptor (LCoD) の有効性を確認するために, 従来手法との比較実験を行った.

4.1 三次元物体ベンチマークと評価尺度

三次元物体ベンチマークには, Princeton Shape Bench-mark Test Sets (PSB)[26] と Engineering Shape Bench-mark (ESB)[9] を用いた.

PSB は, 自動車, 人間, 植物など, 特定の分野に偏らない 907 個の三次元物体が, 90 個のクラスに分類されている. 三次元物体の形状類似検索の研究において, 基本的な検索性能を評価するベンチマークとして用いられている. PSB での評価方法には, Leave-one-out 法を用いる. 907 個の三次元物体から 1 個を検索質問として選び出し, 残り 906 個を検索対象として検索することを, 全ての三次元物体が検索質問となるように繰り返し, 各検索質問ごとに評価尺度の値を計算する.

ESB は, パイプやベアリングなど機械部品の三次元物体 800 個が, 42 個のクラスに分類されている. 3DCAD における部品検索など, 機械設計分野への応用を想定した検索性能を評価できる. ESB の評価方法には, PSB と同様に, Leave-one-out 法を用いる.

PSB, ESB とともに, ベンチマーク全体の評価尺度の値には, 各検索質問の評価尺度の平均値によるマイクロ平均を用いる. クラスごとに評価尺度の平均値を計算し, それらの平均値を全体の評価尺度の値とするマイクロ平均では, 各クラスが少数の三次元物体で構成される場合に, 評価尺度の値に偏りが生じる[17]. 比較実験で使用する PSB と ESB は, 各クラスが少数の三次元物体で構成されるため, マイクロ平均を選択した.

検索性能の評価尺度には, 三次元物体の形状類似検索の研究において一般的に使用される, 再現率 (Recall) - 適合率 (Precision) 曲線, First Tier (FT), Second Tier (ST), Nearest Neighbor (NN), Discounted

Cumulative Gain (DCG), E-measure (E)を用いた[26].

Recallは検索結果の網羅性を, Precisionは正確性を表す値である. このRecallを横軸に, Precisionを縦軸にとりプロットしたものがRecall-Precision曲線である. RecallとPrecisionは, 一般的にトレードオフの関係にあり, Recall-Precision曲線が右上に向かって伸びるほど, 検索性能が優れていると言える.

FT, STは, それぞれ, R-Precision, Bull's Eye Performanceとも呼ばれる, 検索結果上位での適合率を表す値である. 検索質問が属するクラスに, 三次元物体が C 個含まれている場合, FTは上位 $C-1$ 位までに, STは上位 $2(C-1)$ 位までに, 検索質問と適合する三次元物体が, いくつ表れているかを示す値である.

NNは, P@1とも呼ばれ, 検索結果第1位が, 検索質問と適合する三次元物体であることを示す値である.

DCGは, 検索質問と適合した三次元物体が, 検索結果上位に表れているかを示す値である. DCGでは, 検索質問との適合度を点数に置き換え, 検索結果上位に表れるほど高い得点を, 下位に表れるほど低い得点を与え, その累積値を計算する. 三次元物体の形状類似検索の研究では, 検索結果最下位での値が使用されることから, 検索結果全体での適合度を示す値となる. Eは, RecallとPrecisionの調和平均であり, 三次元物体の形状類似検索の研究では, 上位32位でのRecallとPrecisionの値が使用される. Eは, 正確性と網羅性を総合的に表した値である.

FT, ST, NN, DCG, Eいずれも, 値が大きくなるほど, 検索性能が優れていると言える. これら評価尺度の詳細は, Shilaneらの研究で述べられている[26].

