

三维模型搜索引擎的研究进展^{*}

张 菁, 沈兰荪

(北京工业大学 信号与信息处理研究室, 北京 100022)

摘 要: 首先简要介绍了三维模型搜索引擎的发展概况, 然后讨论了三维模型搜索引擎的关键技术及相关的研究进展, 最后指出了三维模型搜索引擎进一步的发展前景。

关键词: 三维模型; 搜索引擎; 相似度匹配; 形状描述符

中图法分类号: TP391 文献标识码: A 文章编号: 1001-3695(2007)02-0010-03

Survey of 3D Model Search Engine

ZHANG Jing, SHEN Lan-sun

(Laboratory of Signal & Information Processing, Beijing University of Technology, Beijing 100022, China)

Abstract: This paper briefly addresses the development of 3D model search engine. The principle of the development of 3D model search engines is presents. At last we point out the trends of the development 3D model search engine.

Key words: 3D Model; Search Engine; Similarity Matching; Shape Descriptor

近年来, 随着三维技术的发展, 一种新型的多媒体数据——三维模型得到了广泛地应用。不同于传统的文本、音频等多媒体数据, 生动、形象、逼真是三维模型的突出特点, 这一特点使其在机械制造、娱乐游戏、生物化学、电子商务、艺术欣赏、考古应用以及虚拟现实等众多领域中有了长足的发展^[1]。计算机图形学和模型数字化工具的进步为构造高质量的三维模型创造了条件, 而因特网为传播这些模型提供了手段。今天三维模型的应用正在渗透到人们工作和生活的各个方面, 如何在网络环境中实现对三维模型的高效检索是目前搜索技术研究的热点^[2]。

1 三维模型搜索引擎的发展概况

传统搜索引擎的工作方式是对网页上的文本建立索引, 返回用户的查询结果是与这些索引相一致的页面^[3]。所谓搜索引擎(Search Engine), 是一种利用自动搜索技术, 通过对网络资源进行标记, 以提供检索服务的工具软件^[4]。多媒体搜索引擎索引的是图像、视频或音频等数据。这些多媒体数据除了文本外, 大多是非结构化的数据。目前, 大多数多媒体搜索引擎均支持使用关键字对多媒体数据进行搜索, 这是一种精确检索。多媒体数据则通过手工标注建立索引, 与用户输入的关键字进行精确匹配^[5]。该方式检索过程实现简单, 可直接获取语义信息, 结果精确; 但是手工标注不仅费时费力, 而且难以准确反映媒体内容, 容易受到标注者个人主观性的影响。

现在许多研究者都致力于基于内容检索的研究^[7], 这是一种相似度检索。例如 IBM 公司的 QBIC 就是一个通过图像

的颜色、纹理和轮廓等进行查询的图像检索系统^[6]。这种方式在预处理阶段就对媒体库中的图像资源提取颜色、纹理和形状等特征信息, 然后存放到特征库中; 在用户提出查询请求后, 系统对用户输入的查询对象进行特征提取并与特征库中的参考特征向量进行相似度匹配; 最后按照相似程度大小排序并返回到检索列表中^[7]。由于基于媒体内容进行检索, 具有较强的客观性, 但从底层特征提取高层语义较为困难, 检索结果不够精确。目前这种技术尚在发展中^[8]。

三维模型通常具有颜色、纹理和形状等特征。用户搜索三维模型时, 一般对颜色和纹理信息使用较少, 这主要是因为在实际应用中, 人们很容易对其颜色和纹理进行重新渲染。用户感兴趣的主要是三维模型的形状, 形状是三维模型最为本质的特征, 可以作为区分模型的主要依据。通常认为, 形状匹配是三维模型检索的主要方式。实际应用中三维模型的空间位置、大小和方向都是变化的, 对三维模型形状特征的描述必须考虑到三维模型的几何变换问题^[8]。例如, 不能因为待检索模型的旋转、缩放和平移就错误判定其属于不相似模型。显然, 三维模型形状描述符(Shape Descriptor, SD) 必须满足几何不变性, 即对模型的平移、旋转、缩放等几何变换具有不变性。

