

3次元CG画像のKAZE特徴を用いた2次元画像をクエリとする 3次元形状検索システム

相良英樹^{†1} 高野茂^{†1} 岡田義広^{†1}

概要: 近年、グラフィックスデバイスの発達や3次元幾何学データに関するセンサの普及が急速に進んでおり、それに伴い3次元形状データを利用したマルチメディアコンテンツも数を増しつつある。そうしたコンテンツや、それを新しく構成するための素材となる3次元形状データを探し出す技術の必要性も高まりつつある。本研究では様々な視点方向から3次元形状データを描画して3次元CG画像を生成し、KAZE特徴を用いてクエリとして入力された2次元画像との類似度を計算し、画像を用いた3次元形状データの類似検索を行うシステムを提案する。

キーワード: 3次元形状データ、類似検索、Bag-of-Features、KAZE特徴

3D Shape Data Retrieval System Using KAZE Feature and Accepting 2D Image as Query

HIDEKI SAGARA^{†1} SHIGERU TAKANO^{†1}
YOSHIHIRO OKADA^{†1}

Abstract: Recently, with the development of graphics devices and widely use of 3D geometrical sensors, the quantity of multimedia data including 3D shape data has been increasing. Therefore, the technological needs for searching those data to create multimedia contents are expanding. In this paper, the authors propose a 3D shape data retrieval system using KAZE image feature and accepting a 2D image as a query for the search.

Keywords: 3D shape data, Similarity retrieval, Bag-of-Features, KAZE Features

1. はじめに

近年、グラフィックスデバイスの発達や3次元幾何学データに関するセンサ(Kinect、ステレオカメラ、3Dスキャナなど)の普及が急速に進んでおり、それに伴い3次元形状データを利用したマルチメディアコンテンツも数を増しつつある。コンテンツの増加に伴い、膨大なそれらコンテンツの中からユーザが希望するものを高速に探し出す技術や、コンテンツを新しく構成するためにユーザがその素材となるデータを高速かつ効率良く探し出す技術の必要性も高まっている。

本研究では、データベース上の3次元形状データを描画し、その描画画像およびクエリとなる画像から得られた局所特徴量を用いて類似検索を行う3次元形状データ検索システムを提案する。そして、その検索システムを実装し、実際に検索精度を検証する。

2. システムの説明

本章では、画像をクエリとした3次元形状データ検索システムを説明する。本システムは複数枚の画像をクエリとすることにより、対象となる3次元形状データの外見的特徴を考慮した直感的な検索を目指している。

3次元形状をレンダリング(描画)して画像を得ることにより、3次元形状から画像特徴を抽出することができる。本章では、まず3次元形状を描画して2次元画像を得る過程を説明する。次にそれらの画像から特徴を抽出し類似度計算のためのヒストグラムを構築する方法、およびそれらを比較して3次元形状検索を実現する方法を説明する。

^{†1}九州大学, 福岡県福岡市西区元岡 744 番地
Kyushu University, Motooka 744, Nishi-Ku, Fukuoka, 819-0395, Japan