三次元物体の特徴量の従来手法には, ShapeDistributionD2(D2)[21], Surflet-Pair Relations Histogram (SPRH)[34], Spherical Harmonics Descriptor (SHD)[10], Light-Field Descriptor (LFD)[3], DESIRE Descriptor (DESIRE)[32]を用いた. これら, 従来手法の計算には, 著者らにより公開されている実行ファイルを用いた. さらに, LCoDと同様に抽出したDIFTから, Bag-of-Visual-Wordsを計算したDBoVWとも比較を行う. Visual Wordは, 訓練データセットにPrinceton Shape Benchmark Training Sets[26]を用いて, k -means法により生成した. Visual Wordの数は, 1,000個以上では, 増加させても検索性能に大きな影響がないことが知られている[15]. これは, DBoVWも同様であり, Lianらの知見から, DBoVWのVisual Wordの数は1,500個とした[15]. DBoVWは, LCoDと同様に, スパースなベクトルとなることから, Power Normalizationを行った後にL2ノルムで正規化した. DBoVWの相違度には, ユーク

リッド距離を用いた. DBoVWと比較することで, Visual Wordを使用せず, 局所特徴量の相関を三次元物体の特徴量とするLCoDの有効性を確認する.

4.2 Princeton Shape Benchmark

Princeton Shape Benchmark Test Sets (PSB)は, 特定の分野に偏らず907個の様々な形状の三次元物体が, 90個のクラスに分類されている. 三次元物体の形状類似検索の研究において, 基本的な検索性能を評価するベンチマークとして用いられる.

図4は, PSBにおける, 各特徴量の検索性能を, Recall-Precision曲線で表したものである. 全てのRecall-Precisionの組み合わせで, LCoDが, 従来手法を上回っていることがわかる. 有意差を確認するため, Recallの値が $\{0.0, 0.1, \dots, 1.0\}$ におけるPrecisionの平均値である11点平均適合率に対して, Holm法による多重比較を行った. 各手法のLCoDとの p 値は, DBoVWで0.0297, その他の手法では0.0001を下回る小さな値となった. いずれの p 値も, 有意水準の $p = 0.05$ を下回っており, LCoDは, 従来手法と比較して, 高い適合率を得られることがわかる.

表1は, PSBにおける, 各特徴量の検索性能を, FT, ST, NN, DCG, Eで表したものである. 各評価尺度の値は, マイクロ平均の値であり, 括弧内の値は標準偏差である. 全ての評価尺度で, LCoDが, LFDやDESIREなどの従来手法と比較して, 大きな値となっている. 有意差を確認するため, 各評価尺度に対して, Holm法による多重比較を行った.

FTにおける各手法のLCoDとの p 値は, DBoVWで0.028, その他の手法では0.0001を下回る値となった. また, STにおける各手法のLCoDとの p 値は, DBoVWで0.1193, その他の手法では0.0001を下回る

図4 PSBでのRecall-Precision曲線による検索性能比較

表1 PSBでのFT, ST, NN, DCG, Eによる検索性能比較

Method	FT	ST	NN	DCG	E
LCoD	0.4708(0.2874)	0.5792(0.3075)	0.7189(0.4496)	0.7119(0.2439)	0.3341(0.2221)
DBoVW	0.4390(0.2777)	0.5568(0.3004)	0.7144(0.4517)	0.6946(0.2397)	0.3219(0.2093)
DESIRE	0.4045(0.2803)	0.5133(0.3188)	0.6582(0.4743)	0.6631(0.2529)	0.2944(0.2051)
LFD	0.3790(0.2816)	0.4869(0.3160)	0.6582(0.4743)	0.6433(0.2552)	0.2790(0.2001)
SHD	0.3099(0.2543)	0.4127(0.3073)	0.5535(0.4971)	0.5846(0.2520)	0.2402(0.1889)
SPRH	0.2798(0.2302)	0.3802(0.2780)	0.5105(0.4999)	0.5551(0.2410)	0.2277(0.1821)
D2	0.1872(0.1967)	0.2723(0.2586)	0.3638(0.4811)	0.4622(0.2205)	0.1608(0.1537)

値となった。LCoDは、検索結果上位において、従来手法よりも、高い適合率を得られることがわかる。一方、DBoVWとは、同程度であることがわかる。

NNにおける各手法LCoDとの p 値は、DBoVWで1.0、DESIREで0.047、LFDで0.047、その他の手法では0.0001を下回る値となった。従来手法との p 値は、有意水準の $p=0.05$ を下回っているが、DBoVWとは同等となった。LCoDは、検索結果第1位での検索性能では、DBoVWと同等であると言える。

