

# 基于图像辅助的三维模型形状检索

刘杰洪<sup>1,2</sup> 黄 惠<sup>1</sup>

<sup>1</sup>(中国科学院深圳先进技术研究院 可视计算研究中心 深圳 518055)

<sup>2</sup>(中国科学院大学深圳先进技术学院 深圳 518055)

**摘 要** 随着 3D 建模技术的快速发展, 互联网上可用的 3D 模型库出现了爆炸式增长, 越来越多的 3D 模型可以方便地通过网络下载使用。这直接促使了 3D 形状检索技术的发展, 即给出特定的搜索信息, 要求系统搜索出符合要求的、相似的 3D 模型。文章提出了一种新的 3D 形状检索方法, 以 3D 模型作为输入, 系统将会从模型数据库中自动检索出与输入形状最相似的模型。对于给定的输入模型以及数据库中的每一个模型, 首先由计算机生成多幅在不同视角下的 2D 草图; 然后, 应用 Gabor 滤波器对每一幅 2D 草图提取图像上的局部特征, 并对特征进行量化, 从而得到代表该图像特征的直方图, 这样对于每一个 3D 模型将得到多个代表该模型的直方图; 最终, 通过对比两个模型之间直方图的相似性, 可以得到它们的相似性值, 从而检索出与输入模型最相似的模型。文章所提出的方法通过采取 2D 图像分析方法提取能反映 3D 模型的特征并计算出模型之间的相似性值。经过测试, 在一些公开的数据集上得到了较好的效果。

**关键词** 3D 形状检索; 2D 草图; 局部特征; 相似性值

**中图分类号** TG 156 **文献标志码** A

## Image-Assisted 3D Model Shape Retrieval

LIU Jiehong<sup>1,2</sup> HUANG Hui<sup>1</sup>

<sup>1</sup>( Visual Computing Research Center, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China )

<sup>2</sup>( Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China )

**Abstract** In recent years, due to the rapid development of 3D modeling technology, 3D model databases have been increasingly available on the Internet. More and more 3D models can be easily downloaded through the Internet. This has directly led to the development of 3D shape retrieval technology, in which the system needs to return a similar 3D model according to the user requirement. This paper presents a new 3D shape retrieval method, which takes a 3D model as input and the system automatically returns some models that are most similar to the input shape from the model database. For a given input model and every model in the database, first, generated a magnitude of 2D sketch images of the model from different perspectives by the computer. Next, for

收稿日期: 2017-02-10 修回日期: 2017-03-27

基金项目: 国家自然科学基金项目(61502471、61402459); 广东省自然科学基金研究团队(2015A030312015); 深圳市基础研究学科布局(JCYJ20151015151249564)

作者简介: 刘杰洪, 硕士, 研究方向为计算机图形学; 黄惠(通讯作者), 研究员, 研究方向为可视计算, E-mail: hui.huang@siat.ac.cn.

each 2D sketch image generated, the algorithm applies Gabor filter to extract the local features of the image, and quantifies the features in order to obtain a histogram representing the sketch image. For each 3D model, we then obtain a number of histograms representing the model. Thus, by comparing the histograms of every two models, we can compute the similarity value between the two models, and so retrieve the most similar shape to the input shape. In brief, the proposed method is capable of extracting effective features of 3D model through 2D image analysis method and evaluating the similarity between models. Experimental results show that the proposed algorithm performs well on some public datasets.

**Keywords** 3D shape retrieval; 2D sketch; local features; similarity value

## 1 引言

目前, 3D 模型库在工程、娱乐、医学影像等各个领域都出现了爆炸式增长, 越来越多的 3D 模型可以通过网络方便地下载使用。由于日渐丰富的 3D 数据, 形状检索、形状分析已经越来越受关注。其中, 3D 形状检索算法以一个 3D 模型作为查询输入, 要求寻找相似的形状模型作为输出。然而, 由于 3D 模型不同于 2D 图像, 如何简单地表示一个 3D 模型的复杂结构, 成为 3D 形状检索中一个关键的挑战。

由于在 2D 图像上提取特征的技术已经发展得较为成熟, 因此将其应用到 3D 形状检索技术中成为一个较为热门的思路。但由于 2D 图像信息与 3D 形状信息具有本质上的差别, 直接对 3D 模型作投影将会丢失大量有效的 3D 信息。因此, 如何通过 2D 图像完整地保留 3D 形状信息, 是解决 3D 形状检索问题的核心思路之一。

在目前较为先进的方法中, 基于 2D 草图的 3D 形状检索方法都需要寻找一个最佳的视角, 以用来匹配在该视角下 3D 模型在 2D 平面上生成的轮廓线条。然而这些方法所要寻找的最佳视角并非一定存在, 即不能保证存在一个视角, 在该视角下 3D 模型在 2D 平面上生成的轮廓线条与 2D 草图的匹配程度最高。同时, 许多寻找这些最佳视角的方法都需要人为有监督地挑选出

来。当数据量较大时, 这是一项较为耗费精力的工作。

本文提出采用词袋模型, 无需挑选出最佳视角, 只要在模型的包围球上均匀采样一定数量的视角, 并基于“如果从各个角度看两个模型都是相似的, 那么这两个 3D 模型是相似的”这个假设, 通过现时较先进的线图技术生成 3D 模型在不同视角下的 2D 草图。然后对生成的 2D 线图进行特征的提取和量化, 从而得到该 2D 图像的特征直方图。因此, 对于每两个 3D 模型, 我们的方法都分别得到不同视角下的多个特征直方图。通过对比模型之间的特征直方图, 可以计算出它们的相似性值, 从而能够检索出数据库中与输入形状最相似的模型。

## 2 相关工作

近年来, 3D 形状检索已经吸引了越来越多研究者的兴趣, 有关这方面的研究成果也相对比较丰富。基于形状描述子的不同表示, 可将形状检索的方法分为 3 大类: (1) 基于特征的方法; (2) 基于图的方法; (3) 基于几何的方法。

### 2.1 基于特征的方法

特征能够表示 3D 模型的几何和拓扑性质, 因此计算和对比模型的这些特征可以区分出不同的 3D 模型。不同的特征又可以分为 3 个子类:

①全局特征；②全局特征的分布；③局部特征。其中子类①和②的方法是用一个  $d$  维的长向量代表一个形状，这样每一个 3D 形状都是高维空间上的一个点，因此两个相似的形状在高维空间上是两个距离较近的点，通过解决  $k$  近邻问题就可以检索出  $k$  个最相似的 3D 形状。而基于局部特征的方法是用不同的特征描述子来描述 3D 模型上的每一个局部区域。目前这方面的研究成果较多：Zhang 和 Chen<sup>[1]</sup>通过计算体积、面积、傅里叶变换系数以及其他统计数据作为全局特征；Paquet 等<sup>[2]</sup>以包围盒和基于小波的描述子来进行 3D 形状的匹配，从而检索出相似的形状模型；Osada 等<sup>[3]</sup>提出了 D2 形状分布，该分布是基于模型表面上随机点之间的距离、角度、面积以及体积的计算，通过分析分布的相似性来得到形状的相似性，实现 3D 形状检索；Pu 等<sup>[4]</sup>通过计算并对比模型的一系列 2D 切片来查询两个 3D 模型的相似性，而 2D 切片的相似性对比是通过二维的 D2 形状分布函数来实现的；Shum 等<sup>[5]</sup>利用一个球坐标系统将 3D 模型的表面曲率映射到单位球上，通过搜索球形旋转空间，计算两个曲率分布之间的距离，以此作为两个模型之间的相似性距离；Zaharia 和 Preteux<sup>[6]</sup>提出了一种 3D 形状频谱描述子，该描述子被定义为形状索引值的直方图。该方法通过对比描述子得到模型之间的相似性距离。