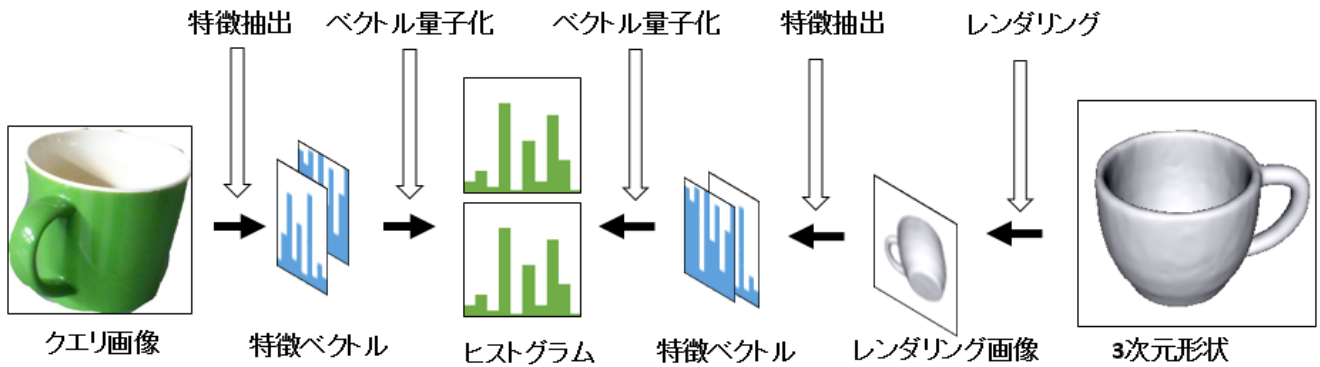


図 1 画像をクエリとした三次元形状検索システムの概略図

図 1 は本研究において提案するシステムの概略図である。まず、データベース上の 3 次元形状をレンダリング(描画)して画像を生成する。3 次元形状から生成した画像とクエリ画像それぞれから抽出した画像特徴について類似度の比較を行うことによって画像をクエリとした 3 次元形状検索を実現する。画像特徴の類似度の比較のために、画像特徴をベクトル量子化によってヒストグラムの形にコーディングする。このベクトル量子化は Bag-of-Features という手法を用いて行う。

以降の節では、各過程についての詳細を述べる。

2.1 三次元形状のレンダリング(描画)

3 次元形状を描画して画像を得る過程について説明する。本システムでは、画像特徴を利用した類似検索を実現するべく、データベース上の 3 次元形状から複数の画像を描画する。この際、次に記す 2 つのステップを経る必要がある。

2.1.1 スケールと位置の正規化

3 次元形状データは、その図形的なスケールや基準となる原点位置がデータ毎に異なる場合がある。そのため、後のレンダリング処理を簡単化すべく、またスケールに対して不変な特徴抽出を実現するため、スケールと原点位置の正規化を行う。

正規化のため、本システムでは 3 次元形状の頂点を包含するような球のうち、最小の半径となるような球(最小包含球)を求める。そのような球を全探索によって求める場合の時間計算量は 3 次元形状の頂点数 V とすると $O(V^4)$ となり莫大な時間がかかるため、今回のプログラムでは Welzl の考案した *move-to-front heuristics* で求めるアルゴリズム [3]を用いる。このアルゴリズムでは最小包含球を線形時間で求めることができる。

そのようにして求めた最小包含球の中心を原点位置、またすべての 3 次元形状データについて最小包含球の半径が一定となるようにスケールを調整することにより、スケールと位置の正規化を行う。

2.1.2 多視点からのレンダリング

画像をクエリとした検索において、画像の視点方向を問わず検索を実現するためには、形状のレンダリングを多くの方向から行う必要がある。そのため、本システムでは 1 つの形状に対して描画する角度を変えながら複数枚の画像を描画する。

カメラアングルを定めるための方向ベクトルを生成する方法について述べる。カメラアングルは、形状を包含するような球を考えた時、なるべく偏りなくその球面上にカメラが配置されることを目指したい。そこで本システムでは、カメラアングルを定める方向ベクトルを以下のような式によって、近似的に様に球面上に分布するアングルを生成する。これは、球面上に螺旋を描き、その螺旋上に黄金比 Φ にまつわる角間隔で点を置くものである。具体的には、アングル数を n とすると、 i 番目のアングルの方向ベクトル \vec{A}_i ($0 \leq i < n$) は、

$$\begin{aligned} \vec{A}_i &= (r_i \cos(i\Phi), r_i \sin(i\Phi), z_i) \\ \Phi &= (3 - \sqrt{5})\pi \\ z_i &= \left(1 - \frac{1}{n}\right)\left(1 - \frac{2i}{n-1}\right) \\ r_i &= \sqrt{1 - z_i^2} \end{aligned}$$

