

# 基于随机漫步图匹配的三维 CAD 模型相似评价\*

唐韦华, 刘玉生\*

(浙江大学 CAD & CG 国家重点实验室, 浙江 杭州 310058)

**摘要:** 为了提高三维 CAD 模型检索的效率, 提出了基于随机漫步图匹配的相似评价算法。使用图转化函数消除特征依赖图的末端节点; 使用概率归一化函数消除特征间的不合理干扰, 实现了对三维 CAD 模型的高效相似评价。实验结果表明, 该方法具有较好的有效性和较高的效率。

**关键词:** 三维 CAD 模型; 相似评价; 随机漫步; 图匹配

**中图分类号:** TH164; TP391.72

**文献标识码:** A

**文章编号:** 1001-4551(2010)07-0043-04

## Similarity assessment of 3D CAD model based on random walks graph matching

TANG Wei-hua, LIU Yu-sheng

(State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310058, China)

**Abstract:** Aiming at enhancing the efficiency of 3D CAD model retrieval, a new similarity assessment algorithm based on random-walks was proposed. Graph translation function was used to eliminate the sink nodes in feature dependency graph. Probability normalization function was used to eliminate the unreasonable disturbance among features. 3D CAD models were efficiently retrieved. Experimental results demonstrate that the method is efficient and effective.

**Key words:** 3D CAD model; similarity assessment; random walks; graph matching

## 0 引 言

在机械工程领域, 超过 75% 的设计活动是变异式设计或基于案例的设计<sup>[1]</sup>, 设计重用可以有效地提高此类设计的效率。三维 CAD 模型含有大量的设计知识, 重用它们可以有效地支持设计重用。随着三维 CAD 模型数量的急剧增长, 三维 CAD 模型的检索受到了越来越多的关注。

基于图的检索方法正是三维 CAD 模型检索中非常重要的一个研究分支。目前常见的方法有基于 Reeb 图的方法<sup>[2]</sup>、基于骨架图的方法<sup>[3]</sup>、基于特征依赖图的方法<sup>[4]</sup>等。然而, 受限于图匹配的复杂度, 现有的这类方法效率往往很低。为了满足检索的需求, 如何提高三维 CAD 模型相似评价的效率引起了大家

的关注, 如 Li 等提出的基于知识通过预筛选缩减匹配空间的方法<sup>[5]</sup>; Ramesh 等提出的将加工特征分解成 7 个向量进行比较的方法<sup>[6]</sup>; Srinivas 等提出的基于特征类型抽象层次的方法<sup>[7]</sup>; Iyer 等提出的基于成组技术的方法<sup>[8]</sup>; Bai 等提出的基于层次骨架图的方法<sup>[9]</sup>等。虽然这些算法在一定程度上改善了相似评价的效率, 但在最差情况下其复杂度仍为指数级, 并未从根本上解决该问题。

本研究针对三维 CAD 模型检索的需要, 提出一种基于随机漫步图匹配<sup>[10]</sup>的三维 CAD 模型相似评价算法。其目标是通过图转换函数消除特征依赖图中的末端节点; 通过概率归一化函数消除特征间的不合理干扰; 凭借随机漫步图匹配较低的时间复杂度达到高效的相似评价。

收稿日期: 2010-01-15

基金项目: 浙江省基金重点资助项目(Z107497)

作者简介: 唐韦华(1983-), 男, 四川成都人, 硕士研究生, 主要从事三维 CAD 模型检索与重用方面的研究。E-mail: mew2\_149@163.com

通信联系人: 刘玉生, 男, 研究员, 博士生导师。E-mail: ysliu@cad.zju.edu.cn

# 1 基于随机漫步的三维 CAD 模型相似评价

## 1.1 基于图变换函数的末端节点消除

随机漫步图匹配算法虽然具有较高的效率,但是难以直接应用于三维 CAD 模型的相似评价中。这是因为三维 CAD 模型有大量的末端特征,反映在特征依赖图上就是大量的末端节点(即出度为 0 的节点),而该算法无法匹配具有末端节点的图。为此,本研究通过图转换函数将一对待匹配的特征依赖图转换为另一对不含末端节点,且保持同构性的新图,从而将原始图的匹配转化为新图的匹配。具体地说,对任意给定图  $G = \{E, V\}$ , 设  $s_i$  是原图中的末端节点,则转换函数  $g(G) = \{V', E'\} = G'$  可定义如下:

$$\begin{aligned} V' &= V \cup \{p\}, p \text{ is an added node} \\ E' &= E \cup \{e'_1, e'_2, e'_3 \dots e'_s\} \cup \{(p, p)\}, e'_i = (s_i, p) \end{aligned} \quad (1)$$

