

基于三维实体特征的机械零件搜索引擎设计^①

黎养明, 仲梁维

(上海理工大学 机械工程学院, 上海 200093)

摘要: 随着各种三维 CAD 系统在产品设计和制造领域的普及应用, 基于内容的三维模型检索技术应运而生。介绍了三维模型搜索引擎的发展概况, 三维模型搜索引擎的关键技术。着重介绍三维实体模型中的归一化, 特征提取, 相似性比较的算法。

关键词: 三维模型; 搜索引擎; 归一化; 特征提取; 相似度匹配

Search Engine Design of Entity Feature Based on Three-Dimensional Mechanical Parts

LI Yang-Ming, ZHONG Liang-Wei

(College of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: With the popularity of 3D CAD systems in product design and manufacturing industry, the techniques of content-based 3D model retrieval become necessary. This paper addresses the development of 3D model search engine. It presents the principle of the development of 3D model search engines and highlights three-dimensional solid model normalization, feature extraction, similarity comparison algorithm.

Keywords: model; search engine; normalization; feature extraction; similarity matching

人类生活在一个三维的物质世界中, 摄影技术只能获得物体的平面影像, 而无法记录物体的三维几何特征。计算机中三维实体模型的一种最基本的表示方式是利用大量多边形来逼近物体的三维外形。如何利用它们的几何形状特征来组织数据库, 方便找到所需要的三维实体模型就成了一个值得研究的问题。

1 三维搜索引擎的发展现状和应用前景

三维实体模型搜索可以分为基于文本和基于内容。三维实体模型搜索引擎的设计思想是通过搜索网上和数据库中的三维特征数据, 向用户提供文本输入和三维实体模型的查询接口, 采用基于关键字和基于几何形状特征实现对实体模型的搜索。随着三维实体特征建模技术的日益成熟和各种 CAD 软件的广泛应用, 三维机械零件模型的数量聚集性的增长。已有三维实体模型数据资源的充分利用, 可以大大减轻新模型设计中工作量, 同时三维数据的流通也可以促进在

各领域的应用: (1)加工成本估计; (2)现有零件重用; (3)开发先进智能的 PDM; (4)开发基于 Internet 搜索引擎; (5)应用于机器人视觉技术。

2 三维搜索引擎的主要构成

三维实体搜索引擎的总体框架应由三部分组成: 一个对应数据库或文件中记录每个三维实体模型所提取的特征数据的模块; 提取输入三维实体几何特征数据的模块; 将提取的输入三维实体特征在数据库或在网络中进行查找匹配和将搜索的结果的输出模块。为方便用户使用, 一个三维实体模型搜索系统应包括多方式的界面设计。多方式是指允许用户以不同的手段进行搜索, 包括利用关键字搜索, 利用已存的三维实体模型搜索等等。由于利用关键字进行匹配的算法和传统的字符串匹配算法是一致的, 因此文中将讨论利用已有的三维实体模型进行搜索的算法。

^① 基金项目:上海市(教委)重点学科建设项目(J50503)

收稿时间:2010-07-15;收到修改稿时间:2010-08-12

3 三维实体模型特征的提取

为了提取的数据便于计算,用户更直观的看到模型之间的相似度关系,对模型进行必要的归一化预处理。三维实体模型的文件数据中通常只包含实体模型的表面几何特性(顶点坐标、法向量、拓扑连接等)和外观特性(实体颜色、材料属性等),没有适合自动匹配的特征数据。如何合理地描述三维实体模型^[1](即特征数据提取)成为三维模型搜索首先要解决的问题,也是三维实体模型搜索的难点。特征数据必须满足:易于表达和计算;方便相似性匹配计算;特征数据必须具有惟一性,即不同的模型对应的特征数据应不相同。

3.1 三维实体模型进行坐标系归一化预处理

目前三维实体模型相似性的预处理^[2]包括坐标系归一化预处理、坐标系平移变换归一化、坐标系比例变换归一化和坐标系旋转归一化三个预处理方面。

3.1.1 坐标系平移变换归一化

在平移变换前后,三维实体模型的特征要保证不变,由于实体模型的形心是一个确定点位置,只要通过将坐标系原点移到三维实体的形心位置就可以实现平移变换归一化处理。这里求出的 T 是三维实体模型的形心,将三维实体模型的坐标原点移动到形心 T 点,假如原来实体模型的顶点集合为 p ,计算三维模型形心的方法如公式:

$$T = \frac{\sum_{i=1}^n p_i}{n}, p_i \in P \quad (1)$$

顶点集合 p 形成新的顶点集合 $P = \{ p \mid p = p - T, p \in P \}$ 。

3.1.2 坐标系比例变换归一化

三维实体模型的特征要保证在缩放变换前后不变,需要计算大小变换因子来变换三维实体模型。假如旋转因子为 s , 其中 px, py, pz 为个坐标轴单位长度的单位值,比例因子的求取如下公式:

$$s = \sqrt{(p_x^2 + p_y^2 + p_z^2)} / 3 \quad (2)$$

计算出比例变换后的点集 $P = \{ p \mid p = ps, p \in P \}$ 。

3.1.3 坐标系比例变换归一化坐标系旋转归一化

坐标系旋转归一化处理是三维实体模型坐标系归一化处理中最难解决的问题之一。在线性代数中,对于二次曲线通常可以利用一个单位正交矩阵做一定角度的旋转变换,把二次曲线转化为标准形式。类似地,我们先计算所有顶点(组成三维实体特征的三角片面的顶点与形心之间的距离)的最大距离,利用形心与最大

距离点形成向量,然后对最大距离向量或向异性变换的方法进行旋转归一化(例如标准椭圆以长轴作为 X 轴)。假设旋转矩阵为 $R1$,求取旋转矩阵的公式如下(3):

$$R1 = \frac{\sum_{i=1}^n p_i^T p_i}{\sum_{i=1}^n \|p_i\|^2}, p_i \in P \quad (3)$$

$R1$ 是个 3 阶的实对称矩阵,实对称矩阵不同特征值的特征向量相正交,因此该矩阵的特征根是 3 个不同的正实数。求出矩阵的特征根,以降序排列 3 个特征根。然后求出 3 个特征根对应的特征向量,并将每个特征向量变为单位向量,最后 3 个特征根的单位向量组成旋转矩阵三阶矩阵 R 。利用 MATLAB 计算矩阵的特征向量与特征, D 为特征值组成的对角矩阵, R 为所求的旋转矩阵即正交矩阵有正交变换长度不变性, M 为特征值对角矩阵。使用 $[R,D]=eig(M)$ 求得的旋转矩阵 R 变换三维实体模型中的点,计算出旋转变换后的点集 $P = \{ p \mid p = pR, p \in P \}$ 。

通过平移向量 T 、大小变换因子 S 和旋转变换矩阵 R 的变换,得到新的顶点集合 P 完成预处理。(这些计算用 MATLAB 编辑成 m 文件函数,再转换成 $c++$ 代码)

3.2 三维实体模型特征数据的提取

三维实体模型特征数据的提取需要考虑以下三个方面:第一,能够反映三维实体模型的内在主要特性;第二,特征数据尽量与实体模型的平移、大小和旋转变换无关,因为三维实体模型在空间中的位置、方向和大小是任意的;第三,对于退化的三维实体模型,特征数据也是稳定的,噪声的存在和实体模型细分对其稳定性影响很小。下面对特征数据提取进行简单的说明,根据不同几何实体形体表面顶点间的相互关系呈现出不同的分布情况,将一个任意的三维实体模型中复杂的特征提取转换成相对简单的几何形状概率分布问题。几何形状分布特征数据主要利用实体模型表面随机顶点间的几何特征关系,特征关系如实体模型表面截面顶点的曲率概率分布、任意三个随机顶点间面积的概率分布等不同的统计特征有着不同的特点。而且经过预处理之后可以满足几何不变性。由于几何形状分布方法用实体模型表面顶点的特征关系替代了实体模型几何体的特征,满足不同的模型呈现出不同的曲线特性,以此作为特征数据具有惟一性。该方法计算简单,具有不变性,对实体模型边界的一些小扰动具有一定的鲁棒性。例如表示模型表面上任意两点间的欧氏距 D (模型表面三个随机顶点间的三角面积、形心至任意顶点间的距离、模型表面四个随机顶点间的体积等)的

概率分布曲线。提取模型的表面特征,对随机采样点的均匀分布要求很高,因此需要对模型进行细分。细分算法比较简单,将三角面片的重心找到,然后把重心与原来三角形的三个顶点组合形成 3 个新三角片,三角面片继续循环找重心 N 次形成 3 的 N 次幂个小三角面片。在模型细分后,按照等截面原则在模型表面选取随机点。随机点数量的增多,形状分布曲线将趋向于一个稳定值。最能反映三维实体的形状分布曲线或考虑多形状分布曲线的综合,可以使相似度计算准确性提高。