となる。

今回のシステムではアングル数 64、画像の解像度は 256×256 とした。

2.2 画像特徴の抽出

データベース上のすべての 3 次元形状データに対して前節のレンダリングを行って画像を得た後、その画像に対して KAZE 特徴のアルゴリズムを適用し、特徴点の検出および各特徴点に対する局所特徴ベクトルの記述を行う。

KAZE とは Pablo Aantarrilla らによって考案された 2 次元画像に対する特徴検出および記述のアルゴリズムである

[1]。このアルゴリズムは、局所特徴抽出アルゴリズムである SIFT [2]を改良したものであり、ガウシアンフィルタの代わりに非線形拡散フィルタを用いてエッジの保存を行い、元の画像の特徴をより良く保って特徴抽出を行うというものである。

このアルゴリズムによってキーポイント 1 つあたり 64 次元の特徴ベクトルが抽出される。

2.3 Bag-of-Features による特徴ヒストグラム構築

データベース上の 3 次元形状データすべてから画像を描画し局所特徴ベクトルを抽出した後、それらの局所特徴ベクトルからデータの類似度比較のために用いる特徴ヒストグラムを作成する。本研究のシステムにおいては、Gabriela Csurka らによって考案された Bag-of-Features [4] という手法を用いて特徴ヒストグラムを作成する。この手法は、データベース上のデータから抽出した局所特徴ベクトルをクラスタリングして分類し、任意の 1 つの特徴ベクトルをクラスタ数 k 通りにコーディングすることによってベクトル量子化を行い、1 つのデータから抽出される複数の特徴ベクトル各クラスタに属する特徴ベクトルを数えあげることによって、1 つのデータを k 次元の特徴ベクトルで表現するというものである。

まず特徴ベクトルのコーディングのために、データベース上の画像から抽出された特徴ベクトルを特徴空間上でクラスタリングする。クラスタリングには k -means++法[5]を用いる。この方法は、通常の k -means 初期値の選択方法を以下のような手続きに従って行うものである：

1. データ点からクラスタ中心をランダムに 1 つ選ぶ。
2. 各データ点 x に対し、 x と x に対して最も近いクラスタ中心との距離 $D(x)$ を求める。
3. 新しいクラスタ中心をもう 1 つ選ぶ。先に求めた $D(x)$ の二乗である $D(x)^2$ を重みとした重み付き確率分布を用いてランダムに新しいクラスタ中心を選ぶ。
4. クラスタ中心が k 個選ばれるまで 2. と 3. を繰り返す。

今回のシステムではクラスタ数 $k=256$ とした。

2.4 特徴ヒストグラムの構築

各画像 I について、クラスタ c に属する特徴ベクトルの個数が $H_I(c)$ となるようなヒストグラム H_I を作成する。ヒストグラムのビン数は先のクラスタリングで設定したクラスタ数 k と同じになる。

1 つの形状からレンダリングされた画像のヒストグラムをすべて足し合わせて 1 つのヒストグラムを作る。そのあと、すべての形状のヒストグラムの面積が一定になるよう、ヒストグラムの面積を正規化する。

クエリは複数枚の画像からなり、これらの画像からもデータベース側のレンダリング画像と同様にヒストグラムの

構築・足し合わせ・正規化を行う。

上の 3 次元形状をレンダリングした画像から得られたヒストグラムと、クエリ画像から作成したヒストグラムの類似度を比較する。

ヒストグラムの比較には Minkowski の L_1 距離を用いる。これによってビン数 n である 2 つのヒストグラム $H_A(c)$ とヒストグラム $H_B(c)$ の距離 d は以下のような式で表される。