## 2.2 基于图的方法

基于特征的方法只是单纯地考虑了模型的几何属性，而基于图的方法则尝试通过连接图来展示模型各个部分之间是怎样连接起来的。这类方法将一个 3D 模型抽象成为一个低维的连接图，并通过图中节点、边的连接关系反映模型的拓扑关系。这方面的研究成果有：Hilaga 等<sup>[7]</sup>提出了一种拓扑匹配技术，通过多分辨率 Reeb 图(MRGs)来计算两个模型之间的相似性距离；Sundar 等<sup>[8]</sup>通过采取匹配图的技术来匹配模型之

间的骨架结构，从而得到两个模型之间的相似性距离；而 Tam 和 Lau<sup>[9]</sup>则同时利用了拓扑和几何特征以提高形状检索的性能。

## 2.3 基于几何的方法

在形状检索中，基于几何的方法又可以分为 4 个子类：①基于视角的方法；②基于体量误差的方法；③基于加权点集的方法；④基于形变的方法。其中，基于视角的方法的核心思想是：如果两个模型从各个角度看都是相似的，那么这两个模型是相似的。Löffler<sup>[10]</sup>对每一个模型都计算出一个由许多 2D 图像组成的描述子，对于输入的 2D 查询图像，通过与计算所得的描述子进行对比，从而检索出相似的 3D 模型；Ansary 等<sup>[11]</sup>提出了通过在单位球上均匀采样 80 个视角来减少搜索空间，并以 2D 图像作为查询输入，用一个概率方法来提高 3D 形状的检索性能。基于体量误差的方法由 Novotni 和 Klein<sup>[12]</sup>提出：基于计算一个物体与另一个物体的偏移外壳序列之间的体量误差，以此用于 3D 形状匹配的几何相似性。这种方法的缺点是它们的不相似性测量是不对称的，并且不服从三角不等式。而在基于由加权的 3D 点组成的形状描述子方法中，Shamir 等<sup>[13]</sup>提出了一种由加权点的层次结构组成的形状描述子，以表示一种球形形状的近似；Tangelder 和 Velkamp<sup>[14]</sup>的加权点集则由一些具有高曲率值的点组成，并由这种点集构成他们的形状描述子。基于形变的方法<sup>[15,16]</sup>通过对比一对 2D 形状来测量一个形状准确配准到另一个形状所需的形变次数，并以此作为形状间的距离测量。但这种类似于形状拟合或者形状演化的方法较难在 3D 形状检索中取得较好的效果。

以上的 3 大类方法主要是基于刚性 3D 形状检索，而对于非刚性的 3D 形状检索，主要方法包括局部特征方法、拓扑结构方法、等长不变的全局几何属性方法、直接形状匹配方法等。由于非刚性 3D 形状检索问题不在本文的研究范畴，

故不一一展开详述。此外, 由于深度学习领域的快速发展, 近年来也出现了将深度网络应用到 3D 形状检索中<sup>[17-19]</sup>。但在本质上, 这些应用到深度学习的方法都可以归类到上述的基于几何方法分类中。

在现有的方法中, 大多数都是通过 3D 形状描述子来提取 3D 模型的特征, 也有部分通过 3D 模型的 2D 投影来提取所需特征。而在本文方法中, 将 3D 模型转换为多幅 2D 模拟草图, 这些草图不仅是 3D 模型的简单投影, 而且是包含模型表面丰富几何信息的图像, 故可利用现有成熟的 2D 图像特征提取技术将 3D 形状的潜在特征挖掘出来。与传统的 3D 形状描述子方法相比, 我们提出的方法能够省去传统 3D 形状描述子复杂繁琐的数学计算代价, 而且有效地结合 2D 图像先进的特征提取技术, 很好地提高了 3D 形状检索性能。

### 3 本文算法原理

本文提出的方法采用了词袋模型 (Bag of Features)<sup>[20]</sup>。首先, 在模型的包围球上均匀采样一定数量的视角, 并通过线图技术生成 3D 模型在所有采样视角下的 2D 草图。然后, 基于“如果从各个角度看两个模型都是相似的, 那么这两个 3D 模型是相似的”这个假设, 对模型在不同视角下生成的 2D 线图都进行特征的提取和量化, 从而得到该 2D 图像的特征直方图。因此对于每两个 3D 模型, 我们的方法都分别得到不同视角下的多幅特征直方图。通过对比模型之间的特征直方图, 可以计算出它们的相似性值, 从而能够检索出数据库中与输入形状最相似的模型。

#### 3.1 多视角下的 2D 草图生成

在两个模型的朝向基本一致的情况下, 如果从多数角度观察这两个模型都是相似的, 那么可以认为这两个模型是相似的。基于这个观点, 我

们提出通过采样出单位球上均匀分布的视角点 (图 1), 并在各个视角点 (沿着该视角点到模型中心的连线方向) 下, 利用线图技术对 3D 模型进行投影生成不仅带有 3D 模型轮廓信息还包含模型内部线条信息的 2D 图像。采样的视角点数量以及分布情况对最终的 3D 形状检索性能有一定的影响。在优化后的实验中, 视角点的采样数量为 102 (相应选取原因在后文 3.5 节中), 基本上能覆盖 3D 模型各个方向角度的视图。虽然这些图像是由计算机自动生成的, 但由于这种通过线图技术生成的 2D 图像与人手画的 2D 草图较相似, 故将这些图像简称为 2D 草图。在我们所采用的线图技术<sup>[21]</sup>中, 其生成的线条类型有多种 (图 2), 而我们主要采用了以下 4 种类型: (1) 轮廓线, 描绘的是 2D 封闭边界的线条; (2) 边界线, 描绘的是 2D 开口边界的线条; (3) 闭合轮廓线, 描绘的是由 3D 模型上法线与视角方向正交的点所组成的轮廓线; (4) 暗示的轮廓线, 描绘的是 3D 模型表面细节所构成的线条。

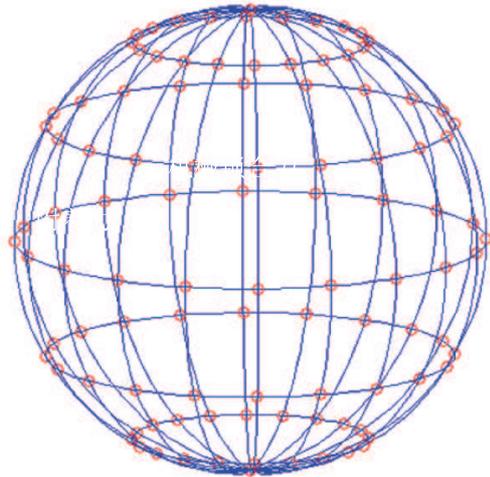


图 1 单位球上均匀分布的相机视角

Fig. 1 Uniformly distributed camera views on a unit sphere

#### 3.2 特征词汇表的生成

对于在 3.1 节得到的多视角下的 2D 草图, 本文方法对每一幅 2D 图像都进行特征的提取并得到一个特征词汇表, 这为之后用直方图表示

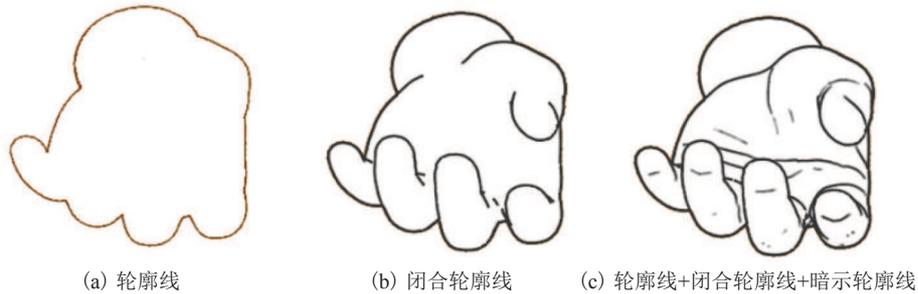


图2 计算机通过线图技术自动生成的几种2D草图

Fig. 2 Several 2D sketches automatically generated by the computer using line drawing

