

多特征动态融合的三维模型检索方法

郑 赢¹ 周明全^{1,2} 耿国华¹ 高原¹

(西北大学信息科学与技术学院 西安 710127)¹ (北京师范大学信息科学与技术学院 北京 100875)²

摘 要 提出一种基于二维正交投影图像的多特征动态融合的三维模型检索方法。首先计算三维模型的二维正交投影图像,然后提取二维正交投影图像的投影直方图和 Zernike 矩特征,通过加权求和在输出层融合,得到总体上模型间的相似度。每种特征采用动态权值,针对输入查询模型的不同,根据用户反馈自动更新知识库中的权值。实验表明,该方法在提高检索准确率的同时,也能保证检索效率。

关键词 三维模型检索,多特征,动态权值,反馈

3D Model Retrieval on Multi-feature Dynamic Integration

ZHENG Ying¹ ZHOU Ming-quan^{1,2} GENG Guo-hua¹ GAO Yuan¹

(Dept. of Computer, School of Information Science and Technology, Northwest University, Xi'an 710127, China)¹

(Dept. of Computer, School of Information Science and Technology, Beijing Normal University, Beijing 100875, China)²

Abstract Presented a method by itergrating multi-feature base on two-dimensional orthogonal projection to improve retrieval accuracy. Firstly, we obtained two-dimensional orthogonal projections of the 3D models, and then two features which among two-dimensional orthogonal projections were used and each feature has a weight. These features were syncretized on output layer. By plusing the two results of the features with their weights, the whole result could be got. According to users' feedback, the system changed the weights of each feature dynamically. Experiments show retriving accuracy and efficiency have a greater improvement than that on single feature.

Keywords 3D model retrieval, Multi-features, Dynamic weights, Feedback

人类视觉系统的感受部分是视网膜,它就相当于一个三维采样系统。三维模型的可见部分投影到网膜上,人们根据投影到视网膜上的二维的投影图像来对该模型进行三维理解。也就是说,人们通过比较两个三维模型的外观来辨别三维模型之间的相似性。

计算机视觉是当前新兴的一门学科,指用摄像机和计算机代替人眼对目标进行识别,通过把三维模型的影像作为输入源,即输入计算机的就是三维物体的二维图像,用这些二维图像去重建三维模型^[1]。计算机视觉研究本来就是和人类视觉研究密切相关的学科,结合以上提到的两点,一是可以获得三维模型的二维投影图像,二是用二维图像能够重建三维模型,于是可以推导出这样一个结论:可以通过比较三维模型的二维投影图像来获取三维模型的相似性。

我们对这类基于计算机视觉的三维模型几何相似性比较算法进行了深入的研究,实验发现基于二维投影图像特征的检索准确率好于其他多数基于内容特征的检索准确率,但同时也发现使用单一特征很难大幅提高检索准确率。Bustos^[2]的实验也证明:采用单一特征很难保证检索系统对于任意模型都有很好的检索准确率。恰当地融合不同的形状特征可使

它们优势互补,能更好地描述三维模型特征。因此本文在用户反馈的基础上,采用融合基于投影图像的多特征来提高检索准确率。

本文在提出一种多视角的二维投影图像获取方法的基础上,计算用户输入模型的投影图像,然后提取投影直方图^[3]和 Zernike 矩特征^[4]并融合,与数据库^[5]中的模型进行相似性比较,文中提出的权值调整算法则根据用户反馈信息对相应权值自动更新。权值调整算法是一个收敛的算法,若反馈次数达到一定程度,权值会趋于一个稳定分布,检索性能将不再变化。我们已经通过实验验证了此方法的合理性和有效性,也证明了该方法对于三维模型检索所要求的平移、旋转和缩放不变性具有较好的鲁棒性。

1 多特征动态融合的三维模型检索方法

1.1 投影图像的生成

三维模型的二维投影变换可以分为两种:平行投影变换(正交投影)和中心投影变换(透视投影)^[6],如图 1 所示。本文采用平行投影方式,计算三维模型三视图上的二维投影图像。