文献[9] 中, 普林斯顿大学提供了一种三维模型搜索引擎的研究方案, 这是一种基于文本关键字、二维和三维形状及其组合的查询方式, 如图 1 所示。方案中对基于二维和三维形状的查询采用了球面调和函数作为形状描述符, 该描述符具有几何变换不变性, 适合三维模型的相似度匹配。但是对用户绘制的具有内部细节的草图匹配较差。基于形状的人机接口使用起来还是比较麻烦, 用二维草图和三维草图进行查询的效率也不如直接用关键字查询。

文献[10] 中, 台湾大学的三维对象检索系统提供了一种基于文本关键字、二维草图(可包含内部细节) 或选择现有模型进行查询的研究方案。相似度是通过对比三维模型的二维

收稿日期: 2006-03-14; 修返日期: 2006-05-17
基金项目: 国家自然科学基金资助项目(60472036, 60402036);
北京市基金资助项目(4042008); 教育部博士点基金资助项目
(20040005015)

投影视图和二维草图的 MPEG-7^[11] 基于区域的描述符得到的 (图 2)。

文献[12]中,亚历山大的一个在线商业系统通过调整模型属性的权值进行相似度匹配(图3)。这些属性包括形状、颜色和缩放比例。该商业系统拥有4 500个三维模型,其中许多模型仅是在方向上有所不同。

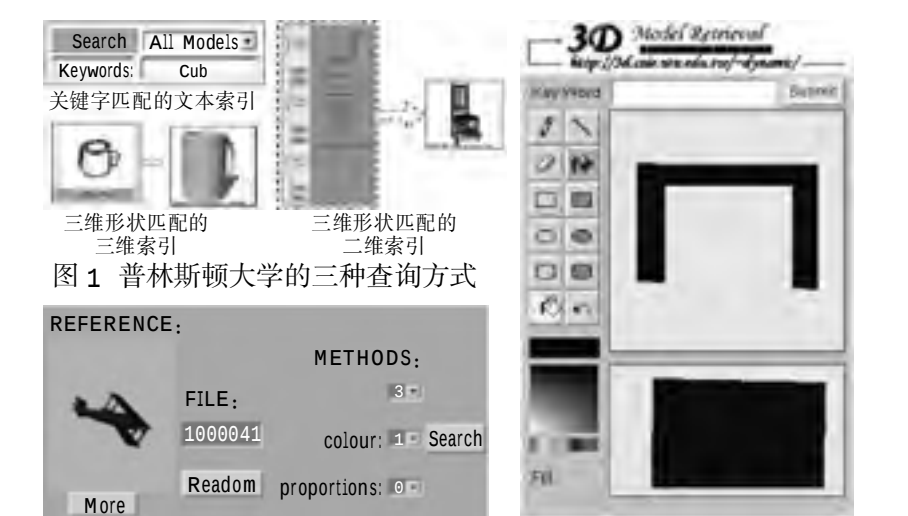


图 3 在线商业系统“Alexandria” 图 2 台湾国立大学的查询方式

2 三维模型搜索技术

三维模型搜索引擎是一个在网络中搜索三维模型的软件。它通过网络爬虫 (Crawler) 搜索网上的三维模型资源,并将这些资源保存到本地数据库中。用户可以通过文本对三维模型进行描述,也可以手工绘制三维模型草图表达所需的三维模型形状。搜索引擎既可以通过匹配关键字检索用户描述的三维模型,也可以通过匹配模型草图的形状特征检索三维模型。因此,三维模型搜索引擎的设计思想是通过网络爬虫搜索网上三维模型数据,向用户提供文本输入和绘制三维模型草图的查询接口,采用基于关键字和基于形状特征实现对模型的检索^[13]。

2.1 三维模型搜索引擎的组成

三维模型搜索引擎的总体架构^[9]由三部分组成(图4):网络爬虫和三维模型库构成了获取部分,这部分负责从网络中搜索可用的三维模型数据;分析索引模块和索引元数据库构成了分析部分,目的是为了对三维模型数据建立索引,并将索引保存在数据库中;查询接口、查询处理和匹配模块则构成了匹配部分,对用户查询进行匹配并返回结果。