DCGにおける各手法LCoDとの p 値は、DBoVWで0.16841、その他の手法では0.0001を下回る値となった。Eにおける各手法LCoDとの p 値は、DBoVWで0.34917、その他の手法では0.0001を下回る値となった。DCGは検索結果全体での適合度を、Eは正確性と網羅性を総合的に表した値である。LCoDは従来手法と比較して、総合的に検索性能が優れていることがわかる。一方、DBoVWとは、同程度であることがわかる。

LCoDは、各評価尺度の値で、LFDやDESIREなどの従来手法を上回っており、有意な差が見られた。一方、DBoVWとの間には、FTでのみ有意な差が見られた。LCoDは、検索結果上位では、DBoVWを上回る適合率を得られるが、それ以外では、DBoVWと同等の検索性能であると言える。しかし、LCoDには、特徴量の抽出にVisual Wordを必要としないという利点がある。DBoVWの検索性能は、次節のEngineering Shape Benchmarkを用いた比較実験で示すように、Visual Wordの生成に使用する訓練データベースの内容に影響を受ける。LCoDは、Visual Wordを使用しないことから、訓練データベースを必要とせず、安定して従来手法よりも優れた検索性能を得られる。

表2は、各手法における、三次元物体1個あたりの特徴量計算と相違度計算の実行時間を表したものである。実験に使用したPCのスペックは、OSがWindows 7、CPUがIntel Core i5-3450 CPU 3.10GHz、メモリが16GBである。

特徴量計算では、LCoDは、DBoVWと比較して、

表2 特徴量計算と相違度計算の実行時間比較

Method	特徴量計算 [sec]	相違度計算 [sec]
LCoD	1.71667	0.03545
DBoVW	4.09919	0.02439
DESIRE	0.08042	0.02343
LFD	0.48198	0.04133
SHD	0.80974	0.03705
SPRH	0.08332	0.02619
D2	0.07284	0.02551

図5 ESBでのRecall-Precision曲線による検索性能比較

約1/2程度の実行時間で特徴量を計算できている。DBoVWは、ヒストグラムの生成において、全ての局所特徴量で、最近傍のVisual Wordを探索する必要があり、その処理に多くの時間を要する[30]。これに対して、LCoDは、Visual Wordを必要とせず、直接、局所特徴量同士の積のみで計算できる。一方、従来手法と比較すると、LCoD、DBoVWともに、多くの実行時間を要している。LCoD、DBoVWに共通して計算時間を必要とする処理は、局所特徴量のSIFT記述子の抽出である[30]。この問題に対しては、SIFT記述子の抽出処理にGPUを用いることなどが考えられる[27]。また、DBoVWでは、この他に前処理とし

図6 PSBにおいてLCoDとLFDでバイクを検索質問とした検索結果 (a)LCoD, (b)LFD

て、Visual Wordを生成するためのクラスタリング処理も必要となる。

相違度計算では、各手法で、実行時間に大きな差はない。従来手法のLFDは、複数生成したシルエット画像より抽出した特徴量ごとに相違度計算を行うため、その他の手法と比較して、わずかに、実行時間を要する。以上から、LCoDで、従来手法と同程度の実行時間で検索を行うには、局所特徴量の抽出処理を高速化する必要がある。

図6は、PSBにおいて、バイクの三次元物体を検索質問とした場合の、LCoDとLFDによる検索結果上位5件の三次元物体である。緑色のチェック印が付与された三次元物体は、検索質問に適合したものであり、赤色のバツ印が付与された三次元物体は、検索質問に不適合なものである。LCoDが、同じ複数視点の二次元画像にもとづく手法であるLFDと比較して、検索質問に適合する三次元物体を検索できていることがわかる。

4.3 Engineering Shape Benchmark

Engineering Shape Benchmark (ESB)は、機械部品の三次元物体800個が、42個のクラスに分類されている。3DCADにおける部品検索など、機械設計分野への応用における検索性能を評価することができる。