2D 图像做准备。在提取图像局部特征之前，对图像做一个简单的关键点检测。对于图像中关键点的获取，本文方法只是简单地在图像上每行每列等间距地采样出一定数量的图像点作为特征点，所得到的关键点在图像上呈规则的网格分布。之后在每一个关键点上提取局部特征的操作。

为了提取图像的局部特征：首先，对整幅2D 图像作傅里叶变换，得到2D 图像频域上的表示；然后，在频域上对其应用 Gabor 滤波器；最后，对卷积后的图像进行傅里叶逆变换，得到滤波后的空域上的图像表示，并计算响应的幅值作为滤波结果。我们的方法依赖于图像的局部特征描述子，该描述子能提取图像中局部小区域最必要的信息。由于我们要进行特征提取的图像对象是计算机自动生成的模拟草图，这样的草图实际上是由几种类型的轮廓线(外轮廓线、边界线、闭合轮廓线、暗示式的轮廓线)组成，而最能反映图像特征的就是这些线条的频率与方向，且不同类型轮廓线的频率是不一致的(暗示式的轮廓线要比闭合轮廓线的频率高)。而我们方法中所采用到的 Gabor 滤波，它的主要特点在于只对频域图像中符合滤波器的频率范围和方向范围的内容具有较大的响应值，而对于那些不在该范围的图像内容，Gabor 滤波器将会对其过滤去掉。同时，我们设置了不同方向以及频率的滤波器，以能提取图像中不同方向和频率的局部特征。图3

为2D 草图与滤波后图像的对比以及局部特征提取。对于本文所采用到的 Gabor 滤波器，其在频域上的定义为：

$$g(u, v) = \exp\left\{-2\pi^2 \left[ (u_\theta - \omega_0)^2 \sigma_x^2 + v_\theta^2 \sigma_y^2 \right]\right\} \quad (1)$$

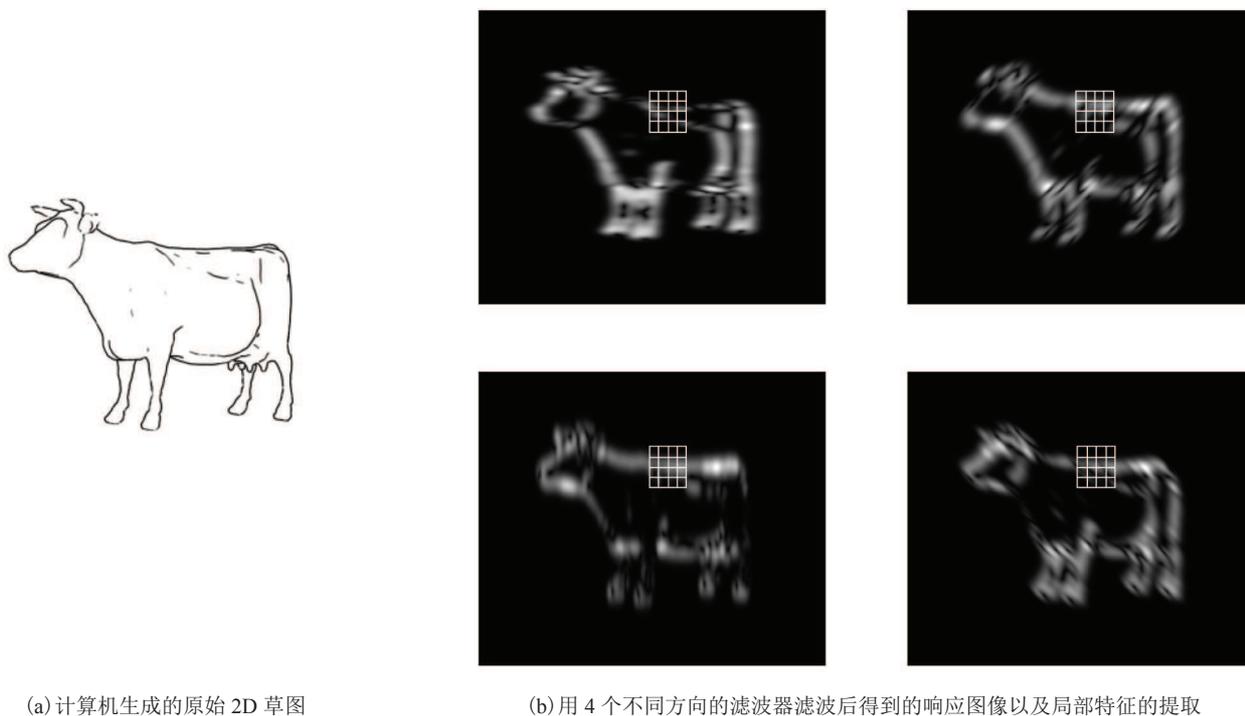
其中， $(u_\theta, \omega_0)$ 是坐标系旋转  $\theta$  角度后得到的坐标表示； $\omega_0$  表示峰值响应的频率； $\theta$  是滤波器的方向； $\sigma_x$  是频率带宽； $\sigma_y$  是角度带宽。用这样形式的 Gabor 滤波器对频域上的2D 图像进行卷积，将会过滤掉与该滤波器的频率以及方向不一致的许多局部内容，最终只保留与该滤波器的频率以及方向相似的局部内容(因为在这些区域上，卷积后的响应值较大)。

为计算特征空间的变换，设置了不同方向的  $n$  个 Gabor 滤波器。对每一幅2D 图像，都用这  $n$  个不同方向的 Gabor 滤波器对其进行卷积，因此得到  $n$  幅卷积后的响应图像：

$$R_i = \left\| \text{idft} \left[ g_i \times \text{dft}(I) \right] \right\| \quad (2)$$

其中， $I$  表示输入的2D 草图； $\text{dft}$  和  $\text{idft}$  分别表示傅里叶变换和逆变换； $g_i$  表示不同方向的 Gabor 滤波器。

在卷积后的每一幅响应图像上，我们在等距采样出来的关键点位置上提取局部特征。本文方法将特征设定为以关键点为中心，边长为  $k$  的一个  $k \times k$  正方形小网格。由于对每一幅图像都用不同方向的  $n$  个 Gabor 滤波器进行卷积，故在



(a) 计算机生成的原始 2D 草图

(b) 用 4 个不同方向的滤波器滤波后得到的响应图像以及局部特征的提取

图 3 原始 2D 草图与滤波后的响应图像

Fig. 3 Initial 2D sketch image and filtered response images

每一个关键点位置上都可以提取出  $n$  个  $k \times k$  的小网格特征。之后对这  $n$  个  $k \times k$  的正方形小网格合并后转化为一个  $n \times k \times k$  的长向量。如果一幅图像上有  $t$  个关键点被采样出来, 那么该图像能被提取出来的特征个数为  $t$ , 每一个特征为一个  $n \times k \times k$  的长向量。其中向量里每一维的值实际上是不同方向 Gabor 滤波器对原图像进行滤波后, 在关键点附近的响应值。同时, 对于原图像中没有任何线条的区域, 虽然在该区域能采样出关键点, 但实验中将从这些区域提取出来的特征向量全部赋值为 0。对滤波后的图像提取局部特征的示意图见图 4。

对数据库中所有的 3D 模型在不同视角下生成的 2D 草图都进行了特征提取后, 可对这些局部特征构建一个特征词汇表。具体而言, 由于提取出来的局部特征实际上是一个  $n \times k \times k$  的一维长向量, 对所有的长向量进行  $k$ -means 聚类, 那么每一个聚类中心就是特征词汇表中的一个核心

