立体形状検索システム

堀越 力*, 浜野輝夫*, 石井健一郎*

A 3D shape retrieval system

HORIKOSHI Tsutomu, HAMANO Teruo, ISHII Kenichiro

[著作抄録] NTT ヒューマンインタフェース研究所では、立体形状検索システムの開発を行っている。このシステムは、立体物の形状を超二次関数という数式で記述しており、ユーザが普段立体を表現するために使う言葉やスケッチにより立体形状を指示するだけで、それに類似した画像を検索することができる。この超二次関数は、少ないパラメータで多様な形状を表現でき、言葉と結び付けやすいといった特徴を有する。この関数パラメータを用いることで、三次元形状を客観的な数値データに変換することができ、言葉による変形操作を可能にしている。本稿では、超二次関数の特徴を述べ、検索システムの検索方法について紹介する。

[著者付与キーワード] 画像データベース, 三次元, ユーザインタフェース, 超二次関数, 中間メディア, 三次元形状インデックス, スケッチ, 言葉

[Author Abstract] NTT Human Interface Laboratories have developed a new 3D shape retrieval system. This system enables users to access a 3D database with descriptive words or 2D rough sketches, both of which are common descriptive media. 3D shapes are described by superquadrics which are used as an intermediate language between users and computers. Superquadrics can describe a wide variety of shapes with just a few simple parameters. These parameters can be used instead of descriptive words. By using these parameters, 3D shapes can be objectively represented by numerical data structures, and a modeled 3D object can be easily deformed by verbal description. This paper discusses the applicability of superquadrics and gives and gives an outline of the system.

[Keywords by Author] image database, 3-dimensional, user interface, superquadrics, intermediate language, 3D shape index, rough sketch, descriptive words

* NTT ヒューマンインタフェース研究所 マルチメディア処理研究部
(〒238-03 横須賀市武1-2356) Tel. 0468(59)2877

NTT Human Interface Laboratories Human and Multimedia Laboratory
(2356, Take 1-chome, Yokosuka-shi, 238-03)

1 ま え が き

画像データベースで立体物を検索する場合、対象物の分類番号等による検索を行い、表示された複数視点での画像をもとに検索者が頭の中で立体形状を想像するといった方法が一般的である。近年のコンピュータグラフィックス(CG)による三次元表示技術の発達にも関わらず、この技術はほとんどデータベース技術に応用されていない。それは、CG技術が主に形状表現のリアリティを目的とし、検索といった観点での開発ではないことが主な原因であると考えられる。つまり、立体物を立体として表示する技術は確立されつつあるが、立体を指定し検索できるような、立体形状の蓄積・検索技術は未確立であった。

さて、普段人が立体形状を相手に伝える場合、言葉やスケッチなどが使われる。つまり、立体形状による検索にも、このような言葉やスケッチが使えることが望ましい。そうすれば、検索時のわずらわしい検索式を必要とせず、ユーザフレンドリなインタフェースを構築することができるようになる。

一般にコンピュータに言葉やスケッチを理解させることは容易ではないが、“中間メディア”を考えることで、疑似的にユーザフレンドリなマン・マシン・コミュニケーションを行うことが可能になる。“中間メディア”とは、人間の三次元表現と、コンピュータ内の三次元表現とを結び付けるものであり、以後、“三次元インデックス”と呼ぶことにする。そして、この三次元インデックスを用いることにより、実際の物体の立体形状と言葉やスケッチなどを結び付けることができるのである。

NTTヒューマンインタフェース研究所では、効率よく、しかも簡単なインタフェース、あたかも、コンピュータが人間であるかのようなコミュニケーションの実現を目指している。そこでの研究の一つとして、立体形状検索技術の研究がある^{(1),(2)}。ここでは、超二次関数^{(3),(4),(5)}を三次元インデックスとして用いることで、立体物をこの数式

で記述し、数式と言葉やスケッチとを結び付けている。そして、立体形状検索システムを作成し、言葉による立体形状の指示や手書きスケッチによる“こんな形の何々”といった検索手段が可能であることを示す。

2 超二次関数

この超二次関数とはどんな関数であるか。まず、次式のような関数から考えることにする。

$$x = \cos^{\epsilon}(\theta), \quad y = \sin^{\epsilon}(\theta) \quad (1)$$

ここで、べき乗 ϵ が1ならば(1)式は円を極座標*で表現した式である。そこで、このべき乗が実数であるとする、図1に示すように ϵ が0に近づくにつれ、円の四隅がだんだんと角張ってきて正方式に近い形ができてくる。また ϵ を1より大きくしていくと、円の四隅が窪んできて、星形に近づいてくる。つまり、この ϵ の値をいろいろ変化させることで円や四角形、さらには、その中間的な形も表現できることがわかる。