図5は、ESBにおける、各特徴量の検索性能を、

Recall-Precision曲線で表したものである。全てのRecall-Precisionの組み合わせで、LCoDが、従来手法を上回っていることがわかる。有意差を確認するため、Recallの値が $\{0.0, 0.1, \dots, 1.0\}$ におけるPrecisionの平均値である11点平均適合率に対して、Holm法による多重比較を行った。各手法のLCoDとの p 値は、DBoVWで0.0002、その他の手法では0.0001を大きく下回る値となった。いずれも有意水準の $p=0.05$ を下回っていることから、LCoDは、機械部品を対象とした検索において、従来手法よりも高い適合率を得られることがわかる。

表3は、ESBにおける、各特徴量の検索性能を、FT, ST, NN, DCG, Eで表したものである。各評価尺度の値は、マイクロ平均の値であり、括弧内の数値は標準偏差である。LCoDは、全ての評価尺度で、最も大きな値となっている。各評価尺度での有意差を確認するため、Holm法による多重比較を行った。

FTにおける各手法のLCoDとの p 値は、DBoVWで0.00434、その他の手法では0.0001を大きく下回る値となった。STにおける各手法のLCoDとの p 値は、いずれの手法も0.0001を大きく下回る値となった。いずれも有意水準の $p=0.05$ を下回っていることから、LCoDは、検索結果上位において、従来手法やDBoVWよりも、高い適合率を得られることがわかる。

表3 ESBでのFT, ST, NN, DCG, Eによる検索性能比較

Method	FT	ST	NN	DCG	E
LCoD	0.5334(0.2682)	0.6928(0.2787)	0.8888(0.3144)	0.8128(0.1813)	0.4384(0.2072)
DBoVW	0.4891(0.2585)	0.6308(0.2796)	0.8788(0.3264)	0.7860(0.1853)	0.4067(0.2075)
DESIRE	0.4522(0.2616)	0.5933(0.2768)	0.8600(0.3470)	0.7715(0.1883)	0.3740(0.1906)
LFD	0.4382(0.2617)	0.5841(0.2866)	0.8562(0.3508)	0.7566(0.1932)	0.3565(0.1851)
SHD	0.4397(0.2576)	0.5803(0.2832)	0.8350(0.3712)	0.7558(0.1916)	0.3660(0.1887)
SPRH	0.4434(0.2655)	0.5681(0.2940)	0.8750(0.3307)	0.7598(0.1903)	0.3638(0.2095)
D2	0.3454(0.2115)	0.4714(0.2607)	0.7937(0.4046)	0.6911(0.1918)	0.2901(0.1690)

図7 ESBにおいてLCoDとLFDでポスト部品を検索質問とした検索結果(a)LCoD, (b)LFD

NNにおける各手法のLCoDとの p 値は, DBoVWで1.0, DESIREで1.0, LFDで0.7663, SHDで0.0349, SPRHで1.0, D2で0.0001を下回る値となった. NNは, 検索結果第1位が, 検索質問と同じクラスに属する三次元物体であるかを表す評価尺度である. 検索結果第1位では, LCoDは, DESIREやLFDといった従来手法と, 同程度の検索性能であることがわかる.

DCGにおける各手法LCoDとの p 値は, DBoVWで0.04169, その他の手法では0.0001を下回る値となった. Eにおける各手法LCoDとの p 値は, DBoVWで0.00783, その他の手法では0.0001を下回る値となった. LCoDは, 従来手法やDBoVWよりも, 総合的に高い検索性能を得られることがわかる.

LCoDは, ESBにおいても, 従来手法よりも優れた検索性能を得られることがわかった. また, PSBの場合と比較して, LCoDとDBoVWとの各評価尺度の値の差が大きくなっている. DBoVWでは, Visual Wordを生成するために, 訓練データセットを必要とするが, BoVWによる手法では, 訓練データセットの内容により, 検索性能が変化する問題がある[15]. 実験で訓練データセットに使用したPrinceton Shape Benchmark Train Setsには, 機械部品モデルが含まれていないため, PSBの場合と比較して, 各評価尺度の値が低下したと考える.

ESBの他に, 一般に公開されている, 機械部品形状の三次元物体データセットは存在しないため, ESBに適した訓練データセットで, Visual Wordを生成することができなかった. このように, 検索対象となる三次元物体と, 形状が類似した三次元物体による訓練データセットが用意できない場合に, BoVWによる手法では, 十分な検索性能を得ることができない. 一方, LCoDは, Visual Wordを使用しないことから, 訓練データセットを必要とせず, このような問題は発生しない. LCoDは, BoVWによる手法と比較して,

特定の分野に偏った特殊な三次元物体データセットにおいても, 安定して優れた検索性能を得ることができる.