该函数具有如下性质:①  $G_1$  和  $G_2$  同构  $\Leftrightarrow G'_1$  和  $G'_2$  同构,匹配  $G_1$  和  $G_2$  等价于匹配  $G'_1$  和  $G'_2$ ;②  $G'$  中不存在末端节点,可直接应用随机漫步图匹配算法匹配。因此将  $G$  转换为  $G'$  匹配是合理的。

证明如下:

(1)  $G'$  中的节点可分为两类:原图中的节点  $n$  和新增的节点  $p$ ,可分别证明其不是末端节点。

对原图中的节点  $n$ ,由  $G'$  定义可知:

$$D_{\text{out}}(G', n) = \begin{cases} D_{\text{out}}(G, n) & n \text{ is not a sink node} \\ D_{\text{out}}(G, n) + 1 & n \text{ is a sink node} \end{cases} \quad (2)$$

式中  $D_{\text{out}}(G, n)$ —节点  $n$  在图  $G$  中的出度。

又:

$$D_{\text{out}}(G, n) = \begin{cases} k (k > 0) & n \text{ is not a sink node} \\ 0 & n \text{ is a sink node} \end{cases} \quad (3)$$

所以,  $\forall n, n \in V, D_{\text{out}}(G', n) > 0$ 。

对新增的节点  $p$ ,根据  $G'$  定义有  $D_{\text{out}}(G', p) = 1$ 。

因此  $G'$  中没有末端节点。

(2) 因  $G_1, G_2$  同构,所以  $G_1 = PG_2P^T$ , 由定义知:

$$G'_1 = \begin{bmatrix} G_1 & M \\ 0 & 1 \end{bmatrix}, G'_2 = \begin{bmatrix} G_2 & N \\ 0 & 1 \end{bmatrix} \quad (4)$$

式中  $M, N$ —值为 0 或 1 的行向量。

记  $G'_1$  和  $G'_2$  的节点分别为  $u$  和  $v$ ,有:

$$\begin{aligned} m_i &= 1 \Leftrightarrow u_i \text{ is a sink node} \\ n_j &= 1 \Leftrightarrow v_j \text{ is a sink node} \end{aligned} \quad (5)$$

令  $P' = \begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix}$ , 有:

$$P'G'_2(P')^T = \begin{bmatrix} PG_2P^T & PN \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} G_1 & PN \\ 0 & 1 \end{bmatrix} \quad (6)$$

由于  $P$  是同构变换矩阵,任意  $N$  中元素  $n_j$  为:

$$n_j = k_i = \begin{cases} 0 & v_j \text{ is not a sink node} \\ 1 & v_j \text{ is a sink node} \end{cases} \quad (7)$$

式中  $i$ —在变换  $P$  下与  $v_j$  匹配的  $u_i$  的下标;  $k_i$ — $PN$  中的第  $i$  个元素。

那么  $v_j \text{ is a sink node} \Leftrightarrow u_i \text{ is a sink node} \Rightarrow k_i = n_j = m_i \Rightarrow PN = M$ , 所以有:

$$PG'_2(P')^T = \begin{bmatrix} G_1 & M \\ 0 & 1 \end{bmatrix} = G'_1 \quad (8)$$

因此  $g(G)$  没有改变原始图的同构性。

## 1.2 基于概率归一化函数的不合理干扰消除

随机漫步图匹配算法里节点的特征量取决于该节点在图中的整体拓扑地位,因此只要有新节点加入,其他节点的特征量一定会改变。在特征依赖图中,这个性质造成了特征间的不合理干扰。为此,本研究基于特征量  $x^*$  的线性变化性质,通过概率归一化函数将随机漫步所求出的特征量进行归一化,消除该干扰。

线性变化性质可描述如下:设有图  $G = \{E, V\}$ , 若图  $G' = \{E', V'\}$  满足  $V' = V \cup \{n_{\text{new}}\}$ ,  $E' = E \cup \{(p_{\text{new}}, n_{\text{new}})\} = 1$ , 则:

$$\forall n \in G: p_{\text{new}} \text{ 到 } n \text{ 非连通} \Rightarrow x_{n'}^*/x_n^* = k \quad (9)$$

