

DOI:10.3969/j.issn.1001-4551.2021.04.008

三维 CAD 模型检索方法的优化研究*

田 柳^{1,2}, 徐赐军^{2*}, 黄 浩¹, 张绍峰¹, 陈 磊¹

(1. 武汉科技大学 机械自动化学院, 湖北 武汉 430081; 2. 湖北理工学院 机电工程学院, 湖北 黄石 435003)

摘要:随着三维 CAD 模型库模型数量的增加,会导致其检索效率大幅度降低,针对这一问题,提出了一种基于机械结构形态的三维 CAD 模型检索优化方案。采用数据分组聚合技术,分析了三维模型相似性特征,提出了模型相似性比较矩阵、模型相似链等聚类方法,简化了模型检索流程;通过对分布式计算框架工作原理的分析,融合基于机械结构形态的三维 CAD 模型检索技术,提出了面向机械结构模型特征的分布式计算方法,构建了高效的三维 CAD 模型检索系统;最后,以美国普渡大学开发的 ESB 模型库为基础数据,进行了实验验证。研究表明:分布式计算技术的应用可以有效地优化基于机械结构形态的三维模型检索方法,解决三维模型检索的算力瓶颈问题,为机械结构三维模型分布式检索提供实践方法。

关键词:机械结构形态;三维模型检索;分布式计算;相似链

中图分类号:TH128

文献标识码:A

文章编号:1001-4551(2021)04-0453-06

Optimization for 3D CAD model retrieval method

TIAN Liu^{1,2}, XU Ci-jun², HUANG Hao¹, ZHANG Shao-feng¹, CHEN Lei¹

(1. School of Machinery and Automation, Wuhan University of Science and Technology, Wuhan 430081, China;

2. School of Mechanical and Electronic Engineering, Hubei Polytechnic University, Huangshi 435003, China)

Abstract: As the number of 3D CAD model library models increased, the retrieval efficiency could be greatly reduced. Aiming at this problem, a 3D CAD model retrieval optimization scheme based on mechanical structure shape was proposed. Data grouping and aggregation technology was used to analyze the similarity characteristics of 3D models, and clustering methods such as model similarity comparison matrix and model similarity chain were proposed to simplify the model retrieval process. Through the analysis of the working principle of the distributed computing framework and the integration of the 3D CAD model retrieval technology based on the mechanical structure form, a distributed computing method oriented to the characteristics of the mechanical structure model was proposed to construct an efficient 3D CAD model retrieval system. Finally, the basic data of the ESB model library developed by Purdue University was used for experimental verification. The results show that the application of distributed computing technology effectively optimizes the 3D model retrieval method based on the mechanical structure shape, solves the computing power bottleneck problem of 3D model retrieval, and provides a practical method for the distributed retrieval of mechanical structure 3D model.

Key words: mechanical structure form; three-dimensional model retrieval; distributed computing; similarity chain

0 引 言

CAD 技术已广泛地应用于工程设计的各个领域,尤其是在机械行业,CAD 建模已成为机械产品设计、制造等各个环节的重要步骤。

随着企业的产品更新,企业拥有 CAD 模型数量也在日益增加^[1]。三维 CAD 模型的重用技术应用满足了设计人员重新绘图和建模等需求,提高了产品的设计效率。

徐赐军等^[2]结合机械结构形态相关知识,提出了基

收稿日期:2020-09-15

基金项目:国家自然科学基金资助项目(51405145)

作者简介:田柳(1997-),男,湖北恩施人,硕士研究生,主要从事智能 CAD 方面的研究。E-mail:pointbanana@163.com

通信联系人:徐赐军,男,博士,副教授。E-mail:xucijun818@163.com

于机械结构形态的三维 CAD 模型检索方法;XU L 等^[3]针对复杂的三维模型,基于局部细节特征的相似性比较来评估 CAD 模型,提高了三维模型的检索效率,但是部分模型检索的准确率有所下降;TASHIRO S 等^[4]采用具有尺度和旋转不变性的兴趣点来提取超向量编码特征,实现了对三维模型的检索;SIDDIQU A A 等^[5]开发了一种自动编码器,进行了语义建模以实现三维模型的检索。

随着大数据应用的发展,出现了分布式计算、分布式服务器集群、分布式存储等技术,大大改善了大数据的处理效率,分布式技术已成为了各研究领域的热点。LI B 等^[6]运用分布式检索技术,基于 Hadoop 建造的检索系统,使用户可以从海量的信息检索数据中找到所需的答案;郭飞等^[7]提出了一种基于 HDFS 和传统关系型数据库相结合的图像及其特征数据的存储结构设计方法,实现了海量图像信息的快速存储和读取,降低了图像检索时间;杨进展^[8]部署了一个具有 HDFS 和 MapReduce 分布式三维模型管理系统;廖先富^[9]基