4 三维实体模型特征相似度匹配算法

利用提取的特征数据采取几何形状分布曲线计算其相似度。几何形状分布曲线是一种简单而有效的三维实体特征的相似性度量算法。其主要是通过形状分布函数即采用物体表面上随机采样点间几何特征的概率分布作为形状比较的基础。它主要的优势是简便性,这种算法直接从物体表面获取特征数据就可以进行相似性的比较,具有较高的鲁棒性。对于模型可以先考虑其对称型,对于一个及其不对称的模型和一个对称度很高的模型我们认为其相似度很小,由于当今主流的三维软件,其实体表面都是以三角片来趋近于模型,例如一个长方体与一个正方体它们是对称,我们定义所有三角片法矢量和为其对称性的描述,我们可以很容易知道他们都由十二个三角片组成,法向矢量和为零,当我们将正方体与正方体进行模型的归一化处理后,两个模型的矢量和的夹角为零,两者对称度很高,据常识来看,也认为长方体与正方体比较相似,如果两模型的法矢量和的空间最小角度大于四十五度,不再计算,认为其相似度太低。定义了如下形状函数来测量三维模型,在高等代数中,计算实体的体积时有一种叫截面法,对于三维实体的相似度我们也以横截面的曲率分布线来计算,两个模型的同等截面进行曲率分布线的匹配计算,对于横截面曲线我们采取尽可能选取多的点用任意多段直线来逼近,然后找到它们的形状中心。对同等截面进行形状分析比较,最后画出统计直方图计算其相似度。其适用于各种三维实体文件格式。该方法的缺点是用一条曲线来表征实体的形状,这是不够的。例如两个模型外形相近但内部细节不同的三维实体具有较为相似的形状分布曲线,因此需要增加一些信息来反映实体的几何特征看作是另一种形状分布曲线,进而提高算法对不同模型的区分能力。经过对截面的研究,注意到截面可以分为两种情况。图 1、图 2 为两种实体截面图(有中空和实心)根据这

种情况可将横截面分成两类:一类是中空截面穿过实体模型,称为 H 类;一类是实心截面模型,称为 S 类。

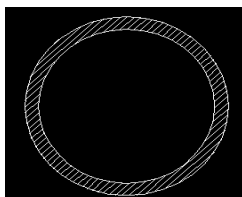


图 1 中空截面

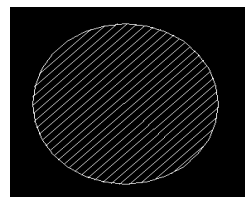


图 2 实心截面

在实验中发现仅通过计算模型的这两类截面对在整个截面中的百分比,就可以发现模型间的相互关系。根据形状分布曲线的不同,利用曲线比较求出其距离差值。综合形状分布曲线与截面百分比的关系,可得到其三维模型的相似度值。

5 基于三维实体特征搜索的具体实现步骤

利用上面介绍的特征提取和相似度算法,具体步骤可如下:

5.1 获得模型的三角片数据

STL 文件格式是由 3D SYSTEM 公司于 1988 年制定的一个接口协议, STL 文件由多个三角形面片的定义组成,每个三角形面片的定义包括三角形各个定点的三维坐标及三角形面片的法矢量。三角形顶点的排列顺序遵循右手法则。STL 文件有 2 种类型: ASC 域格式和二进制格式 ASC 域格式。STL 模型是以三角形集合来表示物体外轮廓形状的几何模型。例如将三维实体模型输出为 VRML 或 STL 等数据格式(例如 CATIA 软件的模型,零件文件另存为 STL 文件格式,采取 ASCII 形式的 STL 文件)。G 近似于实体模型 S, $G = \{t_1, t_2, \dots, t_n\}$ (t_i 为一个三角片数据),将 CAD 的实体模型三角片化。

5.2 在近似实体模型 G 的表面生成均布的随机点

将 G 读入计算机中,遍历实体模型表面的所有三角片,计算每一个三角形的面积 $A(t_1)$,将已经访问到的三角形面片的面积和记录在一个数组中。该数组存放三角形累积面积,该数组的维数即为该模型的三角面片个数。(c++ 语言读取 STL 文件代码

```

BOOL LoadSTL File (LPCTSTR stlfile) {FILE f; if ((f = fopen(stlfile, "r")) == NULL) return false; char str[80]; vertex 3 v1, v2, v3; facet 3 f; while (fscanf(file, "%s", str) == 1) {if (strcmp(str, "normal") == 0) {v1 = (vertex 3) malloc(sizeof(vertex)); v2 = (vertex 3) malloc(sizeof(vertex)); v3 = (vertex 3)

