

一个相关反馈三维模型深度图像检索算法*

王小凤, 耿国华, 张 飞

(西北大学信息科学与技术学院 计算机系, 西安 710127)

摘要: 基于图像的三维模型检索比形状特征和拓扑结构特征具有更易于实现且检索效果好等优点, 但现有基于图像的三维模型检索方法存在一些问题, 如没有考虑三维深度信息、所提特征不能全部表达三维模型信息且不能实现用户交互操作等。对此提出一种基于深度图像的三维模型相关反馈检索算法: 首先提取三维模型深度图像边界方向直方图和 Zernike 矩特征; 然后利用特征距离度量进行三维模型检索, 并采用相关反馈技术实现分类器的构造和模型的标注; 最后利用调整好的权值分类器进行检索和标注。实验表明, 该算法实现了三维模型的相关反馈检索和标注, 提高了检索的精确性, 并增强了系统的实用性, 为进一步进行语义检索打下基础。

关键词: 三维模型; 深度图像; 边界方向; Zernike 矩; 相关反馈

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 1001-3695(2012)06-2350-04

doi:10.3969/j.issn.1001-3695.2012.06.093

3D model relevance feedback retrieval algorithm based on range image

WANG Xiao-feng, GENG Guo-hua, ZHANG Fei

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

Abstract: In 3D model retrieval area, compared to shape features and topological structure features, the method based on image has the advantages that it is easy to realize and the retrieval results are good. But the existing image retrieval methods in 3D model do not consider the depth information and the features cannot express all information of 3D model. This paper proposed the 3D-model relevance feedback retrieval algorithm based on range images. Firstly, this algorithm computed the features of edge direction histogram and Zernike moment of 3D model's range images, then the feature distance was used to retrieval 3D model, and relevance feedback was used to build the classification and label the 3D model in database, at last the trained classification was used to retrieval and label the 3D model. The experiments show that the algorithm improves the accuracy and practicability of system, and realizes the relevance feedback retrieval and the 3D model's label which lay the foundation for semantic retrieval.

Key words: 3D model; range image; edge direction; Zernike moment; relevance feedback

0 引言

三维模型是继音频、图像与视频之后的第四类数字媒体, 从几何角度再现了真实世界中物体形状的全面信息, 可以广泛应用于文物保护、虚拟现实、远程教育和分子生物学等领域。随着网络技术的发展及三维数据获取设备价格的降低, 出现了大量可共享的三维模型数据。因此, 研究三维模型检索 (content-based 3D retrieval) 技术, 帮助用户快速准确地获取符合心意的三维模型, 实现资源重用, 成为当前的热点研究问题。

三维模型检索中, 相似性比较是关键问题之一。近年来, 国内外学者在这方面做了大量的研究工作, 按照三维模型特征描述的信息主要分为以下三类^[1]: 基于形状的检索、基于拓扑结构的检索、基于图像比较的检索。相比基于形状特征和基于拓扑结构的检索, 基于图像比较的方法^[2,3]有如下优点:

a) 将三维模型间的相似性比较转换为二维模型间的相似性比较。

b) 借助成熟的 2D 图像检索技术进行三维模型检索。

c) 易于实现且检索效果较好。

但是, 这些方法存在以下几个问题:

a) Min 法^[2]和 Chen 法^[3]将三维模型投影到二维平面, 丢失了模型的空间信息。

b) Min 法对 2D 图像作特征提取时只考虑了图像的轮廓信息, 忽略了其内部信息。

c) Chen 法虽然计算了图像的内部信息 (二值图像), 但没有考虑图像的轮廓信息。

d) 这些方法在用户觉得检索结果不满意时无能为力, 没有用户加入参与功能。

深度图像实质就是带有距离信息的视觉图像, 每个像素存放的不是灰度或颜色信息, 而是三维模型表面到成像面的距离信息。因此深度图像直接反映了景物表面的三维几何信息。Bosche 等人^[4]研究了深度图像在三维 CAD 建筑领域中的自动识别和检索技术。Adán 等人^[5]研究了基于深度图像的复杂场