其中,获取部分的网络爬虫由一个 Server 和多个 Clients 组成。Server 负责维护已访问和待处理 URLs 的 Hash 表、管理站点访问权限并负责对 Client 的操作进行调度;Client 同时启动 80 ~100 个进程对网络中的三维模型进行爬取 (Crawl) 搜索并存入本地数据库中^[9] (图 5)。

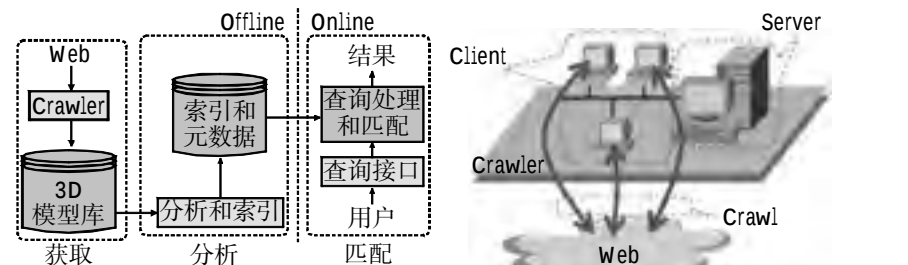


图 5 Crawler 的组成结构

分析部分的主要任务是提取三维模型的索引。由于用户可以通过文本、三维模型草图或模型的二维视图草图三种方式提交查询,因此三维模型的索引有文本索引、三维索引和二维

索引。文本索引可以从爬虫搜索到的模型网页以及模型文件本身提取。三维索引和二维索引的提取需要预先将模型文件格式统一转换为图形格式,其中三维索引可以直接从模型文件中提取,而二维索引还需要将模型文件投影为二维视图图像后才可以提取。这些索引分别被保存到相应的数据库中。

当用户提出查询请求时,由匹配服务器根据用户的查询请求分别与索引库中的模型索引进行文本匹配、三维形状匹配和二维形状匹配,按照匹配的相似度将相似的三维模型返回给用户^[9]。获取与分析部分通常离线进行,匹配部分则在用户查询提出后在线运行。

2.2 三维模型的相似度匹配

三维模型的相似度匹配采用关键字或形状特征匹配两种方法。关键字匹配是将用户输入的文本与文本数据库的索引进行匹配。文本索引源有三维模型的文件名、网页上的文本、URL 路径、网页上下文、网页标题、Wordnet^[14] 中同义词和上位词。形状特征匹配^[14]的过程是:首先计算查询草图与库中模型形状特征之间的距离,然后根据数据库模型和查询草图的邻近度分类,最后将距离最近的前 K 个模型返回(图 6)。

下面进一步讨论几种常用的匹配方法:

(1) 文本匹配方法。该方法比较的是文本索引的相似度。文本索引是通过计算词频/逆文档频数 (TF/IDF)^[15~17] 得到的。TF/IDF 实际上是一个词加权向量 ($vi = \sum wi \times ti$),相似度是两个模型的词加权向量的夹角。词加权 wi 是根据词在本文档和所有文档中出现的频率来定义的,公式为 $tf \times \log(N/df)$,其中 tf 为本文档的词频, N 为文档数量, df 为词在所有文档中的频率。根据 wi 的公式,文档集中包含某一词条的文档越多,说明它区分文档类别的能力越低,其权值越小;而文档中某一词条出现的频率越高,且该词在其他文档出现的频率越低,说明它区分文档类别的能力越强,其权值越大。当 df 定义为文档中包含该词的文档数时,称为 TF/IDF Log Words 法,而 df 为文档中包含该词的总次数时,称为 TF/IDF Log Occur 法^[18]。当用户通过输入关键字检索三维模型时,根据输入关键字在文本索引中的权值确定其查询结果,权值大小决定了每个模型文件在返回结果中出现的位置。

(2) 形状匹配法。该方法通过计算查询草图与数据库中模型形状特征的距离进行相似度匹配。形状特征,也称为形状描述符,相似度通常是两个形状特征向量的欧式距离^[19]。公式如下:

$$D(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

(1)