$$d(H_A, H_B) = \sum_c |H_A(c) - H_B(c)|$$

この距離 d がより小さくなるデータ同士が、より類似したデータであると考えられる。

3. 実験

これまでに、3 次元形状データ検索システムについて説明した。このシステムを実際に実装し、検索実験を行った。本章では、検索実験とその結果について説明する。

3.1 実験用データベース

実験において、データベースとして 2 つの 3 次元形状データベースを用いた。どちらのデータベースも McGill Shape Benchmark[6]からのデータで構成されている。一方は MSB のデータのうち "Ants", "Crabs", "Spectacles", "Hands", "Humans", "Octopuses", "Pliers", "Snakes", "Spiders", "Teddy-bears" の 10 クラス 255 データからなるデータベースであり、もう一方は "Tables", "Cups", "Chairs", "Airplanes", "Dolphins", "Birds", "Four-limbs", "Dinosaurs", "Fishes" の 9 クラス 202 データからなるデータベースである。これらの概要を次ページの図 2 に示す。

3 次元形状モデル検索システムの評価指標として、"First tier", "Second tier", "Top match" を使用した [7]。これらの定義は以下に示す。ただし、質問モデルの属するクラスを C_q 、 C_q が含むモデルの数を k とする。

● First tier

この指標は、質問モデルを除く C_q 内のモデルのいくつが、検索結果の上位 $k-1$ 位以内に入ったかの割合を示す。

$$\text{First tier} = \frac{\text{Top}(k-1)\text{match}}{k-1}$$



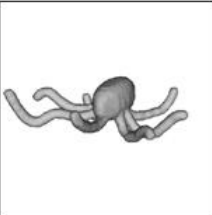
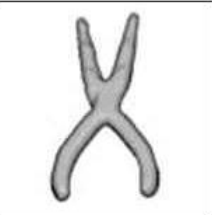

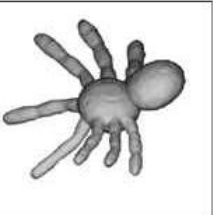
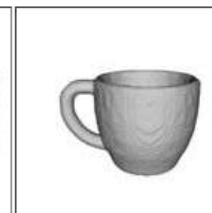
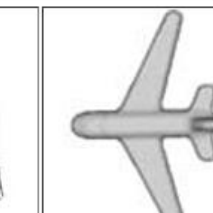
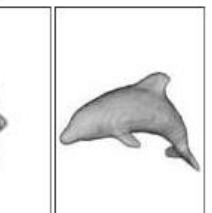
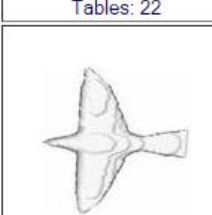

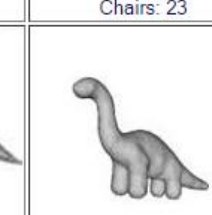
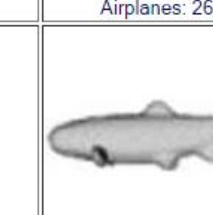
				
Ants: 30	Crabs: 30	Spectacles: 25	Hands: 20	Humans: 29
				
Octopuses: 25	Pliers: 20	Snakes: 25	Spiders: 31	Teddy-bears: 20
				
Tables: 22	Cups: 25	Chairs: 23	Airplanes: 26	Dolphins: 12
				
Birds: 21	Four-limbs: 31	Dinosaurs: 19	Fishes: 23	

図 2 データベース 1 (上 2 行) とデータベース 2 (下 2 行)。
(クラス名横の数字はクラス毎のデータ数)