式中  $p_{\text{new}}-n_{\text{new}}$  惟一的前驱节点;  $k$ —常数。

该性质说明若新加入的特征  $f_{\text{new}}$  依附于特征  $f_p$ , 那么任意与  $f_p$  不连通的特征  $z$  在原图和新图中的特征量成常数比。定义概率归一化函数  $f(x_1^*, x_2^*) = [(x_{1f}^* i, x_{2f}^* i)]$  如下:

$$\begin{aligned} x_{1f}^* i &= \frac{\text{Min}(|x_1^* i|, |x_2^* i|)}{\text{Max}(|x_1^* i|, |x_2^* i|)} \cdot x_1^* i \\ x_{2f}^* i &= \frac{\text{Min}(|x_1^* i|, |x_2^* i|)}{\text{Max}(|x_1^* i|, |x_2^* i|)} \cdot x_2^* i \end{aligned} \quad (10)$$

使用该函数对特征量进行归一化,就可以消除特征图中新特征对原有特征的不合理干扰。对  $x^*$  线性变化性质的证明如下:

令集合  $V_o = V - \{p_{\text{new}}\}$ , 则  $V_o$  可划分为两部分:

$$V_o = V_{o1} + V_{o2}$$

$$V_{o1} = \{n \in V_o \mid \text{there is no path from } p_{\text{new}} \text{ to } n\} \quad (11)$$

$$V_{o2} = \{m \in V_o \mid \text{there is a path from } p_{\text{new}} \text{ to } m\}$$

则有  $\forall n \in V_{o1}, \forall m \in V_{o2}, \text{there is no edge}(m, n)$ 。

因为若存在边  $(m, n)$ , 则在  $p_{\text{new}}$  到  $m$  的路径末尾添加

边 $(m, n)$ 可得到 $p_{\text{new}}$ 到 $n$ 的路径,与 $V_o$ 的定义矛盾。

另一方面,令 $W = [w_{ij}]$ 满足:

$$w_{ij} = \begin{cases} 1/D_{\text{out}}(G, i) & \text{there is an edge}(j, i) \text{ in } G \\ 0 & \text{edge}(j, i) \text{ doesn't exit in } G \end{cases} \quad (12)$$

则求 $x^*$ 和 $x'^*$ 分别等价于求解这两个方程组:

$$\begin{aligned} (E - d \cdot W)x^* &= \frac{(1-d)}{N} \mathbb{I} \\ (E - d \cdot W')x'^* &= \frac{(1-d)}{N+1} \mathbb{I} \end{aligned} \quad (13)$$

因为 $E' = E \cap \{(pa(n_{\text{new}}), n_{\text{new}})\}$ , 则:

$$D_{\text{out}}(G', i) = \begin{cases} D_{\text{out}}(G, i) & i \neq p_{\text{new}} \\ D_{\text{out}}(G, i) + 1 & i = p_{\text{new}} \end{cases} \quad (14)$$

按照定义,可以得到:

$$\forall i \in G: p_{\text{new}} \notin pa(i) \Rightarrow \forall j \in G, W_{ij} = W'_{ij} \quad (15)$$

进一步得出:

$$\forall i \in G, \forall j \in G: i \in V_{o1}, j \in V_{o2} \Rightarrow W_{ij}' = W_{ij} = 0 \quad (16)$$

$\forall i \in V_{o1}$ , 记 $W$ 和 $W'$ 中第 $i$ 行向量对应的方程分别为 $E_{o1}$ ,  $E_{o1}'$ , 那么由上可知它们的变量系数相同, 且均只含 $V_{o1}$ 中的节点概率。同时,  $E_{o1}$ 和 $E_{o1}'$ 的常数项分别是 $\frac{(1-d)}{N} \mathbb{I}$ 和 $\frac{(1-d)}{N+1} \mathbb{I}$ 。记 $E_{o1}$ 的解为 $X_{o1}$ , 若令 $k = \frac{N}{N+1}$ , 那么 $X_{o1}' = X_{o1} \cdot k$ 显然是 $E_{o1}'$ 的解, 得证。

### 1.3 三维 CAD 模型相似评价算法

通过图转换函数消除特征依赖图的末端节点和通过概率归一化函数消除特征间的不合理干扰后, 本研究分析三维 CAD 模型的特征依赖图, 发现其具有以下特点:

(1) 节点分为基体节点和普通节点两类。有且仅有一个基体节点;

(2) 基体节点表示该模型的基体特征, 具有重要的设计语义;

(3) 节点的类型属性表征对应的特征的类型, 具有重要的设计语义;