收稿日期: 2011-10-11; **修回日期:** 2011-11-25 **基金项目:** 国家自然科学基金资助项目 (60873094); 陕西省自然科学基金资助项目 (2011JQ8001, 2010JQ8011); 教育部虚拟现实工程中心 2009 年度开放基金资助项目 (MEOBNUEVRA200903); 陕西省教育厅资助项目 (12JK0730, 09JK738)

作者简介: 王小凤 (1979-), 女, 陕西渭南人, 副教授, 博士, CCF 会员, 主要研究方向为数据挖掘、三维模型检索、音频检索等 (xfwang@nwu.edu.cn); 耿国华 (1955-), 女, 教授, 博导, 博士, 主要研究方向为智能信息处理、数据库、数据挖掘等; 张飞 (1984-), 男, 硕士, 主要研究方向为图形图像可视化。

景中三维物体的识别和检索。Gao 等人^[6]提出用加权二部图进行三维模型匹配的算法,首先选出待检索模型具有代表性的二维视图并给出相应的初始权值,然后对这些权值作进一步修正并与目标模型进行匹配,并采用 Zernike 矩特征进行二维图像距离度量。

相关反馈实际上就是通过人的主观判断和理解来干预模型语义的表达。系统提供给用户一个交互式界面,用户通过自己的判断和理解在此界面上作出主观选择和评价,系统从用户的选择和评价中获取信息,然后通过这些信息来提高模型检索的精确度。相关反馈技术近些年来引起了人们的广泛关注。

近年来,如何在三维模型检索中引入反馈机制用于解决语义鸿沟问题引起了研究人员的兴趣。Elad 等人^[7]将支持向量机应用于三维模型检索相关反馈,其相似性度量采用欧氏均方距离。Leifman 等人^[8]把学习特征子空间的概念应用于三维模型检索,然后结合偏离判别式分析和线性判别式分析方法建立相关反馈机制。Novotni 等人^[9]则对一些典型反馈算法的性能进行分析和比较,并应用于三维模型检索,对不同统计学习理论的反馈性能比较,支持向量机可以得到最好的反馈性能。而 Leng 等人^[10]的工作则是在支持向量机基础上,结合主动学习理论,进一步提高检索性能。

针对现有基于图像检索方法存在的问题,本文结合相关反馈技术提出了一种基于相关反馈的三维模型深度图像检索方法。该方法能够得到一种既表现深度图像的轮廓特征,又不失其整个区域内部特性并具有人机交互功能的检索方法。

1 基于相关反馈的三维模型深度图像检索流程

首先提取深度图像边界方向直方图和区域特征 Zernike 矩描述子特征,并融合这两个特征进行距离度量来进行三维模型检索;然后用户对检索模型进行反馈,算法根据反馈信息构造训练分类器,同时将用户的反馈信息在库中对相应的模型进行标注;接着系统继续检索反馈标注过程,直至检索结果符合用户需求,则完成了相关反馈模型检索,并进行了相关模型的语义标注;最后成功构造模型底层特征分类器,利用这些分类器在检索过程中作进一步学习和标注。

基于相关反馈的三维模型深度图像检索处理流程如图 1 所示,后面每一步骤进行详细介绍。

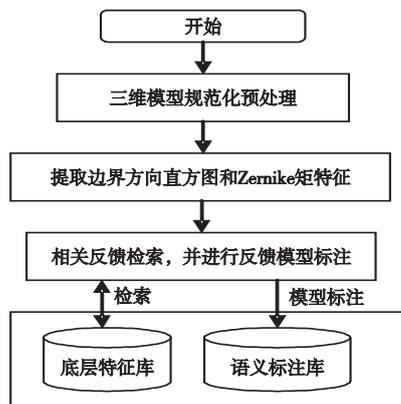


图1 具有相关反馈功能的三维模型深度图像检索流程

1.1 三维模型的预处理

由于三维模型具有很高的自由度,为了得到投影方向、大小和位置相同的投影图像,在特征提取前,必须对三维模型进行规范化预处理。

本文分别对三维模型的旋转不变性、缩放不变性和平移不变性采用面积加权的 PCA 法、模型表面坐标点归一化和采用局部坐标系表示三维模型方法进行变换处理,可表示为如下几何变换:

$$\tau: \tau(v) = s^{-1} \times F \times R \times (Q - G) \quad (1)$$

其中: s 为缩放系数; F 为翻转矩阵; R 为 PCA 变换; Q 为三维模型表面顶点; G 为重心。

1.2 深度图像特征提取

本文采用基于正交投影的六视图投影法来获得现存三维模型的深度图像。采用文献[11,12]中的方法获取三维模型深度图像。图 2 为普林斯顿三维模型库中蚂蚁的六视图投影。图 3 为普林斯顿模型库中的三维蚂蚁模型对应的投影图像。其中, X^- 表示投影到 X 轴负方向, X^+ 表示投影到 X 轴正方向, Y 和 Z 轴相同,正负方向根据局部坐标系的坐标方向确定。

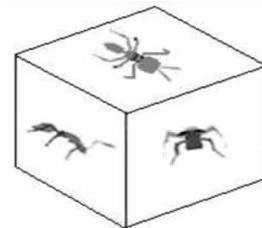


图2 普林斯顿模型库中的蚂蚁投影

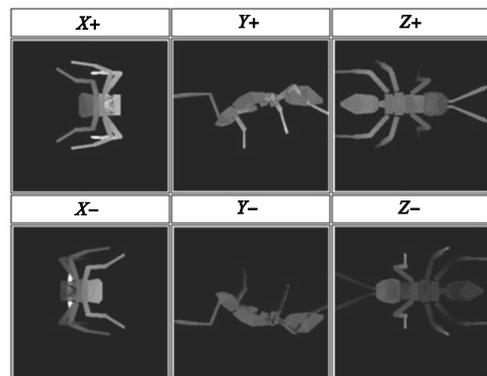


图3 普林斯顿模型库中的蚂蚁投影

1.2.1 边界方向直方图

Canny 算子^[13,14]是目前效果最好的算子。本文首先采用 Canny 算子进行图像的边缘检测,然后在 Canny 算子得到的边缘图像基础上计算边界方向直方图。对每个边界点,根据图像中该点的梯度方向计算该边界点处的法向量方向角^[15]:

$$\theta = \arctan\left(\frac{g(i,j+1) - g(i,j-1)}{g(i+1,j) - g(i-1,j)}\right) \quad (2)$$

其中, $g(i,j)$ 为经过高斯滤波处理后图像的深度值,以 5° 为范围对边界方向角进行划分,得到一个 72 级的方向直方图 $H[i]$ 。边界方向直方图具有平移不变性。对边界方向直方图进行归一化处理,如式(3),得到了缩放不变性。

$$H[i] = H[i] / S \quad (3)$$

其中, S 为图像的面积。

边界方向直方图不具有旋转不变性。为了解决这个问题,对直方图进行平滑处理:

$$H_s[i] = \sum_{j=i-k}^{i+k} H[j] / (2k + 1) \quad (4)$$

其中: $H[j]$ 是原始直方图; $H_s[i]$ 是平滑后的方向直方图;参数 k 决定平滑度,本文取 $k = 2$ 。

1.2.2 Zernike 矩描述子

Zernike 矩是基于深度图像区域形状的特征描述子,它的

基是正交径向多项式,相比于不变矩,它具有易构造高阶矩、旋转不变性、对噪声不敏感和冗余性小等特性,并且通过变换,它还可具有平移不变性。因此,本文提取三维模型深度图像的 Zernike 矩来表述深度图像的区域特征。

本文选择 0 ~ 12 阶的 Zernike 矩作为深度图像的区域特征。由于各阶 Zernike 矩代表不同的物理意义,在数值上相差也很大,因此需要对这 13 个 Zernike 矩进行归一化处理,使得每个分量在距离计算中起到同等的作用。

1.3 支持向量机学习算法

支持向量机(support vector machine, SVM)是根据维数理论与结构最小化原则,构造具有最大分量间隔的超平面、最小化泛化误差上限的一种基于统计学习的机器学习方法。它将向量映射到一个更高维的空间里,在这个空间里建立一个最大间隔超平面,在分开数据的超平面的两边建有两个互相平行的超平面,建立方向合适的分隔超平面使两个与之平行的超平面间的距离最大化。