- Second Tier

この指標は、First tier と類似の指標だが、検索結果の上位 $2(k-1)$ 位以内に入った割合を示す。

$$\text{Second tier} = \frac{\text{Top } 2(k-1)\text{match}}{k-1}$$

- Top match

この指標は、検索結果の第一位が C_q 内のモデルかどうかを表す指標である。 C_q 内のモデルだったら 1、それ以外であれば 0 となる。

3.2 既存手法との比較

この実験では、従来より一般的に使用されている 3 次元形状モデルの検索手法である D2 法との比較を行った[7]。D2 法は、3 次元形状モデルの表面においてランダムに 2 頂点を数組取ってきて、その 2 頂点間の距離のヒストグラムを作成し、そのヒストグラムを特徴として扱う手法である。D2 法では、単純な形状のモデルほど特徴的なヒストグラムを示す。よって球や立方体などのプリミティブと呼ばれる

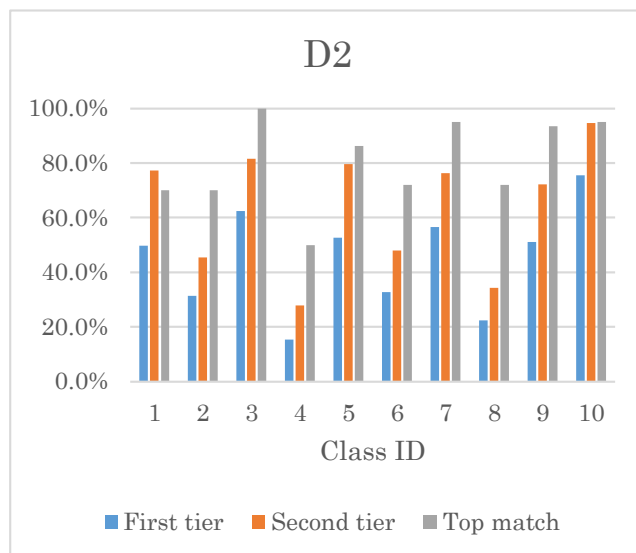
形状の検索に優れている。

実験方法は、まずデータベース内の全ての 3 次元形状モデルを質問モデルとして検索を行い、前述の指標“First tier”、“Second tier”、“Top match”を計算する。次に、それらの 3 指標の値について、クラス内での平均と、データベース内全体の平均を取り、それぞれを比較する。検索実験の結果を次 2 ページの表 1・表 2 に示す。

各項目について、概して提案システムは D2 法のスコアを上回るため、結果の改善がみられる。よって、形状の検索において KAZE 特徴の利用は有効であると考えられる。

D2

Class ID	Class	First tier	Second tier	Top match
1	ants	49.7%	77.2%	70.0%
2	crab	31.4%	45.4%	70.0%
3	spectacles	62.5%	81.5%	100.0%
4	hands	15.3%	27.9%	50.0%
5	humans	52.6%	79.7%	86.2%
6	octopuses	32.7%	48.0%	72.0%
7	pliers	56.6%	76.3%	95.0%
8	snakes	22.3%	34.3%	72.0%
9	spiders	51.2%	72.3%	93.5%
10	teddy	75.5%	94.7%	95.0%
average.		45.0%	63.7%	80.4%



提案手法

Class ID	Class	First tier	Second tier	Top match
1	ants	61.5%	87.9%	100.0%
2	crabs	74.5%	88.0%	100.0%
3	spectacles	95.7%	99.8%	100.0%
4	hands	59.5%	77.4%	95.0%
5	humans	73.8%	87.2%	100.0%
6	octopuses	25.7%	37.7%	72.0%
7	pliers	94.2%	99.5%	100.0%
8	snakes	38.7%	49.7%	84.0%
9	spiders	65.6%	94.8%	93.5%
10	teddy	97.6%	100.0%	100.0%
average.		68.7%	82.2%	94.5%