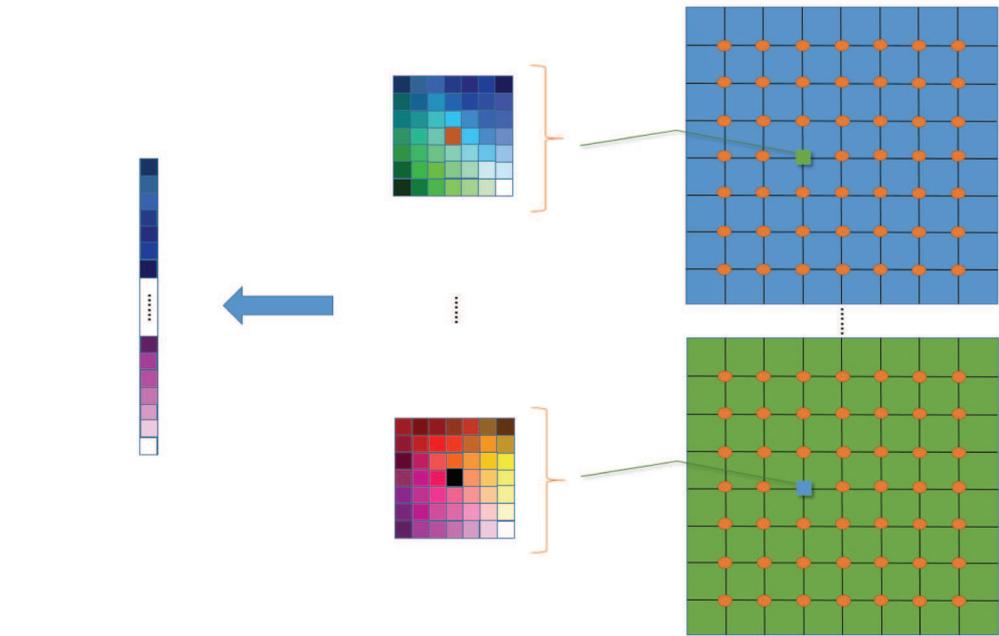
词。聚类的操作实际上是对提取得到的局部特征进行归纳总结, 筛选去掉异常或无效的特征, 形成一批具有较好代表性的局部特征。特征词汇表的构建将为后续的 2D 图像特征量化提供了准备。图 5 为对特征向量聚类的过程。

### 3.3 特征量化

为了进一步简化每一幅图像的特征表示, 对提取出来的特征进行量化。对每一幅图像提取出来的  $t$  个特征, 分别计算其到特征词汇表中每一个核心词的距离:

$$d_{ij} = \arg \min_j \|x_i - c_j\| \quad (3)$$

其中,  $x_i$  是某个关键点上提取出的特征向量;  $c_j$  是特征词汇表中第  $j$  个核心词。这将得到距离该特征长向量最近的一个特征核心词, 此时, 将原图像上该特征核心词的出现次数递增一次。对所有提取出来的特征向量进行类似的操作后, 将得到关于原图像的一个特征直方图。该直方图的横轴为特征词汇表中的核心词的个数, 纵轴为某一



(a) 在每一个关键点上，对  $n$  幅滤波后的图像提取得到的局部特征合并成一个  $n \times k \times k$  的向量  
 (b) 在以关键点为中心， $k \times k$  的网格上提取局部特征  
 (c) 在滤波后的图像上均匀采样得到的关键点位置

图 4 对滤波后的图像提取局部特征

Fig. 4 Extract local feature from the filtered image

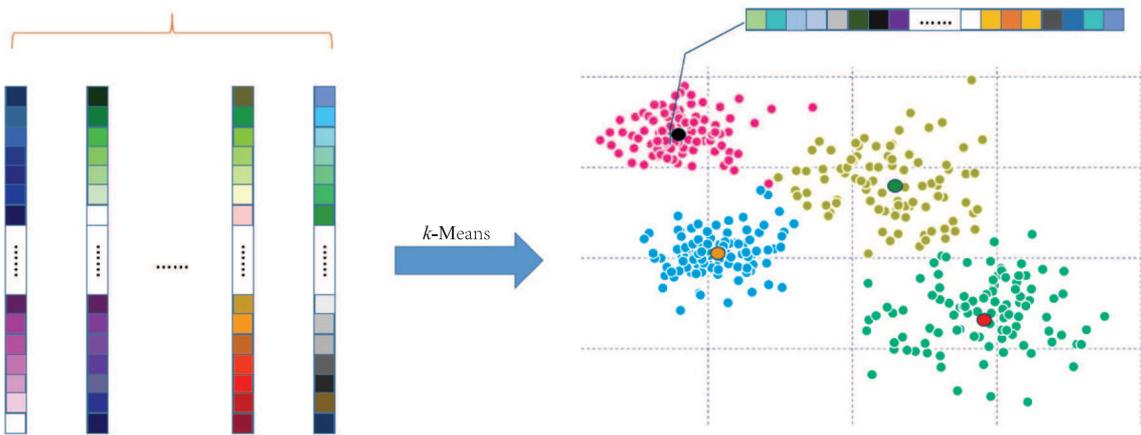


图 5 对提取的特征向量聚类

Fig. 5 Clustering the extracted feature vectors

个特征核心词在原图像上出现的次数。因此，可将  $t$  个  $n \times k \times k$  的特征长向量量化成一个长度为特征词汇表中核心词个数的向量。其中，向量里每一维的数值表示的是某一特征核心词在原图像上出现的次数。

为了提高后续的检索效果和加快检索的速

度，通过采用“词项频率-逆向文档频率”(tf-idf)模型以突出特征直方图中第  $j$  个特征词的重要性。主要的思想是，字词的重要性随着它在文件中出现的次数成正比增加，但同时也会随着它在语料库中出现的频率成反比下降。根据“词项频率-逆向文档频率”模型，通过公式(4)来计算

每一个特征词的权重  $w_j$ 。

$$w_j = \left( \frac{h_j}{\sum_i h_i} \right) \log \left( \frac{N}{f_j} \right) \quad (4)$$

其中,  $N$  代表文档总数 (即 2D 图像的总数);  $f_j$  代表特征词  $j$  在整个语料库中出现的频率; 而  $h_j$  本身代表的是特征词  $j$  在该文档 (即 2D 图像) 中出现的次数。于是, 对于每一幅 2D 图像, 都有一个长度为特征词汇表核心词个数的向量, 其中向量的某一维上的数值代表的是某一特征词的权重  $w_j$ 。图 6 是特征进行量化操作的示意图。

### 3.4 3D 模型检索

对于一个给定的输入 3D 模型, 在从数据库中检索出与之最为相似的形状之前, 需要对该输入模型进行以下几个步骤操作: 首先, 对该模型在  $\nu$  个不同的视角下生成模拟的 2D 草图; 然后, 分别对这  $\nu$  幅 2D 模拟草图进行关键点的检测与特征提取操作; 最后, 对提取出来的特征进行量化并得到一个长度为特征词汇表核心词个数的向量, 其中该向量的某一维上的数值代表的是某一特征词的权重  $w_j$ 。这时, 对于输入的 3D 模型, 得到  $\nu$  个这样的权重向量  $w_i (i \in \nu)$ , 此部分在线上查询时完成; 而对于数据库中的每一个 3D 模型, 得到  $\nu$  个类似的权重向量  $w_j (j \in \nu)$ , 此

