

工学硕士学位论文

基于内容与草图的三维模型检索方法

伊明

哈尔滨理工大学

2020 年 4 月

国内图书分类号: TP301

工学硕士学位论文

基于内容与草图的三维模型检索方法

硕士研究生: 伊明

导 师: 高雪瑶

申请学位级别: 工学硕士

学 科、专 业: 计算机科学与技术

所 在 单 位: 计算机科学与技术学院

答 辩 日 期: 2020 年 4 月

授予学位单位: 哈尔滨理工大学

Classified Index: TP301

Dissertation for the Master Degree in Engineering

3D Model Retrieval Method Based on Content and Sketch

Candidate:	Yi Ming
Supervisor:	Gao xue Yao
Academic Degree Applied for:	Master of Engineering
Specialty:	Computer Science and Technology
Date of Oral Examination:	April, 2020
University:	Harbin University of Science and Technology

哈尔滨理工大学硕士学位论文原创性声明

本人郑重声明：此处所提交的硕士学位论文《基于内容与草图的三维模型检索方法》，是本人在导师指导下，在哈尔滨理工大学攻读硕士学位期间独立进行研究工作所取得的成果。据本人所知，论文中除已注明部分外不包含他人已发表或撰写过的研究成果。对本文研究工作做出贡献的个人和集体，均已在文中以明确方式注明。本声明的法律结果将完全由本人承担。

作者签名：伊明

日期：2020 年 4 月 18 日

哈尔滨理工大学硕士学位论文使用授权书

《基于内容与草图的三维模型检索方法》系本人在哈尔滨理工大学攻读硕士学位期间在导师指导下完成的硕士学位论文。本论文的研究成果归哈尔滨理工大学所有，本论文的研究内容不得以其他单位的名义发表。本人完全了解哈尔滨理工大学关于保存、使用学位论文的规定，同意学校保留并向有关部门提交论文和电子版本，允许论文被查阅和借阅。本人授权哈尔滨理工大学可以采用影印、缩印或其他复制手段保存论文，可以公布论文的全部或部分内容。

本学位论文属于

保密 ☐，在 年解密后适用授权书。

不保密 ☒。

(请在以上相应方框内打√)

作者签名：伊明

日期：2020 年 4 月 18 日

导师签名：高彦斌

日期：2020 年 4 月 18 日

基于内容与草图的三维模型检索方法

摘 要

随着三维模型广泛地应用于我国制造加工技术和数字化建模技术，三维模型的分类与检索已经成为一个重要的研究方向。由于三维模型内在的复杂性，本文采用三维模型降维的方式，利用二维视图作为三维模型的检索条件，降低检索成本。本文基于视图的算法包括利用二维草图检索与利用投影视图检索。考虑到普通用户未必拥有可以直接检索的二维视图或者三维模型本体，所以本文第一部分利用手绘草图完成粗滤检索。第二部分利用甄选出的二维视图进行详细分类与检索。利用双重检索达到迅速且准确的三维模型检索的目的。

手绘草图检索部分首先对三维模型按照固定投影的方法获取二维视图集。在二维视图集中选取轮廓最长的 24 幅二维视图作为三维模型的特征视图集。然后将手绘草图进行降噪处理，并构建全局视图特征和二维形状分布直方图的加权集合作为集成特征描述子。利用集成特征描述子对手绘草图和二维视图进行相似性评估来检索三维模型。实验结果表明：本文中算法的检索性能优于大多数其它基于手绘草图的三维模型检索算法，对区域特征和边界轮廓特征具有更高的辨识度。在检索出的视图选取出理想的三维模型，作为下一步骤中视图检索的条件。

二维视图检索部分首先采用球体将选取出的三维模型完全包住，获取三维模型固定投影下的二维投影视图集；然后采用 Apriori 算法选出其中的典型视图，将典型视图作为卷积神经网络的输入；在 AlexNet 模型的基础上进行参数调整，并将其用于三维模型分类与检索；最后选取正向传播和反向传播相结合的方式对卷积神经网络进行训练，以提高其泛化性能。实验表明，该方法能够提高三维模型分类与检索的准确性和效率。

关键词 三维模型检索；二维草图检索；集成特征描述子；Apriori 算法；卷积神经网络

3D Model Retrieval Method Based on Content and Sketch

Abstract

With wide application of three-dimensional (3D) models in China's manufacturing, processing technology and digital modeling technology, it becomes an important research direction to classify and retrieve 3D models. 3D model has inherent complexity. The method of reducing the dimension of 3D model is adopted in this paper. Two-dimensional (2D) image is used as the condition of searching 3D model, in order to decrease retrieval costs further. 2D image retrieval algorithm in this article consists of 2D sketch retrieval algorithm as well as 2D view retrieval one. Ordinary users may not have 2D view and 3D model to be retrieved directly. Rough filtering retrieval is performed with rough-hand sketch in the first part of this article. The selected 2D views are applied to detail classification and retrieval in the second part. This paper utilizes double search to achieve the goal of retrieving 3D models fast and accurately.

The set of 3D model's 2D views is obtained according to the fixed projection method in freehand sketch retrieval part. From 2D view set, 24 two-dimensional views with the longest contours are selected as the feature view set of 3D model. Then, the hand-drawn sketch is denoised. A weighted set of global view features and 2D shape distribution histogram is constructed as the integrated feature descriptor. The integrated feature descriptor is used to evaluate the similarity between hand-drawn sketches and 2D views in order to retrieve 3D models. Experimental results show that the performance of the algorithm in this paper is better than those of other 3D model retrieval methods based on hand-drawn sketches. It has higher recognition of regional features and contour ones. An ideal 3D model is selected from the retrieved views as the condition of view retrieval in next step.

Firstly, a sphere is used to envelop the selected 3D model completely in 2D view retrieval part. The purpose is to obtain a set of 2D projection views for 3D

model along a fixed perspective. Secondly, Apriori algorithm is adopted to select typical views as the input of convolutional neural network. Based on AlexNet model, parameters are adjusted and used for 3D model classification and retrieval. Thirdly, convolutional neural network is optimized by combining forward propagation and back one in order to improve its generalization performance. Experiments show that this method can improve accuracy and efficiency of 3D model classification and retrieval.

Keywords retrieving 3D models, 2D sketch retrieval, integrated feature descriptor, Apriori algorithm, convolutional neural network

目 录

摘 要.....	I
Abstract	II
第 1 章 绪论.....	1
1.1 课题研究的背景和意义.....	1
1.2 国内外研究现状.....	2
1.3 课题的来源及研究内容.....	5
1.3.1 课题来源.....	5
1.3.2 课题的主要研究内容.....	5
1.4 本文章节安排.....	6
第 2 章 三维模型检索系统总体框架.....	8
2.1 检索系统体系框架.....	8
2.2 检索系统的处理流程.....	10
2.2.1 手绘草图检索架构.....	10
2.2.2 基于内容的检索架构.....	10
2.3 投影技术.....	11
2.3.1 固定视角投影.....	12
2.3.2 最佳视角投影.....	14
2.4 CNN 基本原理	15
2.4.1 卷积层.....	16
2.4.2 下采样层.....	17
2.4.3 全连接层.....	18
2.5 本章小结.....	19
第 3 章 手绘草图与三维模型预处理.....	20
3.1 三维模型数据集.....	20
3.2 手绘草图降噪.....	20
3.2.1 聚点消除.....	21
3.2.2 闭合检查.....	22
3.2.3 冗余笔画删除.....	22
3.3 甄选视图集.....	22

3.3.1 基于草图检索视图集	22
3.3.2 Apriori 算法	23
3.3.3 基于二维视图的检索视图集	23
3.3.4 基于多视角点云模型的检索视图集	25
3.4 本章小结	27
第 4 章 特征提取描述符及手绘草图检索	28
4.1 全局视图特征描述符	28
4.1.1 Zernike 矩特征描述符	28
4.1.2 Fourier 描述符	29
4.2 二维形状分布	30
4.3 集成描述符	31
4.4 检索系统设计	32
4.4.1 基于草图的检索模块	33
4.4.2 基于内容的检索模块	34
4.5 草图检索实验及结果分析	35
4.6 本章小结	35
第 5 章 基于内容的三维模型分类与检索	39
5.1 CNN 模型的构建	39
5.2 AlexNet 模型调参	39
5.3 点云模型的架构	42
5.4 AlexNet 模型训练与测试	42
5.5 相似性计算	49
5.6 本章小结	50
结论	51
参考文献	52
攻读硕士学位期间所发表的学术论文	56
致谢	57

第 1 章 绪论

1.1 课题研究的背景和意义

随着 3D 扫描与建模技术的发展,三维模型在日常生活中的应用也越来越普及。因此,越来越多的研究人员将精力投入到三维模型的研究中,其中,设计一个有效的检索系统帮助用户从大型三维模型数据库中找到所需的三维模型已经成为计算机图形学领域中的一个研究热点。检索三维模型最常用的方法是基于内容的方法与基于手绘草图的方法。

对于基于内容的方法而言,三维模型的分类是检索的前提。传统的三维模型分类方法是通过人工进行分类的,分类过程存在主观性强、实用性差等缺点,并且人工长时间工作会伴有错误的发生,所以需要开发一种更加智能与具有自动分类功能的三维模型分类器。近些年,机器学习技术取得了突飞猛进的发展,机器学习中的卷积神经网络在语音识别和图像处理方面有着独特的优越性。一些学者自然而然地想到将深度学习应用在三维模型的分类与检索上,其优点在于输入数据通过逐层卷积,样本特征转化成高层特征表示,用此抽象特征作模型的分类和检索更符合人类的视觉习惯。本文提出一种基于卷积神经网络的三维模型分类检索方法,通过固定角度投影的方法获得三维模型的二维视图集,采用卷积神经网络模型实现对三维模型的分类与检索。

然而考虑到普通用户未必拥有适当的三维模型,而绘制三维模型一般需要一定的专业技术。通过实例检索三维模型虽然很准确,但是从实际应用角度考虑,并不太实用。另一种方法是使用二维手绘草图作为查询输入,用户可以通过快速绘制二维草图来描述目标三维模型。基于草图的检索方法是一种比较合适的检索方式,但是二维手绘草图仅仅是简单的表示,其仅包含原始三维模型的部分信息。因此,实现基于二维手绘草图的检索比基于实例的检索更具有挑战性。

基于草图检索最大的优势在于:草图方便用户手绘,提高了普通用户对三维模型检索系统的使用能力。用户绘制三维模型视角的不固定及绘制风格差异可能较大,可能导致草图跟三维模型投影的形状完全不同。因此,很多研究工作就围绕将二维草图与三维模型进行匹配而展开。总之,基于手绘草图的三维模型检索在科学研究与实际应用中都具有重大的研究意义。

基于草图检索三维模型的方法总体上可分为两类：基于草图重建三维模型的检索方法和基于草图匹配三维模型投影视图的检索方法。现有常见的三维模型重建方法，需要使用一些复杂算法，而且只能够重建一些简单模型，因此利用二维草图重建三维模型检索的方法效率不高，且检索复杂模型的结果不够理想。基于草图匹配三维模型投影视图的检索方法与基于实例的三维模型检索方法类似，都是将三维模型从多个视角投影成二维视图，并利用草图与二维视图进行相似性计算，该方法顺利地把利用二维草图检索三维模型的问题转换为利用二维草图检索二维视图的问题，具有更高的实用性。

1.2 国内外研究现状

针对基于内容的三维模型分类与检索的算法，目前国内外学者进行了很多研究，它们都是将三维模型的特征信息作为输入，且根据现有的检索研究依据输入的特征信息划分类别，所以分类的研究进展就是检索的研究进展。其研究分为基于骨架、基于统计特征、基于几何结构分析和基于视图四大类。本课题主要研究基于视图的三维模型分类与检索。

在基于骨架的研究方面，林佼采用凸壳和有向包围盒进行三维模型的分割，利用得到的三维模型骨架特征向量进行相似性匹配^[1]。Hilaga 采用了测地线距离构建多分辨率 Reeb 图的方式，实现对三维模型的相似性评价^[2]。刘志通过骨架形状上下文特征对渲染图像进行分类，并提取角度结构特征进行相似度匹配^[3]。

在基于统计的研究方面，李月峰针对形状分布特征提取过程中，可能存在复杂局部表面模型采样过程偏移的问题，提出以三维模型的余弦值作为另一种统计特征^[4]。Zhuang 提出一种 3D 形状描述符和权重组合优化方案，并将其用于检索三维模型，根据描述符对应的距离直方图上的 L1 距离的加权和来判断模型之间的相似性，利用权重组合优化方案来提高三维模型数据库的检索精度^[5]。

在基于几何结构分析方面，杨军将点云数据通过占用网格规范化计算转化为二值 3D 体素矩阵，通过附加正则化项的随机梯度下降算法提取体素矩阵的特征，再通过共享权重的旋转增强对训练集进行数据增广并以此对模型标签进行预测^[6]。赵宇晨提出融合特征拓扑结构射影不变量和空间统计分布特征的方法。通过模型数据预处理完成各模型的姿态标准化和缩放标准化，然后采用空间区域划分和随机分布采样方法，构建三维模型的点特征拓扑结构射影不变量进行三维模型检索，提高三维模型检索的准确率与鲁棒性^[7]。Tao 提出了基于分布分配算法的 CAD 模型相似度计算方法，将 CAD 模型转换成邻接图。利用邻接矩

阵表示 CAD 模型。通过计算两个模型的邻接矩阵之间的相似度，来度量模型之间的相似性^[8]。

很多研究人员将精力放在基于视图的三维模型检索算法研究上。投影视图数量的选择是基于视图投影的一个难点，尽量保证对三维模型的特征充分提取，又不会产生冗余视图，造成计算量过大。最简单的视图集合是三维模型的三视图。Min 在三维模型上利用固定投影获取了正视图、俯视图、侧视图等三视图，然后利用这些二维视图进行相似性比较，但是这类视图集合包含三维模型特征较少，检索性能较差^[9]。由于深度学习在图像识别领域取得了不错的效果，一些学者将深度学习引入三维模型的检索之中。通过投影三维模型的方式得到二维视图集合，将视图集合作为深度学习的输入，用于检索与分类。这种处理方法也可以归为基于视图的三维模型检索算法。例如秦飞巍将深度学习引入三维模型的检索之中，他选择光场描述符提取出的 Zernike 特征作为深度神经网络的输入进行学习，分类效果比较理想^[10]。石民采用 Apriori 算法甄选出多视图集中的典型视图，比较三维模型对应的典型视图判断三维模型之间的相似度^[11]。冯元力提出一种多分支卷积神经网络下的三维模型识别方法。该方法将模型的球面全景图从多个角度展开，创建多幅视图作为三维模型的识别特征，识别效果较优^[12]。Bao 用足够大的圆柱体包围三维模型，圆柱体上的点与三维模型的圆心进行相连，穿过三维模型的面片数作为像素值。展开圆柱体得到三维模型的全景视图，然后利用深度学习进行特征提取，用于三维模型的分类^[13]。Kalogerakis 将不同视角下阴影视图和深度视图作为三维模型特征，并用卷积神经网络对视图特征进行学习^[14]。Chen 提出一种基于光场描述符(Light Field Descriptor, LFD)的检索算法。此算法利用正 12 面体包围模型，在正 12 面体的 10 个非对称顶点设置 10 个不同的光场。投影获取 100 幅视图，提取视图集的 Zernike 矩作为三维模型检索的特征^[15]。虽然视图集合能够充分表达三维模型的特征，但是每次提取出 100 幅视图会产生大量冗余，不利于计算机的计算。为了避免产生冗余特征，一些学者提出：先通过投影产生大量视图，然后选择出具有代表性的视图。Cyr 提出一种基于形状相似性的方法。首先通过不同的角度投影得到大量的投影视图，然后通过聚类分析得到有明显差异的视图集合，用选出的视图集来代表三维模型^[16]。因此本课题提出一种新的视图投影方法和卷积神经网络的结合方法。

基于草图检索三维模型的方法总体上可分为两类：基于草图重建三维模型的检索方法和基于草图匹配三维模型投影视图的检索方法。现有常见的三维模型重建方法，需要使用点的深度坐标计算和复杂算法等，而且只能够重建一些

简单模型^[17,18]。基于草图匹配三维模型投影视图的检索方法，主要包含三个阶段的研究内容：首先是进行手绘草图的识别。其次是获取三维模型的二维投影视图集合。最后是提取二维投影图像与手绘草图的特征集合，二者进行相似性比较。针对这三个阶段，不少学者都已作了相关研究工作。针对每个阶段的工作，本文作出总结。

首先，在草图的识别阶段，Fang 提出一种新的热形状描述符，使用自动编码器进行特征处理^[19]。Paul 提出的算法能够识别出几何图形，提高了识别草图阶段的准确率^[20]。Hammond 提出的算法能够利用简单几何图形的识别器进行复杂草图识别^[21]。Xie 提出一种基于非线性深度学习的检索算法，通过网络模型识别出距离最小的样本^[22]。Li 发明一种智能化的草图识别工具，可以通过监督学习正确地区分草图的形状描述^[23]。Ye 也开发了一种新的基于场景的草图识别系统^[24]。Cao 提出一种基于三视图的草图识别方法，该方法能够对草图进行建模处理^[25]。Li 通过概率框架对三维模型的投影视图进行整合，再与草图特征进行比对^[26]。