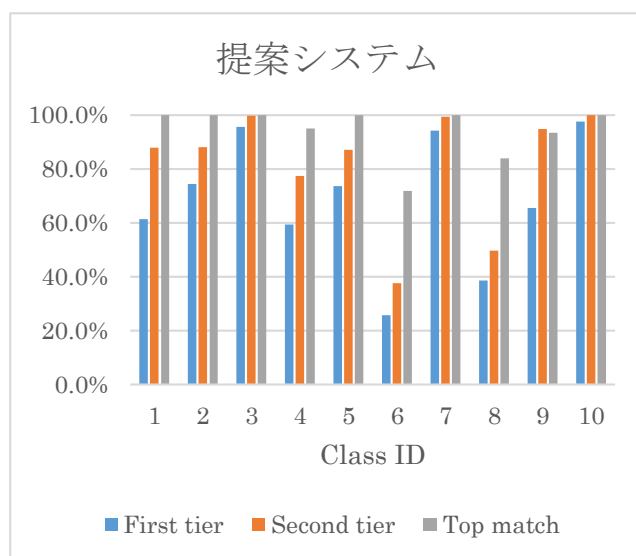
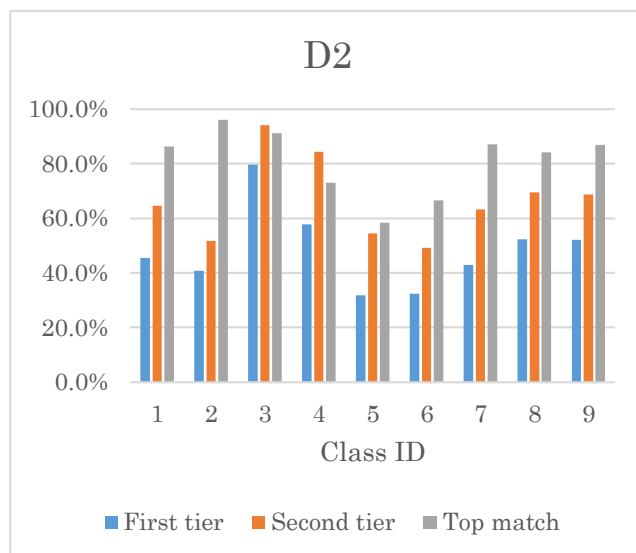


表 1: データベース 1 に対する
検索精度の比較

D2

Class ID	Class	First tier	Second tier	Top match
1	table	45.5%	64.7%	86.4%
2	cup	40.8%	51.7%	96.0%
3	chair	79.6%	94.1%	91.3%
4	airplane	57.8%	84.3%	73.1%
5	dolphin	31.8%	54.5%	58.3%
6	bird	32.4%	49.3%	66.7%
7	four	43.0%	63.3%	87.1%
8	dinosaur	52.3%	69.6%	84.2%
9	fish	52.2%	68.8%	87.0%
average.		48.4%	66.7%	81.1%



提案手法

Class ID	Class	First tier	Second tier	Top match
1	table	35.7%	51.9%	77.3%
2	cup	76.2%	87.8%	100.0%
3	chair	76.7%	91.3%	100.0%
4	airplane	75.8%	94.5%	96.2%
5	dolphin	45.5%	57.6%	83.3%
6	bird	41.2%	66.7%	81.0%
7	four	73.1%	91.9%	90.3%
8	dinosaur	40.6%	69.9%	78.9%
9	fish	81.8%	96.0%	100.0%
average.		60.7%	78.6%	89.7%