部分是离线完成的。因此, 可以得到一个关于输入 3D 模型与数据库中任意一个 3D 模型的相似矩阵  $m$  (矩阵的维数为  $\nu \times \nu$ )。定义  $m_{ij}$  表示的是为输入模型在第  $i$  个视角下生成的模拟草图与数据库模型在第  $j$  个视角下生成的模拟草图之间的相似性值, 而该相似性值的计算公式为:

$$m_{ij} = \sum_{k \in s} w_{ik} \times w_{jk} \quad (5)$$

其中,  $s$  代表权重向量  $w_i$  和  $w_j$  拥有的共同项项集合, 即只对两幅图像中共同拥有的特征词的权重进行相乘累加。如果两幅图像没有共同的特征词, 那么它们的相似性值为 0。而对于两个 3D 模型之间的相似性值度量, 可以根据已经得到的任意两个模型在  $\nu$  个视角下的模拟草图之间的相似性度量矩阵  $m$  来计算。计算公式如 (6) 所示。

$$\frac{\sum_i \max_{j \in \nu} (m_{ij})}{\nu} \quad (6)$$

该计算公式表明, 对于输入的 3D 模型  $a$  以及模型库中任意的一个 3D 模型  $b$ , 在每一个视角  $\nu_i$  下, 都能找出一个视角  $\nu_j$ , 使得  $b$  在视角  $\nu_j$  下生成的模拟草图与  $a$  在  $\nu_i$  视角下生成的模拟草图较为相似。因此, 输入模型  $a$  的所有视角都有数据库模型  $b$  的某一个视角与之较为相似, 之后将所有视角下这些基于 2D 图像的相似性值累加并求

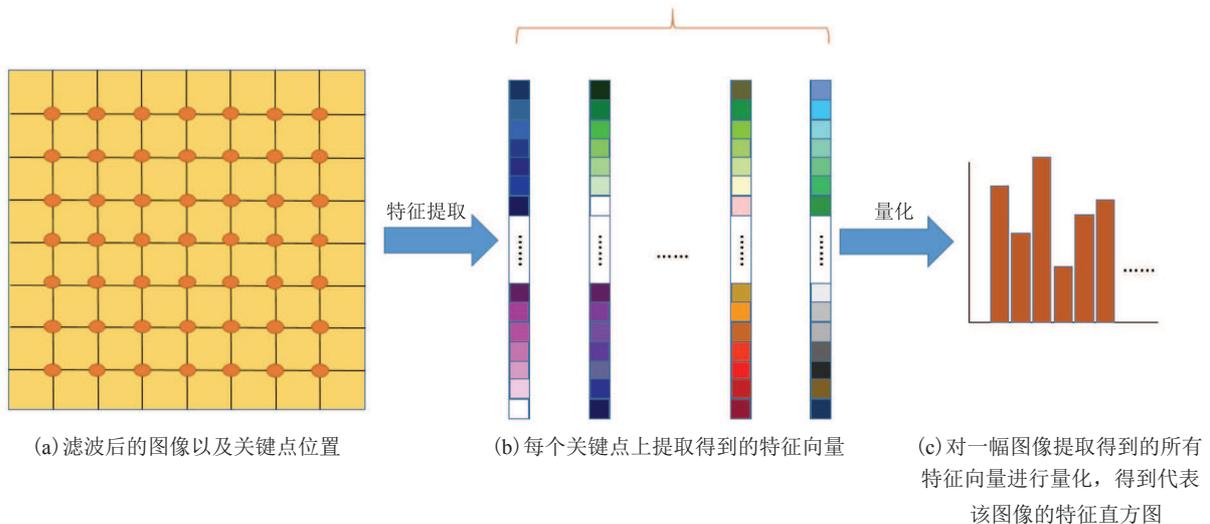


图 6 特征量化示意图

Fig. 6 Feature quantization schematic diagram

平均得到 3D 模型  $a$  和  $b$  之间的相似性值。最后在模型数据库中，对每一个模型都计算它与输入模型的相似性值，返回前  $n$  个相似性值最大的模型即为检索结果。相似性值计算过程如图 7 所示。

### 3.5 算法优化

在我们的系统中，对形状检索性能起关键性作用的参数较多。对比优化前后的系统发现，选择错误的参数将会严重地降低系统的检索性能。为了进一步提高算法性能，我们每次都只改变系统中的某一个参数，固定其他参数不变，以观察检索性能的变化。在每一次的参数调整中，记录下使系统具有最佳检索性能的参数值，并在其他参数调整过程中固定该参数值不变，直至所有的参数都调整完毕且系统的检索性能不再提高为止。对于每一个特定的数据库，调整的参数如下。

(1) 滤波器的带宽、频率、方向：由于本文算法需要在 2D 图像上提取局部特征，滤波器的带宽、频率参数决定了 2D 图像上哪些符合特征的局部内容会被保留，而不符合的特征将会被过滤掉。同时，由于我们的 2D 图像来自计算机对

3D 模型自动生成的多视角下的 2D 草图(2D 草图中既包含频率较低的闭合轮廓线、外轮廓线，也包含频率较高的暗示轮廓线)，因此适当的滤波器的频率值与带宽值将会有效地提高系统的检索性能。而滤波器的方向决定了滤波器的种类，若方向过少，则只能提取纵向和横向的特征；若方向过多，则会提取出太多杂乱和重复的特征。优化后的滤波器的带宽  $B=0.03$ ，滤波器的频率  $f=0.15$ ，滤波器方向种类数  $n=4$ 。

(2) 关键点个数、特征核心词的个数：关键点的个数指的是 3.2 节中在 2D 图像上等间距采样出来的网格点个数。过少的关键点容易忽略掉图像上的有用信息，而过多的关键点则会导致所提取的特征存在冗余。特征词汇表的容量(特征核心词个数)指的是在对数据库中所有的 3D 模型生成的 2D 草图提取局部特征后，对局部特征进行  $k$ -Means 聚类时，类别的总数  $k$ 。聚类总数  $k$  越大，在特征量化过程中，图像之间的直方图差别也越大，这将有利于形状检索性能的提升，但过大的聚类总数  $k$  也具有较高的计算代价。故我们的聚类总数  $k$  将综合这两方面进行选择。优化

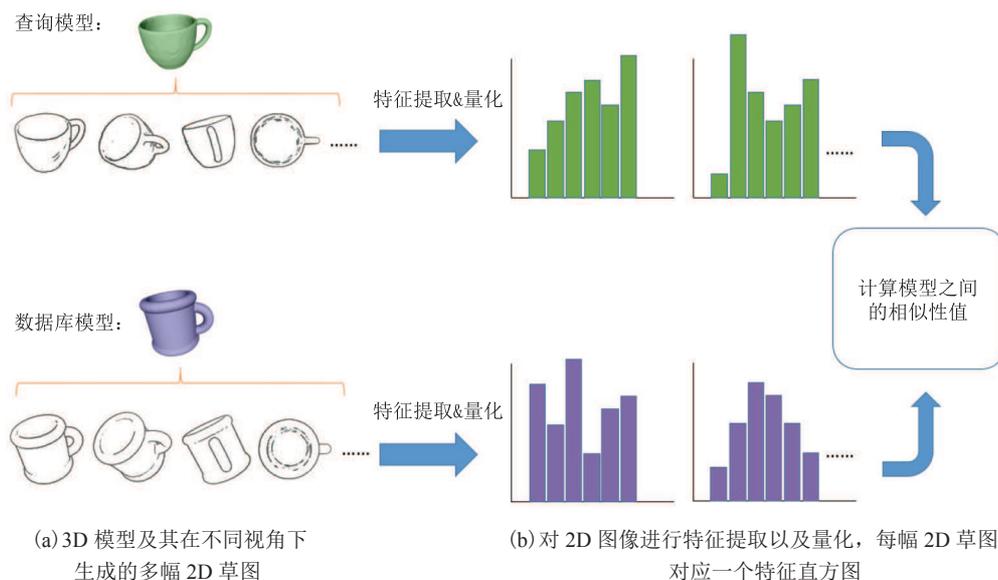


图 7 3D 模型之间的相似性值计算

Fig. 7 Computation of similarity value between 3D models

后的图像关键点个数  $k=1\ 024$ , 特征核心词个数  $t=1\ 000$ 。

(3) 采样视角的数量、草图中的线条类型: 采样视角数量指的是 3.1 节中单位球上均匀分布的视角点数量。计算机将会在这些不同的视角点下, 对模型生成不同视觉方向的 2D 草图, 因此采样视角点的数量直接决定了 3D 模型向 2D 特征转化过程中的效率。过少的视角点必然导致所生成的 2D 草图不能很好地反映 3D 模型的结构特征。草图中的线条类型指的是 3.1 节中采用线图技术生成的不同的轮廓线类型。由于不同的线条类型包含 3D 模型不同的结构特征, 故我们通过调整线条类型, 使提取的 2D 特征尽可能地包含 3D 模型本身的信息而又不会引入过多的重复信息。优化后的采样视角的数量为 102, 草图中线条类型种类数为 4。