于 Hadoop 的 HDFS 文件存储功能与 Django 技术,构建了一个云端分布式三维数据存储系统。

随着模型库三维 CAD 模型数量的增多,单台计算机算力不足的问题成为检索的瓶颈。三维 CAD 模型检索系统检索需要的时间越来越多,分布式计算成为优化三维 CAD 模型检索方法的首选。

针对三维 CAD 模型检索效率问题,笔者采用机械结构形态知识^[10]及分布式计算技术,研究协同设计环境下三维 CAD 模型检索方法的优化。

1 三维模型检索框架

基于机械结构形态和分布式计算的三维模型检索系统以机械结构形态为主要特征,以分布式计算技术为主要工具。

基于机械结构形态和分布式计算的三维 CAD 模型检索方法,其框图如图 1 所示。

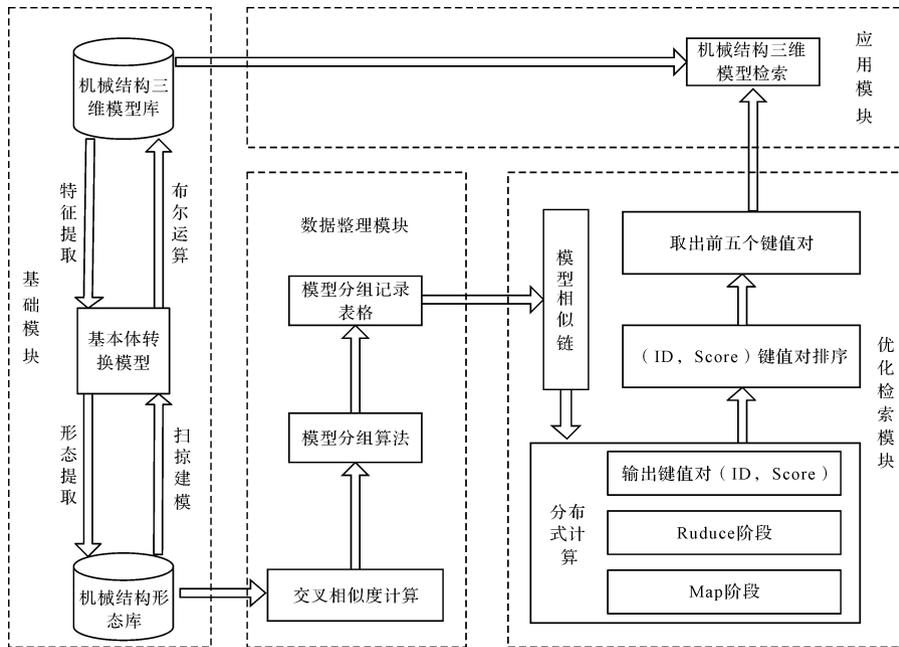


图 1 三维模型检索框架

图 1 中:

- (1)基础模块主要功能是通过基本体转换模型实现机械结构三维模型和机械结构形态的相互转换;
- (2)数据整理模块主要是对基础模块中的机械结构形态库进行处理,以获取用于后续分布式计算的数据表,主要研究机械结构形态库中的模型相似度分组方法,并将模型的分组结果存储于数据库中;
- (3)优化检索模块主要是分布式计算技术在三维模型检索中的应用,即对数据处理模块产生的数据进行处理;

(4)应用模块主要根据模型检索需要选择合适的指标实现三维模型的检索。

分布式计算是基于机械结构形态的三维模型相似度比较方法,结合 MapReduce 分布式计算编程模型,将模型相似性比较过程分为 Map 阶段和 Reduce 阶段,并经过分布式计算输出模型 ID 与相似度的键值对,获取模型 ID 所对应的机械结构形态和三维模型。

因此,图 1 中的三维模型检索框架是基于机械结构形态的三维 CAD 模型检索方法,该方法采用数据整理模块和优化检索模块对其检索算法进行优化。

2 模型相似链

基于机械结构形态的三维 CAD 模型检索方法,可大幅度提升模型检索的效率,但是在三维模型检索过程中模型库中的所有模型都需要参与相似度计算,大大增加了计算量。机械结构模型库往往存在许多只有少许细节不同的机械结构,若将这一系列的机械结构模型分成一组,构成模型相似链,则当遇到类似的机械结构进行检索时,待检索模型仅与该集合中每个模型进行相似度计算,可大大减少不必要的比较过程和计算。