直接匹配法^[20]是一种形状匹配法,其相似度等于模型表面对应点距离的总和。这种方法在进行几何相似性匹配时,需要将两个模型以重心为基准点平移对齐,相似度计算简单,但是对齐操作是一个耗时的工作,对于大型数据模型库和实时系统来说,使用这种方法的效果不好(图 7)。

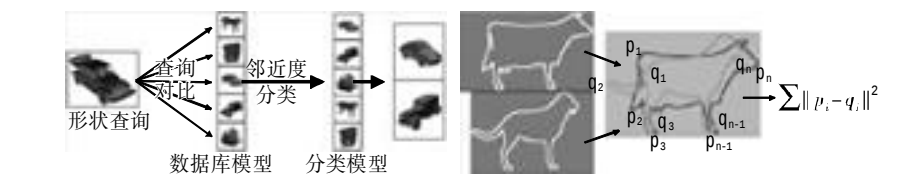


图 7 直接匹配法

在实时系统中比较实用的方法是 将三维模型表示为形状

描述符, 模型的距离为形状描述符的距离, 常用的形状描述符^[21~25]有:

形状直方图(Shape Histogram)。使用一组同心球作为直方图划分网格, 统计每个网格内的模型顶点数量, 然后通过计算直方图之间的欧氏距离来获得三维模型的几何相似度。

形状分布(Distribution Shape)。使用两点之间的距离作为几何函数, 计算三维模型的几何分布直方图, 然后通过计算几何分布直方图的范数距离来获得三维模型的几何相似性。

球面调和函数(Spherical Harmonics)。首先计算三维模型所有采样顶点在每个球面经纬度网格中的最大半径距离, 然后计算最大半径距离的球面调和函数系数, 最后通过球面调和函数系数比较来获得三维模型的几何相似性。

三维调和函数向量(3D Harmonics Vector)。首先对三维模型进行顶点采样, 计算一组同心球面与三维模型相交进行体素化, 然后计算每个同心球相交的球面调和函数特征向量, 最后通过计算球面调和函数特征向量之间的相似距离来获得三维模型的几何相似性。

扩展高斯图像(Extended Gaussian Images)。首先计算三维模型的扩展高斯图像, 然后计算扩展高斯图像的球面调和函数特征向量, 最后通过球面调和函数特征向量之间的相似距离来获得三维模型的几何相似性。

(3) 球面调和函数法。由于在实际应用中三维模型常常会有旋转、平移和缩放等几何变换, 因此形状描述符应满足几何变换不变性。球面调和函数法由于其特征向量容易计算和对比, 且具有几何变换不变性, 适合用于检索三维模型。获取三维模型的形状描述符需要预先将三维模型体素化, 然后提取其二维形状特征。具体步骤是首先计算出以三维模型重心为球心的最小包围球, 以该球半径的 $1/R$ 为单位将这个模型投射到 $2R \times 2R \times 2R$ 的网格中, 然后将网格中心与三维模型的重心对齐。每个格子的数值非 0 即 1, 如果任何一个格子内部, 存在三维模型表面上的点, 就将该格子的数值设为 1, 反之将该格子的数值设为 0。这样, 就得到了三维模型的体素化表示, 再将网格限制在一系列半径的球壳上, 每个球壳均对应一个二值化的球坐标方程:

$$f(r, \theta, \phi) = \text{Voxel}(r \sin(\theta) \cos(\phi) + R, r \cos(\theta) + R, r \sin(\theta) \sin(\phi) + R) \quad r \in [0, R], \theta \in [0, \pi], \phi \in [0, 2\pi] \quad (2)$$

最后将每个方程分解为球面调和函数的和, 组合系数就是一组具有几何变换不变性的特征向量, 即

$$f_r(\theta, \phi) = \sum_m f_r^m(\theta, \phi), \quad f_r^m(\theta, \phi) = \sum_{n=-m}^m a_{mn} \sqrt{\frac{(2m+1)(m-|n|)!}{4\pi(m+|n|)!}} P_{mn}(\cos \theta) e^{in\phi} \quad (3)$$