值得说明的是, 对于视觉采样方法, 我们是这样做的: 对一个单位球按照纬度等分原则得到两极点之间共 10 个纬度间隔一致的圆环, 然后分别对每个圆环都按照经度等分原则, 得到 10 个经度间隔一致的视角点位置, 最后在两个极点位置分别增加一个视角点, 得到  $10 \times 10 + 2 = 102$  个视角点。另外, 生成草图时, 原始 3D 模型的摆放位置和角度对检索结果具有一定的影响, 影响的大小主要决定于采样视角点的多少以及分布情况。这是因为, 若视角点的数量较少且分布得极其不均匀, 即使两个 3D 模型较相似, 但由于初始的摆放位置不同, 也较难从少量的 2D 草图中找到两者相似的情况, 导致检索性能下降。但只要采样视角数足够多且视角点分布得当, 则模型初始摆放位置对检索性能影响较小。这是因为当模型在不同视角下的 2D 草图数量足够多时, 相似的两个 3D 模型之间总能找到它们在不同视角点下最相似的匹配视图。但考虑到过多的视角点会大大地增加计算代价, 检索速度也相应地下降。因此, 对于采样视角数量选择, 我们是权衡

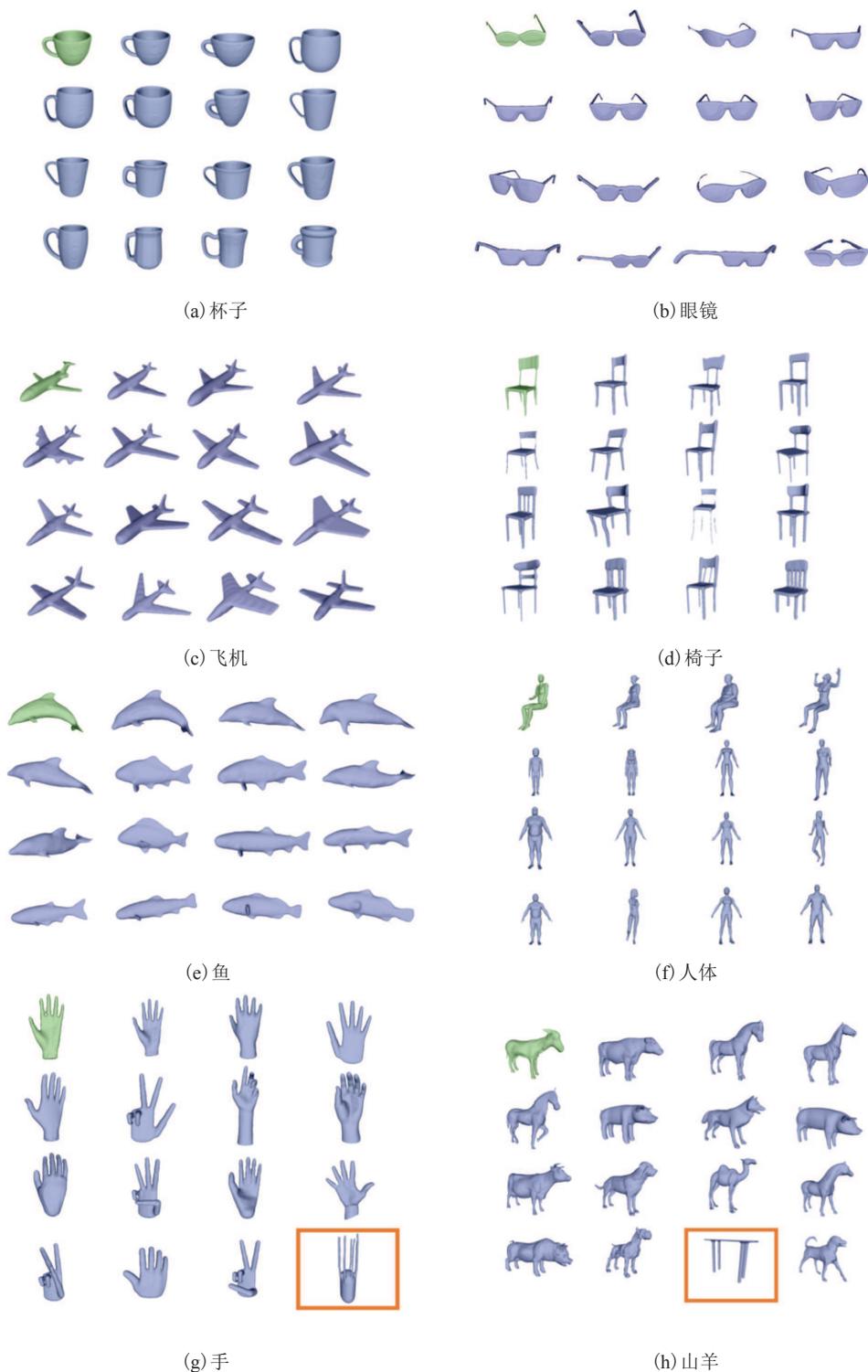
了视角点分布与计算代价这两方面的因素来进行抉择的。

## 4 实验结果与分析

### 4.1 3D 模型检索结果示例

本部分所采用的数据库来自 SHREC2014 中的 WMB<sup>[22]</sup>部分的数据。该数据共包含 20 个不同的类别(动物、人体、花瓶、杯子和眼镜等), 每一个类别均有 20 个不相同的 3D 模型(同一类别的模型之间具有一定的相似性), 因此整个数据库共有 400 个 3D 模型。同时, 由于这些模型均由无缝表面表示, 模型没有缺陷孔或者间隙, 因此这批模型被称为水密模型。将本文提出的 3D 模型形状检索方法在该数据库上进行测试。具体来说, 在每一次检索过程中, 首先从每一个类别里都挑出一个 3D 模型作为查询的输入模型; 然后, 当系统接收到输入模型后, 将会在包含 400 个模型的数据库中计算每一个模型与该输入模型的相似性值(相似性值越高代表两个模型越相似); 最后, 按照相似性值从大到小排序, 返回相似性值最大的前 15 个模型作为检索结果。在本部分实验中, 特征计算所耗费的时间约为 2 h, 而检索一个模型所需时间约为 3~5 s。

图 8 为 3D 形状检索结果。其中, 图 8(a)~(f) 返回的检索结果均与输入模型同属一个类别, 同时, 相似性值越大的模型确实与查询的输入模型越相似。这说明本文提出的方法具有一定的有效性和可行性。图 8(f) 中的输入模型为一具坐着的人体, 返回的检索结果中相似性值较大的那些模型都是与输入模型的动作姿势较为相似的人体模型, 但后面的几个模型出现了动作姿势不大相似的情况(站立或腿部交叉)。图 8(g)、(h) 返回结果中都出现了一个模型与输入模型不属于同一类别的情况(如章鱼的爪、4 个脚的桌子)。这是由于这些模型在较多的视角下生成的



注：绿色模型为查询的输入模型；蓝色模型为从数据库中检索返回的模型；从左至右、从上至下相似性值不同程度递减，相似值最大的是查询输入模型本身，结果实例中已忽略显示；橙色方框标注的模型表示该模型与输入模型不属于同一个类别