模型相似链是指彼此之间相似度较高的机械结构的模型集合,并以模型编号表示。当待检测机械结构模型与相似链中的某个机械结构模型相似度较高时,说明待检测机械结构模型与相似链中其他机械结构模型相似度高的可能性较大;相反,当待检测机械结构模型与相似链中的某个机械结构模型相似度较低时,说明待检测机械结构模型与相似链中其他机械结构模型相似度低的可能性较大。

因此,只将待检测机械结构模型与相似度高可能性大的相似链中的模型进行比较,不与相似度低可能性大的相似链中的模型进行比较,可以大大减少比较次数。

模型相似链基于所有模型的相似性比较矩阵的相似度进行分组,模型相似性比较矩阵即是记录模型库中三维模型之间的相似度。

n 个机械结构三维模型的相似性比较矩阵如图 2 所示。

ID	a_1	a_i	a_j	a_n
a_1	1	S_{1i}	S_{1j}	S_{1n}
.....	1
a_i	S_{i1}	1	S_{ij}	S_{in}
.....	1
a_j	S_{j1}	S_{ji}	1	S_{jn}
.....	1
a_n	S_{n1}	S_{ni}	S_{nj}	1

图 2 模型相似性比较矩阵

a_i —第 i 个模型的 ID 编号; S_{ij} —第 i 个模型与第 j 个模型的相似度

本文采用基于机械结构形态的三维模型相似度计算方法,构建三维 CAD 模型的相似性比较矩阵,以便于模型相似链构建过程中进行遍历和筛选数据。

若将所有模型 ID 组成集合 A ,令 A 中有 n 个元素, ID 分别为 a_1 、 a_2 、 a_3 ... ,则数据库模型相似链构建流程如图 3 所示。

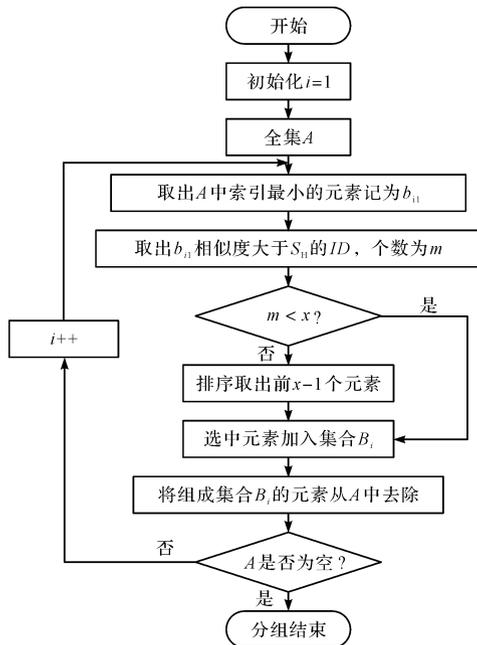


图 3 数据库模型相似链构建过程

$X = \lfloor \sqrt{n} \rfloor$ 向下取整,即小于或等于 \sqrt{n} 的最大整数; B_i —第 i 条相似链, B_{ij} —第 i 条相似链的第 j 个元素; S_H —表示模型相似链元素构建的相似度阈值,一般取 0.6; \sqrt{n} —表示相似链规模的控制值,其目的是保证相似链构建效率,能够有效控制一条相似链的长度,如果一条相似链过长,会导致在后续对相似链所有元素进行相似度计算时计算量过大

3 分布式计算

分布式系统是多个处理机通过通信线路互联而构成的松散耦合的系统,通过网络相互连接传递消息与通信后并协调行为,且组件之间彼此交互以实现一个共同的目标;把需要进行大量计算的工程数据分割成小块,由多台计算机分别计算,然后上传运算结果,并将结果统一合并产生数据结论。

分布式计算过程如图 4 所示。

该模型将输入工作划分为多个小块,并行计算,并将得出的结果进行汇总,最终产生计算结果。其中, map_1 表示待检索模型与第一条相似链中元素进行相似度计算,计算得出的 $map(id, score)$ 即为模型 id 和相对应的相似度。

首先是待检索模型与每个相似链的相似度计算,将每组的运算过程为一个 Map 阶段,每个 Map 可并行进行计算;然后将 Map 的运算结果自动上传到 Reduce,Reduce 将所有键值对传入全局变量散列表 N 中,散列表 N 是用来记录键值映射关系的数据结构,可存储三维模型 ID 与相似度的映射关系;最后通过对 N 中的数据进行值排序得到相似度高的三维模型 ID ,