对于整个三维模型库来说,某个具体类别模型只是众多模型中的极小部分,该类模型和其余模型之间是少对多的关系;有时返回给用户的检索结果中相关模型很多,不相关模型较少,也存在该类模型与其他模型是多对少的关系。因此本文采用加权支持向量机^[16](weight support vector machine, WSVM)进行模型训练检索。该方法通过引入两个惩罚权重因子分别对正类和负类样本的错划程度进行分开惩罚,达到控制样本非平衡分布对分类的影响效果。

$$\varphi(w, b) = \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^l \mu_+ \xi_{i+} + \sum_{i=1}^l \mu_- \xi_{i-} \right) \quad (5)$$

满足约束条件:

$$\begin{aligned} y_i(w \cdot x_i + b) - 1 + \xi_i &\geq 0 \quad i=1, 2, \dots, l \\ \mu_+ + \mu_- &= 1 \\ 0 \leq \mu_+, \mu_- &\leq 1 \end{aligned} \quad (6)$$

其中: C 和 ξ_i 分别是惩罚参数和松弛变量; ξ_{i+} 为正类样本; ξ_{i-} 为负类样本; μ_+ 是正类样本的惩罚权重因子; μ_- 是负类样本的惩罚权重因子,满足上式中的约束条件。

从以上公式可以看到,对于某类样本误判的惩罚力度最终是由惩罚参数 C 和惩罚权重因子的乘积共同决定。

本文将加权支持向量机应用于三维模型检索相关反馈技术中,实现三维模型检索中的相关反馈。

1.4 基于相关反馈的检索算法

相关反馈是一种指导性学习技术,用于提高系统的检索能力。三维模型检索中的一次完整的相关反馈过程如图 4 所示。

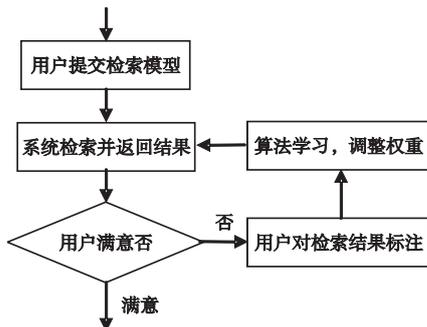


图4 相关反馈流程

具体算法如下:

a) 提取库中三维模型边界方向直方图和 0 ~ 12 阶 Zernike

矩特征,并对这些特征归一化即得到三维模型特征库。

b) 用户选择一个三维模型 Q ,提取模型 Q 的边界方向直方图和 0 ~ 12 阶 Zernike 矩特征并进行归一化。

c) 采用传统方法对查询模型进行检索,并将前 Nrt 个模型返回给用户。

d) 用户对返回的 Nrt 个模型进行标记,算法将用户对模型的标注信息标注在模型库中。

假设 $x_i (i=1, 2, \dots, Nrt)$ 为检索返回的模型(其中 i 为检索返回的第 i 个模型, Nrt 为检索返回模型数目), $y_i \in \{+1, -1\}$ 是用户标记矩阵。若 x_i 被标记为与检索模型相关的模型,则 $y_i = +1$,并将所有 $y_i = +1$ 的模型集合称为正类,记为 I^+ ;反之,若 x_i 和检索模型不相关,则 $y_i = -1$,并将所有 $y_i = -1$ 对应的 x_i 的集合称为负类,记为 I^- 。

e) 准备 SVM 样本集 (x_i, y_i) :

$$x_i \in I^+ \cup I^0, y_i = \begin{cases} +1 & x_i \in I^+ \\ -1 & x_i \in I^0 \end{cases} \quad (7)$$

f) 若正类样本数目远多于负类样本数目,则给正类样本施加一个惩罚权重因子 μ_+ :

$$\mu_+ = 1 - \text{正类样本数} / \text{检索结果总数 } N$$

同样也给负类样本施加一个惩罚权重因子 μ_- :

$$\mu_- = \text{正类样本数} / \text{检索结果总数 } N$$

反之,若负类样本数目远多于正类样本数目,则给正类样本施加一个惩罚权重因子 μ_+ :

$$\mu_+ = \text{负类样本数} / \text{检索结果总数 } N$$

同样也给负类样本施加一个惩罚权重因子 μ_- :