其次，针对三维模型投影二维图像的研究，与基于内容的三维模型分类与检索相同。

最后，对于提取草图和二维投影图像的特征描述符已经有很多研究，主要分为两个大类：提取全局特征和提取局部特征。在提取全局特征的过程中，很多学者采用二维视图的外围轮廓作为特征描述。Dai 使用编码量化的全局特征作为神经网络的输入，在 SHREC 中取得较好的成果^[27]。Shao 对手绘草图与二维投影图像的轮廓特征进行矢量化变换，对形成的变换特征进行相似性比较^[28]。Zou 提出一种基于特征袋的描述符，比较手绘草图与二维投影视图之间的相似性^[29]。而在提取局部特征的过程中，首先将视图的内部分成若干个区间，在每个区间内进行编码，用以获取区间信息。Yoon 利用扩散的张量特征在每个区间内进行编码，利用直方图提取的特征计算草图与二维投影图像的相似性^[30]。谢晓华利用上下文信息填补语义鸿沟，能够在用户意向不明确的情况下给出更加合理的推荐^[31]。Cao 利用基于风格敏感的三维模型检索方法和三视图用户草图，开发了一个新的三维模型检索原型系统，实现了用户友好的图形界面^[32]。Wu 利用骨架结构图表达二维视图的内部结构信息，描述手绘草图与二维投影视图的空间关系^[33]。Li 使用视点熵分布来衡量视觉的复杂性，并采用形状上下文的方法比较草图与二维投影图像的相似性^[34]。Li 提出一种基于自适应视图聚类 and 语义信息的三维模型检索框架，提出了基于熵的三维信息的复杂性度量方法^[35]。Li 提出了一种新的基于语义树的算法。首先建立三维语义树，树中的每个叶节

点都包含一组语义信息，对应一个三维模型，然后测量三维语义树与手绘草图的语义相关性，度量手绘草图与三维模型之间的相似性^[36]。同时，一些学者结合全局特征和局部特征的优势，进行混合方面的研究。Aono 利用 Zernike 矩和 HOG 特征的融合提出一种新的基于草图的检索算法^[37]。Lei 提出一种包含外围轮廓特征与内部骨架特征的三维模型检索算法^[38]。Li 提出一种混合二维投影视图的全局特征和局部特征相结合的复合特征，其中全局特征采用形状字符代表物体的外部边界形状，局部特征通过方向梯度直方图表示内部细节^[39]。Mao 提出的算法将特征向量定义为全局特征和局部特征的复合，保证了描述子的全面性^[40]。而在实际应用过程中，可以发现只考虑单一特征来评估手绘草图与三维模型之间的相似度是不可靠的。其中全局特征描述忽略了手绘草图和二维投影视图内部所包含的模型信息，丢失了大量的内部特征。局部特征描述对于轮廓特征描述不足，只考虑了内部区域的特征，不能获取全局结构特征。因此，将两者复合进行研究，可以提高三维模型检索的准确率。

1.3 课题的来源及研究内容

1.3.1 课题来源

本课题来源于国家自然科学基金(编号：61502124)、黑龙江省自然科学基金(编号：F201420)和黑龙江省普通高校基本科研业务费专项资金(编号：LGYC2018JC014)。

1.3.2 课题的主要研究内容

通过阅读大量的文献，对国内外三维模型检索方法的起源、发展和趋势有了深刻的理解，分析并总结目前所面临的问题与不足，对三维模型的检索方法进行了仔细分析和深入研究，本文主要研究内容为以下几个部分：

1. **草图预处理** 基于手绘草图匹配三维模型投影视图的检索方法中，由于绘画者受绘画水平所限制，可能会造成手绘草图噪声太多的问题。如草图端点处容易造成的冗余笔画，笔画与笔画之间连接不完整，存在交叉过连或者空隙，以及绘画者抖动时造成聚点等。故本文采用判断曲率值大小的方法消除冗余笔画，用偏移角度和大小的方法判断闭合检查，利用 Jin 提出的滑动窗口消除聚点，降低草图噪声，有利于后续检索的准确性^[41]。

2. **三维模型投影方式** 由于直接基于模型的检索方式计算量过大,所以本文提出了基于二维图像的三维模型检索方法,在草图检索和二维视图的检索中,采用固定投影的方式,利用二维视图表示三维模型,降低检索数据量,获取的投影视图集特征能够尽量地保证三维模型的特征不丢失。

3. **集成描述子特征提取** 在手绘草图检索三维模型的算法中,为了解决单一描述符对二维视图特征提取不全面的问题,本文在全局视图描述符和二维形状分布的基础上进行集成,得到一个能更好地表示二维视图特征的描述符。即本文提出的集成描述子,该描述子对区域特征和边界轮廓具有更高的辨识度。

4. **手绘草图初步检索** 本文利用 ModelNet-40 数据库作为算法的训练测试库,利用固定投影的方式进行模型降维,利用集成描述子提取三维模型的特征并存入数据库,提取输入草图的特征并与数据库中特征进行初步检索,得到所需实例模型或者近似实例模型,为后续进一步检索作准备。

5. **CNN 进行模型训练与分类** 针对传统三维模型分类检索方法,本文在二次检索过程中,利用 AlexNet 模型作为三维模型检索的深度学习模型,将三维模型固定投影所得视图甄选出一部分作为 CNN 的输入,先进行分类以减少检索的数量。

6. **基于内容二次检索** 第二次检索即为基于内容的检索,利用 Euclidean 距离进行模型特征检索,求出所需模型,两次检索结果的结合既避免了基于实例方法可能出现的没有所选模型情况与基于草图方法可能出现检索效果不好的情况。

1.4 本文章节安排

本文基于二维图像对三维模型检索进行研究,检索算法包括基于二维手绘草图的检索、基于二维视图的模型检索和基于多视角点云切片的模型检索。其中,手绘草图检索为粗略检索。首先对草图进行降噪处理,然后对三维模型固定投影获取二维视图集,接着利用集成描述子获得手绘草图和二维视图的集成特征,最后利用距离公式检索出最相近的 8 个三维模型。上述过程为第一次检索,将第一次检索的结果作为基于二维视图模型检索的输入,将选中的模型的二维视图集作为 CNN 的输入,得到该类模型的分类,以降低检索的数据量,提高检索速率。然后利用距离公式,求得最佳检索结果。论文共 5 章,各章的内容结构安排如下:

第 1 章,绪论。本章首先简要介绍了关于三维模型检索的研究背景及意义,

并论述了目前基于二维视图的三维模型在国内外的研究现状，以及目前存在的问题，然后介绍了本文课题来源及主要研究内容，最后对本文的章节安排进行陈述。

第 2 章，三维模型检索系统总体框架。检索系统包括两次检索，包括手绘草图检索架构与二维视图的实例检索，并针对每一步进行详细介绍。然后介绍三维模型降维的投影技术，详细解析 CNN 网络模型的结构。

第 3 章，手绘草图和三维模型预处理。本章主要提出解决草图噪声对检索结果造成的影响，包括冗余笔画，闭合检查，消除聚点等，提高检索性能。介绍了本文实验所需的数据集合，详细介绍了基于草图的模型检索、基于二维视图的模型检索和基于点云的模型检索需要作的预处理，以及如何进行投影变换。

第 4 章，特征提取描述符及草图检索。本章提出由全局视图描述符和二维形状分布结合的集成描述子，用于提取手绘草图与二维视图的特征，最后通过实验验证该描述符对区域特征和边界轮廓具有更高的辨识度，解决单一描述符对二维视图特征提取不全面的问题，具有更高的检索性能。

第 5 章，基于内容的三维模型分类与检索。主要包括基于二维视图的检索方法和基于多视角点云投影的方法。利用第一次检索得到的结果作为第二次检索的输入，精确检索结果，利用 CNN 先进行模型的分类，降低检索时间的消耗，再进行模型检索。然后对本文提取的两种基于视图的检索方法进行有效性的测试。

最后对全文的工作进行总结并对未来发展进行了展望。

第 2 章 三维模型检索系统总体架构

三维模型检索方法以基于内容的检索方式为主，即用现有的三维模型检索三维模型。然而普通用户未必有现成的三维模型，导致基于内容的检索方法局限性较大。本文检索系统首先研究的基于手绘草图的检索方法可以更加方便的表达用户的检索意图。在初次检索过程中，使用手绘草图与三维模型的二维视图特征集进行相似性的比较，得到多个检索模型。用户可以按照需求选择一个三维模型进行基于内容的二次检索，降低由于用户手绘草图表达不明确或者检索性能较差造成的检索结果不明确。基于手绘草图检索和基于内容检索结合的方法具备二者的优点，避免了二者的缺点。本章描述基于草图和内容所设计的体系框架，介绍了后续用到的主要技术，包括投影技术和深度学习技术。

2.1 系统体系框架

本文实现的检索系统包括基于手绘草图的三维模型检索和基于内容的三维模型检索。基于手绘草图的三维模型检索主要包括手绘草图的绘制、手绘草图降噪、手绘草图特征提取、三维模型固定投影、投影视图集的甄选过程、投影视图集的特征提取和相似度比较。基于内容的三维模型检索主要包括初次检索模型结果选取、投影视图集的甄选过程、AlexNet 模型训练和三维模型分类与检索。图 2-1 所示为系统的总体框架。

1. 手绘草图降噪。用户在检索系统上绘制二维草图，可能由于用户绘画功底的不同导致出现噪声，本文处理噪声的方式包括删除冗余笔画，闭合检查，聚点消除等，降噪能更明确的表达用户的意图，方便后续草图的特征提取及草图检索。

2. 手绘草图特征提取。本文提出了由全局视图特征和二维形状分布直方图结合的集成描述符作为新的描述符，全局视图特征是 Zernike 矩和傅里叶描述符结合而成，用此描述符提取手绘草图的特征检索三维模型。

3. 三维模型固定投影及投影视图集的甄选过程。直接用二维草图检索三维模型，由于维度的问题导致特征相差较大，本文的方法是进行三维模型的降维处理，即通过投影的方式将三维模型变成二维视图的表示，具体算法详见 3.3.1 节。

4. 投影视图集的特征提取。利用轮廓长度甄选出三维模型投影之后的视图集，利用集成描述子进行视图集特征提取，该视图集的特征即表示三维模型的

特征，为后续草图检索三维模型作准备。

5.相似度比较及检索模型选取。利用距离公式计算出与草图最相近的 8 个检索模型，此为粗滤检索，并在这 8 个模型里选取出用户最想要的模型，为二次基于内容的检索作准备。

6.投影视图集的甄选过程。在二次检索的过程中，本文采用的是基于二维视图的模型检索，在投影的视图集中采用 Apriori 算法选出特征最明显的视图集合作为甄选视图集，将甄选视图集的视图作为卷积神经网络的输入，训练 CNN 分类器。

7.AlexNet 模型训练和三维模型分类与检索。首先对输入三维模型进行分类，以降低检索花费的时间，然后在分类类别中利用距离计算公式进行检索。

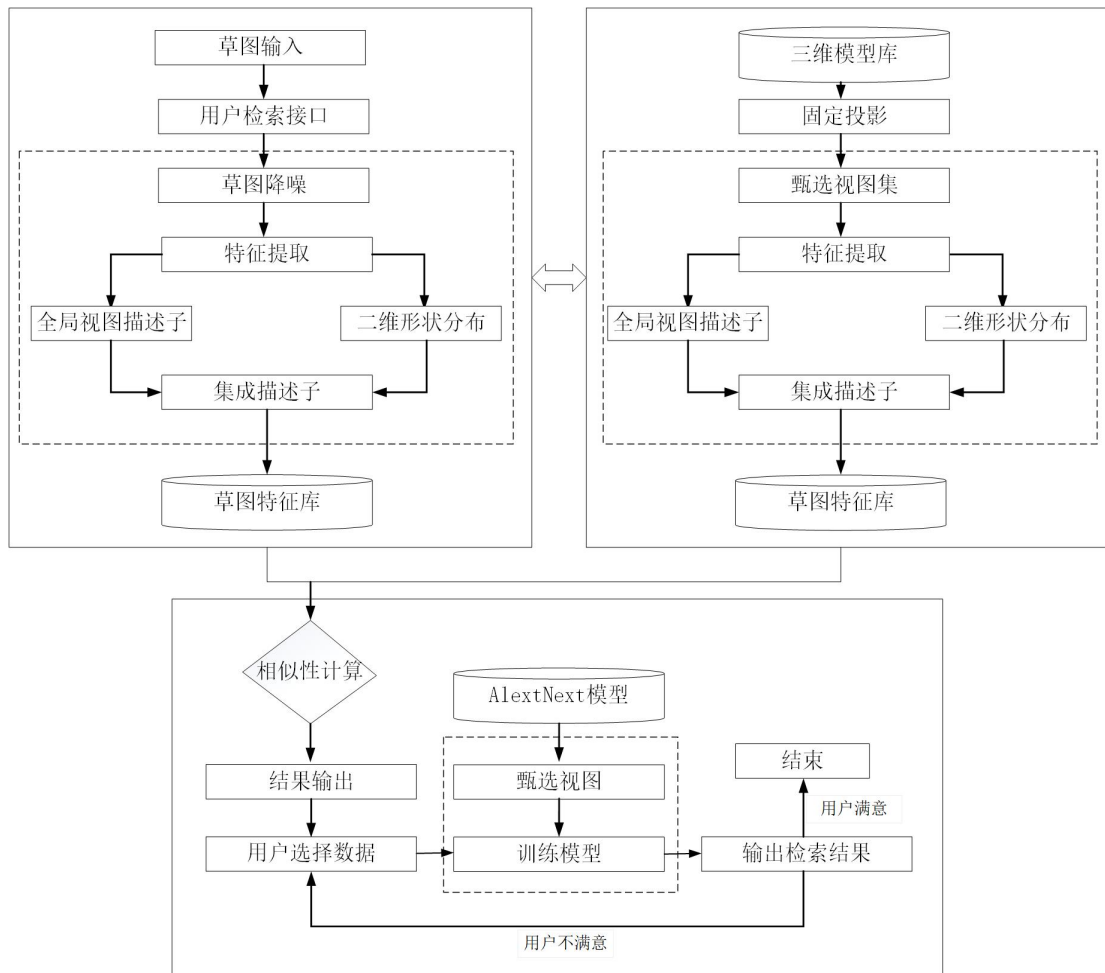


图 2-1 三维模型检索系统框架

Fig.2-1 Framework of 3D model retrieval system

2.2 手绘草图检索架构

2.2.1 基于草图检索架构

基于手绘草图的三维模型检索过程分为在线检索与线下检索两个阶段，如图 2-2 所示。在线下检索过程中，首先利用固定投影的方法对三维模型进行降维处理，获取三维模型库的二维视图集，然后利用最长轮廓长度甄选出特征最突出的 24 张视图作为最优视图集。最后利用集成描述子提取最优视图集的全局视图特征与二维形状特征，加权得到 53 维集成描述特征，并存入数据库。在线上用户检索阶段，首先用户进行人工手绘草图，然后对手绘草图进行降噪等预处理操作，接着利用集成描述子提取集成特征。再与数据库内保存的三维模型的集成特征进行相似度比较。按相似度从大到小依次排序，相似度最高的前 8 个模型显示在检索界面，为后续第二次检索作输入准备。

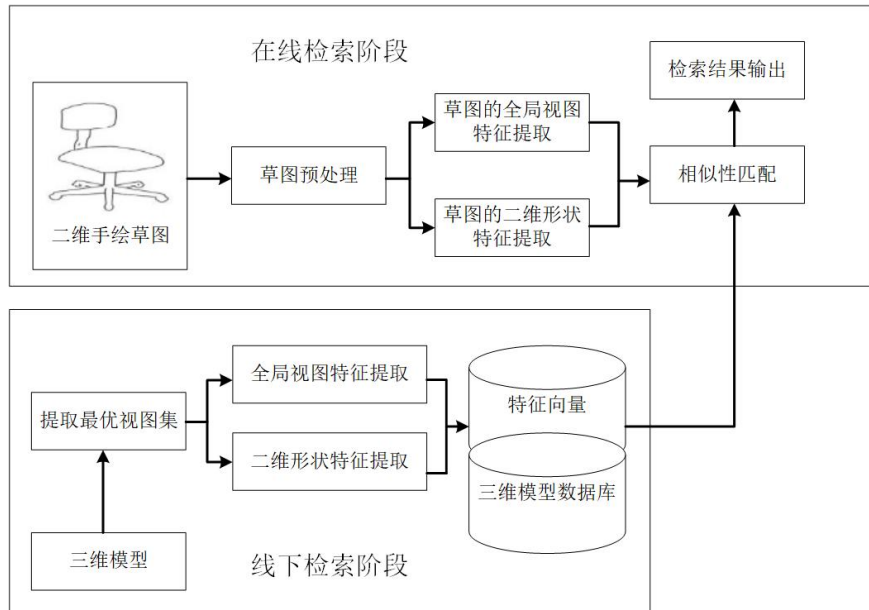


图 2-2 手绘草图检索架构

Fig.2-2 Architecture of freehand sketch retrieval

2.2.2 基于内容的检索架构

基于内容的检索是检索系统的再次检索过程，草图检索是比较粗滤且快速的检索，本文利用第二次精确检索提高检索精度。如图 2-3 所示，在草图检索

阶段得到实例模型，首先利用固定投影在三维模型的 62 张视图中采用 Apriori 算法甄选出特征最明显的视图集。然后在训练阶段，利用生成的二维视图对改进的 AlexNet 模型进行有监督训练。接着在测试阶段，将测试集中的二维视图作为训练 AlexNet 模型的输入，其输出就是改进后的模型分类标签，最后利用距离计算公式计算相似性，如果输出结果用户不满意，可以重复二次检索。