将这样一系列特征向量组合起来, 形成一个特征矩阵, 该矩阵就是形状描述符, 可以用来比较不同三维模型之间的形状差异(图 8^[13])。

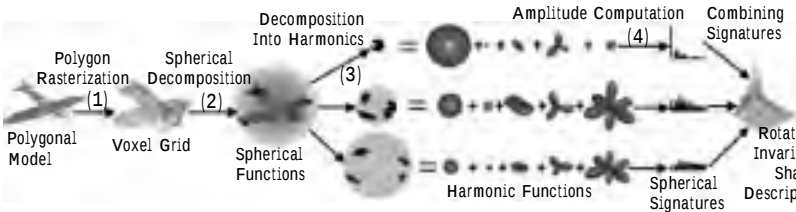


图 8 三维草图的球面调和函数法

使用二维草图与三维模型的二维投影视图进行匹配时, 需要提取二维形状描述符。其提取过程是: 首先对二维图像的边

界轮廓进行欧式变换, 以二维图像中心为圆心获取最小包围圆; 给定不同半径求出一组圆周函数; 将圆周函数展开为三角函数的和; 由于在相同频率内进行几何变换时不会改变幅值, 将圆周函数的特征值定义为由三角函数的幅值组成的序列; 组合不同的特征值从而得到边界轮廓的二维形状描述符。相似度为二维形状描述符的欧式距离^[9](图 9)。

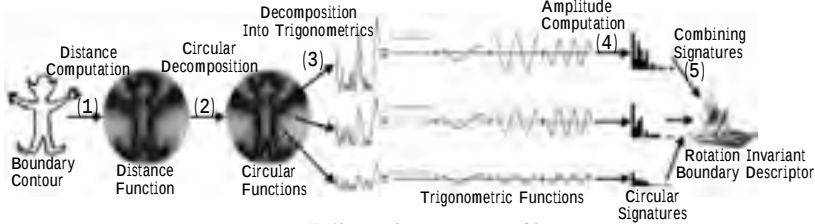


图 9 二维草图球面调和函数法

3 展望

本文首先对三维模型的发展概况进行了简要介绍, 然后对三维模型搜索引擎的关键技术与研究进展进行讨论, 并以球面调和函数为例, 详细讨论了形状特征匹配的方法。显然, 三维模型搜索引擎技术的研究, 将极大地推动网上智能化检索的进展, 提高人们生活与工作的质量。

目前看来, 三维模型搜索引擎在人机交互^[26]和匹配方法上还亟待改进。在人机交互上, 首先需要进一步提高人机接口的易用性, 如增加用户输入指南、用户草图后序处理、匹配时可加入内部细节等。另外, 亟待开发一种新的三维模型集成交互式绘图工具^[27], 使用户在绘制三维草图或二维草图上不必经过专业培训就能自由绘制。在匹配方法上, 可以加入颜色、纹理和结构属性进行组合匹配。为了提高查询速度, 可以增加一些常用的模板和训练集^[28]或进行部分形状匹配等, 相似度匹配方法的改进研究也是一个重要方面。

构建用户个性化分析模块, 实现个性化智能搜索, 则是另一个更具挑战性的课题。我们生活的世界是一个三维世界, 随着三维技术的发展, 可以看到多媒体数据的表示一定会逐渐由三维描述替换现有的二维描述方式, 对于三维模型搜索技术的研究具有远大前景, 是一个值得我们去深入研究和探索的前沿课题。

参考文献:

[1] Funkhouser, P Min, M Kazhdan, et al. A Search Engine for 3D Models[J]. ACM Transactions on Graphics, 2003, 22(1): 83-105.

[2] N Iyer, K Lou, S Janyanti, et al. Three Dimensional Shape Searching: State-of-the-art Review and Future Trends[J]. Computer-Aided Design, 2005, 37(5): 509-530.

[3] TREC. Text Retrieval Conference[EB/OL]. <http://trec.nist.gov>, 2005-12.

[4] Anthes H G. Search Engines: The Future[EB/OL]. <http://www.computerworld.com/softwaretopics/software/story/0,10801,91841,00.html>, 2004-04.

[5] C L Sable, V Hatzivassiloglou. Text-based Approaches for the Categorization of Images[C]. Paris: Proc. of Research and Advanced Technologies for Digital Libraries, 1999. 19-38.