图 8 3D 形状检索结果

Fig. 8 3D shape search results

2D 草图与输入模型生成的 2D 草图较为相似而导致的。

#### 4.2 算法优化前后对比

在这一部分, 我们将对比本文算法优化前后的检索性能。同样采用的是包含 20 个类别, 每个类别均有 20 个 3D 模型的水密模型数据库。分别将优化前后的系统应用到该数据库进行测试对比。具体来说, 对数据库中每一个模型都做查询测试, 当系统接收到一个查询的输入模型后, 将在整个数据库中进行检索, 并返回一定数量(前 20、40、60、80 个)的最相似的模型作为检索结果。为了量化地对比不同方法的差异, 采用准确率和召回率两项指标来评定方法的性能。其中, 准确率指的是在检索返回结果中, 跟查询输入模型的类别一致的模型个数与返回结果总数的比率; 召回率指的是在检索返回结果中, 跟查询输入模型的类别一致的模型个数与查询输入模型的类别所包含的模型总数的比率(该数据库共有 20 个类别, 每一个类别有 20 个模型, 因此查询输入模型的类别所包含的模型总数恒为 20)。对每一个模型都计算以该模型作为查询输入时返回结果的准确率和召回率, 并在整个数据库中求平均值, 结果如表 1 所示。

从表 1 可知, 与优化前的系统相比, 经过算法优化后的系统在检索性能方面有了明显的提高, 验证了错误或不当的参数选择将会严重地降低系统的检索性能。同时也进一步说明了本文提出的 3D 模型形状检索方法具有一定的有效性与可行性, 即我们的方法根据输入模型返回的检索

结果是符合要求的。

#### 4.3 与其他方法的对比

在这一部分, 我们将优化后的本文算法与近年其他方法进行对比。所采用的是一个数据规模更大的 3D 模型库, 该数据库来自于 SHREC2014 中的 PSB (PRINCETON)<sup>[23]</sup>、WMB<sup>[22]</sup>、MSB<sup>[24]</sup> 混合的数据。该数据库包含 1 515 个 3D 模型, 总共 69 个类别, 每一个类别所包含的模型数量不相等, 相应的模型示例见图 9(由于篇幅限制, 只展示了部分有代表性的数据例子)。由于该库中同一类别的模型之间的差异比水密模型数据库的更明显, 因此本部分的对比实验与 4.1、4.2 节相比挑战难度更大, 对系统的检索性能要求更高。我们将本文经优化后的方法应用到该数据库上进行测试, 重新生成针对该数据库的特征词汇表, 并将测试结果与其他方法进行对比。采用的评价指标依然是准确率和召回率, 但由于该数据库每个类别所包含的 3D 模型数量都不相等, 故不能采用 4.2 节中的计算方法。取而代之的是: 首先, 固定召回率的取值范围(0.05 至 1, 每次递增 0.05); 然后, 在每一个固定的召回率下, 都计算每一个模型作为输入查询时, 所返回的检索结果中的准确率; 最后, 在整个数据库中计算所有模型在该召回率下的平均准确率。其中准确率与召回率的定义与 4.2 节相同。图 10 是本文方法与 Chen 等<sup>[25]</sup>和 Li 等<sup>[26]</sup>的结果对比图。由图 10 可知, 在该数据库的 3D 模型形状检索中, 本文的方法明显优于 Chen 等<sup>[25]</sup>和 Li 等<sup>[26]</sup>的方法, 这进一步说明了本文提出的 3D 模型形

表 1 返回的检索结果为 20、40、60、80 个时的准确率和召回率

Table 1 Precision and recall after 20, 40, 60, 80 retrieved items

系统状态	准确率 (%)				召回率 (%)			
	前 20	前 40	前 60	前 80	前 20	前 40	前 60	前 80
优化前	58.55	34.33	24.87	19.97	58.55	68.66	74.61	79.88
优化后	67.38	38.49	27.15	21.13	67.38	76.99	81.45	84.51
理想值	100.00	50.00	30.00	25.00	100.00	100.00	100.00	100.00

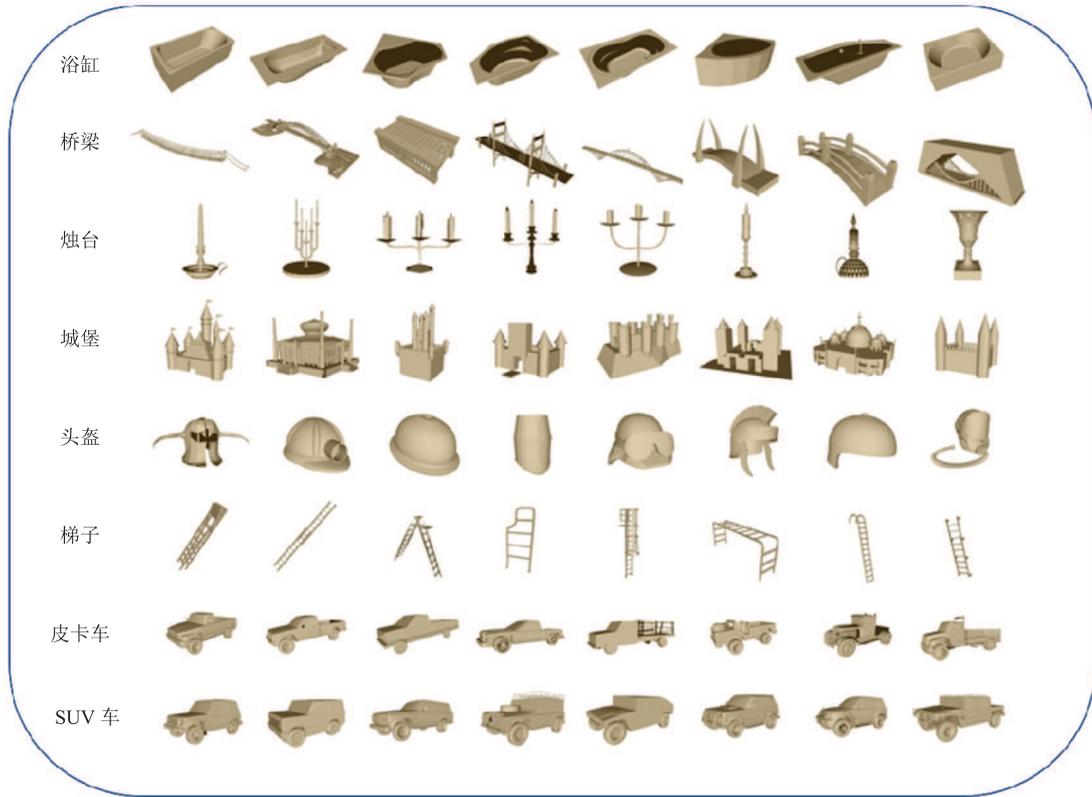


图9 PSB、WMB、MSB混合数据库中的部分示例

Fig. 9 Some examples in the PSB, WMB, MSB mixed database

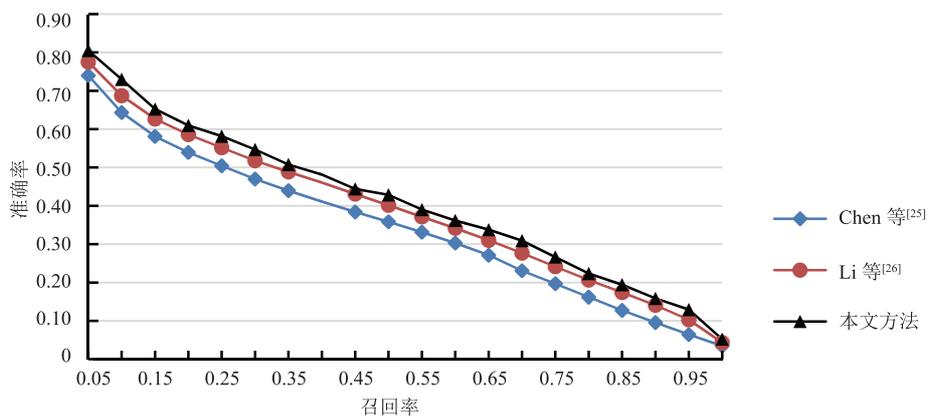


图10 与其他方法的结果对比

Fig. 10 Comparison with other methods

状检索方法的有效性与可行性。在本实验中，本文方法的特征计算所耗费的时间约为 20 h，检索一个模型所需时间约为 30 s。而 Chen 等<sup>[25]</sup>和 Li 等<sup>[26]</sup>的方法框架总体上与本文方法有较大的不同，他们主要侧重在各种不同的三维特征的计算