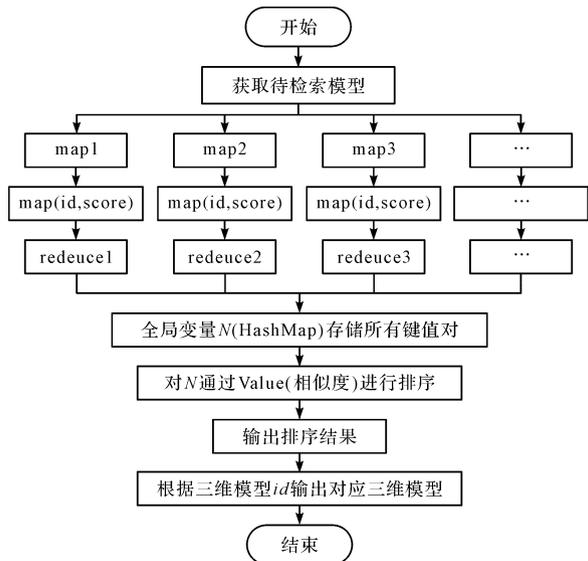


图 4 分布式计算过程

通过 ID 映射到机械结构形态库和三维模型形态库,输出检索到的三维模型,完成检索。

Map 函数主要逻辑运行调用相似度计算方法,经过相似度计算后,输出 (id, score), 即 (序号, 相似度) 键值对; 经过 MapReduce 框架自动将键值对数据传入 Reduce 阶段。

令模型数据库共有 m 个模型相似链 B_i , 且令初始值 $i = 1, j = 1$, 则 Map 阶段运算过程如下。

步骤 1: 当 $i \leq m$ 时, 将待检索模型与 B_i 中第 j 个元素进行相似度计算, 即 S_{ij} ; 否则过程终止;

步骤 2: 若 $S_{ij} \geq S_H$ (S_H — 相似度上限值, 一般取 0.6), 则待检索模型可以与模型相似链 B_i 其他模型进行相似性比较, 即该链为可用链, 进行下一步; 若 $S_i \leq S_L$ (为相似度下限值, 一般取 0.2), 则待检索模型不与模型相似链 B_i 中其他模型进行相似性比较, 即该链为不可用链, $i = i + 1$ 返回步骤 1; 若 $S_L \leq S_{ij} \leq S_H$, 则当 $j < 3$ 时 $j = j + 1$, 返回步骤 1; 当 $j > 1$ 时 $i = i + 1, j = 1$, 返回步骤 1;

步骤 3: 将可用链中所有模型与待检索模型进行相似度计算, 计算出的键值对 (id, score) 上传至 Reduce 处理;

步骤 4: 若 $i > m$, 则结束 Map 运算, 否则返回步骤 1。

4 实验验证

此处以美国普渡大学开发的 ESB 模型库为基础数据进行实验验证。ESB 中包含 866 个 STL 格式的工程模型, 笔者以其中任意 1 个模型作为输入, 以检索出模型库中相似的模型。

实验采用的 PC 机参数为: Intel® Core (TM) i7-

7700HQ@ 2.80 GHz CP, 分布式 Slave1 内存 2 GB; Slave2 内存 2 GB; Master 内存 2 GB。

以 ESB 模型库^[11] ganter_din6304 号模型机械结构形态为例, 其相似度计算结果如表 1 所示。

表 1 机械结构形态相似度

模型序号	模型图	机械结构形态	相似度
ganter_din6304 - m20 - 110 - f1			0.89
ganter_din6304 - m8 - 50 - f1			

表 1 中, 元素值越大说明三维模型之间的相似度越高。

同理, 在三维模型数据库构建过程中, 基于机械结构形态的三维模型相似度比较算法, 获得的模型库中三维模型的相似性表矩阵, 如图 5 所示。

ID	1	2	3	4	5	6	7	8
1	1	0.89	0.43	0.36	0.3	0.47	0.49	0.4
2	0.89	1	0.72	0.36	0.39	0.33	0.02	0.26
3	0.43	0.72	1	0.5	0.76	0.43	0.72	0.78
4	0.36	0.36	0.5	1	0.43	0.66	0.52	0.63
5	0.3	0.39	0.76	0.43	1	0.39	0.39	0.77
6	0.47	0.33	0.43	0.66	0.39	1	0.63	0.26
7	0.49	0.02	0.72	0.52	0.39	0.63	1	0.22
8	0.4	0.26	0.78	0.63	0.77	0.26	0.22	1

图 5 模型相似性比较矩阵(部分)