[6] R C Veltkamp, M Tanase. Content-based Image Retrieval Systems: A Survey[R]. Utrecht University, 2000. 33-43.

[7] Y Rui, T S Huang, S F Chang. Image Retrieval: Current Techniques, Promising Directions and Open Issues[J]. Journal of Visual Communication and Image Representation, 1999, 10(3): 39-62.

1996, 29(8) : 1347-1357.

[5] M A Fischler, J M Tenenbaum, H C Wolf. Detection of Roads and Linear Structures in Low Resolution Aerial Imagery Using a Multi-source Knowledge Integration Technique[J] . Computer Graphics and Image Processing, 1981, 15(3) : 201-203.

[6] D Terzopoulos, A Witkin, M Kass. Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion[J] . Artificial Intelligence, 1988, 36(1) : 91-123.

[7] M Kass, A Witkin, D Terzopoulos. Snake: Active Contour Models [J] . International Journal of Computer, 1987, 1(4) : 321-331.

[8] J A Sethian. Numerical Algorithms for Propagating Interfaces: Hamilton-Jacobi, Equations and Conservation Laws[J] . J. Diff. Geom, 1990, 31(1) : 131-161.

[9] S Osher, J A Sethian. Fronts Propagating with Curvature-dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations[J] . Journal of Computational Physics, 1988, 79(1) : 12-49.

[10] J S Weszka. A Survey of Thresholding Techniques[J] . Computer Graphics and Image Processing, 1978, 7(2) : 259-265.

[11] 边肇祺. 模式识别[M] . 北京: 清华大学出版社, 2000. 96-108.

[12] K R Castleman. 数字图像处理[M] . 朱定刚, 等. 北京: 电子工业出版社, 1998.

[13] J C Rajapakse, J N Giedd, J L Rapoport. Statistical Approach to Segmentation of Single-channel Cerebral MR Images[J] . IEEE Trans. on Medical Imaging, 1997, 16(2) : 176-186.

[14] K Held, *et al.* Markov Random Field Segmentation of Brain MR Images[J] . IEEE Trans. on Medical Imaging, 1997, 16(6) : 878-886.

[15] S Z Li. Markov Random Field Modeling in Computer Vision[M] . Tokyo: Springer-Verlag, 1995.

[16] R M Haralick, L G Shapiro. Image Segmentation Techniques[J] . Computer Vision, Graphics and Image Processing, 1985, 29(1) : 100-132.

[17] D H Ballard, C M Brown. Computer Vision[M] . Prentice Hall, Englewood Cliffs, 1982.

[18] J B A Maintz, M A Viergever. A Survey of Medical Image Registration[J] . Medical Image Analysis, 1998, 2(1) : 1-36.

[19] J W Clark. Neural Network Modeling[J] . Phys. Med. Biol. , 1991, 36: 1259-1371.

[20] S Haykin. Neural Networks: A Comprehensive Foundation(1st ed) [M] . New York: Macmillan, 1994.

[21] G Valli, *et al.* Artificial Newral Networks for the Segmentation of Medical Images[D] . Toronto, Ontario, Canada: McGill University, 1994.

[22] D L Wang, D Terman. Locally Excitatory Globally Inhibitory Oscillator Networks[J] . IEEE Transactions on Neural Networks, 1995, 6(1) : 283-286.

[23] D L Wang, D Terman. Image Segmentation Based on Oscillatory Correlation[J] . Neural Computing, 1997, 9(4) : 805-836.

[24] N Shareef, D L Wang, R Yagel. Segmentation of Medical Images Using LEGION[J] . IEEE Transactions on Medical Imaging, 1999, 18(1) : 74-91.

作者简介:

何晓乾(1981-), 男, 山东曹县人, 硕士研究生, 主要研究方向为数字图像处理; 陈雷霆(1966-), 男, 重庆人, 教授, 博士, 主要研究方向为数字图像处理、三维图形技术; 沈彬斌(1980-), 女, 上海人, 硕士研究生, 主要研究方向为数字家庭网络; 房春兰(1974-), 女, 四川成都人, 讲师, 硕士, 主要研究方向为医学图像处理、计算机图形学。

(上接第 12 页)

[8] itendra Datta, Jia Li, James Z Wang. Content-based Image Retrieval: Approaches and Trends of the New Age[C] . Proc. of International Workshop on Multimedia Information Retrieval, ACM, 2005. 253-262.