上，计算代价比较小，特征计算所耗费的平均时间也比本文的稍短。

#### 4.4 本文方法的局限性

本文方法依赖于线图技术对不同视角下的 3D 模型生成的 2D 草图质量，适用于刚性模型的

检索, 同时适合于表面细节丰富的模型数据库的检索。对于极其简单、表面细节不多的光滑模型, 本文方法对其生成的 2D 草图也只是一些零碎的线条, 这不利于图像特征的提取, 导致每一个 3D 模型所能获得的特征表示相当稀疏, 而且模型之间的特征表示差异不大, 直接影响了 3D 模型的检索性能。同时, 由于本文方法只考虑用不同视角下的 2D 线图来代表 3D 模型的结构特征, 没有添加语义上的辅助, 对于两个外观较为相似, 但属于不同类别的模型(如飞机和展开翅膀的小鸟), 本文方法检索性能还有待提高。此外, 由于本文方法没有考虑模型的骨架连接属性, 对于非刚性模型的检索(如查询输入模型与被检索模型同属一个类别, 但动作姿势有所不同), 本文方法仍需要进一步的改善。

## 5 总结与展望

在本文的工作中, 提出了一种 3D 形状检索方法, 并在公开的数据集上经过测试对比, 取得了较好的检索性能表现。本文方法从“如果从各个角度看两个模型都是相似的, 那么这两个 3D 模型是相似的”这个假设出发, 用不同视角下计算机自动生成的 2D 模拟草图来表示 3D 模型的结构特征, 并结合 Gabor 滤波器以及  $k$ -Means 聚类算法, 采用较为成熟的 2D 图像特征提取与量化的方法, 将 3D 模型转化为多个特征直方图, 最后通过对比两个 3D 模型的多个特征直方图得到二者之间的相似性值。本文 4.1、4.2、4.3 节中的结果示例与对比实验能够有效地说明本文方法确实能在刚性的 3D 形状检索中取得较好的性能。

在将来, 我们希望能够对目前方法的一些局限进行改善。首先, 可以考虑添加 3D 模型的骨架连接属性, 该属性信息的补充能够使动作姿势有所不同但与查询输入模型属于同一个物体的模

型仍然具有较高的相似性值。其次, 还可以考虑结合机器学习的方法。若能在检索之前首先从语义上识别出查询输入模型的类别属性, 然后再在更精准的范围内检索符合要求的模型, 将会大大提高 3D 形状检索性能。同时, 由于本文算法在检索之前的离线处理过程中(特征提取和特征量化)所耗费的计算时间比较多, 尤其是在数据库规模较大的情况下, 我们希望通过采取并行计算的方法来加速特征提取和特征量化过程; 而对于检索过程的计算速度优化, 若能如上述所说, 在语义上首先精准地识别出查询输入模型的类别, 将可减少与数据库中绝大部分其他不相关的模型之间的相似性计算, 从而大大提高检索过程的计算速度。

## 参 考 文 献

- [1] Zhang C, Chen T. Efficient feature extraction for 2D/3D objects in mesh representation [C] // International Conference on Image Processing, 2001: 935-938.
- [2] Paquet E, Rioux M, Murching A, et al. Description of shape information for 2D and 3D objects [J]. Signal Processing Image Communication, 2000, 16(1-2): 103-122.
- [3] Osada R, Funkhouser T, Chazelle B, et al. Shape distributions [J]. ACM Transactions on Graphics, 2002, 21(4): 807-832.
- [4] Pu J, Liu Y, Xin G, et al. 3D model retrieval based on 2D slice similarity measurements [C] // International Symposium on 3D Data Processing, 2004: 95-101.
- [5] Shum HY, Hebert M, Ikeuchi K. On 3D shape similarity [C] // Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition, 1996: 526-531.
- [6] Zaharia T, Preteux FJ. 3D-shape-based retrieval within the MPEG-7 framework [C] // Proceedings of SPIE 4304, Nonlinear Image Processing and Pattern Analysis XII, 2001: 133-145.

- [7] Hilaga M, Kohmura T, Kunii TL, et al. Topology matching for fully automatic similarity estimation of 3D shapes [C] // Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 2001: 203-212.
- [8] Sundar H, Silver D, Gagvani N, et al. Skeleton based shape matching and retrieval [C] // Proceedings of the Shape Modeling International 2003, 2003: 130.
- [9] Tam GKL, Lau RWH. Deformable model retrieval based on topological and geometric signatures [J]. IEEE Transactions on Visualization & Computer Graphics, 2007, 13(3): 470-482.
- [10] Löffler J. Content-based retrieval of 3D models in distributed web databases by visual shape information [C] // IEEE International Conference on Information Visualization, 2000: 82-87.
- [11] Ansary TF, Vandeborre JP, Mahmoudi S, et al. A bayesian framework for 3D models retrieval based on characteristic views [C] // Proceedings of the 17th International Conference on Pattern Recognition, 2004: 139-146.
- [12] Novotni M, Klein R. A geometric approach to 3D object comparison [C] // Proceedings of the International Conference on Shape Modeling & Applications, 2001: 167-175.
- [13] Shamir A, Scharf A, Cohenor D. Enhanced hierarchical shape matching for shape transformation [J]. International Journal of Shape Modeling, 2003, 9(2): 203-222.
- [14] Tangelder JWH, Veltkamp RC. Polyhedral model retrieval using weighted point sets [C] // Proceedings of the Shape Modeling International, 2003: 119.
- [15] Basri R, Costa L, Geiger D, et al. Determining the similarity of deformable shapes [J]. Vision Research, 1998, 38(15-16): 2365-2385.
- [16] Cohen I, Ayache N, Sulger P. Tracking points on deformable objects using curvature information [C] // Proceedings of the Second European Conference on Computer Vision, 1992: 458-466.
- [17] Xie J, Fang Y, Zhu F, et al. Deepshape: deep learned shape descriptor for 3D shape matching and retrieval [C] // Computer Vision and Pattern Recognition, 2015: 1275-1283.
- [18] Wang F, Kang L, Li Y. Sketch-based 3D shape retrieval using convolutional neural networks [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1875-1883.
- [19] Zhu ZT, Wang XG, Bai S, et al. Deep learning representation using autoencoder for 3D shape retrieval [J]. Neurocomputing, 2014, 204(C): 41-50.
- [20] Eitz M, Richter R, Boubekeur T, et al. Sketch-based shape retrieval [J]. ACM Transactions on Graphics, 2012, 31(4): 1-10.
- [21] Judd T, Durand F, Adelson E. Apparent ridges for line drawing [C] // Proceedings of ACM Siggraph, 2007: 19.
- [22] Veltkamp RC, Ter Harr F. SHREC 2007 3D Retrieval Contest [Z]. Utrecht: Utrecht University, 2007.
- [23] Shilane P, Min P, Kazhdan M, et al. The princeton shape benchmark [C] // Proceedings of the Shape Modeling International 2004, 2004: 167-178.
- [24] Siddiqi K, Zhang J, Macrini D, et al. Retrieving articulated 3D models using medial surfaces [J]. Machine Vision and Applications, 2008, 19(4): 261-275.
- [25] Chen Q, Fang B, Yu YM, et al. 3D CAD model retrieval based on the combination of features [J]. Multimedia Tools and Applications, 2015, 74(13): 4907-4925.
- [26] Li B, Johan H. 3D model retrieval using hybrid features and class information [J]. Multimedia Tools and Applications, 2013, 62(3): 821-846.