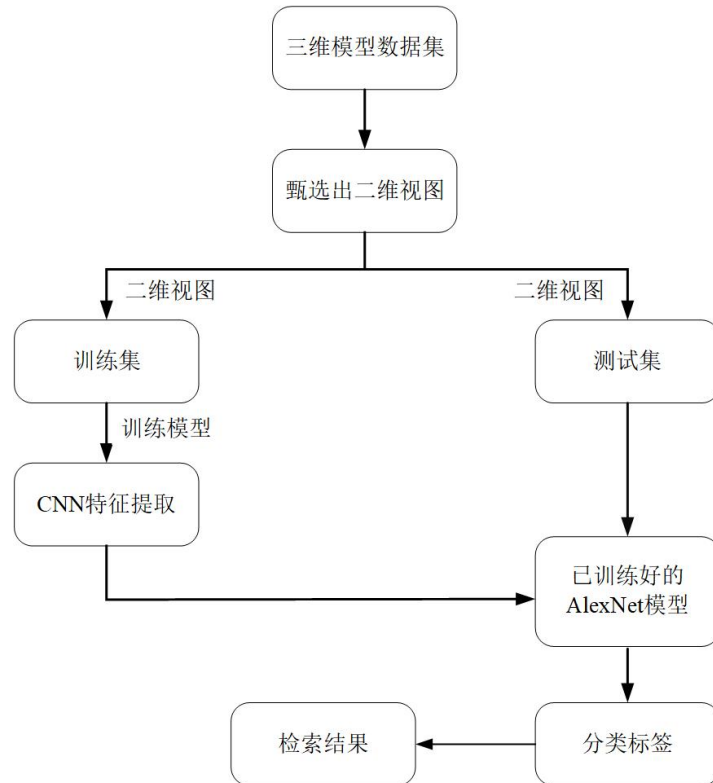


图 2-3 基于内容的检索架构

Fig.2-3 Architecture of content-based retrieval

2.3 投影技术

针对基于手绘草图的三维模型检索算法，国内外学者在这方面作了大量的研究工作，主要包括两方面，一是选择三维模型的最佳视角，二是设计合适的特征描述符。

由于三维模型比二维草图多一维，所以不能直接进行二者的相似度匹配，并且不同角度下三维模型的投影视图可能完全不同，所以需要尽可能多的选择投影视图。为了进行二维草图与三维模型的正常检索，首先应当对三维模型进

行降维处理，生成二维视图。在绘制草图的过程中，用户会下意识地在大脑里虚构出一个三维模型，并且选择一个自身意识里最佳的视角，将二维视图渲染出来，绘制出草图。因此，手绘草图收到用户主观意识的决定，而每个人选择的角度可能是不同的，这也是为什么不能只选择一个或者少数个角度进行投影。因此，一个好的视角选择应当覆盖绝大多数用户对三维模型的绘画角度，并且需要考虑检索效率，即视图不能过多。视图如果过多会造成特征的冗余，增加三维模型检索系统的计算负载。考虑到上述几点，一个优秀的三维模型检索算法对三维模型的投影角度选择是非常重要的。

2.3.1 固定视角投影

在将三维模型投影的过程中，需要确定渲染的方法，一般分为线条图、侧影图、轮廓图等方法，如图 2-4 所示为三种渲染方式的不同外观视图。根据不同的需求选择适当的投影算法也是有必要的。



图 2-4 三维模型投影效果

Fig.2-4 3D model projection effect

在投影的选择上大多数选择的都是固定投影，本文基于草图的三维模型检索和基于内容的三维模型检索都是采用固定投影的方式，固定投影的方式就是选择固定的角度和固定的投影数量进行投影。该方法具有采样全面、容易等优点。如图 2-5 所示：

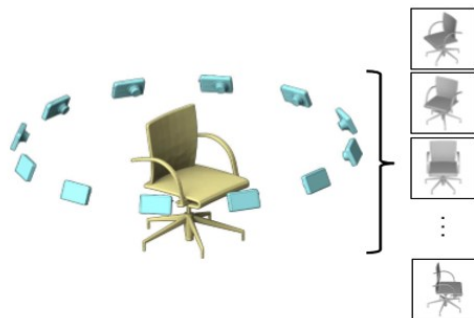


图 2-5 三维模型投影方法

Fig.2-5 3D model projection method

对于每一个三维模型，Su 在水平面以上 30 度角位置，每隔 30 度放置一个

虚拟摄像机，指向三维模型的中心，该投影方法可以生成 12 张视图。对大多数模型具有一定的适应性，因为大多数用户都是从模型的正面去渲染视图^[42]。

Chen 提出一种基于光场描述符(Light Field Descriptor, LFD)的检索算法。此算法利用正 12 面体包围模型，在正 12 面体的 10 个非对称顶点设置 10 个不同的光场。投影获取 100 幅视图，提取视图集的 Zernike 矩作为三维模型检索的特征，如图 2-6 所示。

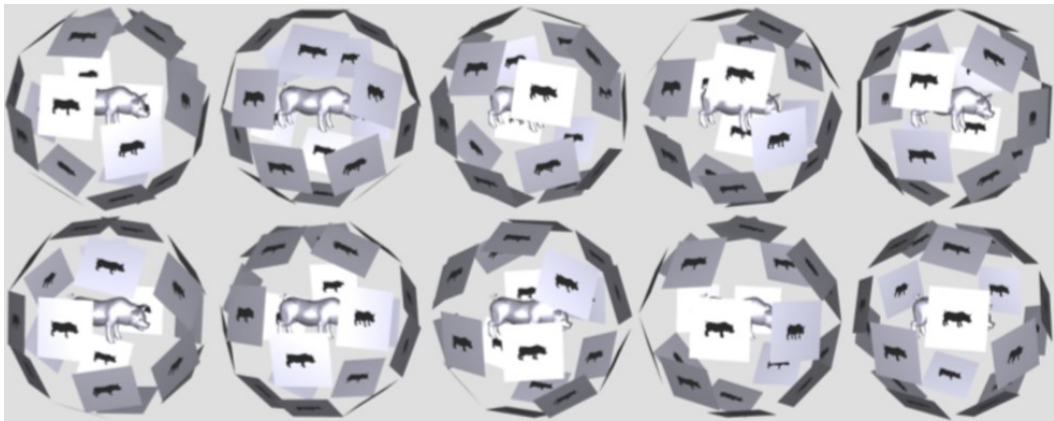


图 2-6 LFD 投影方法展示

Fig.2-6 Display of LFD projection method

Katayama 提出了由多个组件组成的三维模型匹配方法。从不同的角度提取了三维模型的特征向量，并加以预测和匹配模型，有效地评估了两个模型之间的差异^[43]。石民利用正八十面体的 42 个顶点作为投影视角，将一个三维模型表示为二维投影视图集 $I = \{i_1, i_2, \dots, i_{42}\}$ ；然后采用 Apriori 算法甄选出其中的典型视图；最后通过比较三维模型的典型视图来计算模型之间的相似性，如图 2-7 所示。

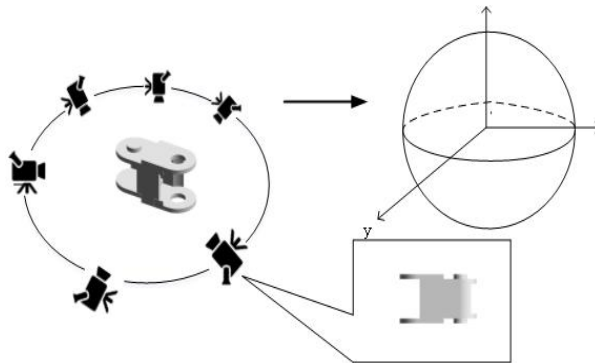


图 2-7 Apriori 算法投影展示

Fig.2-7 Projection presentation of Apriori algorithm

在上述三种方法中，Chen 的方法获取视图特征较全面，然而太过冗余，降低计算效率；Su 的方法考虑到模型的方向性，石民的方法在大量视图的基础上选取特征突出的视角，该方法比较可取。

2.3.2 最佳视角投影

刘玉杰提出一种基于融合信息熵的投影方法，该方法通过熵来度量投影视图的信息含量。熵是对对象或者系统混乱程度的度量，表示平均信息含量。二维视图可以通过熵值直接表现出视图的信息含量。作者通过多投影的方式获取完整的视图集，将视图集中每一个视图的投影面积与所有视图投影面积和的比值作为视图的熵，熵越大视图信息量越多，视图被分辨出来的机会越大^[44]。如图 2-8 所示为床的三维模型在不同角度下计算二维视图的熵值。可以验证 σ 越大，表示视图的投影面积相对越大，含有信息量越多，越容易辨认。反之 σ 越小，表示视图的投影面积相对越小，含有信息量越少，越不容易辨认。

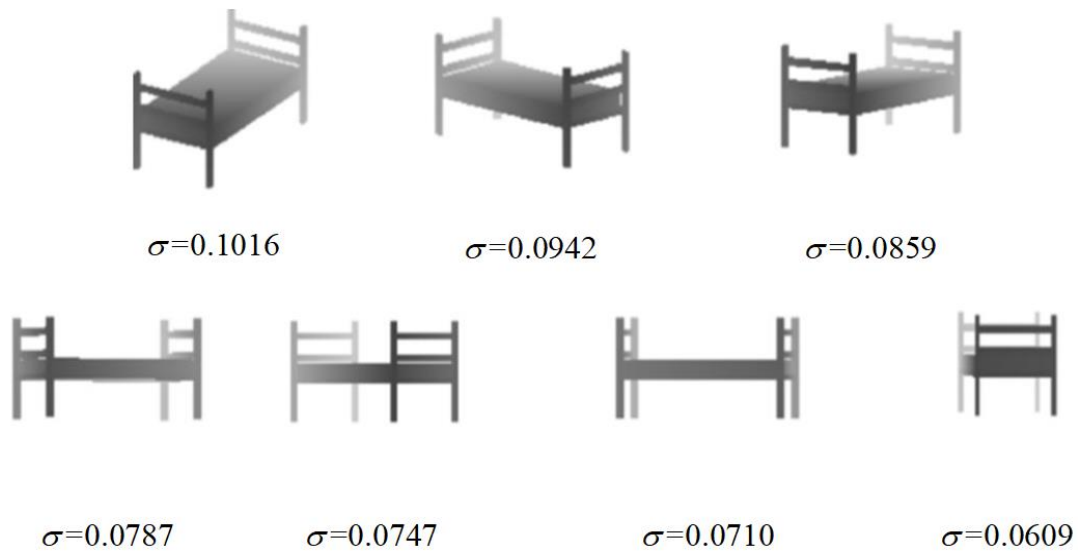


图 2-8 床的信息熵值

Fig.2-8 Information entropy of bed

最佳视角投影的方法在于选择最优视图，针对某些草图检索三维模型具有不错的效果，但是也产生一些问题。用户在绘制草图时，通常会在脑海中想象出三维模型的样子，然后按照个人偏好选择某一角度进行绘制草图，利用草图检索三维模型，绘画时主观性较强，并且不同的模型绘画时的角度选择也是不同的，因此最佳视角选择算法得到的视图与用户的绘制草图出入可能太大。另

外, 最佳视图的选择是一个难点, 不同的模型可能不合同一个选择算法。

本文基于草图的三维模型检索方法也采用固定投影的方法, 先获得三维模型的多视角投影视图集, 该视图集共 62 张视图。然后通过轮廓长度甄选出特征最显著的 24 张视图作为模型检索的最优视图集。算法在 3.3 节中详细介绍。而本文基于内容的三维模型检索方法采用的也是固定投影的方法, 在 62 张视图集的基础上采用 Apriori 算法甄选出其中的典型视图集, 作为卷积神经网络的训练输入, 该算法在 3.3.1-3.3.3 节中详细介绍。基于多视角点云的检索方法在 3.3.4 节中介绍。

2.4 CNN 基本原理

机器学习算法来源于 20 世纪 60 年代, 其中卷积神经网络(CNN)是 Hubel 在研究猫的视觉皮层细胞时提出的。CNN 本质上是一种深度前馈人工神经网络。在之后的 20 年, K.Fukushima 提出并实现了一个新的卷积神经网络模型, 随后很多学者开始投入研究并在原模型上进行改进。随着卷积神经网络的模型层数加深, 卷积神经网络渐渐归类到深度学习。在近些年, 深度卷积神经网络在图像识别、声音处理等方面有独特的优越性, 吸引着越来越多的学者投身其中。卷积神经网络引入了局部连接, 权值共享和下采样等方法, 大大降低了需要训练的参数量, 降低了网络的复杂度, 提高了训练效率。

卷积神经网络的网络结构主要包括稀疏连接和权值共享两部分, 稀疏连接在 BP 神经网络的基础上进行改进, 大大的降低了网络模型的参数规模等级, 进而降低了计算的负载量, 对于实际开发应用具有极高的价值。权值共享使特征提取时不用考虑局部特征, 并且共享参数。经过训练得到的卷积神经网络模型的参数量将极大地减少, 降低了计算复杂度。

卷积神经网络的模型结构包括特征提取层和特征映射层两个层级结构^[45]。每一层的特征提取层是利用上一层与输入的神经突触全连接来获取局部特征。当上一层被成功提取后, 与它相连接的神经突触及局部接受域的位置关系也随之确定。在特征映射层, 一般使用非线性函数 Relu 函数作为激活函数, Relu 函数具有比线性函数训练速度快等优点。卷积神经网络的结构选择, 使映射具有位移不变性。CNN 在识别图像时, 它的优点就是能够对图像像素进行直接卷积实现特征选择与提取, 这是一种与人脑处理系统近似的处理模式。也就是说直接以原始数据作为输入数据, 在大量的训练中自动学习。

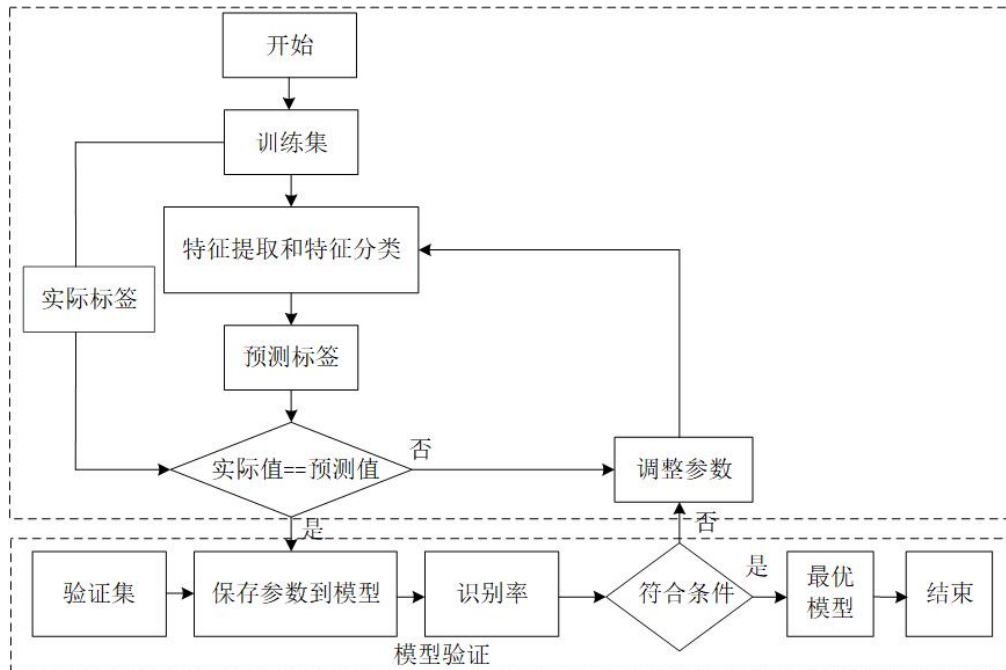


图 2-10 CNN 算法流程图

Fig.2-10 Flow chart of CNN algorithm

图 2-10 为 CNN 的训练过程, CNN 训练包括前向传播阶段和反向传播阶段。在前向传播阶段, 对训练数据进行特征提取和特征分类, 对预测标签下的数据采用梯度下降的方法调整参数。同时计算网络损失, 将调整好的训练参数保存在模型文档中。在反向传播阶段, 对模型文档中训练好的 CNN 网络模型进行验证, 将测试数据输入到调用参数的模型中。通过模型分类与人工标签的匹配量来判断 CNN 模型的分类准确度, 用户不满意可以从头进行模型训练, 如果用户满足则输出最优模型。

2.4.1 卷积层

卷积层是神经网络和卷积神经网络的核心网络层级, 输入层的数据特征经过卷积层的卷积过程, 从低维特征变为高维特征。卷积时需要利用一个固定大小的卷积核, 然后从不同的位置对上一层的结果特征图进行特征提取, 每个卷积核相当于一个固定大小的滤波器, 每次与卷积层数据等大小的矩阵进行内积的操作。卷积核按照步长规定大小进行移动, 在卷积过程中还有个很重要的因素, padding 大小, padding 大小用于增强边缘特征。可知卷积过程与卷积核、步长、padding 项相关。

在卷积神经网络的识别过程中，卷积层的每一个卷积核都对上一层的特征结果进行再次特征提取，同一类别模型提取出的高维特征差异较小，有利于卷积神经网络的后续类别分类。在本文的训练过程中由于颜色不是类别分类的影响因素，所以先进行图像的灰度化，避免卷积过程中颜色对特征的影响。选择卷积神经网络模型时，一般可以根据数据库模型的数量、卷积层提取的特征与目标特征差异的大小进行选择。本文采用经典 AlexNet 模型，5.2 节详细介绍 AlexNet 模型优化过程，得到更有意义的特征表达。

2.4.2 下采样层

卷积神经网络结构中下采样层也叫池化层，即 Pooling 项，卷积层和下采样层的操作大致相同。不同的是卷积层可以进行反卷积，反卷积可以使特征越变越大。而下采样层经过卷积核池化操作能提取相应位置的特征向量，特征维度越来越小。卷积神经网络模型对输入的图像进行特征提取时，使用下采样的操作主要有以下三点优势：(1)可以去除部分噪声，降低提取图像特征与目标图像特征的差异；(2)增强图像特征的稳定性，可以对发生形状变化、位置变化的图像保证其不变性；(3)下采样层能够通过降低维数来减少计算量，提升特征提取的效率，防止出现过拟合。因此，下采样层的操作对后续获得的高维特征向量起着决定性的作用。