图 5 中, 模型 1 表示 ganter_din6304 - m20 - 110 - f1, 模型 2 表示 ganter_din6304 - m8 - 50 - f1。

模型库相似链存储表如图 6 所示。

	1	2	3	4	5	6	7	9	10
1	1	2	9	18	26	27	36	42	231
2	3	5	7	8	13	33	34	97	
3	4	6	12	24	33	89	96		
4	10	11	25	39	46	98	118	119	
5	44	74	129	145	164	175	176		

图 6 相似链(部分)

由图 6 第 1 行数据可以看出, 1 号模型与 2 号模型在同 1 条相似链中。

图 6 中行表头表示相似链 ID, 列表头表示相似链中的元素索引, 表中数据表示模型 id。

图 5 模型相似性比较矩阵生成图 6 相似链的过程如下: 第 1 行中, 1 为第 1 条相似链的第 1 个元素, 继续遍历得到与 1 号模型相似度大于 0.6 的模型 ID, 分别为 2 号, 9 号等模型, 故集合 [1, 2, 9, 18...] 为 1 条相似链; 第 2 行中, 1, 2 均被剔除, 故第 2 条相似链从 3 开始, 集合 [3, 5, 7, 8...] 为 1 条相似链, 继续生成多条相似链。

在进行和模型检索时,只需根据流程先和1号、2号模型进行相似度计算,即可确定该相似链是否可用。若为可用相似链,则与后续9号、18号等模型继续进行相似度计算;若为不可用链,则直接跳过该链,继续判断下一条相似链是否可用。

例如:待检索模型为1号模型,1号模型与1号相似链相似度为1,相似度大于0.6,则1号相似链为可用链,可与该相似链后续元素继续比较。1号模型与2号、3号、4号相似链相似度均小于0.6,则这3条相似链为不可用链,简化检索流程,以提高检索效率。

本文算法与文献[2]算法查准率与查全率曲线图如图7所示。

图7显示,本文算法的查准率查全率与文献[2]算法基本一致,说明基于机械结构形态三维模型检索

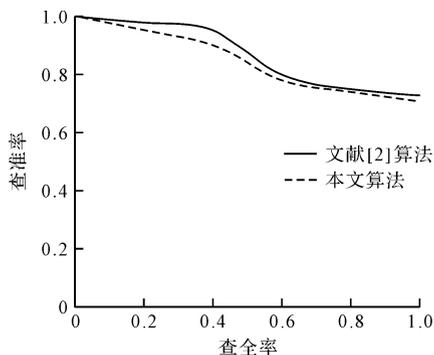


图7 查准率-查全率曲线图

的准确率较高。

笔者随机选取3个模型 ganter_din6305 - M20、Nozag_bfi_60,以及 winkel_ap,分别使用文献[2]检索算法和本文算法进行比较,其结果如表2所示。

表2 检索算法的结果比较

输入模型	算法	检索结果(前5位)					时间
	文献[2]算法						7.3 s
	本文算法						1.7 s
	文献[2]算法						6.8 s
	本文算法						1.5 s
	文献[2]算法						5.4 s
	本文算法						1.2 s

由表2可知:本文算法与文献[2]算法所获检索结果大致相同,但是本算法的检索时间明显较短,检索效率较高。

由此可见,通过模型相似性比较矩阵和相似链等技术的应用,三维模型的检索效率可以获得提升。

5 结束语

三维模型的重用是提升机械产品开发效率的重要途径,三维模型检索是实现三维模型重用的重要方法;但是随着三维CAD模型库模型数量的增加,导致检索效率大幅度降低。

为了保证检索效率,本文针对三维模型检索的需

求,分析了基于机械结构形态的三维CAD模型检索的特点,提出了模型相似性比较矩阵、相似链两个概念,以简化模型检索过程;并给出了模型相似性比较矩阵、相似链的计算以及使用方法;结合分布式系统架构提出的分布式计算方法,实现了基于机械结构形态的三维CAD模型检索分布式计算系统的搭建;通过简化模型检索过程和使用集群代替单机以解决单机算力瓶颈问题的方法,完成了对三维CAD模型检索方法的优化。

实例验证的结果表明:该方法能够在保证检索精度的条件下,显著减少三维模型的检索时间,检索效率获得了提升。

(下转第463页)

本文引用格式:

田柳,徐赐军,黄浩,等. 三维CAD模型检索方法的优化研究[J]. 机电工程,2021,38(4):453-457,463.

TIAN Liu, XU Ci-jun, HUANG Hao, et al. Optimization for 3D CAD model retrieval method[J]. Journal of Mechanical & Electrical Engineering, 2021, 38(4):453-457,463.