到稿日期:2009-08-12 返修日期:2009-10-15 本文受国家自然科学基金重点项目(60736008),国家 863 重点基金项目(2008AA01Z301),国家自然科学基金项目(60873094),西北大学研究生自主创新基金项目(No. 08YZZ33),陕西省自然科学基金(SJ08F29)资助。

郑 赢(1984-),男,硕士生,主要研究方向为图形图像处理、可视化技术研究,E-mail: qq258338468@163.com;周明全(1953-),教授,博士生导师,主要研究方向为计算机可视化技术、生物特征识别、中文信息处理;耿国华(1954-),教授,博士生导师,主要研究方向为智能信息处理、计算机可视化技术;高原(1975-),女,博士生,主要研究方向为数据库、图像数据挖掘。

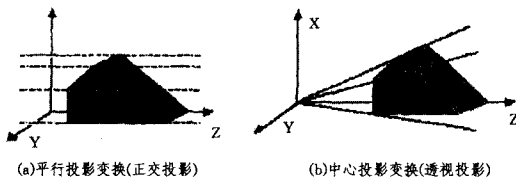


图1 三维模型投影变换

三维模型数据将点、面信息分别存储,简单地将数据投影到平面上,获得的只是离散的点,不仅丢失大量的模型特征信息,而且从中难以提取特征。要得到投影图像,必须按照模型面片将投影后的面片顶点之间的间隙进行填充。

现以经过预处理的3D模型在XOY平面上的投影为例(模型Z轴数据为零),其对应的点数据集为 $3D_point$,面片(以三角面片为例)数据集为 $3D_face$ 。实现描述见算法1。

算法1 ProToImg

输入: $3D_point, 3D_face$

输出:二值投影图像 $Image$

Step1 计算 $3D_point$ 中X,Y轴上的最大值 x_max, y_max 及最小值 x_min, y_min ;

Step2 计算X,Y任意两点之间差的非零最小值 $\Delta x, \Delta y$;

Step3 用式(1)、式(2)计算待生成图片的宽 $width$ 和高 $height$,申请存储空间 $M(width, height)$

$$width = \frac{x_max - x_min}{\Delta x} \quad (1)$$

$$height = \frac{y_max - y_min}{\Delta y} \quad (2)$$

Step4 用公式(3)转化 $3D_point$ 中的X,Y:

$$X' = \frac{X}{\Delta x}, Y' = \frac{Y}{\Delta y} \quad (3)$$

Step5 While($3D_face$ 存在未遍历面片){

Step5.1 读取对应的点数据 $p_1(x_1, y_1), p_2(x_2, y_2), p_3(x_3, y_3)$;

Step5.2 计算得到3点中最值 $x_pmin, y_pmin, x_pmax, y_pmax$;

Step5.3 判断以 $(x_pmin, y_pmin), (x_pmax, y_pmax)$ 为对角点构成的矩形区域中的点是否落在原三角面片区域中,若是,则记 $M(a, b) = 1$,其中 $a \in [x_pmin, x_pmax], b \in [y_pmin, y_pmax]$;

Step6 保存图像数据 $Image = M$;结束

本文提出的投影图像生成算法 ProToImg,能根据三维模型的精细程度得到适当的 $width$ 和 $height$ 的投影图像。图2是蚂蚁模型用文中算法在三视图方向上的投影图像。图2中(a)是模型,(b),(c),(d)分别是正视、俯视、侧视3个方位的投影图像。

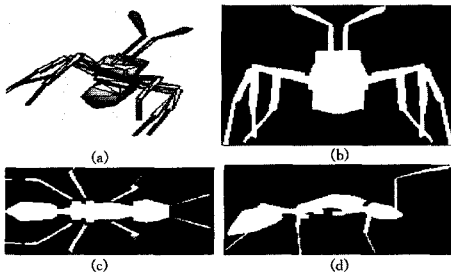


图2 投影图像

1.2 特征提取

在获得三维模型三视图投影图像的基础上,分别对图像

提取 Zernike 矩及投影直方图特征。

1.2.1 Zernike 矩特征

对投影图像进行比例、平移变换,将图像的中心移到坐标原点,将图像像素点坐标映射到单位圆内。 N 是图像沿坐标轴的像素数,则 P 阶 Zernike 矩特征按式(4)提取:

$$Z_{pq} = \frac{p+1}{\pi(N-1)^2} \sum_{x=1}^N \sum_{y=1}^N v_{pq}(r, \theta) f(x, y) \quad (4)$$

$$r = (x^2 + y^2)^{1/2} / N, \theta = \tan^{-1}(y/x) \quad (5)$$

式中, $f(x, y)$ 代表图像的亮度, $r(r \leq 1)$ 为像素点到原点的距离, θ 为 r 与 x 轴的夹角。 $v_{pq}(r, \theta)$ 是在单位圆内正交的 Zernike 多项式,设为 p 阶 q 重,表达式见式(6)。

$$v_{pq}(r, \theta) = R_{pq}(r) e^{jq\theta} \quad (6)$$

$$R_{pq}(r) = \sum_{s=0}^{(p-|q|)/2} (-1)^s \frac{(p-s)!}{s! (\frac{p-2s+|q|}{2})! (\frac{p-2s-|q|}{2})!} r^{p-2s} \quad (7)$$

式中, $n=0, 1, 2, \dots, 0 \leq |q| \leq p, p-|q|=even, i = \sqrt{-1}$ 。

本文在三维模型的3个投影图像上各提取20维的 Zernike 矩特征,然后合成60维的 Zernike 矩,用 PCA 降维,选取相关性最大的前20维作为3D模型的特征。

1.2.2 投影直方图特征

以投影图像的中心点为原心,使用10个不同半径的同心圆等距离切分图像。如图3所示,对三视图方向的投影图像进行切分。

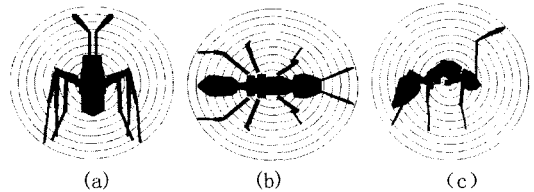


图3 环切投影图像

在图像 j 中,提取的特征中某一向量 i 描述为:统计圆环 i 内的像素个数 A_i^j ,计算每幅图像各自的最大圆环面积,然后相加,以二者的比值作为特征向量。例如, A_i^j 表示第 j 幅图像中第 i 个环内的像素个数, C_j 是指在第 j 幅图像当中最大圆环的面积,投影图像 j 的特征表达见式(8)。

$$F_i^j = \frac{A_i^j}{\sum_{j=1}^3 C_j} \quad (8)$$

文中计算了模型三视方向上投影图像的投影直方图特征,合成并利用 PCA 降维得到最终特征 F 。

经过实验发现文中两个特征有如下特点:

1) 投影直方图特征对模型的退化具有较好的鲁棒性,如对 Stanford Bunny,一个带有洞的兔子^[6]进行检索,依然能够保证准确率。但是,该特征描述模型比较粗糙,且对噪声较敏感;

2) Zernike 矩特征相关性小,冗余性小,抗噪声能力强,在表达图像的形状特征时,其低阶矩特征表达图像目标的整体信息,高阶矩特征表达图像目标的细节信息。

为了进一步在单一特征基础上提高检索准确率,本文对两个特征进行加权融合,实现优势互补。

1.3 动态融合特征与相似性比较

对用户给定的输入查询,不同特征所表现出来的检索性能是有差异的。文中通过对不同特征分配不同权值来度量这

种差异,并通过输出层的加权求和得到总体的相似度。根据用户的反馈结果,自动调整特征权值^[7-10]。在特征融合前,因为提取的不同特征所在的值域范围差距比较大,所以要先按式(9)对特征值进行高斯归一化。

$$f_i' = \frac{(f_i - \mu)}{3\sigma} + 1/2 \quad (9)$$

式中, f_i 指对特征 i 提取的特征值, μ 指对应特征 i 的特征值的均值, σ 指对应特征 i 的特征值的标准差,归一化后得到的特征值 $f_i' \in [0, 1]$ 的概率大于 99%。