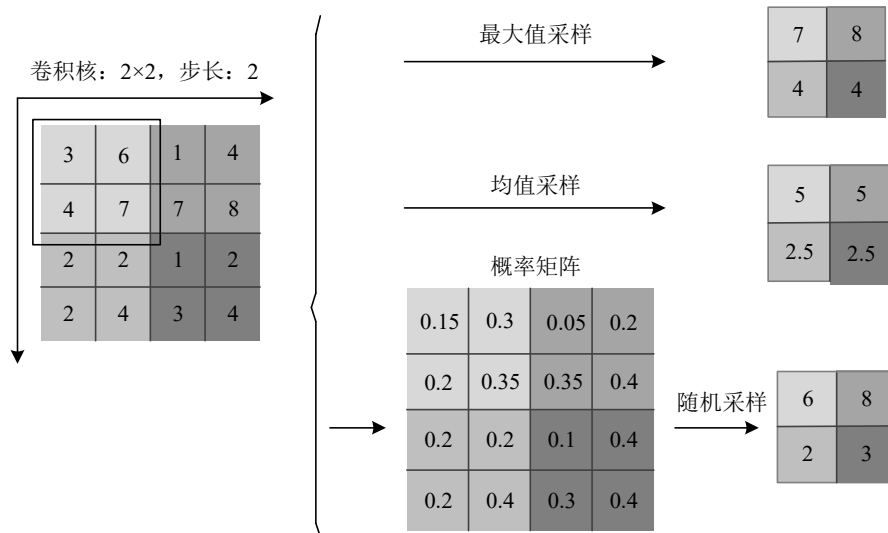


图 2-11 三种采样方式

Fig.2-11 Three sampling methods

常用的下采样方法包括最大值采样、均值采样、随机采样。其中，最大值采样是选择图像特征中卷积核大小区域内的最大灰度值，然后作为采样结果进行记录保存；均值采样是选择图像特征中卷积核大小区域内的灰度值平均值，然后将平均值作为采样后的结果进行记录保存；随机采样是指在卷积核大小区域内随机分配一个概率矩阵，并且概率矩阵的和为 1，然后将采样得到的概率值和对应的灰度值进行内积求和，最后得到随机采样值记录保存，图 2-11 为三种采样方式示意图。

三种下采样的方法在保留图像特征损失很小的前提下，有效地降低了特征的维数，减少了后续过程中的计算量，提高特征提取的效率。然而相对于其他方式而言，最大值采样方法可以更大限度地保留原有特征的梯度信息，与目标特征的差距更小，所以在当下研究中使用广泛，因此本文的采样方法采用最大值采样。

2.4.3 全连接层

全连接层在卷积神经网络中担任分类器的功能，如果说卷积层提取的是局部特征，全连接就是把上一层的局部特征通过权值矩阵进行组装，由于该层利用了所有的局部特征，所以叫作全连接层。

全连接层可调节 CNN 模型中大部分的参数，使经过全连接层后提取的特征向量降低到相对合适的维数，不会产生过大的计算量。

在不同的分类识别任务中，损失函数的选择也是不同的，单标签分类任务一般选择 Softmax 函数作为损失函数，如公式(2-1)所示， p_{nk} 表示概率， K 表示特征向量：

$$E = \frac{1}{N} \sum_{n=1}^N \log(P_{nk}), K \in [0, 1, \dots, K-1] \quad (2-1)$$

而在多标签分类任务中，通常选择 Sigmoid 函数作为目标的概率函数，如公式(2-2)所示， p_n 表示多标签任务下的概率， \hat{p}_n 表示多标签任务下概率的预测值：

$$E = \frac{-1}{N} \sum_{n=1}^N [p_n \log \hat{p}_n + (1 - p_n) \log(1 - \hat{p}_n)] \quad (2-2)$$

而在实数值回归的分类任务中，通常选择欧式距离作为损失概率函数，如公式(2-3)所示， y_n 表示实际距离， \hat{y}_n 表示预测距离：

$$E = \frac{1}{N} \sum_{n=1}^N \| \hat{y}_n - y_n \|_2^2 \quad (2-3)$$

本文主要研究基于视图的三维模型分类任务，也就是分类对象是二维视图，该任务属于单标签任务。因此，本文选择 Softmax 函数作为损失函数。利用 Softmax 损失函数将得到的特征向量进行归一化，得到的值可归一化为分类概率。概率值最大的类别作为分类训练器输出的最终结果。如公式(2-4)所示， K 表示一组特征向量， $e^{x_{nk}}$ 表示当前概率的激活值， p_{nk} 表示当前概率。

$$p_{nk} = \frac{e^{x_{nk}}}{\sum_{l=0}^{K-1} e^{x_{nl}}}, K \in [0, 1, \dots, K-1] \quad (2-4)$$

2.5 本章小结

本章主要介绍了三维模型检索算法的两部分内容，一部分是完整的检索流程，包括基于草图的三维模型检索流程以及基于内容的三维模型检索流程，分析了各部分模块的功能与作用，给出了流程图。另一部分是介绍了投影的主流方式，包括固定投影和最佳角度选择投影等技术，通过投影达到三维模型降维的目的，降低三维模型与二维草图的语义鸿沟，为后续的草图检索三维模型作了预处理的准备。其次介绍了 CNN 的网络结构与工作原理，在后续的基于内容检索三维模型算法中详细介绍具体网络模型。

第 3 章 手绘草图与三维模型预处理

3.1 三维模型数据库

在三维模型分类检索任务中，最常用的数据集是 ModelNet，它提供了两个版本：一个 10 类的数据集和一个 40 类的数据集，是世界上分类检索任务中最常见的三维模型的集合。ModelNet 数据集使众多学者和研究人员能够客观地评估自己与其他人的算法。正因为如此，高质量的数据集聚集了大量先进的研究，竞争并努力实现最佳识别效果。在本文中利用 ModelNet-40 作为训练集与测试集，该数据集如表 3-1 所示：

表 3-1 ModelNet-40 数据集

Table 3-1 ModelNet-40 sets of data

种类	类别	训练集	测试集
1	airplane	626	100
2	bathtub	106	50
3	bed	515	100
4	bench	173	20
...
...
37	tv_stand	267	100
38	vase	475	100
39	wardrobe	87	20
40	xbox	103	20

3.2 手绘草图降噪

手绘草图降噪在基于草图的三维模型检索系统中是不可缺少的一个步骤，由于用户绘画水平的限制，手绘草图能否正确地表达用户的绘画意图也是三维模型正确检索的条件之一。手绘草图降噪的目的是将用户绘画时不小心或者无意识造成的聚点，封闭处的缝隙，提笔或者收笔造成的毛刺消除，防止手绘草图过于粗糙，导致手绘草图离用户原本意图差异较大。本文中采用的草图降噪

算法包括：聚点消除、闭合检查和删除冗余笔画。

3.2.1 聚点消除

聚点出现的原因是由于抖动等因素造成的小区域笔画聚集，本文采用 Jin 提出的滑动窗口来判断是否有聚点出现，判断在窗口内是否有重复出现的采样点。若出现则存在聚点，并予以消除，如图 3-1 所示为消除聚点的过程。

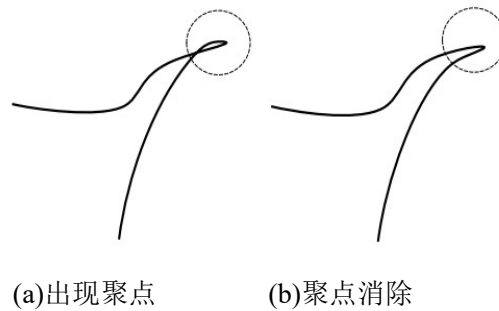


Fig. (a) Gathering point Fig. (b) Gathering point elimination

图 3-1 聚点消除视图

Fig.3-1 The view of eliminating cluster points

3.2.2 闭合检查

在用户绘制一个封闭曲线时，可能会造成起点与终点交叉或者分离的情况，如图 3-2 所示。在闭合曲线上，采样点的偏移角度累加和一定是 2π 。若和大于 2π ，则出现(a)的交叉情况，记录交叉点 p ，删除 p 到两端点的冗余线段。若和小于 2π ，则出现(b)的分离情况，计算两端点的中间结点 p ，将 p 作为两端点的闭合中点。

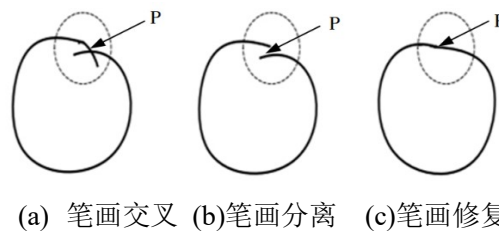


Fig.(a) Stroke cross Fig.(b) Stroke separation Fig.(c) Stroke repair

图 3-2 闭合检查视图

Fig.3-2 Closed check view

3.2.3 冗余笔画删除

本文采用判断曲率值大小的方法消除冗余笔画。在检索过程中，笔画 L 采样后可表示为 $L = \{p_i | 0 \leq i < n\}$ ，假定曲率值 $\varphi=0.8$ 。算法分如下两步：

(1)记录绘制笔画的首部位置和尾部位置。计算 L 的最小外包矩阵的对角线长度 $length$ 。(2)计算笔画中曲率最大的点，如果该点的曲率大于设定的曲率 ($\varphi=0.8$)，并且该点到达首部或者尾部的长度小于 $length*0.1$ ，则从这个点开始删除直到首部或者尾部，算法效果如图 3-1 所示：



图(a)冗余笔画

图(b)处理后笔画

Fig. (a) Stroke redundancy Fig. (b) Stroke after processed

图 3-3 消除冗余笔画视图

Fig.3-3 Eliminate redundant stroke views

3.3 甄选视图集

数据库中的每个三维模型在空间中都具有任意位置、方向和比例，因此，有必要在投影前对其归一化。本文采用主元分析法(PCA)通过平移、旋转和缩放将所有三维模型从模型数据库转换至同一三维坐标系。对于给定的三维模型，首先通过遍历三维模型外围三维坐标来确定该模型的中心坐标。然后，将中心转移到三维坐标系的原点。接下来，通过标准 PCA 方法和对称技术将三维模型旋转到最佳方向。最后，将所有三维模型缩放为相同的单位尺寸。

3.3.1 基于草图检索视图集

在对 ModelNet-40 数据库中的三维模型规范化之后，需要将它们投影为与草图进行匹配的二维视图。如图 3-4 所示，本文首先将每个三维模型设置在一个虚拟球体的球心，并将虚拟摄像机放在该模型上方。然后，将模型以每步 30 度(在经度和纬度方向上)旋转 360 度，以生成 62 张二维视图。

投影选择方面的研究表明：人类选择绘制二维视图的角度可能与几个简单

的因素相关，比如模型可见部分的轮廓长度和投影区域的相对面积等。本文在这 62 张视图中，使用轮廓长度评价二维视图包含的信息量。选取其中轮廓长度最长的 24 张视图作为基于草图检索三维模型的最优视图集。

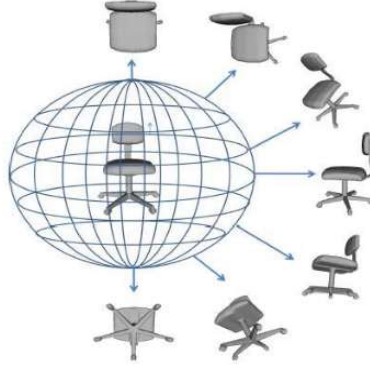


图 3-4 二维视图的选择

Fig.3-4 Selection of 2d view

3.3.2 Apriori 算法

Apriori 算法在数据挖掘领域和计算机信息领域具有广泛的应用，主要包括如下 3 个基本概念。

1.事务：事务是数据集的子集，不同事务的全体构成事务数据库。

2.关联规则：形如 x 出现，则 y 必出现的表达式，即 $x \rightarrow y$ ，其中 x 和 y 是不相交的项集，即 $X \cap Y \neq \emptyset$ 。

3.支持度和关联规则的置信度：设 x, y 是项集的子集， X 是事务集中包含 x 的事务的数量， Y 是事务集中包含 y 的事务的数量， N 是事务集中所有事物的数量，则 $x \rightarrow y$ 的支持度 s 如公式(3-1)， $x \rightarrow y$ 关联规则的置信度 c 如公式(3-2)。

$$s(x \rightarrow y) = \frac{X \cup Y}{N} \quad (3-1)$$

$$c(x \rightarrow y) = \frac{X \cup Y}{X} \quad (3-2)$$

4.频繁集：支持度大于最小支持度的项集即为频繁集。

Apriori 算法的主要思想是通过对事务数据库进行多次扫描，首先找出所有支持度大于指定支持度的项集即为频繁项集；然后通过频繁项集产生新的关联规则重复扫描，此过程不断重复直到不产生新的频繁项集。

3.3.3 基于二维视图的检索视图集

本文中基于二维视图的检索视图集，需要利用 Apriori 算法。Apriori 算法的具体步骤如下所示：

1.构建事务数据库，二维视图间的相似性度量如公式(3-3)所示。

$$Sim(i, i') = 1 - dis(D_i, D_{i'}) \quad (3-3)$$

$dis()$ 是两视图 i 与 j 对应特征描述子 D_i 和 D_j 的距离，将其在区间[0,1]作归一化操作。如果度量大于规定的最小相似度，则 i 与 j 相似。

2.对于数据库中的每一幅视图，统计所有特征相似的视图并作为一条事务存入事务数据库。如表 3-2 所示。

表 3-2 事务数据库

Table 3-2 Transaction database

事务	项集
1	i_1, i_4, i_9, i_{11}
2	i_1, i_5, i_9
..	...
62	i_{62}

3.对事务数据库进行第一轮扫描，筛选出支持度大于最小支持度 d_1 的投影视图作为频繁集 L_1 ，并将频繁集 L_1 作为输入候选项集，在扫描了 L_1 筛选出支持度大于最小支持度 d_2 的投影视图作为频繁集和候选项集 L_2 。同理，不断筛选候选项集，直到不会再生成频繁集为止。本文需要的典型视图即为所有频繁项集中关联规则置信度最小的视图。甄选典型视图的算法流程如下：

算法输入：事务数据库 $data$ ，最小支持度阈值 $min_support_value$

算法输出：三维模型的典型视图集 $collection$

1. $L_1 = 1$ 项频繁项目集
2. for ($i=2$; L_{i-1} ; $i++$) do
3. $Can_k \leftarrow Apriori_gen(L_{i-1}, min_support_value)$: 构建候选集 Can_k
4. for 数据库 $data$ 中的事务 t do
5. $Cant \leftarrow Subset(Can_k, t)$: 提取包含在事务 t 中的候选集 $Cant$
6. for $Cant$ 中的候选项 Can' do
7. $Can'.Count++$: 直到不再产生新的频繁集 Can' 为止
8. end for

9. $L_i = \{Can' \in Can_k | Can'.Count \geq min_support_value\}$
10. end for
11. $L \cup = L_i$
12. end for
13. collection SeekMinConf(L): 计算置信度, 提取典型视图集

甄选完典型视图之后, 利用二维典型视图集 collection 来表示对应的三维模型。三维模型之间的相似度计算就可以转换为各自模型的典型二维视图集之间的计算。然而不同种类的三维模型典型视图集的规模可能不同, 本文采用贪心算法解决典型视图集规模不一致的问题。设三维模型 P 和 Q 对应的视图集为 collectionP 和 collectionQ, 算法流程如下所示:

第一步: 每次提取三维模型 P 的一幅典型视图 $c_p^i \in collection_p$, 遍历 collectionQ 视图集中视图, 计算出最相近的视图 c_Q^j , 则两幅视图的距离函数如公式(3-4)所示:

$$D(c_p^i, c_Q^j) = \min(dis(D_{c_p^i}, D_{c_Q^j})) \quad (3-4)$$

其中 n 为三维模型 Q 的视图集规模, $D_{c_p^i}$ 和 $D_{c_Q^j}$ 分别是 c_p^i 和 c_Q^j 对应的特征,

D 越小视图相似度越高。

第二步: 循环重复第一步, 直到其中一个三维模型的所有视图都匹配成功, 其中匹配成功视图不重复匹配, 三维模型 P 和 Q 的相似度如公式(3-5)所示:

$$Sim(P, Q) = 1 - \sum_{n=1}^N D \quad (3-5)$$

其中, $N = \min(s, r)$, s 和 r 是两个三维模型对应视图集的规模, D 是相似度距离。

3.3.4 基于多视角点云视图的检索视图集

本文提出了一种基于多视角点云视图的三维模型分类方法, 该方法由点云模型经过切割、投影等操作生成二维视图, 再由 AlexNet 模型学习二维视图的具体特征来实现三维模型的分类。首先对 ModelNet-40 数据集中的模型进行点云化; 然后对于数据集中的每个模型, 延三视图平面分别截取一部分; 接着将这三个部分投影到不同的平面上, 这样从每个模型中可以获取三个二维图像; 最后将这三个二维图像作为深度卷积神经网络的输入。

ModelNet-40 数据库提供了以对象文件格式(OFF)作为多边形网格的三维模型。然而, 本文的架构需要点云来生成切片, 所以本文执行了一个转换器转换阶段, 将 OFF mesh 表示转移到点云数据中, 这个过程包括将每个网格放置在一个 3D 截断的二十面体(镶嵌球体)中, 然后将一个虚拟摄像机放置在每个顶点指向球体的中心。捕获每个虚拟摄像机的快照, 保持从每个视图生成部分点云。这些点云合并成一个完整的点云模型。这些点云模型将被切片处理, 然后作为 CNN 的输入进行训练, 并使用 ModelNet-40 提供的相应数据进行测试, 图 3-5 说明了上述转换过程。