図7は, ESBにおいて, ポスト部品の三次元物体を検索質問とした場合の, LCoDとLFDによる検索結果上位5件の三次元物体である. 緑色のチェック印が付与された三次元物体は, 検索質問に適合したものであり, 赤色のバツ印が付与された三次元物体は, 検索質問に不適合なものである. LCoDと同様に, 複数視点の二次元画像にもとづく特徴量のLFDが, 検査結果上位三位までに, 検索質問に似て非なる三次元物体を提示しているのに対して, LCoDは, 検索結果上位5件全てで, 検索質問に適合した三次元物体を提示できている. LCoDは, LFDと比較して, 機械部品の三次元物体の形状類似検索において, 優れた検索性能であることがわかる.

5. おわりに

本論文では, 三次元物体から生成したDepth Buffer画像の局所特徴量の相関による, 新しい三次元物体の特徴量Local Feature Correlation Descriptor(LCoD)を提案した. LCoDでは, まず, 複数視点で生成したDepth Buffer画像から局所特徴量を抽出する. そして, 各Depth Buffer画像ごとに, 局所特徴量の相関行列を求め, 相関行列の要素を並べたものを, 三次元物体の特徴量とする.

LCoDは, Princeton Shape Benchmark Test Sets[26]を用いた比較実験において, 検索結果上位での適合率を表すFirst Tierで0.4708, 検索結果全体の適合度を表すDocument Cumulative Gainで0.7119となり, Light-Field Descriptor[3]やDESIRE Descriptor[32]といった従来手法よりも, 優れた検索性能を示した.

LCoDは, 従来手法よりも高い検索性能を得られるが, 局所特徴量の相関行列を三次元物体の特徴量とす

るため、次元数が大きくなる問題がある。LCoDでは、18方向からレンダリングしたDepth Buffer画像から、局所特徴量としてSIFT記述子を抽出した。SIFT記述子の次元数は128次元であることから、LCoDの次元数は $18 \times (128 \times (128+1)) / 2 = 148,608$ 次元と、非常に大きなものとなる。今後の課題として、LCoDの次元数を抑えるために、SURF記述子[1]など、SIFT記述子よりも次元数の小さい局所特徴量を用いることや、主成分分析で次元削減を行うことなどが考えられる。

謝辞

本研究は日本学術振興会科学研究費補助金(特別研究員奨励費)の助成、総務省戦略的情報通信研究開発推進制度(SCOPE)地域ICT振興型研究開発(112306001)、ならびに科学研究費基盤研究(C)23500119の支援を受けて行われた。