$$\mu_- = 1 - \text{负类样本数} / \text{检索结果总数 } N$$

g) 用 WSVM 对训练样本进行学习并构造分类器 $f(x)$ 。

h) 对模型库中每一模型 I_i ,求

$$\text{score}(I_i) = -f(x_i) \quad (8)$$

i) 对所有模型的 score 从小到大排序,返回结果。

j) 继续 d) 反馈循环,直至用户满意检出目标模型为止。则智能检索算法训练完成。

2 实验分析

本文实验数据包含颅骨、文物、人体、植物、头部、椅子沙发、桌子等共计 3 020 套三维模型数据。其中有西北大学可视化扫描的模型,也有从普林斯顿大学的三维模型数据库和互联网上下载的数据。

对各种模型应用本文算法进行检索可看出,本文算法具有较好的检索结果。图 5 为效果图,用户点击反馈之后的界面如图 6 所示。

对三维模型沿 X, Y, Z 轴随机地进行平移、旋转和缩放后,进行检索。鲁棒性测试结果如图 7 所示。从图 7 中看到,本文算法对三维模型的缩放、旋转和平移具有较好的鲁棒性。将本算法与文献[2,3]进行比较,结果如图 8 所示。从中可以看出,本文算法检索的准确率优于文献[2,3]。

3 结束语

本文提出了一种基于相关反馈的三维模型检索算法。该算法实现了三维模型的人机交互检索,采用支持向量机作为智能分类器。实验表明,本算法对三维模型的检索具有较好的检

索效率和精度。

三维模型底层特征和高层语义特征之间有一定的语义鸿沟,基于底层特征的检索结果往往不能满足用户需求,因此基于语义的三维模型检索是现在的研究热点,但其首要问题是模型的正确标注。本文实现了三维模型的半自动语义标注,为进一步研究语义检索打下了基础。

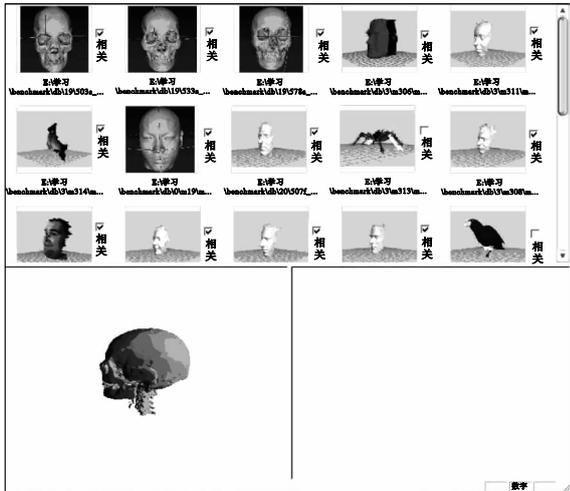


图5 颅骨模型检索效果

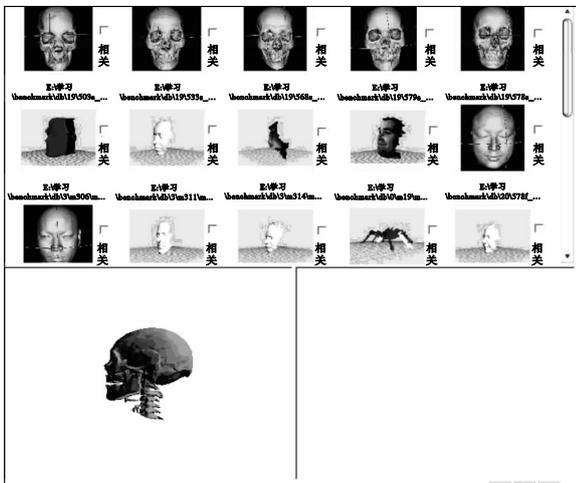


图6 相关反馈之后的检索效果

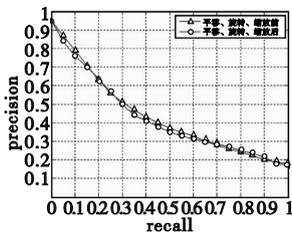


图7 平移、旋转和缩放处理前后效果对比

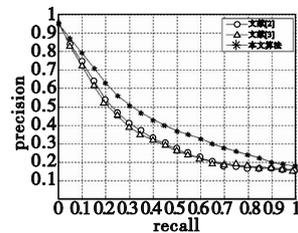


图8 三种算法检索效果比对曲线

参考文献:

- [1] 周明全, 耿国华, 韦娜. 基于内容图像检索技术[M]. 北京: 清华大学出版社, 2007.
- [2] MIN P, CHEN J, FUNKHOUSER T. A 2D sketch interface for a 3D model search engine[C]//Proc of ACM SIGGRAPH Conference. New York: ACM, 2002: 138.
- [3] CHEN Ding-yun, TIAN Xiao-pei, SHEN Yu-te, et al. On visual similarity based 3D model retrieval [J]. Computer Graphics Forum, 2003, 22(3): 223-232.
- [4] BOSCHE F, HAAS C T. Automated retrieval of 3D CAD model objects in construction range images[J]. Automation in Construction, 2008: 499-512.
- [5] ADÁN A, MERCHÁN P, SALAMANCA S. 3D scene retrieval and recognition with depth gradient images[J]. Pattern Recognition Letters, 2011, 32(9): 1337-1353.
- [6] GAO Yue, DAI Qiong-hai, WANG Meng, et al. 3D model retrieval using weighted bipartite graph matching [J]. Image Communi, 2011, 26(1): 39-47.
- [7] ELAD M, TAL A, AR S. Content based retrieval of vrml objects—An iterative and interactive approach[C]//Proc of the 6th Eurographics Workshop on Multimedia. New York: Springer-Verlag, 2001: 107-118
- [8] LEIFMAN G, MEIR R, TAL A. Semantic-oriented 3D shape retrieval using relevance feedback [J]. The Visual Computer, 2005, 21(8): 865-875.
- [9] NOVOTNI M, PARK G J, WESSEL R. Evaluation of kernel based methods for relevance feedback in 3D shape retrieval [C]//Proc of the 4th International Workshop on Content-based Multimedia Indexing. 2005: 325-334.
- [10] LENG Biao, QING Zheng, LI Li-qun. Support vector machine active learning for 3D model retrieval [J]. Journal of Zhejiang University Science A, 2007, 8(12): 1953-1961.
- [11] <http://shape.cs.princeton.edu/search.html> [EB/OL].
- [12] 张飞. 三维模型特征提取和相关反馈技术研究[D]. 西安: 西北大学, 2010
- [13] CANNY J. A computational approach to edge detection [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1986, 8(16): 679-698.
- [14] 章志勇, 潘志度. 三维模型几何相似性比较的研究[D]. 杭州: 浙江大学, 2005.
- [15] MIN P, HALDERMAN J A, KAZHDAN M, et al. Early experiences with a 3D model search engine [C]//Proc of the 8th International Conference on 3D Web Technology. New York: ACM, 2003: 7-8.
- [16] 杨志民, 刘广利. 不确定性支持向量机原理及应用[M]. 北京: 科学出版社, 2007.

下期要目

- ❖ 量子遗传算法研究进展
- ❖ 基于结构化信息源的本体构建方法综述
- ❖ 云数据中心虚拟资源管理研究综述
- ❖ 云计算下的数据存储安全可证明性综述
- ❖ 求解带时间窗车辆路径问题的混沌遗传算法
- ❖ 基于共享最近邻的离群检测算法
- ❖ 一种改进的动态多种群并行差分进化算法
- ❖ 基于 OpenMP 的 multi-critical 分子动力学并行算法优化

- ❖ 基于 OpenMP 的 Winograd 并行矩阵乘法应用研究
- ❖ 可交易瓶颈许可证的交通网络对偶均衡模型
- ❖ PPI 网络的改进谱聚类算法
- ❖ 基于混合势场法的移动机器人路径规划
- ❖ 价格折扣契约下应对突发事件下的闭环供应链协调策略
- ❖ 集合 CHI 与 IG 的特征选择方法
- ❖ 一种基于多终端动态协同的室内定位方法
- ❖ 基于改进遗传—模拟退火算法的公交排班优化研究
- ❖ 基于核聚类的近邻人工免疫网络算法研究
- ❖ 带反馈多传感器模糊最大熵单目标跟踪算法