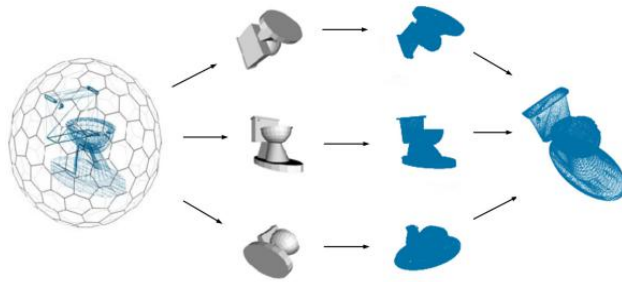


图 3-5 点云模型的生成过程

Fig.3-5 Generation process of point cloud model

本文从三维点云模型中提取切片并投影作为 AlexNet 的输入。为此, 本文加载点云并计算每个轴的中心点。然后基于 XY, XZ 和 YZ 三个平面进行切片, 厚度为模型大小的 5%。位于切片内部的点被投影到坐标轴的平面上, 并生成 500×500 像素的图像。这些图像都是背景为黑色且投影点为白色的二进制图。这个过程如图 3-6 所示。



图(a)模型切片 图(b)提取点云图

(c)模型切片的投影

Fig. (a) Model slice Fig. (b)Extract Point Cloud Fig. (c) Projection of model slices

图 3-6 点云中提取切片的过程

Fig.3-6 The process of extracting slices from a point cloud

由于采样过程产生的点密度不一致, 有些切片中的点非常分散, 因此投影并不能准确地表示切片对象。为了解决这个问题, 我们对每个投影进行后期处

理，其中每个像素进行 10 倍扩张，这个后处理步骤填补了散射点之间的间隙，并产生了更适合的图像表示，如图 3-7 所示。



图 3-7 模型切片与像素扩张

Fig.3-7 Model slicing with pixel expansion

点云视图提取过程针对数据集中存在的每个三维模型，首先将三维模型转换成点云模型，然后提取点云模型三个面上的切片信息，投影后生成三个对应的二维视图。该方法表示允许以二维的方式训练和测试三维模型分类器，它提供了较高的分类准确率。

3.4 本章小结

本章主要介绍了检索系统中用到的数据集 ModelNet-40 及手绘草图在检索前的降噪过程，包括删除冗余笔画、闭合检查和聚点消除等，去除噪声对检索过程造成的干扰。并且提出了基于草图检索算法和基于内容检索算法的视图集选择方法。在草图检索算法中采用固定投影的方法，利用轮廓长度甄选典型视图集。在基于内容的检索算法中包括基于二维视图的检索和基于点云的检索方法，基于二维视图的检索方法利用多角度的投影获取视图集，在视图集中利用 Apriori 算法甄选出典型视图。基于点云的检索方法，利用切片的方式获取模型特征，对特征进行三视图的投影，投影视图作为特征视图集。

第 4 章 特征提取描述符及手绘草图检索

通过固定投影的方法已经可以将三维模型降维成二维视图集，此时直接用草图与二维视图进行像素点的直接匹配是没有意义的，所以需要一种能够表达二维视图和草图特征的描述符。为了能够解决二维视图和草图之间存在旋转和尺度大小的不同，本文需要的特征描述符应当具有平移、尺度和旋转不变等特性。在计算机图形学领域应用过大量的描述符，包括 HOG、SIFT、Zernike 和形状上下文等描述符。本章着重介绍 Zernike 描述符、Fourier 描述符、二维形状分布以及本文提出的集成描述符，本章的集成描述符能够解决单一描述符提取特征不全的问题。

4.1 全局视图特征描述符

本文提出的全局视图特征包含 Zernike 矩和傅里叶描述符。首先，本文提取二维视图的 Zernike 矩，并采用标准矩的方法归一化到(0, 1)范围。然后，取得二维视图的轮廓视图，提取它的一维傅立叶算子，除以直流分量标准化到(0, 1)范围。基于草图的三维模型检索不应该受到草图绘制位置、尺度大小以及旋转角度的影响。

4.1.1 Zernike 矩特征描述符

Zernike 矩定义为单位圆内部的 Zernike 多项式，如公式(4-1)定义一组在单位圆上的复值函数集 $V_{nm}(x, y)$ ，表示在单位圆内的任何平方可积函数。

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (4-1)$$

其中， ρ 表示原点到点(x,y)的矢量长度； θ 是矢量 ρ 与 x 轴逆时针方向的夹角；j 是复数单位；n 和 m 是 Zernike 矩的维数。其中，Zernike 多项式 $R_{nm}(\rho)$ 如公式(4-2)所示。

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \frac{(n-s)}{s! (\frac{n+|m|}{2} - s)! (\frac{n+|m|}{2} + s)!} \rho^{n-2s} \quad (4-2)$$

Zernike 矩就是图像 $f(x, y)$ 在正交基上的投影，连续图像 $f(x, y)$ 的 Zernike 矩定义为公式(4-3)：

$$Z_{nm} = \frac{n+1}{\pi} \iint f(x, y) [V_{nm}(x, y)] dx dy \quad (4-3)$$

对于数字图像 $f(x, y)$ 来说, Zernike 矩如公式(4-4):

$$Z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}(\rho, \theta) \quad (4-4)$$

针对 Zernike 矩进行平移和尺度变换得到具有平移、尺度和旋转不变的图像 $g(x, y)$ 。图像 $g(x, y)$ 的 Zernike 矩定义如公式(4-5)所示:

$$\begin{aligned} Z_{nm} &= \frac{n+1}{\pi} \sum_x \sum_y g(x, y) V_{nm}(\rho, \theta) \\ &= \frac{n+1}{\pi} \sum_x \sum_y f\left(\frac{x}{m_{00}} - \bar{x}, \frac{y}{m_{00}} - \bar{y}\right) V_{nm}(\rho, \theta) \end{aligned} \quad (4-5)$$

其中, $\bar{x} = \frac{m_{01}}{m_{00}}$, $\bar{y} = \frac{m_{10}}{m_{00}}$; m_{01} 表示视图轮廓内所有黑色像素点的横坐标相加的和; m_{10} 表示视图轮廓内所有黑色像素点的纵坐标相加的和; m_{00} 是投影视图中白色像素点的总和; (\bar{x}, \bar{y}) 表示手绘草图的中心。

4.1.2 Fourier 描述符

傅里叶描述是对二维视图的傅里叶变换。在进行傅里叶变换之前, 需要将其转换为一维形状表示。目前, 常用的方法函数包括复坐标函数、曲率函数和重心距离函数。本文采取重心距离函数, 它是通过计算外轮廓边界点到重心的距离来实现的。外轮廓边界点 $\{P\{x_i, y_i\} | i=1 \dots N\}$ 重心定义如公式(4-6)所示:

$$\begin{aligned} x_c &= \frac{1}{N+1} \sum_{i=0}^N x_i \\ y_c &= \frac{1}{N+1} \sum_{i=0}^N y_i \end{aligned} \quad (4-6)$$

根据重心定义重心距离, 如公式(4-7)所示:

$$r(i) = ([x(i) - x_c]^2 + [y(i) - y_c]^2)^{1/2} \quad (4-7)$$

对 $r(i)$ 进行离散傅里叶变换得到傅里叶系数, 如公式(4-8)所示:

$$R(n) = \frac{1}{N} \sum_{i=1}^N r(i) \exp\left(\frac{-j2\pi ni}{N}\right), n = 1, \dots, N \quad (4-8)$$

其中, j 是复数单位, 傅里叶系数 $R(n)$ 除以第一个系数 $R(0)$ 就可以进行归一化得到傅里叶描述符, 如公式(4-9)所示:

$$fourier = \left[\frac{|R(1)|}{|R(0)|}, \frac{|R(2)|}{|R(0)|}, \dots, \frac{|R(n)|}{|R(0)|} \right] \quad (4-9)$$

这个描述符具有平移、尺度和旋转不变特性。其中，平移不变性和旋转不变性是它本身的特性，归一化和绝对值操作保证了尺度不变性。

4.2 二维形状分布

二维形状分布是在三维形状分布的基础上变化得来的。三维形状分布的主要思想是利用形状函数来表示三维模型表面上点对的距离关系、模型内部角度关系和区域面积，构建表达形状特征的统计分布直方图。同理，二维形状分布能够构建出二维视图的统计分布直方图。二维形状主要包括 A3、D1、D2、D3 和 D4 等五种形状函数，如图 4-1 所示。

A3：在模型表面随机选取三个点，三点构成的三角形内角的概率分布构成 A3 形状特征。

D1：连接模型表面任意一点，与模型中心位置产生的距离概率分布构成 D1 形状特征。

D2：在模型表面选取任意两点相连，两点之间距离的概率分布构成 D2 形状特征。

D3：在模型表面选取任意三点，三点构成的三角形面积的平方根的概率分布构成 D3 形状特征。

D4：在模型表面选取任意四点，四点构成的立方体体积的概率分布构成 D4 形状特征。

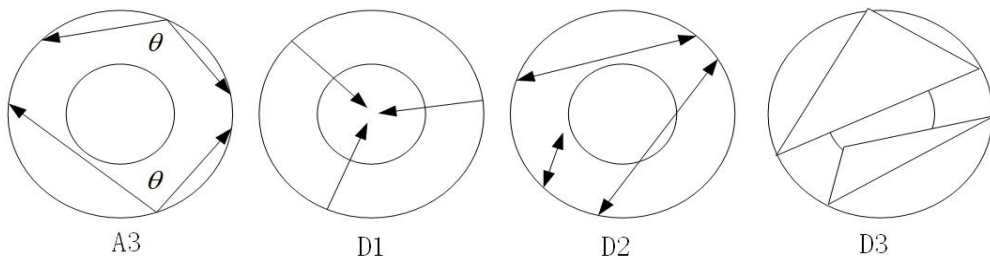


图 4-1 形状描述符

Fig.4-1 Shape descriptor

本文采用效果较好的 D2 形状描述符来描述二维视图的形状区域特征。如图 4-2 所示，本文在三维模型的每个二维视图和输入草图上采用随机采样。采用的样本点 $\{p_1, p_2, \dots, p_n\}$ 不是全部在外部的边界上，每个采样点 p_i 可以位于给定视图

的任意边缘线上。二维形状分布特征提取的算法描述如下：

1. 在绘制草图过程中，记录草图绘制的所有像素点，放入集合 $S = \{p(x_i, y_i) | i = 0, 1, \dots, N-1\}$ ， N 是所有的像素点总数。对像素点进行随机采样，本文选取采样个数为 1024。

2. 针对 1024 个随机采样点，可知二维形状描述符产生的所有点对如公式 (4-10) 所示：

$$Point\ Num = \frac{1024!}{2!(1024-2)!} = 523776 \quad (4-10)$$

计算所有点对之间的欧式距离，如公式 (4-11) 所示：

$$d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4-11)$$

(x_i, y_i) 、 (x_j, y_j) 为任意两个随机点，构成一组点对， d 为两点之间的距离。

3. 统计随机点对之间的距离，构建二维形状分布直方图。统计分布直方图的横坐标范围应该从 0 到点对之间的最远距离。在计算两个直方图之间的相似度时，应该保证直方图区间数目是相同的。对于区间较少的直方图，应该采用归零的方法去填充区间。

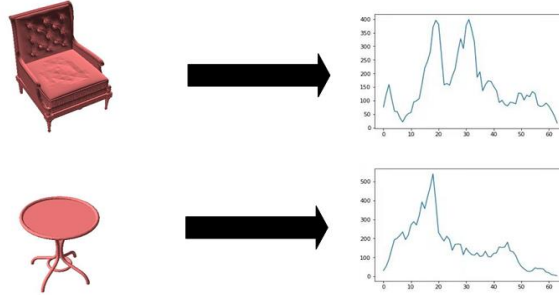


图 4-2 D2 描述符案例

Fig.4-2 D2 descriptor case

4.3 集成描述符

经过大量的实验表明傅里叶描述符中较小的值适用于表示全局视图特征。Zernike 描述符中较大的值适用于表示全局视图特征。本文的傅里叶描述符与 Zernike 描述符均选取经验值^[46]。其中，傅里叶描述符选取前 10 个较小的值；Zernike 描述符使用前 35 阶较大矩。

对于一幅草图和三维模型的二维视图，本文使用向量的 L1 范式去度量傅里叶描述符、Zernike 描述符和二维形状分布之间的距离，如公式 (4-12) 所示。

$$\begin{aligned}
 D_F &= \sum_{p=1}^{10} |F_1(p) - F_2(p)| \\
 D_Z &= \sum_{q=1}^{35} |Z_1(q) - Z_2(q)| \\
 D_D &= |D_1 - D_2|
 \end{aligned} \tag{4-12}$$

其中, $F(p)$ 、 $Z(q)$ 分别表示草图和二维视图的傅里叶变换特征和 Zernike 矩特征。 D_1 、 D_2 分别表示草图和二维视图的二维分布特征。 D_F 、 D_Z 和 D_D 分别为使用傅里叶描述符、Zernike 描述符和二维形状分布的二维草图与投影视图之间的相似性结果。其中, 傅里叶描述符和 Zernike 矩构成本算法的全局视图特征描述符, 如公式(4-13)所示:

$$D_G = D_F + D_Z \tag{4-13}$$

为了解决单一描述符对二维视图特征提取不全面的问题, 本文在全局视图描述符和二维形状分布的基础上进行一定权重的集成, 得到一个能更好地表示二维视图特征的描述符。集成描述符 D 定义为公式(4-14):

$$D = \omega D_G + (1 - \omega) D_D, \omega \in (0, 1) \tag{4-14}$$

ω 为全局视图描述符的权重, 它与二维形状分布的权重相加为 1。草图与二维视图相似度越高, D 越小。

不同用户绘画风格不同, 绘画出的草图可能差异很大, 集成描述子特征也会不一致, 因此本文有如下方法避免检索结果的不准确。首先本文提出的集成描述子具有平移、尺度和旋转不变等特性, 避免了旋转、大小等对检索结果的影响, 因此, 只有绘画风格存在对草图的影响, 而绘画风格存在两方面问题, 1) 绘制者对模型观察角度的不同, 影响集成描述子的特征, 因此本文采用多投影的方式, 尽可能多的考虑不同用户观察角度的差异。2) 本文草图检索对绘制有一定要求, 绘制越精准检索效率越高, 需要检索的次数越少, 反之为了避免绘画能力不足导致的检索不准, 本文系统可以采用多次反复检索, 系统中 8 个模型框, 只要一个能够检索出用户想要的模型(比如说“凳子”), 就可以直接获取凳子模型, 然后进行基于实例的检索(基于实例的检索更加精准)。

4.4 检索系统设计

本文提出的基于草图检索三维模型算法与基于内容检索三维模型算法均在三维模型检索系统中进行验证, 该系统的开发是为了后续研究学习的方便, 缩短算法验证的时间。软硬件环境如下:

操作系统：Window10；开发环境及开发语言：VS2012，C++；数据库：MySQL；几何造型平台：Opengl，Opencv；图形开发库：MFC；处理器：Intel Core i7-4578U CPU 3.00GHz；内存：8GB；硬盘：216GB；

本文在 VS2012 环境下，以 Opengl 和 Opencv 为几何造型平台，以 MFC 为图形界面库开发了一个支持二维草图绘制、支持二维草图与三维模型检索的应用系统。对本文基于草图检索和基于内容检索的方法进行了验证。从 ModelNet-40 模型库中选取 40 个类别的模型用于实验，每个类别选取 100 个模型。三维模型检索系统如图 4-3 所示。



图 4-3 三维模型检索系统

Fig.4-3 3D model retrieval system

4.4.1 基于草图的检索模块

本文检索系统初始化界面就是基于草图的检索界面，草图检索模块如图 4-4 所示，用户步骤包括如下几步：

- 步骤 1.用户在画板中绘制草图。
- 步骤 2.选择检索三维模型库路径。
- 步骤 3.选择草图检索算法。

步骤 4.点击检索，输出初次检索结果。

步骤 5.如果检索结果不满意，重复步骤 1-4。



图 4-4 草图检索系统流程

Fig.4-4 The flow of sketch retrieval system

4.4.2 基于内容的检索模块

基于内容的检索是本系统中的第二次检索，是在手绘草图检索基础上进行检索，基于内容检索模块如图 4-5 所示，检索步骤如下：

步骤 1.用户在检索出的三维模型中选取一个适当的模型，作为第二次基于内容的检索三维模型，这样做的目的是为了普普通用户直接进行基于内容的检索没有合适的三维模型。

步骤 2.将选好的三维模型拖入检索输入界面。

步骤 3.选择检索三维模型数据库。

步骤 4.点击检索，输出基于内容的检索结果。

步骤 5.如果检索结果不满意，重复步骤 1-4。

此外本系统还有一些三维模型操作，比如三维模型手动转动，便于观察，三维模型格式转换等操作。



图 4-5 基于内容检索系统流程

Fig.4-5 The flow of content-based retrieval system

4.5 草图检索实验及结果分析

为验证集成描述子中 ω 对算法性能的影响, 本文针对 ModelNet-40 模型库中的 4000 个模型进行了一系列的实验。

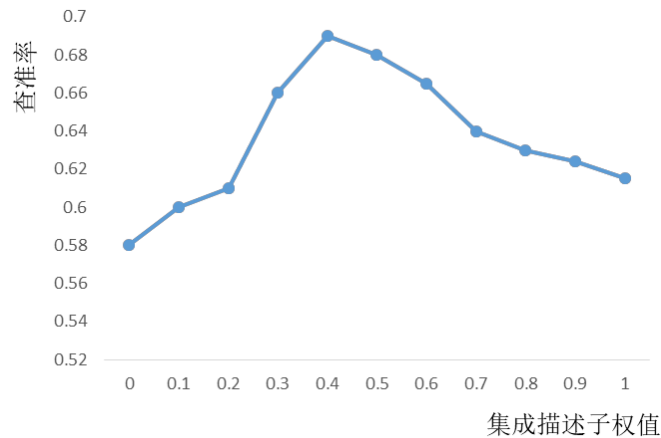


图 4-6 集成描述子权值变化图

Fig.4-6 Weight change diagram of integrated descriptor

由图 4-6 可知, ω 在[0-1]之间以每次 0.1 的速率进行增长。在不同 ω 值时,

记录三维模型在查全率(查全率是:在检索时,检索到的所有模型占据全部模型的比例)为 0.1 时的查准率(查准率是:在检索时,检索到所求模型占据该类模型总数的比例)。 ω 在[0.3-0.6]范围时,集成描述子的检索性能较好; ω 为 0.4 时,检索效果最佳。在后续验证本文提出算法与其它算法的优劣中, ω 取值为 0.4。