超二次関数とは、この(1)式で表現される二つの二次元曲線を三次元空間で垂直に組み合わせた関数であり、次式で表される。

図1 二次元の超二次関数の形状変化

$$\begin{aligned} x &= a_1 \cos^{\epsilon_1}(\eta) \cos^{\epsilon_2}(\omega) \\ y &= a_2 \cos^{\epsilon_1}(\eta) \sin^{\epsilon_2}(\omega) \\ z &= a_3 \sin(\eta) \end{aligned} \quad (2)$$

ここで、 η 、 ω は、それぞれ極座標における緯度、経度を示す、 ϵ_1 、 ϵ_2 は、ともに実数のべき乗を表し、形状パラメータと呼ぶ。この ϵ_1 、 ϵ_2 がともに 1 の場合、(2)式は球を表す式となる。そして、 ϵ_1 、 ϵ_2 はそれぞれ、関数を真横から見たときの形状（ z 軸を含む面での断面形状）および真上から見たときの形状「 x y 平面での断面形状」の角張り具合を変化させるパラメータである（図 2 参照）。また、 a_1 、 a_2 、 a_3 は、各座標軸方向の大きさを表し、スケールパラメータと呼ぶ。そして、これらの関数をパラメータを変えることで、図 3 に示すような基本的な立体形状のみならず、これらの中間的な形状も表現することが可能である。

このような特徴を持つ超二次関数の三次元インデックスとしての有効性を判断するため、次のような条件を設定した。

3 三次元形状インデックスの条件

三次元インデックスは、次のような条件を満た

図 2 超二次関数

さなければならない。

1. 三次元インデックスを画像（つまり物体）から容易に入手できること。
2. 立体形状の類似性の表現能力を持つこと。
3. 言葉と立体形状とが結び付けやすいこと。
4. 表現できる形状の種類が多いこと。

このような条件の下に超二次関数の特徴について詳しく調べていった。

4 超二次関数のインデックスとしての有効性

4.1 インデックス入手のしやすさ

物体からインデックスを取り出すということは、物体の表現に超二次関数を近似し、その近似された関数のパラメータを取り出すことに相当する。

図 2 で示したように、形状パラメータは、関数を二方向から見た形状の角張り具合を独立に変化させる。したがって、入力したい物体の真上および真横から見た輪郭形状それぞれに、(1)式を近似すればよいことになる。

そこで、得られた輪郭形状の座標データ列（ x 、 y ）に対して、超二次関数の最小二乗近似⁽⁵⁾を行う。まず超二次関数(1)式は、(3)式に変形できる。

$$S = (x/a_1)^{2/\epsilon_1} + (y/a_2)^{2/\epsilon_2} = 1 \quad (3)$$

座標点（ x 、 y ）が関数上に存在すれば関数 S の値は 1 となる。そのため、その二乗誤差の和 χ^2 は(4)式により簡単に表すことができる。

$$\chi^2 = \sum (S - 1)^2 \quad (4)$$

つまり、この二乗誤差が最小となるパラメータの組み合わせを求めることで、輪郭形状に最も近似された関数が得られる。すなわち、容易に三次元インデックスを取り出すことができるのである。

例えば、 z 軸および y 軸方向から物体を撮影した場合、真上（ z 軸方向）から撮影した画像から ϵ_2 、 a_2 が求まり、真横（ y 軸方向）から撮影した画像から ϵ_1 、 a_1 が求まるが、 a_3 （ x 軸方向のスケールパラメータ）は両方の画像から求められる。

$\epsilon_1=0.1$
 $\epsilon_2=1.0$
 $\alpha_1=2.0$
 $\alpha_2=2.0$
 $\alpha_3=2.0$

$\epsilon_1=1.0$
 $\epsilon_2=1.0$
 $\alpha_1=2.0$
 $\alpha_2=2.0$
 $\alpha_3=2.0$

$\epsilon_1=2.0$
 $\epsilon_2=0.1$
 $\alpha_1=2.0$
 $\alpha_2=2.0$
 $\alpha_3=2.0$

(a)形状パラメータを変化させた例

$\epsilon_1=0.1$
 $\epsilon_2=0.1$
 $\alpha_1=2.0$
 $\alpha_2=3.0$
 $\alpha_3=1.0$

$\epsilon_1=0.1$
 $\epsilon_2=1.0$
 $\alpha_1=3.0$
 $\alpha_2=3.0$
 $\alpha_3=1.0$