[9] Patrick Min. A 3D Model Search Engine[D] . Princeton University in Candidacy for the Degree of Doctor of Philosophy, 2004. 27-52

[10] D Y Chen, M Ouhyoung. A 3D Object Retrieval System Based on Multi-resolution Reeb Graph[C] . Taiwan: Proc. of Computer Graphics Workshop, 2002. 16-20.

[11] T Sikora. The MPEG-7 Visual Standard for Content Description: An Overview[J] . IEEE Trans. on Circuits and Systems for Video Technology, 2001, 11(6) : 696-702.

[12] E Paquet, M Rioux. Nefertiti: A Tool for 3D Shape Databases Management[J] . SAE Transactions: Journal of Aerospace, 2000, 108: 387-393.

[13] Michael M Kazhdan. Shape Representations and Algorithms for 3D Model Retrieval[D] . Princeton University in Candidacy for the Degree of Doctor of Philosophy, 2004. 56-85.

[14] G A Miller. WordNet: A Lexical Database for English[J] . Communications of the ACM, 1995, 38(11) : 39-41.

[15] T Joachims. A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization[C] . Proc. of Int. Conf. on Machine Learning, 1997. 143-151.

[16] P Min, M Kazhdan, T Funkhouser. A Comparison of Text and Shape Matching for Retrieval of Online 3D Models[D] . Princeton University, 2004. 209-220.

[17] Rocchio. The SMART Retrieval System: Experiments in Automatic Document Processing[M] . Englewood Cliffs, NJ: Prentice-Hall, 1971. 313-323.

[18] G Salton. Automatic Text Processing: The Transformation, Analysis,

and Retrieval of Information by Computer[M] . Addison-Wesley Pub., 1988. 530.

[19] T Saito, J I Toriwaki. New Algorithms for Euclidian Distance Transformation of an *n*-Dimensional Digitized Picture with Applications[J] . Pattem Recognition, 1994, 27(11) : 1551-1565.

[20] Michael Kazhdan, Thomas Funkhouser, Szymon Rusinkiewicz. Shape Matching and Anisotropy[J] . ACM Transactions on Graphics, 2004, 23(3) : 623-629.

[21] R Zeleznik, K Herndon, J Hughes. SKETCH: An Interface for Sketching 3D Scenes[C] . Proc. of SIGGRAPH, ACM, 1996. 163-170.

[22] C Zahn, R Roskies. Fourier Descriptors for Plane Closed Curves[J] . IEEE Trans. on Computers, 1972, 21(3) : 269-281.

[23] K Siddiqi, A Shokoufandeh, S J Dickinson, *et al.* Shock Graphs and Shape Matching[J] . Int. Journal of Computer Vision, 1999, 35(11) : 13-32.

[24] 章志勇. 三维模型几何相似性比较的研究[D] . 杭州: 浙江大学, 2004. 2-6.

[25] 普建涛. 三维模型检索关键技术研究[R] . 北京: 北京大学, 2004. 2-10.

[26] 袁保宗, 等. 新一代(第四代) 人机交互的概念框架特征及关键技术[J] . 电子学报, 2003, 31(Z1) : 1945-1954.

[27] J Assfalg, A Del Bimbo, P Pala. Three-Dimensional Interfaces for Querying by Example in Content-based Image Retrieval[J] . IEEE Trans. Visualization and Computer Graphics, 2002, 8(4) : 305-318.

[28] Z Zhang. Iterative Point Matching for Registration of Free-form Curves and Surfaces[J] . IJCV, 1994, 13(2) : 119-152.

作者简介:

张菁(1975-), 女, 讲师, 博士研究生, 主要研究方向为图像/视频信号处理; 沈兰荪(1938-), 男, 教授, 博导, 主要研究方向为图像/视频信号处理、传输、压缩与应用。