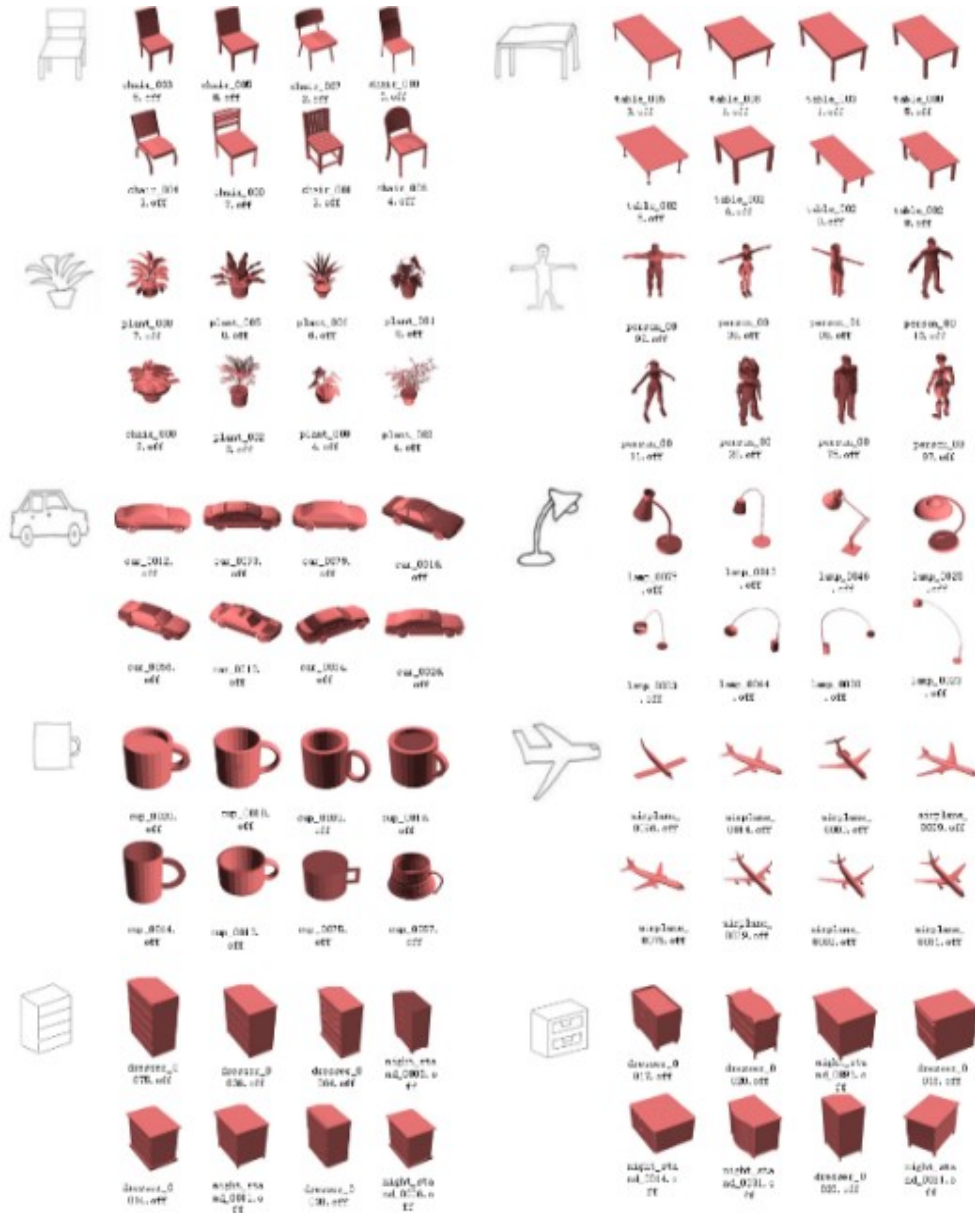


图 4-7 检索结果

Fig.4-7 The retrieval results

图 4-7 为在三维模型检索系统中使用本文基于草图检索算法的一些草图检

索结果示例。图的左侧是手绘草图，右侧为系统中显示的前 8 个检索到的三维模型。这些三维模型是按照与查询草图的相似性大小进行排序的。从图中可以发现：大多数检索结果都可以成功地返回与查询草图相关的三维模型。本文基于草图的算法对使用者绘画有一定要求，但是本文的草图检索相对于精确的 CAD 视图而言，只需要一些相对简单的草图绘制即可，如图 4-7 可以看出草图绘画相对简单，因此通用性相对较强。

即使查询草图包含复杂的内部结构，例如：car 草图，本算法仍然可以获得令人满意的检索结果。然而，仍然存在一些失败的情况，例如：dresser 和 night_stand。在检索结果中，二者出现检索结果混淆的问题。这种混淆的主要原因为：一是两类模型的相似度较高；二是由于草图绘制过于粗糙，影响算法检索性能。

为了进一步验证集成描述子特征提取的性能，本文分别应用全局视图特征和二维形状分布来检索三维模型。此外，将本文的算法与其它三种基于草图的三维模型检索算法进行了比较，包括 LFD、STELA 和 GALIF^[47]。本文采用查全-查准率作为检索评估指标，其结果如图 4-8 所示。

在图 4-8 中，FZ 指的是仅应用本文全局视图特征检索模型的方法；CON 指的是仅应用本文提取的二维形状分布特征来检索模型；FZ+CON 表示本文提取的集成描述子方法。

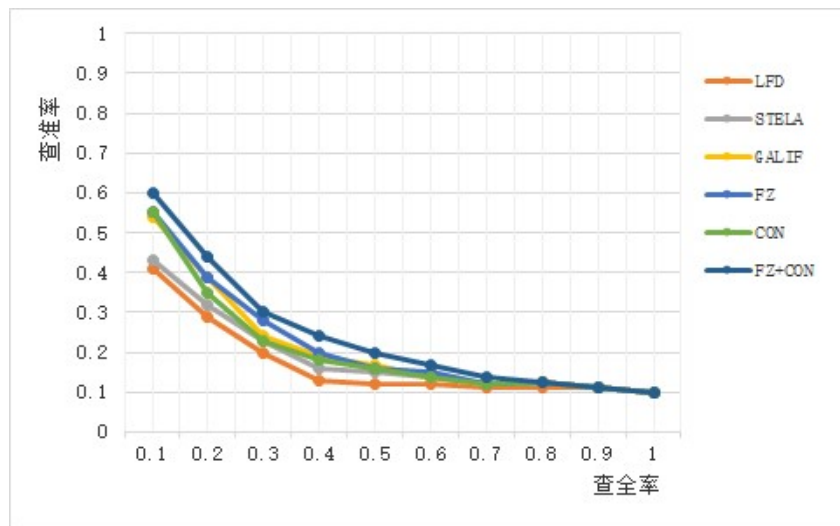


图 4-8 检索方法的性能比较

Fig.4-8 Performance comparison of retrieval methods

从实验结果中可以发现：利用集成描述子提取特征的方法 FZ+CON 要优于其它基于草图的三维模型检索方法。FZ 方法排在第二位，GALIF 排在第三位。

STELA 比第四位的二维形状分布特征 CON 的效果要差,但是优于 LFD。STELA 和 LFD 仅利用输入草图的边界信息,GALIF 方法可以提取查询草图的局部信息,但是这种方法对草图的风格变换高度敏感。其原因是:它提取的特征描述符缺少草图的全局结构信息。相比之下,本文提出的集成描述子既包含全局视图特征又包含边界形状特征,具有更好的辨识度。

4.6 本章小结

本章主要介绍了集成描述子,并详细分析了集成描述子的特征构造,它是由 Zernike 矩和傅里叶描述符组合的全局视图特征与二维形状特征按照一定比例加权得到的。集成描述子作为特征提取的描述符,能够解决单一描述符对二维草图信息描述不全的问题,利用欧式距离计算二维视图与草图之间的相似性,对手绘草图的检索结果进行从小到大的排序显示。本章还着重介绍了三维模型检索系统的应用,包括基于草图的检索过程和基于内容的检索过程两部分。针对基于草图的检索过程为第一步检索,降低了对用户的要求,只需要简单的草图即可完成检索,实验结果也表明本文的方法可以有效地检索相关的三维模型,相比于其它的方法具有更准确的检索结果。但是草图检索相对粗糙,所以根据结果可以使用基于内容的检索,即第二次检索。

第 5 章 基于内容的三维模型分类与检索

5.1 CNN 模型的构建

CNN 网络模型可以同时进行特征提取与分类，相对于传统分类方法，具有适应性强、泛化能力强等优点。相对于 BP 神经网络，具有计算复杂度低，运算效率高等优点。因此，本文选择 CNN 网络模型作为识别二维视图的分类模型，并且对二维视图进行特征提取。

CNN 网络模型基本结构包括两种神经层级，也就是卷积层和池化层。如图 5-1 所示，在卷积的过程中用一个人工设定参数的滤波器 f_x 与输入的二维视图作内积，设置一个偏置项 b_x ，得到卷积后的特征 C_x 。池化的过程是将邻域若干个二维视图特征像素求和，通过变量 w_{x+1} 进行加权，设置一个偏置项 b_{x+1} ，经过激活函数，得到一个缩小数倍的映射特征图 S_{x+1} [48]。

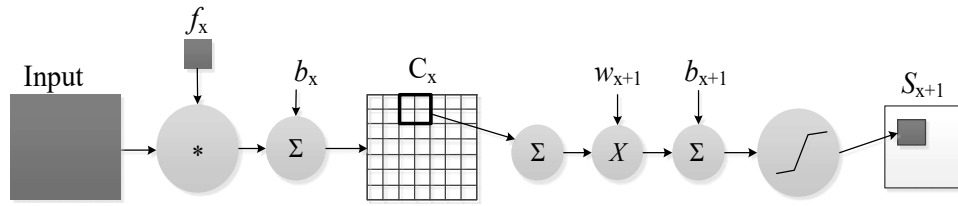


图 5-1 卷积核采样过程

Fig.5-1 Convolution kernel sampling process

考虑到二维视图的识别在科研实验中对准确性有一定的要求。因此，本文在设计识别二维视图的卷积神经网络结构时，不能采用网络复杂度较低的模型，如 3 层的普通卷积网络模型，或者 5 层的 LeNet 模型，虽然速度较快，但是准确率较低。采用层数较深的网络结构。对于卷积神经网络结构的选择，需要判断多少层的卷积层才能使分类识别率较高且速度较快。在深度学习算法中，深度模型的层数越多，模型识别能力越强，但也可能会导致数据出现过拟合的现象。如果模型的层数太少，识别能力太弱，又可能会出现数据欠拟合的现象。这两种情况都会导致训练模型出现极端现象，测试识别率较低。

5.2 AlexNet 模型调参

本文所采用的 AlexNet 模型是一个 8 层网络结构的卷积神经网络模型，其

中包括 5 层卷积层、3 层池化层和 2 层全连接层,如图 5-2 所示。以输入 $227 \times 227 \times 3$ 大小的图像为例,原有 AlexNet 模型在第一层卷积层中利用 96 个大小为 11×11 的卷积核对输入图像数据进行卷积,移动步长为 4,Padding 项为 0,经过第一层卷积层后图像特征为 $55 \times 55 \times 96$;第一层池化层中卷积核大小为 3×3 ,移动步长为 2,经过第一层池化层后图像特征变为 $27 \times 27 \times 96$;在第二层卷积层中利用 256 个大小为 5×5 的卷积核对输入图像数据进行卷积,移动步长为 1,Padding 项为 2,经过第二层卷积层后图像特征为 $27 \times 27 \times 256$;第二层池化层中卷积核大小为 3×3 ,移动步长为 2,经过第二层卷池化层后图像特征为 $12 \times 12 \times 256$ 。

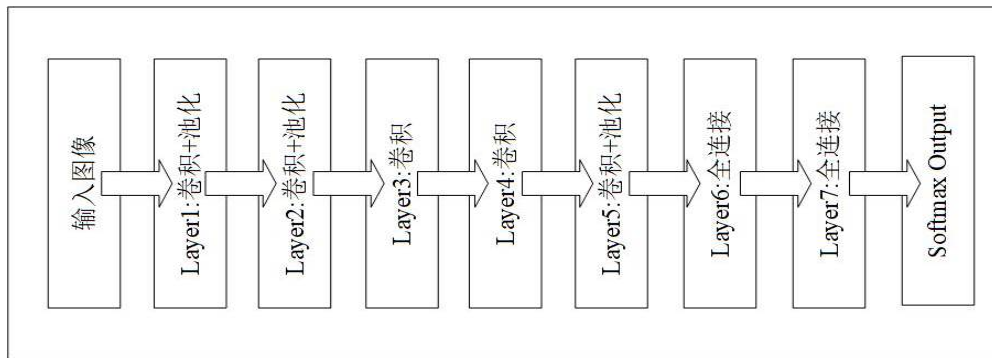


图 5-2 AlexNet 模型结构

Fig.5-2 AlexNet model architecture

在第三层卷积层中利用 384 个大小为 3×3 的卷积核对输入图像数据进行卷积,移动步长为 1,Padding 项为 1,经过第三层卷积层后图像特征为 $13 \times 13 \times 384$;在第四层卷积层中也是利用 384 个大小为 3×3 的卷积核对输入图像数据进行卷积,移动步长为 1,Padding 项为 1,经过第四层卷积层后图像特征为 $13 \times 13 \times 384$;在第五层卷积层中利用 256 个大小为 3×3 的卷积核对输入图像数据进行卷积,移动步长为 1,Padding 项为 1,经过第五层卷积层后图像特征为 $13 \times 13 \times 256$;第三层池化层中卷积核大小为 3×3 ,移动步长为 2,经过第二层卷池化层后图像特征为 $6 \times 6 \times 256$;图像特征经过三层全连接层得到 4096 维的向量,最后通过 softmax 输出分类。

本文对 AlexNet 模型进行参数微调,将卷积核均变为 3×3 大小,移动步长为 1,padding 项为 1。本文参数调整能够利用更多的特征,并保证在卷积层中输入图像经过卷积后大小不发生改变。如图 5-3 所示,改进后的模型第一层卷积层中卷积核个数为 96,卷积核大小为 3×3 ,移动步长为 1,之后对该卷积层进行线性偏移函数的转置和池化操作,能够去掉大量不重要的参数,降低计算量级,来提高模型的鲁棒性,以达到更好的学习效果。

在卷积神经网络中为了加快模型收敛速度，激活函数选取非线性非饱和的 Relu 函数^[49]，Relu 函数比饱和函数 sigmoid 函数和 tanh 双曲正切函数训练速度更快，并且保留了非线性的表达能力，而且在正值部分又具有线性的性质，不会引起非线性导致梯度弥散的现象，更有利于训练深层的网络。因此，本文卷积神经网络模型选择 ReLU 作为激活函数，如公式(5-1)所示。

$$y_{\mu,v}^i = \frac{x_{\mu,v}^i}{(e + \alpha_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (x_{\mu,v}^j)^2)^\beta} \quad (5-1)$$

其中， $x_{\mu,v}^i$ 表示 (μ, v) 处通过卷积和作内积计算得到的激活度量， n 是每一层的卷积核总数，其余参数为超参数。池化的操作选择最大池化，即在卷积层每一组 $2*2$ 的邻域点内取最大值，滑动步长为 1，最后利用局部响应归一化的方法来提高网络的泛化能力得到第一层的特征向量。

同理，第二层卷积层设置卷积核个数为 96，卷积核大小为 $3*3$ ，移动步长为 1，对卷积层进行线性偏移函数的转置和池化操作；第三层与第四层卷积层只有卷积核最大池化的步骤，参数为卷积核大小为 $3*3$ ，移动步长为 1；第五层经过卷积、最大池化和线性偏移函数转置后得到 256 个 $29*29$ 的特征向量；第六层是全连接层，将第五层卷积层得到的特征向量拉伸得到 4096 维的向量；第七层与第六层操作相同；第八层是通过 softmax 输出分类，本文中输出为 40，即 40 分类。对于分类类别判断，如公式(5-2)所示。

$$type = T(\max(p)) \quad (5-2)$$

其中， p 为 $N*1$ 的列向量，该向量的最大值就是最可能的类别。函数 T 为阈值函数，当输入大于函数设定的阈值时，则函数输出目标类别，否则输出 negative。

为了防止过拟合的现象发生，尽量在可允许的范围内保持标签不变的情况下增大数据集，并且在每个全连接层后面添加 dropout 函数，使卷积神经网络模型学习更加鲁棒的特征。