文中将提取的 Zernike 矩特征 $f_1(f_1^1, f_1^2, f_1^3, \dots, f_1^n)$ 和投影直方图特征 $f_2(f_2^1, f_2^2, f_2^3, \dots, f_2^m)$ 标准化到区间 $[0, 1]$ 之间,然后按式(10)融合成新的特征 $F(wf_1, wf_2, wf_3, \dots, wf_n, vf_1, vf_2, vf_3, \dots, vf_m)$ 。

$$F = wf_1 \cdot vf_2 \quad (10)$$

式中, w, v 是对应特征的权值,且 $w + v = 1$,在初始化时可以将所有特征的权值等分。本文提出了基于用户反馈的权值动态调节算法 F_adjust ,描述见算法 2。

算法 2 F_adjust

输入:待检索模型 R

输出:检索出待显示的模型集合 $Query$

初始化:矩特征库 f_1 ,投影直方图特征库 f_2 ,查询融合权值 w, v ;

按式(10)融合生成集合 $F_Combine$;

Step1 对 R 进行特征提取、融合,与 $F_Combine$ 中的数据进行相似度计算、排序,得出检索结果集 $Query$;

Step2 IF 用户不满意当前检索结果

Repeat;

step2.1 获取检索结果模型 M_i 的相关度系数 P_i ($P_i \in [0, 1]$);

Step2.2 用公式 $D'(M_i, R) = \frac{1 - P_i}{0.8P_i} * D(M_i, R)$ 更新模型 R 与 M_i 的欧式距离 $D(M_i, R)$,并按照欧式距离从小到大排列 $Query$ 中的模型;

Until:用户满意检索结果

Step3 调节权值 $w' = \frac{\phi_2}{\phi_1} v, v' = 1 - w'$ 并保存;结束

ϕ_1, ϕ_2 是 $Query$ 集中模型与待检模型 R 对应单特征的欧式距离均值。

F_adjust 中通过用户调节滑动条的方式来反馈各个检索结果的相关程度,提高反馈的精确率。若用户输入的待检索模型未在数据库中,则用初始权值融合特征。

2 系统实现与结果分析

本文中我们结合所提出的检索方法,在 MATLAB7.1 下仿真实现系统原型。实验数据采用普林斯顿大学三维模型库与网络下载模型,共 400 多个动物模型,共分为 11 类,蚂蚁类模型有 108 个。

2.1 实验过程及其分析

系统将上面两种特征融合,初始化权值,输入查询模型,通过调节滑动条来反馈检索结果,一般经过 2 到 4 次的反馈过程,检索结果就不再变化,检索结果如图 4 所示。图中矩形框标记出的模型都是相关度较小的模型。实验得到的查全率与查准率曲线如图 5 所示。

另据实验表明,文中使用单特征和融合特征进行检索,因为对融合特征进行了 PCA 降维处理,单特征和多特征的维数

接近,所以其检索效率相差很小,均小于 1.35 秒。

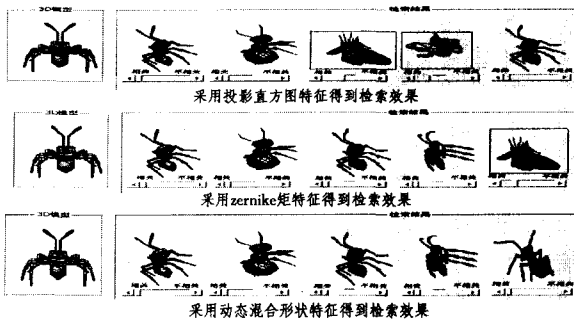


图 4 最终检索结果

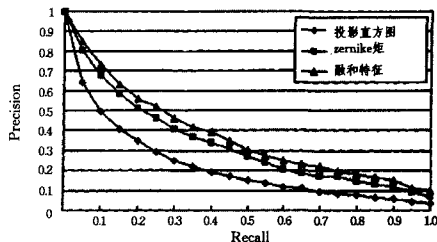


图 5 查全率与查准率曲线

结束语 随着可视化技术、虚拟现实、数字化和三维图像建模的发展,互联网和专用数据库中三维模型数量持续增长,3D 模型检索技术具有很大的研究价值和前景。为了解决单一的形状特征很难保证系统对任意的输入查询模型都具有很好的检索准确率这一问题,本文在获得三维模型投影图像基础上,提取投影直方图与 Zernike 矩特征,通过用户反馈的结果调节权值,在输出层动态融合此特征,用于模型检索。由于 Zernike 矩弥补了投影直方图对噪声敏感,因此 Zernike 高阶矩能够描述模型的细节信息;而投影直方图提高了 Zernike 矩对退化或有洞模型的鲁棒性。实验数据表明,融合后的特征同单特征相比,在保证检索效率前提下,较好地提高了三维模型检索的查全率与查准率。