$\epsilon_1=0.1$
 $\epsilon_2=1.0$
 $\alpha_1=1.0$
 $\alpha_2=1.0$
 $\alpha_3=3.0$

(b)スケールパラメータを変化させた例

図3 様々な形状の超二次関数

カメラから物体までの距離が一定であれば問題ないが、異なる場合は、二つの画像から得られた a_1 の値の違い（例えば、 a_{1-top} 、 a_{2-side} とする）から、大きさの比 ($R = a_{1-top}/a_{2-side}$) を求め、大きさの正規化*を行う。

ここで、当然関数で表現しきれない形状に近似させてもその形には収束しないため、関数で表現できる形状に物体を部品分割し、個々の部品を関数で近似する必要がある。

4.2 立体形状の類似性の表現能力

超二次関数はパラメータの変化とともに連続的に形状が変化する。図4は、形状パラメータを0.1ずつ変化させることで関数の形状が立方体から球に変形する様子を示している。図を見てわかるように、関数パラメータの値が近いものは三次元形状としても類似しており、パラメータ間の距離により三次元形状の類似性を判断できる。

4.3 言葉との結び付けやすさ

超二次関数の特徴として、人間の直感に合った形状変化ができることが挙げられる。そのため、立体形状を表現する言葉とその立体形状を表現している関数パラメータの組とを容易に結びつけることができる。

立体形状に関する言葉は、“形状”、“変形”および“変形の程度”を表現する三種類に分類できる。

まず、形状を表現する言葉に対しては、例えば、球は ($\epsilon_1 = \epsilon_2 = 1$, $a_1 = a_2 = a_3$)、円筒は ($\epsilon_1 \approx 0$, $\epsilon_2 = 1$, a_1, a_2, a_3 は任意) 直方体は ($\epsilon_1 \approx 0$, $\epsilon_2 \approx 0$, a_1, a_2, a_3 は任意) などというように、関数パラメータの組み合わせで記述することができる。ここで、形状パラメータ（つまりべき乗）が0になると、(2)式は空間の点 (a_1, a_2, a_3) を表すことになってしまい、立体として定義できなくなる。

また、変形やその程度を表す言葉に関しては、

図4 関数形状の変化

図5 関数パラメータと変形を表す言葉との結びつき

関数パラメータの変化の方向とその変化量(δ)との組み合わせで対応が付けられる(図5参照)。

例えば“直方体”と“円筒”の間において $\epsilon_2 = 0.1$, $\delta = 0.2$ としたとき, “丸まる”とは“ $\epsilon_2 + \delta$ ”と表現でき, この言葉により, 新しい ϵ_2 の値は0.3になるわけである。また, “角張る”とは“ $\epsilon_2 - \delta$ ”と数値的に対応づけることができる。

そして, 変形の程度に関しても, 例えば“少し”という意味は, 変化量(δ)を小さく設定することであり, “もっと”という表現は, 変化量(δ)を大きく設定することで記述することができる。

4. 4 適用範囲

超二次関数の形状は, 垂直軸に関して左右対称であり, かつ水平面に関して面対称な形状に限定される。しかし, このスケールパラメータの部分を変形関数と置き換えることで非対称な形状, 例えば円錐や四角錐といった形状が表現可能となる。図6は, (2)式の a_1 , a_2 の代わりにそれぞれ先細り関数(この場合, z 軸方向の一次関数) $f_1(z)$, $f_2(z)$:

$$f_1(z) = a_1 (1 - k_1(z/a_3)) \quad (5)$$

$$f_2(z) = a_2 (1 - k_2(z/a_3)) \quad (6)$$

で置き換えた超二次関数の例である。

また, 物体を部品として分割できる場合は, 強引に変形関数で記述しようとするよりは複数の部品の組み合わせとして記述した方が有効である。各部品間の結合関係は, 各部品の重心の位置ベクトルにより記述できる。

三次元インデックスとしては, 変形関数と部品とを組み合わせることにより, 関数で表現できる形状の適用範囲を広げることが重要である。

次に, これを応用した画像検索システムについて具体的操作例を述べる。

5 立体形状検索システム

このシステムでは, 二次元の画像に三次元インデックスを付与しており, 言葉や手書きスケッチ

で立体形状を作成(指示)することにより, それに最も近い形状の画像を捜し出すことができる。

5. 1 画面構成

図7に示すように本システムは, 言葉やスケッチを入力する立体形状入力用と検索画像等を表示する画像表示用ディスプレイの二つで構成されている。画像表示用ディスプレイにおいて, 3Dウィンドウには, 三次元インデックスが表示され, マウスを動かすことにより, 自由にインデックスの向きを変えることができ, 立体形状として把握することができる。関連画像の部分には, コアウィンドウに表示されている物体の他の視点での画像などが表示される。