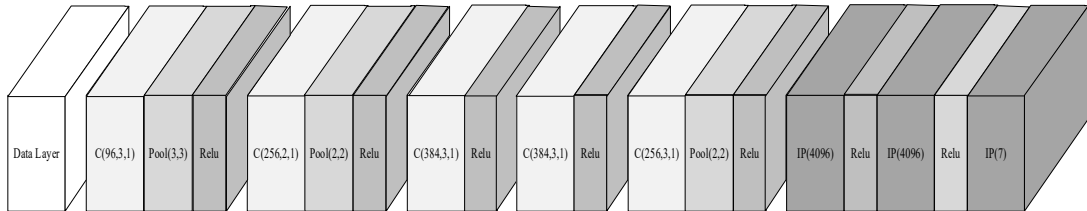


图 5-3 改进的卷积神经网络结构

Fig.5-3 Improved convolutional neural network structure

5.3 点云模型的架构

本文提出一种基于点云视图的三维模型分类方法，该方法也用到了上文提到的调参后的 AlexNet 模型。所有的卷积层都使用了 ReLU 激活函数，分类之前选择平均池化的操作，将一个三维模型的三个切片视图定义为点云视图集，作为卷积神经网络的输入，并且在最后一个连接层的底部加入 softmax 函数进行最后的分类，如图 5-4 所示。

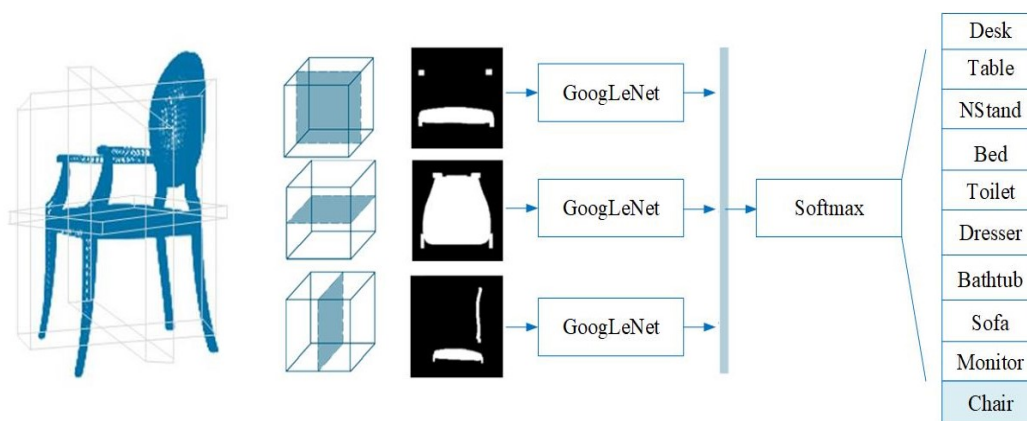


图 5-4 点云检索网络架构

Fig.5-4 Architecture of point cloud retrieval network

本文将整个数据集转换为切片投影视图。在训练和测试阶段，将三张视图提供给卷积神经网络，而不是通过点云。这是为了缩短神经架构的训练和测试时间。

5.4 AlexNet 模型训练与测试

本文采用二维典型视图训练卷积神经网络。在训练前，模型所有的权值都要通过高斯分布进行初始化。卷积神经网络的训练分为正向传播和反向传播两个过程。

在正向传播阶段，本文中样本均为 RGB 彩图，所以输入 x 就是对应 RGB 的三个矩阵。此时，上一层的输出就是下一层的输入，从隐藏层前向传播到卷积层的过程如式(5-4)所示。其中， l 代表层数， $*$ 代表卷积， W 代表卷积核， b 代表偏置， σ 为激活函数，在卷积层为 Relu 函数，在全连接层为 sigmoid 函数。最后得到一个 40 维的向量 y ，表示输入 x 被分为指定类型的概率。

$$a^l = \sigma(z^l) = \sigma\left(\left(\sum_{k=1}^M (a_k^{l-1} * W_k^l) + b^l\right)\right) \quad (5-4)$$

反向传播阶段，通过比较上一阶段得到的输出与本次实验收集的数据集类型标签向量的误差 δ^l ，使用误差代价函数调整权值参数，每一层的误差都可以根据公式(5-5)从反向逐层推导得到。其中 W^l 是由卷积核的权重组成的数组。 \circ 表示矩阵中每个元素对应相乘， $f'(net^{l-1})$ 是加权输入的激活函数的导数。

$$\delta^{l-1} = \delta^l * W^l \circ f'(net^{l-1}) \quad (5-5)$$

以下为本文在训练卷积神经网络模型的公式推导过程：

在卷积过程中，假设第 l 层为卷积层，第 $l+1$ 层为池化层，在第 l 层中第 j 个特征向量 x_j^l 如公式(5-6)所示。其中 u_j^l 表示每一个与卷积核同大小的矩阵向量，表示第 j 个矩阵与卷积核作完内积操作的向量，它是由上一层的特征向量 x_i^{l-1} 和卷积核 k_{ij}^l 作卷积运算，加上一个固定偏置 b_j^l 得到。其中 $f(\cdot)$ 是 ReLU 激活函数，“ $*$ ”表示卷积操作， N_j 是特征视图大小。

$$\begin{aligned} x_j^l &= f(u_j^l) \\ u_j^l &= \sum_{i \in N_j} x_i^{l-1} * k_{ij}^l + b_j^l \end{aligned} \quad (5-6)$$

计算第 l 卷积层的 j 个的残差 σ_j^l 如公式(5-7)所示，其中 $f'(\cdot)$ 为激活函数 $f(\cdot)$ 的偏导数， \circ 表示所有元素累加相乘， $up(\cdot)$ 是上采样计算。

$$\sigma_j^l = \beta_j^{l+1} (f'(u_j^l) \circ up(\sigma_j^{l+1})) \quad (5-7)$$

计算总误差 E 对偏置参数和卷积层参数的偏导数，其中， $(p_i^{l-1})_{uv}$ 为 x_i^{l-1} 中区域块， ∂b_j^l 表示偏移量， ∂k_{ij}^l 表示卷积核。

计算误差 E 对卷积层参数 k 的偏导数，其中 $(p_i^{l-1})_{uv}$ 为 x_i^{l-1} 中区域矩阵， ∂k_{ij}^l 表示卷积核， σ_j^l 为第 l 卷积层的残差，如公式(5-8)所示。

$$\frac{\partial E}{\partial k_{ij}^l} = \sum_{u,v} (\sigma_j^l)_{u,v} (p_i^{l-1})_{u,v} \quad (5-8)$$

计算误差 E 对偏置项 b 的偏导数， σ_j^l 为第 l 卷积层的残差， ∂b_j^l 表示偏移量，如公式(5-9)所示。

$$\frac{\partial E}{\partial b_j^l} = \sum_{u,v} (\sigma_j^l)_{u,v} \quad (5-9)$$

在池化层中，若第 l 层为池化层，则上一层为卷积层。在第 l 层中池化过程如公式(5-10)所示：

其中 $pooling(\cdot)$ 函数为池化层函数， x_i^{l-1} 为上一层卷积后的特征向量， β 是

池化权重大小， b_j^l 是偏置项， $f(\cdot)$ 是激活函数。经过池化操作后可以得到下一层卷积层需要的特征向量 x_i^l 。

$$\begin{aligned} x_j^l &= f(u_j^l) \\ u_j^l &= \beta_j^l \text{pooling}(x_i^{l-1}) + b_j^l \end{aligned} \quad (5-10)$$

第 l 层的池化层残差计算如公式(5-11)所示，其中 conv 函数为计算库中残差计算函数。

$$\sigma_j^l = f'(u_j^l) \circ \text{conv}(\sigma_j^{l+1}) \quad (5-11)$$

在池化层中，误差 E 对权重 β 的偏导如公式(5-12)所示， d_j^l 是特征向量 x_i^{l-1} 池化后的算子。

$$\begin{aligned} \frac{\partial E}{\partial \beta_{ij}^l} &= \sum_{u,v} (\sigma_j^l \circ d_i^l)_{u,v} \\ d_j^l &= \text{pooling}(x_j^{l-1}) \end{aligned} \quad (5-12)$$

通过正向传播和反向传播的训练，得到该模型的权重参数，再利用梯度下降的方法即可得到卷积神经网络的模型。本文经过多轮测试，学习率为 0.001 时，优化效率较高。训练迭代次数为 1400 次，训练达到 1200 次左右时，代价函数基本收敛，训练集上的训练误差曲线如图 5-5 所示。

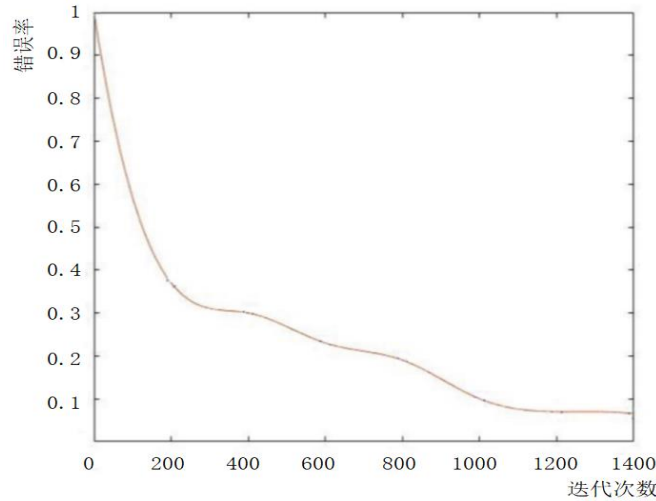


图 5-5 三维模型训练误差曲线

Fig.5-5 3D model training error curve

本文从 ModelNet-40 模型库组建自己的数据库。分别收集了 chair、table、sofa、monitor 等 40 类模型，每类模型 100 个数据样本。随机选取 80 个模型作为训练集，剩下 20 个模型作为测试集。将训练集三维模型通过投影和 Apriori 算法筛选出典型视图集，视图集作为卷积神经网络训练集，为每种模型人工标注标签。训练好的模型参数保存在模型文件中。在测试阶段，剩余 20 个模型在预处理阶段同样采用投影和 Apriori 算法，作为卷积神经网络的输入，输出结果为长度 40 的一维向量。向量中每个元素表示测试结果的分类占比率，其中占比最大项的类别就是测试类别。测试集部分三维模型如图 5-6 所示。



图 5-6 部分三维模型

Fig.5-6 Partial three-dimensional model

表 5-1 给出算法 1: 新 AlexNet 模型+LFD, 算法 2: AlexNet 模型+投影+Apriori 算法, 算法 3: 新 AlexNet 模型+投影+Apriori 算法, 算法 4: 多视角点云切片+新 AlexNet 在测试集上的分类准确率。从表 2 中可知传统的 AlexNet 模型与筛选二维典型视图的方法结合对三维模型的识别准确率达到 90.125%, 而采用参数调整后的 AlexNet 模型与筛选二维典型视图的方法结合对三维模型的识别准

确率高达 93.25%。实验验证了参数调整后的 AlexNet 模型在分类的效果上有了比较明显的改善；又通过 LFD、多视角点云等方法和生成二维典型视图的方法比较可知，AlexNet 算法筛选二维典型视图的方法在三维模型分类中效果更优。

在基于点云的检索方法中，对于影响学习过程的参数，本文将基础学习速率设定为 0.001，每迭代 1000 次将当前学习速率乘以 0.75，训练过程执行了 16000 次迭代。但是，最好的权重集是在迭代 14200 次上产生的。在 ModelNet40 数据集上测试精度为 94.37%。测试结果如表 5-1 和图 5-7 所示。

表 5-1 检索分类准确率

Table 5-1 Accuracy of retrieval classification

编号	类名	数量	算法 1	算法 2	算法 3	算法 4
1	airplane	20	0.7	0.95	0.85	0.9
2	bathhtub	20	0.85	0.85	0.85	0.85
3	bed	20	0.75	0.95	1.0	0.9
4	bench	20	0.7	0.9	0.95	0.85
5	bookshelf	20	0.65	0.85	0.85	0.9
6	boottle	20	0.75	0.85	0.9	0.85
7	bowl	20	0.85	0.85	0.85	0.85
8	car	20	0.85	0.95	0.95	0.95
9	chair	20	0.85	0.95	1	0.9
10	vase	20	0.9	0.95	0.95	0.95
11	wardrobe	20	0.95	0.95	0.95	1
12	xbox	20	0.85	0.85	1	0.9
13	cone	20	0.85	0.8	0.85	0.95
14	cup	20	0.75	0.85	0.95	0.85
15	curtain	20	0.65	0.85	0.85	0.8
16	desk	20	0.6	0.9	0.95	0.95
17	door	20	0.65	0.9	0.9	0.85
18	dresser	20	0.75	0.9	1	1
19	flower_pot	20	0.6	0.9	0.95	0.9
20	glass_box	20	0.75	0.9	0.9	0.85
21	guitar	20	0.6	0.95	0.9	0.9

编号	类名	数量	算法 1	算法 2	算法 3	算法 4
22	keyboard	20	0.75	0.95	1	0.9
23	lamp	20	0.85	0.9	0.9	0.95
24	laptop	20	0.75	0.9	1	0.85
25	mantel	20	0.85	0.9	0.95	0.9
26	monitor	20	0.75	0.9	0.9	0.85
27	night_stand	20	0.75	0.9	0.8	0.9
28	person	20	0.85	0.9	1	0.85
29	piano	20	0.75	0.8	0.9	0.85
30	plant	20	0.75	0.9	1	0.9
31	radio	20	0.75	0.8	0.85	0.85
32	range_hood	20	0.75	1	0.9	0.9
33	sink	20	0.85	0.9	0.9	0.85
34	tv_stand	20	0.65	0.9	0.95	0.85
35	sofa	20	0.65	0.9	1	0.9
36	stairs	20	0.65	0.95	1	0.95
37	stool	20	0.6	0.95	0.95	0.85
38	table	20	0.75	0.9	0.95	0.95
39	tent	20	0.85	0.95	1	0.9
40	toilet	20	0.75	0.95	1	0.95
41	all	800	0.755	0.90125	0.9325	0.89375

方便直观地看到比较结果，本文根据实验结果绘制出分类准确率折线图。如图 5-7 所示，明显看出采用 LFD 算法获取二维视图集的方法分类模型的效果最差，辨识度最低。这是由于 LFD 对于三维模型所有的特征信息一视同仁，无论是相对较明显的特征或是模糊特征，不能精准地描述三维模型的独有特征。

本文算法在作多角度投影操作的基础上，采用 Apriori 算法筛选出 12 幅典型视图，构成卷积神经网络的输入，甄选出能够更完整地描述三维模型的二维视图集。同时，由于视图集规模的降低，实验计算复杂度也大大降低。由实验结果可知此方法对三维模型分类的准确率有了较大的提高。

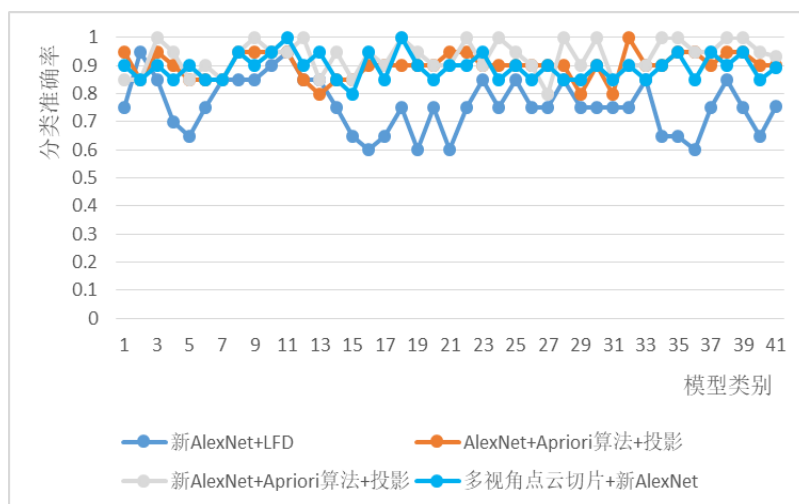


图 5-7 检索分类准确率

Fig.5-7 Accuracy of retrieval classification

本文采用两种检索算法结合的方式，即解决了普通用户没有现成实例的问题，又避免了只有单一草图算法的检索结果存在较大误差的问题，本文就以下实验进行说明，如图 5-8 所示，若用户手中没有适合的三维模型，此时可以先采用基于草图的检索算法，如图为手绘“床”模型的草图，可以看到，结果中有错误检索的类别(红圈中标识)。

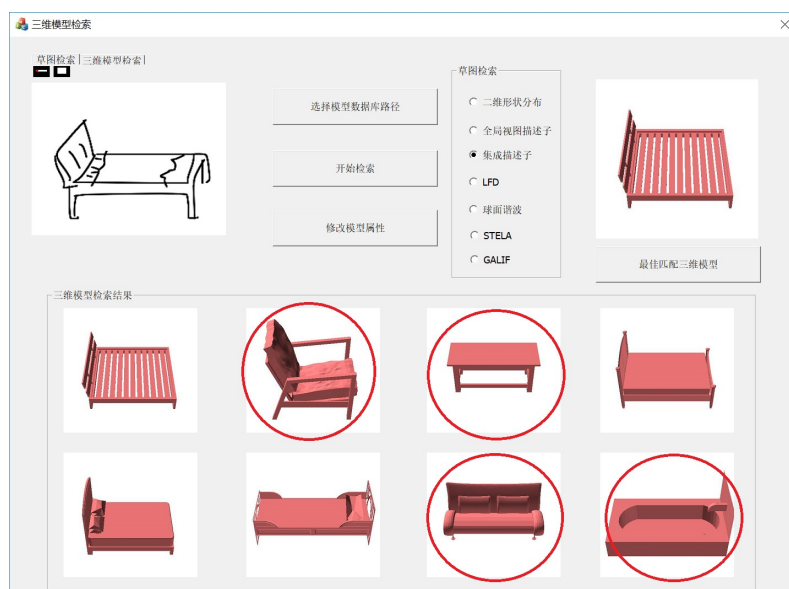


图 5-8 基于草图的检索

Fig.5-8 The flow of sketch retrieval

本文中针对草图认为第一排第四个模型为相对理想模型，可以将其拖到检索系统所示 picture_ctl 位置，利用基于内容检索完成二次检索，从图 5-9 中可以看出检索结果较单一的基于草图检索精度更高