在以后的工作中,将进一步研究如何把三维模型的材料或纹理信息映射到投影图像上,将研究更多的基于三维投影图像的特征,以及特征之间的互补性利用。

参考文献

- [1] Cyr C, Kimia B. 3D object recognition using shape similarity-based aspect graph. Proceedings of IEEE International Conference on Computer Vision, 2001[C]//Eighth IEEE International Conference, Vancouver, Canada, 2001; 254-261
- [2] Bustos B, Keim D A, Saupé D, et al. Automatic Selection and Combination of Descriptors for Effective 3D Similarity Search [C]// IEEE International Workshop on Multimedia Content-based Analysis and Retrieval. Miami, Florida, USA; [s. n.], 2004
- [3] Atmosukarto I, Leow W K, Huang Zhi-yong. Feature Combination and Relevance Feedback for 3D Model Retrieval[C]//Proc. of 11th International Multimedial Modelling Conference(MMM'05). IEEE, 2005
- [4] Saiful A, Josef K, Roland W. Multi-feature based 3D Model Similarity Retrieval [C]// International Conference on Computing and Informatics (ICOCI) Malaysia. 2006

[5] 林金杰, 韦伟, 杨育彬. 面向内容的三维模型数据库设计及其检索系统的实现[J]. 计算机科学, 2008(10)

[6] 崔晨昉. 三维模型检索中关键技术的研究[D]. 杭州: 浙江大学, 2005

[7] 贾骥, 覃征, 卢江, 等. 圆环射线分解投影轮廓的三维模型检索方法[J]. 系统仿真学报, 2006, 18(10): 2847-2851

[8] 周明全, 耿国华, 韦娜. 基于内容的三维模型检索[M]. 北京: 清

华大学出版社, 2007

[9] 高波涌, 莫国良, 张三元. 基于混合形状特征的三维模型检索研究与实现[J]. 计算机应用研究, 2007, 24(3): 221-223

[10] Saiful A, Josef K, Roland W. Multi-feature Intergration With Relevance Feedback On 3D Model Similarity Retrieval[A]// 8th International Conference on Information Integration and Web-based Application & Services 2006[C]. Indonesia, 2006

(上接第 250 页)

等于 x_2 。C 测度度量了一个集合中元素被另一个集合覆盖的程度(从一定程度上反映了集合的收敛性能)。 $C(X_1, X_2) > C(X_2, X_1)$ 说明集合 X_1 在 C 测度上的性能优于 X_2 。

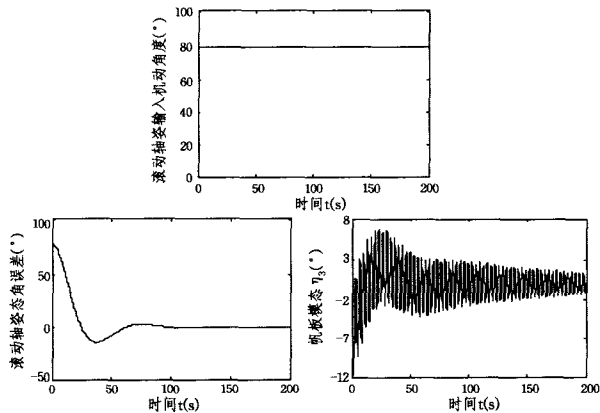


图 6 未作机动路径规划的滚动轴输入机动角度、滚动轴姿态角误差和帆板模态坐标 γ_3 随时间的变化曲线

(2) 分布测度(sd):

$$sd = \sqrt{\frac{\sum_{i=1}^Q (d_i - \bar{d})^2}{Q}} \quad (6)$$

式中, Q 为非支配解集中解的个数, d_i 为解 x_i 到集合中其余解的最短距离, \bar{d} 为所有 d_i 的平均值。sd 越小, 说明非支配解集中解的分布越均匀。