また, **Retrieve**は検索ボタン, **3D Index**は, コアウィンドウの画像の三次元インデックスを3Dウィンドウに表示させるためのボタンである。また, **Words**や**Sketch**をマウスで指示することで, 入力手段として言葉とスケッチのいずれかを選択できる。

5. 2 蓄積画像およびインデックスの内容

画像としては

- ・500×500画素*の原画像
- ・170×170画素の候補画像

の2種類を持ち, インデックスとして,

- ・立体物(あるいは部品)の名前
- ・三次元インデックス(超二次関数パラメータ)
 - (1) 形状(形状パラメータ ϵ_1 , ϵ_2 , 変形関数のパラメータ k_1 , k_2)
 - (2) 大きさ(スケールパラメータ a_1 , a_2 , a_3)
 - (3) 物体の位置(位置ベクトル(x , y , z))
 - (4) 部品の有無(有り: 1, 無し: 0)

で構成される。また, 部品が存在する場合, 各部品についても, 同様のインデックス(1)~(4)およびその部品の名前がさらに付与されることになる。

5. 3 インデックスの作成方法

入力したい物体から三次元インデックスを取り出すためには, まず, その物体に対して真上と真

図6 先細り関数を組み込んだ超二次関数

図7 画面構成図

横の2方向から見た画像を入力する。入力された画像は、それぞれ図7の上面図、側面図のウィンドウに表示され、入力したい物体の輪郭をマウスを使って一筆書きの要領で、人手によりトレースする。このとき、関数で表現できる形状に部品を分割する。例えば、コップの本体と取っ手の輪郭を別々にトレースすることになる。トレースされた各部分の輪郭の座標点の列に対して、4.1で述べたように、それぞれ超二次関数が最小二乗法により近似される。その結果、それぞれの部品の輪郭データをもとに、三次元の超二次関数のパラ

メータが取り出され、その物体のインデックスとして記憶される。

5.4 検索手段およびその結果

コアウィンドウ内にカーソルを移動させ、マウスボタンを押すと、図8(a)に示すポップアップメニュー*が表示される。新規作成を選択すると、図8(b)に示すメニューが現れ、直接立体形状を指示するか、あるいはスケッチにより立体形状を作成することができる。

(a) 変形を表す言葉のメニュー

(b) 新規作成のメニュー

図8 変形を表す言葉のメニュー

図9 スケッチ入力による概略形状の作成

5. 4. 1 スケッチによる検索

例えば図9のように検索しようとする形状をスケッチすると、その座標データ列に二次元の超二次関数が近似され、二つのスケッチから三次元の超二次関数が作られる。この形状をもとに検索を行うと、スケッチに近似してできた関数パラメータの値と蓄積されている各画像の三次元インデックスの値との差の総和が、各パラメータ間の距離

Dとして次式(7)により計算される。

$$D = \sum_{(\text{関数パラメータ})} \{ (\text{入力された関数のパラメータ値}) - (\text{画像に付与されているパラメータ値}) \}^2 \quad (7)$$

そして、距離Dの小さい順に画像を並び替えられ、上位6枚の画像が左側から順に候補画像として表示され、最も距離の小さい画像がコアウィンドウにも表示される。また、コアウィンドウに表示されている物体と同一形状の物体で他の視点か

(a) 概略形状による検索

(b) 候補画像（三番目の画像）の選択

(c) 三次元インデックスを取り出す

(d) 変形（角張る）されたインデックスによる検索

(e) 変形（円筒状、先細り）されたインデックスによる検索

図 10 検索結果

ら撮影された画像が下部のウィンドウに表示される。

また、例えば図 10(a)の左から 3 番目の候補画像にカーソルをマウスにより移動させ、マウスのボタンを押すと、その画像がコアウィンドウに大きく表示され、その関連画像も同時に表示される(図 10(b)参照)。つまり、任意の候補画像をマウスで選択することで、その画像に関する情報を得ることができる。

5. 4. 2 インデックスの変形による検索

検索しようとする物体形状が複雑な場合や細かな部分の形状による検索を行いたい場合は、各部品の断面形状をスケッチするにはかなりの熟練が要求される。そこで、このシステムでは、すでに画像に付けられている三次元インデックスを変形することで、より具体的な立体形状を簡単に作成することができる。