```

```

malloc ( sizeof ( vertex ) ); f = ( facet 3 ) malloc ( sizeof
( facet ) ); // 读取面片的法向量 f scanf ( file , " %1f %1f
%1f" , &(f - > x) , &(f - > y) , &(f - > z) ); f scanf
( file , " %3 s %3 s" ); // 读取三角面片的三个顶点信息
到 v1 ,v2 ,v3 三个顶点计算三边距离为 a、 b、 c， 然
后利用海伦公式 (S=根号下 {s(s-a)(s-b)(s-c)}), 其中
s=(a+b+c)/2) 计算其面积。三角形重心为 (v1+v2+v3)/3。
利用该公式可以细化模型与得到模型表面的随机点。根
据法向量来判断随机点是否属于中空截面。

```

5.3 曲率求取

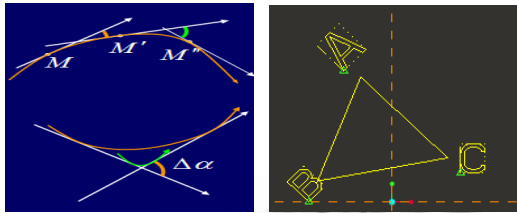


图3 理论上求取曲率 图4 近视求取曲率

图3为理论上的曲率求取，由于在计算中实体表面都是用离散点来逼近的，所以用图4近视求取曲率，利用相邻的三点组成的三角形来求解。设A点坐标为(x1,y1),B点坐标为(x2,y2),c点坐标(x3,y3)。曲线的弯曲程度与切线的转角有关，A点处的转角变化用其中的 Δs 表示 360除以截面边界所取点数量 n(以等同截面取得点多数者为 n，即为最小刻度)。曲率计算公式如下：

$$\frac{y_1 - y_2}{x_1 - x_2} - \frac{y_2 - y_3}{x_2 - x_3} \quad \Delta s \quad (4)$$

对于截面的轮廓线我们用离散点直线相连，当离散点越多时曲线越精确于理论曲线。

5.4 计算形状分布直方图

计算实体模型截面任意的面积和形状曲率参数，对其进行统计，得到一条分布曲线，称为该三维实体的几何形状分布曲线。为了比较不同实体，需要对每一个实体模型的曲线转换为直方图。首先计算所有最大值和最小值，将最大值减去最小值的值等分成 S 个等距离单元，计算落在每个距离单元内的曲率差距个数，根据落入固定距离范围的概率计算出对应的形状分布直方图 h。(用 MATLAB 与 c++混合编程画直方图，x 为组成的数组,假设划分为九个区，部分代码如下： a=[min(x):(max(x)-min(x))/9:max(x)];

```

for i=1:8b(i)=sum(x>=a(i)&x<a(i+1)); end
b(9)=sum(x>=a(9)&x<=a(10));
b1=b/200;
a1=[min(x)+(max(x)-min(x))/18:(max(x)-min(x))/9:max(
x)-(max(x)-min(x))/18]; bar(a1,b1))

```

5.5 使用曲线匹配技术比较分布直方图

对于分布直方图 h1 和 h2，其比较计算公式如下：

$$D(h1,h2) = \frac{\sum_{i=0}^n |h1i - h2i|}{n} \quad (5)$$

由于将实体模型截面分两类,因此在基本方法的基础上加入了权值作为改进。设 A 和 B 表示两个模型, S 和 H 表示两种模型同类截面所形成的直方图, 两直方图的权值距离计算公式如下(6):

$$\frac{1}{1-|\%f_h - \%g_h|} D(A_h, B_h) + \frac{1}{1-|\%f_s - \%g_s|} D(A_s, B_s) \quad (6)$$

下面图5是三个 catia 模型，由于他们都是对称的零件模型，所以他们的对称性向量和均为零，对其三角片的顶点计算，以第一个模型为比较模型，相似度分别为 35.236%和 29.963%。利用以上介绍的算法，其结果具有很高的可信度。

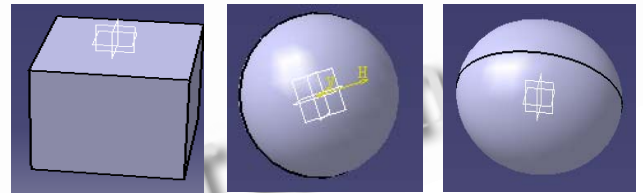


图5 三个互相比对的模型正方体、球体、椭圆柱

6 结束语

由于三维实体模型很难由一个数学公式来表达，而且相互之间的比较具有一定的主观性。基于三维实体特征搜索引擎，采用大量的形状函数分布来表述三维实体的内在，对模型特征的提取与模型实体相似度的比较算法是基于大量的数据统计，该方法具有可靠性和客观性。

参考文献

- 1 崔晨吻,石教英.三维模型检索中的特征提取技术综述.计算机辅助设计与图形学报,2004,16(7):882-889.
- 2 章志勇.三维模型几何相似性比较的研究[博士学位论文].杭州:浙江大学,2004.226.