(3) 宽广度(extent)^[11]:

$$extent = \sqrt{\sum_{i=1}^m |f_{imax} - f_{imin}|} \quad (7)$$

式中, m 为目标的个数, f_{imax} 和 f_{imin} 为非支配解集在目标 f_i 上的最大值和最小值。extent 越大, 说明非支配解集在目标空间上的分布范围越广。

SPEA2 采用与 KDMOEA 相同的群体初始化方法、进化算子和终止条件, 种群规模取为 30, 外部存储器的大小固定为 30, 交配选择中的联赛规模取为 2。为了克服随机性带来的影响, 将 KDMOEA 和 SPEA2 在 1.3 节挠性卫星大角度机动姿态控制问题中各运行 20 次。表 2 给出了两种算法在 sd, extent 测度上 20 次运行结果的平均值。如表 2 所列, 本文算法求得的非支配解集在 sd, extent 测度上的性能均优于 SPEA2。两算法在 C 测度上的平均值分别为 $C(KDMOEA, SPEA2) = 0.1000$, $C(SPEA2, KDMOEA) = 0.1333$, 即它们相互覆盖的程度(收敛性能)接近。由此表明, 本文方法能够搜索到一组与 SPEA2 的收敛性能相当, 而分布更均匀、范围更宽广的非支配解集, 即本文算法具有更高的求解效率。

表 2 两种算法在测度 sd, extent 上的比较

	sd	extent
KDMOEA	0.5311	7.5530
SPEA2	0.9863	6.2176

结束语 本文基于卫星姿态机动快速性和高精度稳定性的要求, 建立了带挠性帆板卫星大角度机动姿态控制问题的多目标优化模型, 并基于一种保持群体多样性的多目标进化算法, 设计了卫星姿态机动路径的规划方案。仿真结果表明, 所提算法运行一次便能够有效地搜索到一组多样性较好的非支配路径解, 供决策者在不同的工作目标下选择, 有效地缓解了快速性与高精度指标之间的矛盾。与未作机动路径规划的情况相比, 本文算法同时优化了卫星调节时间和帆板振动强度两个性能; 与另一种经典的多目标进化算法相比, 它具有更高的求解效率。

参考文献

[1] 刘国汉, 王长龙, 高益军. 卫星姿态动力学与控制(2)[M]. 北京: 宇航出版社, 2005

[2] 靳永强, 刘向东, 侯朝桢. 航天器大角度姿态机动的自适应滑模控制[J]. 北京理工大学学报, 2007, 27(5): 422-426

[3] Boskovic J D, Li S M, Mehra R K. Robust adaptive variable structure control of spacecraft under control input saturation [J]. Journal of Guidance, Control and Dynamics, 2001, 24(1): 14-22

[4] 周黎妮, 唐国金, 李海阳. 航天器姿态机动的自抗扰控制器设计[J]. 系统工程与电子技术, 2007, 29(12): 2122-2126

[5] 朱良宽, 马广富, 胡庆雷. 带有死区非线性输入的挠性航天器姿态机动智能控制[J]. 吉林大学学报, 2008, 38(2): 195-200

[6] 靳瑾, 张景瑞, 刘藻珍. 航天器大角度姿态快速机动控制器参数优化设计[J]. 清华大学学报, 2009, 49(2): 289-292

[7] 申晓宁, 郭毓, 陈庆伟, 等. 一种保持群体多样性的多目标遗传算法[J]. 控制与决策, 2008, 23(12): 1435-1440

[8] 李英波. 挠性充液卫星动力学分析与姿态控制研究[D]. 上海: 上海交通大学, 2001

[9] Deb K, Agrawal R B. Simulated binary crossover for continuous search space [J]. Complex Systems, 1995, 9: 115-148

[10] Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm [R]. Technical Report 103, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland, May 2001

[11] Zitzler E, Thiele L. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach [J]. IEEE Transaction on Evolutionary Computation, 1999, 3(4): 257-271