まず、3D ウィンドウ上部にある **3D Index** という部分をマウスで選択する。するとコアウィンドウに表示されている画像の三次元インデックスが 3D ウィンドウに表れる(図 10(c)参照)。コップの場合、本体と取っ手それぞれが一つの関数で記述されている。コアウィンドウ上でマウスのボタンを押すと図 8 のようなメニューが表示され、変形を表す言葉を選択することができる。ここで、例えば“もっと角張っている”コップを検索した

場合、コアウィンドウの中でマウスのボタンを押すとポップアップメニューが表示されるので、“角張る”という言葉を選択する。すると“角張る”という言葉に対応づけられている関数パラメータ(この場合 e_2)のみが変化できるようになる。変形の程度は、マウスの移動量に対応しており、自由に変えることができる図 11(a)~(c)。そして、ユーザが意図した角張り具合に変形できると、マウスのボタンを再度押すことで変形が終了する。これにより、角張ったコップが新たに作られたことになる。このような操作を繰り返すことにより、目的とした形状を作り上げていく。図 10(d)は、角張ったコップに変形した結果をもとに検索した結果である。図 10(e)は、さらに、“円筒状”、“先細る”という言葉により変形し、検索した結果である。

5. 4. 3 言葉の入力による検索

言葉を直接入力するモードでは、ユーザが
(もっと 角張る)
(もう少し 丸まる)

と言ったように言葉をキーボードから入力する。すると、システム側は、“もっと”、“もう少し”といった言葉に対応する変化量(δ)および“角張る”、“丸まる”に相当する関数パラメータを調べ、各パラメータに変化量を加えた値で、再度関数を表示する。すると、視覚的にインデックスの形状

図 11 インデックスの変形

が言葉に対応した変形を行っているように見える。そして、意図した形状を作成したならば、再度 **Retrieve** の部分をマウスで選択することで図 10 と同様の検索結果が得られる。

6 あ と が き

超二次関数による立体形状検索システムを作成し、その概要を述べた。本システムでは、従来のキーワード形式ではなく、スケッチや言葉など、普段、人が形状を表現するのに使っている手段での検索が可能であり、視点の相違による形状の変化等を考慮することなく、立体形状に関する検索を行うことができる。

現状では、単純に各パラメータ間の直線距離の和を用いて、形状の類似性を判断している。しかし、物体のどの形状に注目するかは、対象物、あ

るいは個人の視点により異なり、その人の意図に合った順番に画像が並べられているとは限らない。よりユーザの意図に合ったデータベースを可能にするには、対象によりパラメータの重み付けを変化させるなどの検討が必要である。

超二次関数による三次元インデックスは、現状では複雑な形状の物体を詳細に表現する能力に乏しいが、概略的な三次元画像として記述し処理させる範囲ならば、インデックスとして良好な性質を持っている。現在、概略形状のみならず、より広範囲な形状記述、入力自動化に関する検討を行っている。

将来的には、音声認識装置との組み合わせにより、ユーザがコンピュータの前で、「こんな形の何々が欲しい」というと、それに類似した画像を探し出してくれるシステムを作ることも可能であろう。

参 照 文 献

- (1) 堀越力, 笠原久嗣. 超二次関数による三次元インデクシング. 電子情報通信学会論文誌 D-II, J73-D-II(10) p.1716-1724 (1990)
- (2) 堀越力, 笠原久嗣. 超二次関数による三次元形状インデクシング. NTT R&D. 40(3) 393-398 (1991)
- (3) Barr, A. Superquadrics and angle-preserving transformations, IEEE Computer Graphics & Applications. 1,p.11-23 (1981)
- (4) Pentland, A. Recognition by parts, SRI Technical Note. No.406, Menlo park, CA (1986)
- (5) Bajcsy, R., Solina, F. Three dimensional object representation revisited. Proc. Int. Conf. on Computer Vision. London England, June 8-11 (1987)

用語説明

画素 二次元デジタル画像の最小要素 (離散点) をいう (例えば, TV の画像は $(500 \sim 700) \times 480$ 画素で構成される)。

マウス 簡易入力装置。画像表示されたカーソルを動かして指示をするので, 操作がたやすい。マウスに付いているボタンを押すことで, カーソルの位置を入力することができる。

ポップアップメニュー マウスのボタンを押すことにより, カーソルの位置に表示されるメニュー。

極座標 平面上の点は座標 (r, θ) で表される。 θ は, x 軸の正の向きと原点からその点までの線分のなす角, r は線分の長さを表す。空間内の点は, 原点からの距離 r と緯度, 経度の三つで位置が定められる。

正規化 ある基準に従って一定の大きさに変換すること。

ベクトル 大きさと方向を持っている量。