参考文献

- [1] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features(surf). *Comput. Vis. Image Underst.*, Vol.110, No.3, pp.346-359, June 2008.
- [2] Benjamin Bustos, Daniel A. Keim, Dietmar Saupe, Tobias Schreck, and Dejan V. Vranic. Using entropy impurity for improved 3d object similarity search. In *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo, ICME '04*, pp.1303-1306, 2004.
- [3] Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen, and Ming Ouhyoung. On visual similarity based 3d model retrieval. *Computer Graphics Forum*, Vol.22, No.3, pp.223-232, 2003.
- [4] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cedric Bray. Visual categorization with bags of keypoints. In *Proceedings of the Workshop on Statistical Learning in Computer Vision, ECCV '04*, pp.1-22, 2004.
- [5] Thomas Funkhouser, Michael Kazhdan, Philip Shilane, Patrick Min, William Kiefer, Ayellet Tal, Szymon Rusinkiewicz, and David Dobkin. Modeling by example. In *ACM SIGGRAPH 2004 Papers, SIGGRAPH '04*, pp.652-663, New York, NY, USA, 2004. ACM.
- [6] 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.
- [7] Masaki Hilaga, Yoshihisa Shinagawa, Taku Kohmura, and Tosiya L. Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp.203-212, 2001.
- [8] Natraj Iyer, Subramaniam Jayanti, Kuiyang Lou, Yagnanarayanan Kalyanaraman, and Karthik Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, Vol.37, No.5, pp.509-530, April 2005.
- [9] Subramaniam Jayanti, Yagnanarayanan Kalyanaraman, Natraj Iyer, and Karthik Ramani. Developing an engineering shape benchmark for cad models. *Computer-Aided Design*, Vol.38, No.9, pp.939-953, 2006.
- [10] Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *Proceedings of the 2003 Eurographics Symposium on Geometry Processing, SGP '03*, pp.156-164, 2003.
- [11] Khronos Group. Opengl-es-the standard for embedded accelerated 3d graphics, 2007. <http://www.khronos.org/opngles/>.
- [12] Khronos Group. WebGL-opengl es 2.0 for the web, 2011. <http://www.khronos.org/webgl/>.
- [13] Bo Li and Henry Johan. 3d model retrieval using hybrid features and class information. *Multimedia Tools and Applications*, pp.1-26, 2011.
- [14] Fei-Fei Li and Pietro Perona. A bayesian hierarchical model for learning natural scene categories. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2 of CVPR '05, pp.524-531, Washington, DC, USA, 2005. IEEE Computer Society.
- [15] 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.
- [16] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal Computer Vision*, Vol.60, No.2, pp.91-110, November 2004.
- [17] Patrick Min. *A 3D Model Search Engine*. PhD thesis, Princeton University, 2004.
- [18] David Nister and Henrik Stewenius. Scalable recognition with a vocabulary tree. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2 of CVPR '06, pp.2161-2168, Washington, DC, USA, 2006. IEEE Computer Society.
- [19] Marcin Novotni and Reinhard Klein. 3d zernike descriptors for content based shaperetrieval. In *Proceedings of the eighth ACM symposium on Solid modeling and applications*, pp.216-225, 2003.
- [20] Ryutarou Ohbuchi, Takahiro Minamitani, and Tsuyoshi Takei. Shape-similarity search of 3d models by using enhanced shape functions. In *Proceedings of the Theory and Practice of Computer Graphics 2003, TPCG '03*, pp.97-104, Washington, DC, USA, 2003. IEEE Computer Society.
- [21] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Transactions on Graphics*, Vol.21, pp.807-832, 2002.

- [22] Florent Perronnin, Jorge Sánchez, and Thomas Mensink. Improving the fisher kernel for large-scale image classification. In *Proceedings of the 11th European Conference on Computer Vision: Part IV*, ECCV '10, pp.143–156, Berlin, Heidelberg, 2010. Springer-Verlag.
- [23] Vanamali Thiruvadandam Porethi, Afzal Godil, Helin Dutagaci, Takahiko Furuya, ZhouhuiLian, and Ryutarou Ohbuchi. Shrec '10 track: Generic 3d warehouse. In *Proceedings of Eurographics Workshop on 3D Object Retrieval 2010*, pp.93–100, 2010.
- [24] Raif M. Rustamov. Laplace-beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, SGP'07, pp.225–233, 2007.
- [25] Konstantinos Sfikas, Ioannis Pratikakis, and Theoharis Theoharis. Contopo: Non-rigid 3d object retrieval using topological information guided by conformal factors. In *Proceedings of Eurographics Workshop on 3D Object Retrieval 2011*, pp.25–32, 2011.
- [26] Philip Shilane, Patrick Min, Michael Kazhdan, and Thomas Funkhouser. The princeton shape benchmark. In *Proceedings of the Shape Modeling International 2004*, SMI '04, pp.167–178, Washington, DC, USA, 2004. IEEE Computer Society.
- [27] Sudipta N. Sinha, Jan-michael Frahm, Marc Pollefeys, and Yakup Genc. Gpu-based video feature tracking and matching. In *Proceedings of Workshop on Edge Computing Using New Commodity Architectures*, EDGE'06, 2006.
- [28] Josef Sivic and Andrew Zisserman. Video google: A text retrieval approach to object matching in videos. In *Proceedings of the Ninth IEEE International Conference on Computer Vision*, Vol.2 of ICCV '03, pp.1470–1477, Washington, DC, USA, 2003. IEEE Computer Society.
- [29] Johan W. Tangelder and Remco C. Veltkamp. A survey of content based 3d shape retrieval methods. *Multimedia Tools Applications*, Vol.39, No.3, pp.441–471, September 2008.
- [30] J. R.R. Uijlings, A. W.M. Smeulders, and R. J.H. Scha. Real-time visual concept classification. *IEEE Transactions on Multimedia*, Vol.12, No.7, pp.665–681, November 2010.
- [31] Dejan V. Vranic. *3D Model Retrieval*. PhD thesis, University of Leipzig, 2004.
- [32] Dejan V. Vranic. Desire: a composite 3d-shape descriptor. In *Proceedings of IEEE International Conference on Multimedia and Expo*, pp.962–965, 2005.
- [33] Dejan V. Vranic and Dietmar Saupe. 3d shape descriptor based on 3d fourier transforms. In *Proceedings of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services*, pp.271–274, 2001.
- [34] Eric Wahl, Ulrich Hillenbrand, and Gerd Hirzinger. Surflet-pair-relation histograms: A statistical 3dshape representation for rapid classification. In *Proceedings of International Conference on 3D Digital Imaging and Modeling*, pp.474–482, 2003.
- [35] Jinjun Wang, Jianchao Yang, Kai Yu, Fengjun Lv, Thomas S. Huang, and Yihong Gong. Locality-constrained linear coding for image classification. In *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR '10, pp.3360–3367, 2010.
- [36] 大淵竜太郎. マルチメディア検索の最先端第7回: 3次元形状の検索. 映像情報メディア学会誌, Vol.64, No.7, pp.967–972, 2010.
- [37] 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.91-D, No.1, pp.23–36, 2008.