(4) 普通节点与基体节点的距离含有重要的拓扑信息, 表达了重要的设计语义。

针对特征依赖图的上述特点, 可得到以下匹配原则:

- (1) 基体节点只能与基体节点匹配;
- (2) 类型属性不相同的节点不能匹配;
- (3) 与基体节点距离不同的节点不能匹配。

按照上述匹配原则, 本研究提出了基于随机漫步图匹配的三维 CAD 模型相似评价算法, 其具体流程如下:

(1) 求解节点特征量。通过 $g(G)$ 将图 $G_1, G_2$ 转

化为 $G'_1, G'_2$ 。对图中的节点 $p$ , 以 $x(j|q)$ 表示点 $q$ 采取跳转的概率, 取常数 $d$ , 以 $x(p|q, l)$ 表示从点 $q$ 跳转到点 $p$ 的概率, 记 $M(e)$ 为边 $e$ 的权重, 令 $x(p|q, l) = \frac{M(e_{qp})}{\sum_{k \in ch(q)} M(e_{qk})}$ , 则概率跳转函数形成的马尔科夫链如下:

$$\begin{aligned} x_p(t+1) &= \sum_{q \in G} x(p|q, j) \cdot x(j|q) \cdot x_q(t) + \\ &\quad \sum_{q \in pa(p)} [1 - x(p|q, j)] \cdot [1 - x(j|q)] \cdot x_q(t) \end{aligned} \quad (17)$$

求解该式的稳态分布, 再以概率归一化函数进行归一化, 即为所需的特征量 $x^*$ 。

(2) 对节点进行分类。遍历 $G'_1, G'_2$ , 记录下基体节点 $n_b$ 到每个节点 $n$ 的最短路径, 记为 $Td(n)$ 。将所有 $Td$ 值相同且特征类型相同的节点归为一类, 记为 $C$ 。

(3) 求解节点间的最佳匹配。为每一个分类 $C$ 构造二部图 $BG^c: BG_{ij}^c = S(a_i, a_j, x_i^*, x_j^*)$ ,  $S$ 表示节点 $i, j$ 的相似度, 其值由属性 $a$ 及特征量 $x^*$ 决定。利用二部图的最佳匹配算法<sup>[11]</sup>求出每个 $BG$ 的最优匹配, 即可得到所有的配对节点。

(4) 计算最终相似度。对所有配对节点的相似度 $S$ 求加权和, 即为最终的相似度。

### 1.4 算法分析

求解特征量的复杂度取决于解 $x^*$ 的方法, 采用 $LU$ 分解法可达到 $o(N^3)$ 。对节点分类的复杂度为 $o(N)$ 。记分类数为 $N_c$ , 分类 $C_i$ 节点数为 $N_{ci}$ , 则 $\sum_{i=0}^{N_c} N_{ci} = N$ , 求最佳匹配的复杂度为 $\sum_{i=0}^{N_c} o(N_{ci})^3 \leq o(N)^3$ 。求 $D$ 的复杂度为 $o(N)$ 。综上, 复杂度为 $o(N^3)$ 。





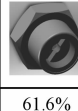
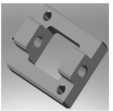

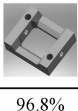









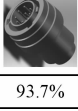
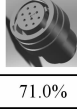

## 2 实验结果及分析

本研究以 ACIS 为几何引擎, MySQL 为数据库系统, 在 Microsoft Visual C++ 2003 开发环境下实现了本研究的算法, 并在 Pentium4 1.6 GHz CPU, 2.0 GB DDR3 内存, 80 GB 硬盘的 PC 机上进行了测试。下面的两组测试结果分别用来测试本研究算法的有效性和效率。

### 2.1 有效性测试

在实验中, 笔者建立了一个包含 236 个三维 CAD 模型及对应特征依赖图的数据库, 以特征类型及特征尺寸为节点属性, 进行相似评价测试。实验结果如表 1 所示(表 1 中左侧表示输入模型; 右侧表示检索的结果, 下方数字表示计算出的相似度)。

表 1 三维 CAD 模型检索结果

输入模型	检索结果模型			
				
a	100%	87.1%	64.7%	61.6%
				
b	100%	96.8%	81.2%	70.4%
				
c	100%	89.3%	72.1%	67.9%
				
d	100%	93.7%	71.0%	65.4%

从表 1 可以看出,本研究的相似评价算法有如下特点:①模型和本身的相似度一定为 100%;②局部特征的差异不会干扰其他特征的相似度,如输入模型 b 与第 2 个检索结果的差别仅在局部特征,因此它们相似度较高;③相似度较好地反映了拓扑和形状的综合信息,如输入模型 c 和第 2 个检索结果拓扑相似,尺寸也较为接近,因此它们具有较高的相似度。