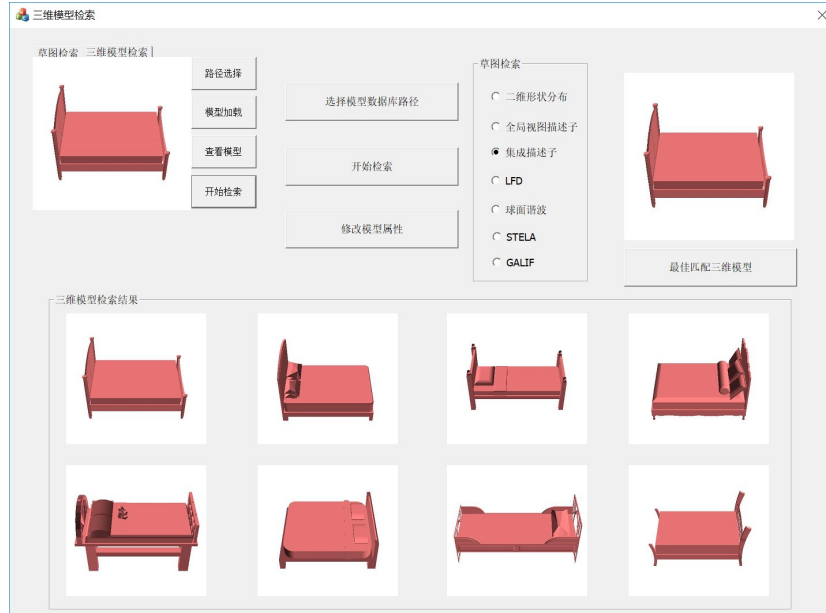


图 5-9 基于内容的检索

Fig.5-9 The flow of content-based retrieval

5.5 相似性计算

本文利用 CNN 实现模型分类。虽然在训练模型阶段费时较长，但是测试阶段需要的时间较短。若按照传统方法，仅对视图特征进行相似性度量是非常耗时的，而采用先分类再检索的方式，耗时大概只有传统方式的 $\frac{1}{40}$ 。

最终通过卷积神经网络提取出的都是一个 4096 维的特征向量，即每一幅二维视图都可以映射为一个特征向量。因此，二维视图之间的相似性比较也就转换成两个特征向量之间的比较。

常见的距离公式有三种，分别是 Manhattan 距离公式、Euclidean 距离公式和 Hausdorff 距离公式。计算两个 n 维的特征向量距离，如公式(5-13)，(5-14)，(5-15)所示，其中 $X = (x_1, x_2, x_3 \dots x_n)$ ， $Y = (y_1, y_2, y_3 \dots y_n)$ 。

Manhattan 距离公式：

$$D(X,Y) = \sum_{i=1}^n |x_i - y_i| \quad (5-13)$$

Euclidean 距离公式:

$$D(X,Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5-14)$$

Hausdorff 距离公式:

$$D(X,Y) = \max_{1 \leq i \leq n} \min_{1 \leq j \leq n} d(x_i, y_j) \quad (5-15)$$

Hausdorff 距离中 $d(x_i, y_j)$ 为两个特征向量间任意两个点之间的距离, 因此可以比较维度不等的向量之间的距离。

在本文中经过 CNN 提取的特征都是 4096 维, 本文在基于内容的检索方法中采用最通用的 Euclidean 距离公式计算模型相似度。两个特征向量距离 $D(X,Y)$ 越小, 证明相似度越高。在检索三维模型时, 将相似度计算结果从小到大进行排序。距离最小的二维视图所对应的三维模型就是目标模型, 即检索算法匹配的最相似模型。

5.6 本章小结

本章首先介绍了卷积神经网络模型的结构与推导过程, 包括经典的 AlexNet 模型, 然后提出了两种基于内容的三维模型分类检索算法。通过第四章的多视角投影与 Apriori 算法的结合甄选出典型视图集, 并在原有 AlexNet 模型基础上进行适当地调参, 使卷积神经网络模型能够利用更多的输入特征。实验结果显示, 该方法分类准确率较高, 达到了 93.25%, 且较好地弥补了传统视图提取方法容易造成数据冗余, 忽略不同的视图特征对三维模型检索效果的影响。将多视角点云检索算法用于三维模型的分类, 分类准确率也达到 89.375%。最后, 利用 Euclidean 距离公式完成检索过程。

结论

随着三维模型在日常生活中的应用越来越普及，快速查找三维模型检索算法逐渐成为国内外研究热点之一。本文描述了三维模型检索技术在国内外的研究现状，对三维模型检索技术的发展和主要方法作了概述。本文主要研究了基于视图的三维模型检索，包括基于草图的检索算法、基于二维视图的检索算法和基于多视角点云的检索算法。依托于本文的检索系统验证算法的优劣，并且按照检索算法的顺序能够降低用户的使用难度，提高检索准确率。

本文在基于草图检索的过程中，首先利用草图降噪的方式去除手绘草图的噪声，改善草图质量，提高识别准确率。提出了一种新的利用轮廓曲线的固定投影方式，避免数据冗余。提出了一种新的描述子，即集成描述子，该描述子有效地利用了视图的区域特征和边界轮廓特征，最后利用距离公式完成检索。实验验证本文方法比其它草图检索算法效果更好。基于草图的检索算法相对于基于内容的算法性能较差，但优势在于便于普通用户检索，不需要拥有三维模型即可检索。因此，本文系统先通过草图检索算法检索出需要的众多模型，在这些模型中选择三维模型进行第二次的精确检索。

本文在基于二维视图的检索过程中，提出一种固定投影加上 Apriori 算法的视图集甄选方法。该方法降低二维视图与三维模型的语义鸿沟，并且能够甄选出三维模型中最明显的特征。运用调参后的 AlexNet 模型对视图集进行训练，进行先分类后检索，能够数十倍的提高检索效率，并且分类准确率远远超过传统方法。本文将该方法作为用户检索的第二次检索，通过草图检索出的模型用作基于二维视图的检索输入，能够避免普通用户直接进行检索时没有适当模型的情况。

本文在三维模型检索算法中还提出了一种基于多视角点云投影的方法，对于三维点云模型进行一定厚度的切片并投影。针对投影点云视图进行像素扩散，再利用 AlexNet 模型对视图集进行训练，该方法的分类准确率略低于基于二维视图的检索方法。但相对于传统方法效果较优，并且由于点云模型相对于三维模型 OFF 文件较小，特征较少，因此检索时能够节省一定的时间，速率较快。

本文利用 Opencv、Opengl 和 PCL 开发了一个三维模型检索系统。该系统集成了本文提出的所有算法，利用 ModelNet-40 完成一系列的实验，验证了本文提出算法的优良性能和检索系统的人性化开发。

参考文献

- [1] 林佼, 李重, 金小刚. 基于凸壳与有向包围盒的骨架提取方法[J]. 计算机辅助设计与图形学学报, 2012, 24(6): 793-798.
- [2] HILAGA M, SHINAGAWA T, KOMURA T. Topology Matching for Fully Automatic Similarity Estimation for 3D Shapes[C]. ACM, 2001: 203-212.
- [3] 刘志, 潘晓彬. 基于渲染图像角度结构特征的三维模型检索方法[J]. 计算机科学, 2018, 45(S2): 261-265.
- [4] 李月峰. 一种基于多特征结合的三维模型检索方法[J]. 计算机科学, 2019, 46(S1): 266-269.
- [5] ZHUANG T, ZHANG X, HOU Z, et al. A Novel 3D CAD Model Retrieval Method Based on Vertices Classification and Weights Combination Optimization[J]. Mathematical Problems in Engineering, 2017: 1-12.
- [6] 杨军, 王亦民. 基于深度卷积神经网络的三维模型识别[J]. 重庆邮电大学学报(自然科学版), 2019, 31(02): 115-122.
- [7] 赵宇晨, 刘涛, 韩燮. 一种基于拓扑结构射影不变量的三维模型检索方法[J]. 中国科技论文, 2017, 12(20): 38-42.
- [8] TAO S Q. CAD Model Retrieval Based on Graduated Assignment Algorithm[J]. 3D Research, 2015, 6(2): 1-11.
- [9] MIN P, CHEN J, FUNKHOUSER T. A 2D Sketch Interface for 3D Model Search Engine[C]. ACM, 2002: 138.
- [10] 秦飞魏. 基于语义的异构三维 CAD 模型检索[D]. 浙江大学, 2014.
- [11] 石民, 张树生, 李亮. 基于二维典型视图的三维 CAD 模型检索算法[J]. 制造业自动化, 2012, 34(9): 82-84.
- [12] 冯元力, 夏梦, 季鹏磊. 球面深度全景图表示下的三维形状识别[J]. 计算机辅助设计与图形学学报, 2017, 29(09): 1689-1695.
- [13] SHI B, BAI B, ZHOU Z. 3D Deep Panoramic Representation for 3D Shape Recognition[J]. IEEE Signal Processing Letters. 2015: 2339-2343.
- [14] KALOGERAKIS E, AVERKIOU M, MAJI S, et al. 3D Shape Segmentation with Projective Convolutional Networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6630-6639.
- [15] CHEN D Y, TIAN X P, SHEN Y T. On Visual Similarity Based 3D Model

- Retrieval[J]. Computer Graphics Forum, 2003, 22(3): 223-232.
- [16] CYR C M, KIMIA B. 3D Object Recognition Using Shape Similarity-Based Aspect Graph[C]. ICCV, 2001: 254-261.
- [17] 刘玉杰. 基于形状的三维模型检索若干关键技术研究[D]. 北京: 中国科学院计算技术研究所, 2006.
- [18] CAO L, LIU J, TANG X. 3D Object Retrieval Using 2D Line Drawing and Graph Based Relevance Feed Back[C]. ACM, 2006: 105-108.
- [19] FANG Y, XIE J, DAI G. 3D Deep Shape Descriptor[C]. Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition. 2015: 2319-2328.
- [20] PAULSON B, HAMMOND T. Accurate Primitive Sketch Recognition and Beautification[C]. International Conference on Intelligent User Interfaces, 2008: 1-10.
- [21] HAMMOND T, DAVIS R. A Sketching Language for User Interface Developers[J]. Computers&Graphics, 2005, 29(4): 518-532.
- [22] XIE J, DAI G, ZHU F. Deep Nonlinear Metric Learning for 3D Shape Retrieval[J]. IEEE transactions on cybernetics, 2018, 48(1): 412-422.
- [23] LI B, LU Y, FARES R. Semantic Sketch-based 3D Model Retrieval[C]. IEEE International Conference on Multimedia and Expo Workshops, 2013: 555-558.
- [24] YE Y X. Applying Deep Learning to Scene Sketch Recognition and 3D Sketch-based 3D Model Retrieval[D]. Texas State University, 2016: 15-18.
- [25] CAO B, KANG Y, LIN S. Style-sensitive 3D Model Retrieval Through Sketch-based Queries[J]. Journal of Intelligent & Fuzzy Systems, 2016, 31(5): 2637-2644.
- [26] LI H, WU H, HE X. Multi-view Pairwise Relationship Learning for Sketch Based 3D Shape Retrieval[C]. IEEE International Conference on Multimedia and Expo (ICME), 2017: 1434-1439.
- [27] DAI G, XIE J, ZHU F. Deep Correlated Metric Learning for Sketch-based 3D Shape Retrieval[C]. Proceedings of The 31th Conference on Artificial Intelligence, 2017: 4002-4008.
- [28] SHAO T, XU W, YIN K. Discriminative Sketch-based 3D Model Retrieval Via Robust Shape Matching[J]. Computer Graphics Forum, 2011, 30(7): 2011-2020.
- [29] ZOU C, WANG C, WEN Y. Viewpoint-Aware Representation for Sketch-based

- 3D Model Retrieval[J]. IEEE Signal Processing Letters, 2014, 21(8): 996-970.
- [30] YOON S M, SCHERER M, SCHRECK T, et al. Sketch-based 3D Model Retrieval Using Diffusion Tensor Fields of Suggestive Contours[C]. International Conference on Multimedia, 2010: 193-200.
- [31] 谢晓华. 基于二维手绘草图的三维形状检索[J]. 集成技术, 2015(2): 22-33.
- [32] CAO B, KANG Y, LIN S, et al. A Novel 3D Model Retrieval System Based on Three-view Sketches[J]. Journal of Intelligent and Fuzzy Systems, 2016, 31(5): 2675-2683.
- [33] WU J, LIU L. Generating Quad Mesh of 3D Articulated Shape for Sculpting Modeling[J]. Journal of Advanced Mechanical Design Systems & Manufacturing, 2012, 6(3): 354-365.
- [34] LI B, LU Y, JOHAN H. Sketch-based 3D Model Retrieval by Viewpoint Entropy-based Adaptive View Clustering[C]. Eurographics Workshop on 3D Object Retrieval and Eurographics Association, 2013: 49-56.
- [35] LI B, LU Y, JOHAN H. Sketch-based 3D Model Retrieval Utilizing Adaptive View Clustering and Semantic Information[J]. Multimedia Tools & Application, 2016, 76(227): 1-9.
- [36] LI B, LU Y, SHEN J. A Semantic Tree-based Approach for Sketch-based 3D Model Retrieval[C]. International Conference on Pattern Recognition. IEEE, 2017: 3880-3885.
- [37] AONO M, IWABUCHI H. 3D Shape Retrieval from A 2D Image as Query[C]. Signal & Information Processing Association Summit and Conference. IEEE, 2012: 1-10.
- [38] LEI H, LI Y, CHEN H. A Novel Sketch-based 3D Model Retrieval Method by Integrating Skeleton Graph and Contour Feature[J]. Journal of Advanced Mechanical Design Systems & Manufacturing, 2015, 9(4): 49-63.
- [39] LI Y, LEI H, Lin S. A New Sketch-based 3D Model Retrieval Method by Using Composite Features[J]. Multimedia Tools & Applications, 2017: 1-24.
- [40] MAO D, YIN H, LI H. Sketch-based 3D Model Shape Retrieval Using Multi-feature Fusion[C]. Proceedings of The 2015 Chinese Intelligent Systems Conference. Springer Berlin Heidelberg, 2016: 395-402.
- [41] JIN X Y, LIU W Y. On-line Graphics Recognition, Proceedings of The 10th Pacific Conference on Computer Graphics and Applications[C], 2002: 256-264.

- [42] SU H, SUBHRANSU M, EVANGELOS K. Multi-view Convolutional Neural Networks for 3D Shape Recognition[C]. 15th IEEE International Conference on Computer Vision, 2015: 945-953.
- [43] KATAYAM K, SATO T. Matching 3D CAD Assembly Models with Different Layouts of Components Using Projections[J]. IEICE Transactions on Information and Systems, 2015, E98D (6): 1247-1250.
- [44] 刘玉杰, 宋阳, 李宗民. 融合信息熵和 CNN 的基于手绘的三维模型检索[J]. 图学学报, 2018, 39(04): 735-741.
- [45] 常亮, 邓小明, 周明. 图像理解中的卷积神经网络[J]. 自动化学报, 2016, 42(9): 1300-1312.
- [46] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection[C]. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, 1: 886-893.
- [47] Eitz M, Hildebrand K, Boubekur T. Sketch-based Shape Retrieval: Benchmark and Bag-of-features Descriptors[J]. Visualization and Computer Graphics, 2011, 17(11): 1624-1636.
- [48] EITZ M, RICHTER R, BOUBEKEUR T. How Do Humans Sketch Objects[J]. ACM Transactions on Graphics, 2012, 31(4): 44.
- [49] DENG S, TIAN Y. Application of New Advanced CNN Structure with Adaptive Thresholds to Color Edge Detection[J]. Communications in Nonlinear Science & Numerical Simulation, 2012, 17(4): 1637-1648.

攻读硕士学位期间所发表的学术论文

- [1] 基于手绘草图集成描述子的三维模型检索, 201910869044. 0, 2019. 09. 16(发明专利, 实质审查中)

致谢

时光荏苒，日月如梭，不知不觉我的研究生生活即将结束，在欢声笑语中我们放飞梦想，心存憧憬。在这段时间里，我成长了很多，不论是学习和生活中都受益很多，很感谢老师、同学和朋友的陪伴与帮助，让我有更大的勇气去迎接新的征程。

首先特别感谢我的导师高雪瑶教授，高老师严以律己、宽以待人、学识渊博、平易近人具有崇高的敬业精神，在科研学习的过程中，遇到问题老师会很耐心的帮我解答，并提供给我一些学习方法，受益很深。同时高老师在我最困难的时候对我不抛弃、不放弃，这种精神鼓舞我重新振作，克服困难，也激励我，让我在未来的日子里有信心面对今后的困难与挫折。同时高老师也为我的毕业论文提供了很多意见和帮助，使我顺利的完成论文。感谢张春祥教授，张老师不仅在学习期间给了我许多学术上的指导与帮助，在生活上也给了我许多的鼓励和关心。

同时我要感谢我的室友，感谢这几年他们的陪伴与鼓励。在学习上他们给予我最大的帮助与肯定，很感动在他们自己科研任务繁重的情况下，还抽出时间帮助我解决问题。在生活上一一起做感兴趣的事情，使我们的研究生生活变得多姿多彩。

最后，我要感谢我的父母。感谢他们一直在生活上和精神上的支持与帮助，作为我坚强的后盾，让我得学习生涯变得无忧无虑，希望我可以变得更加优秀，让父母舒心。

在研究生即将结束的日子，我再次向所有帮助过我的、关心过我的老师、同学、朋友及家人致以真挚的谢意！