(2012年5月30日 受付)

(2012年10月1日 採録)

[問い合わせ先]

〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘1-1

豊橋技術科学大学

立間 淳司

TEL: 0532-44-6764

FAX: 0532-44-6757

E-mail: atsushi@kde.cs.tut.ac.jp

著者紹介

たつ ま あつ し
立間 淳司 [非会員]

2006年豊橋技術科学大学情報工学課程卒。2008年豊橋技術科学大学大学院修士課程情報工学専攻修了。同年ヤフー(株)入社。2010年より豊橋技術科学大学大学院博士後期課程電子・情報工学専攻、現在に至る。情報検索、データマイニングなどの研究に従事。日本学術振興会特別研究員。電子情報通信学会、電気学会、日本データベース学会各会員。

あお の まさ き
青野 雅樹 [非会員]

1981年東京大学理学部情報科学科卒。1984年東京大学大学院理学系研究科情報科学専攻修士課程修了。同年日本アイビーエム(株)入社。1994年、米国レンセラー工科大学コンピュータサイエンス学科Ph.D.課程修了。2003年より豊橋技術科学大学情報工学系教授、現在に至る。情報検索、データマイニングなどの研究に従事。著書に『Javaで学ぶコンピュータグラフィックス』など。電子情報通信学会、人工知能学会、情報処理学会、言語処理学会、日本データベース学会、ACM、IEEE各会員。

3D Object Retrieval based on Correlation of Multi-view Image Local Feature

by

Atsushi TATSUMA and Masaki AONO

Abstract :

In this paper, we propose a new feature vector for 3D object retrieval, which we call Local Feature Correlation Descriptor (LCoD). Given a 3D object, we first render depth-buffer images from multiple viewpoints. We then extract local features from each depth-buffer image. For every depth-buffer image, we compute the correlation matrix of local features, and define the vector as LCoD, which is obtained by the elements of the correlation matrix. Our experiments on the Princeton Shape Benchmark show that LCoD achieves the First Tier of 0.4708, which exhibits higher search performance than conventional techniques.

Keywords : Information Retrieval, Feature Extraction, 3D Object, Correlation, Database

Contact Address : Atsushi TATSUMA

Toyohashi University of Technology

1-1, Hibarigaoka, Tenpakucho, Toyohashi-shi, 441-8580, JAPAN

TEL : 81+532-44-6764

FAX : 81+532-44-6757

E-mail : atsushi@kde.cs.tut.ac.jp