### 2.2 效率测试

在实验中,本研究随机生成了节点数为 20, 40, 80, 普通节点到基体节点最大距离分别为 1, 3, 6 的多个特征依赖图,结果如图 1 所示(纵坐标采用对数时间轴,横坐标表示特征依赖图节点的数目;3 条线记录最大距离分别为 1, 3, 6 的特征依赖图匹配 100 次所需的时间,单位为 s)。

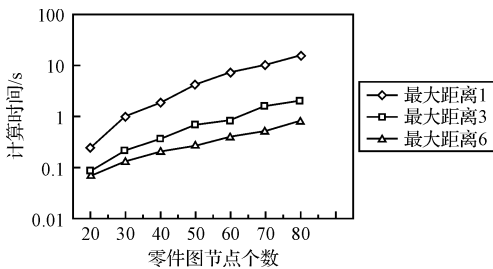


图 1 相似评价算法效率测试图

实验结果表明,在最差情况下,即特征依赖图中普通节点与基体节点的最大距离为 1 时,因为无法按照拓扑距离对节点进行分类,其计算时间与节点个数接近简单的立方关系,远远低于指数级,这说明随机漫步图匹配的高效性得到了保留。随着最大距离的增加,其匹配时间的增长率明显下降。这说明通过引入“按最大距离分组”的原则明显缩减了匹配空间,显著提升了匹配效率。在实际应用中,普通节点到基体节点

的最大距离为 1 的特征依赖图一般比较简单,算法常数因子很小。而在另外两条曲线中,当节点数增长到 80 时,计算时间在 1 s 左右,因此总体上本研究的相似评价算法效率较高。

### 3 结束语

本研究提出了基于随机漫步图匹配的三维 CAD 模型相似评价算法。通过图转换函数消除了特征依赖图中的末端节点;通过概率归一化函数消除了特征间的不合理干扰。该算法具有多项式级的复杂度,实现了对三维 CAD 模型的高效相似评价。实验结果表明,该方法具有较好的有效性和较高的效率。

### 参考文献 (References):

- [1] ULLMAN D G. The Mechanical Design Process[M]. New York: McGraw-Hill,1997.
- [2] BESPALOV D, REGLI W C, SHOKOUFANDEH A. Reeb Graph based Shape Retrieval for CAD[C]. ASME DETC 2003, Chicago, Illinois, United States,2003.
- [3] GAO W, GAO S M, LIU Y S, et al. Multi resolutional similarity assessment and retrieval of solid models based on DBMS[J]. *Computer-Aided Design*,2006,38(9):985 - 1001.
- [4] CICIRELLO V A, REGLI W C. Resolving Non-uniqueness in Design Feature Histories[C]//Proceedings of the Fifth ACM Symposium on Solid Modeling and Applications. New York: [s. n. ],1999:76 - 84.
- [5] LI M, FUH J Y H, ZHANG Y F, et al. General and Partial Shape Matching Approaches on Feature-Based CAD Models to Support Efficient Part Retrieval[C]//ASME Conference Proceedings, 2008. New York: [s. n. ],2008:121 - 130.
- [6] RAMESH M, YIP-HOI D, DUTTA D. Feature based shape similarity measurement for retrieval of mechanical parts[J]. *Journal of Computing and Information Science in Engineering*,2001,1(3):245 - 256.
- [7] SRINIVAS G, FASSE E D, MAREFAT M M. Retrieval of Similarly Shaped Parts from a CAD Database[C]//IEEE International Conference on Systems, Man, and Cybernetics. San Diego: [s. n. ],1998:2809 - 2814.
- [8] IYER S, NAGI R. Automated retrieval and ranking of similar parts in agile manufacturing[J]. *IIE Transactions*, 1997,29(10):859 - 876.
- [9] 白 静,唐韦华,刘玉生,等. 面向实体模型相似评价的层次图生成与高效匹配研究[J]. *计算机辅助设计与图形学报*,2009,21(7):869 - 879.
- [10] MAGGINI G M, SARTI L. Exact and approximate graph matching using random walks[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions*,2005,27(7): 1100 - 1111.
- [11] KUHN H W. The hungarian method for the assignment problem[J]. *Naval Research Logistics Quarterly*,1955, 2(1-2):83 - 97.