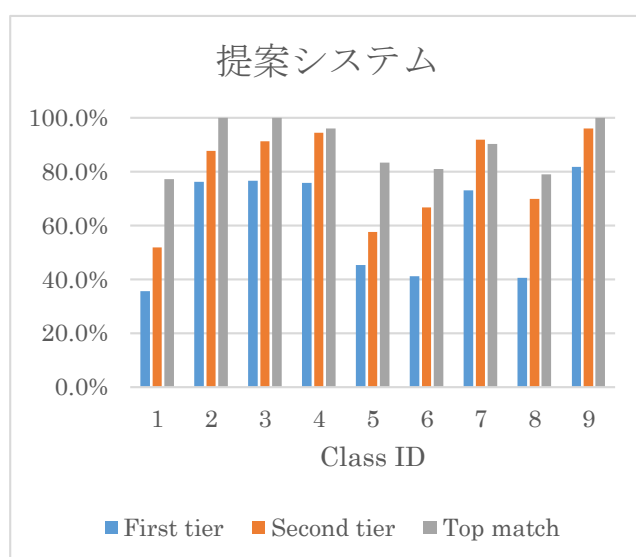


表 2: データベース 2 に対する
検索精度の比較

3.3 実世界からの画像をクエリとした検索実験

実世界において撮影した物体の複数枚の写真をクエリとして、データベースに対する検索実験を行う。今回の実験では、実世界に存在する物体から写真を8枚、アングルを変えながら撮影し、それをクエリとした検索を行った。

以下にクエリとした画像と、そのクエリ画像による結果について前節で導入した First tier, Second Tier, Top match の指標を適用して記す。



図 3: マグカップの写真 8 枚 (Cups クラスとする)



図 4: 恐竜の置物の写真 8 枚 (Dinosaurs クラスとする)

Cups クラスのデータ数 (クエリ除く)	25
First tier	28.0%
Second tier	44.0%
Top match	0

表 4: 検索結果の First tier, Second Tier, および Top match

Dinosaurs クラスのデータ数 (クエリ除く)	19
First tier	10.5%
Second tier	10.5%
Top match	0

表 5: 検索結果の First tier, Second Tier, および Top match

結果として各指標は前節における提案手法のスコアの平均を下回るため、実世界からの写真を用いた検索の精度は低いと考えられ、改善の余地がある。精度が低い原因の一つとして、データベースでのレンダリング画像と実世界の写真とでは照明条件が異なることが考えられる。特に、物体が背景に落とす影にも特徴点が現れるため(図5)、実用には影や背景の除去を要する。



図5: 背景の影に現れる特徴点

4. まとめと今後の課題

4.1 まとめ

本研究では、描画画像およびクエリとなる画像から得られた局所特徴を用いて類似検索を行う3次元形状データ検索システムを提案し、また実際に構築したシステムについて検索の実験を行った。

4.2 今後の課題

本システムではクエリに対する処理が限定的であるため、写真をクエリとした場合の照明条件の違いについての対処を検討している。例えば、データベース上の3次元形状をレンダリングする際にアングルだけでなく照明条件も変更しながらのレンダリング処理を行う、クエリ画像から背景を除去するアルゴリズムの導入をするなどの改善を検討している。また、実用的には複数枚のクエリ画像を得ることが難しい場合も考えられるため、より少ない枚数(理想的には1枚)の画像から検索する方法を検討している。

参考文献

- [1]Pablo. F. Alcantarilla and A. Bartoli and A. J. Davison, ``KAZE features'', Proceedings of the 12th European Conference on Computer Vision - Volume Part VI, pp. 214-227, 2012.
- [2]David Lowe, ``Distinctive image features from scale-invariant keypoints'', International Journal of Computer Vision. Volume 60, pp. 91-110, 2004.
- [3]Emo Welzl, ``Smallest Enclosing Disks (Balls and Ellipsoids)'', New Results and New Trends in Computer Science, pp. 301-303, 1991.
- [4]G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray, ``Visual Categorization with Bags of Keypoints'', ECCV In Workshop on Statistical Learning in Computer Vision, pp. 1-22, 2004.
- [5]David Arthur and Sergei Vassilvitskii, ``k-means++: the advantages of careful seeding'', Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics Philadelphia. pp. 1027-1035, 2007.
- [6]Shape Analysis Group of McGill University, "McGill 3D Shape Benchmark", <http://www.cim.mcgill.ca/~shape/benchMark/>, 2005.
- [7]Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin, ``Matching 3D Models with Shape Distributions'', Int. Conf. on Shape Modeling and Applications, pp. 154-165, 2001.