

中图分类号: TP391  
学科分类号: 081202

论文编号: 1028716 12-S044

# 硕士学位论文

## 基于手绘草图的三维模型检索系统

研究生姓名	张 海 龙
学科、专业	计算机软件与理论
研 究 方 向	智能人机交互
指 导 教 师	谢 强 副教授

南京航空航天大学

研究生院 计算机科学与技术学院

二〇一一年十二月

Nanjing University of Aeronautics and Astronautics

The Graduate School

College of Computer Science and Technology

# **3D Model Retrieval System Based On Freehand Sketch**

A Thesis in

Computer Science and Technology Engineering

by

Zhang Hailong

Advised by

Xie Qiang

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Engineering

December, 2011

---

# 承诺书

本人声明所呈交的硕士学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得南京航空航天大学或其他教育机构的学位或证书而使用过的材料。

本人授权南京航空航天大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本承诺书）

作者签名：

日 期：

## 摘 要

随着计算机图形学技术的发展以及三维模型在各领域内的广泛应用，如何利用自由的表达方式快速准确的从数据库中检索设计过程中需要的三维模型变得尤为重要。目前，三维模型检索研究在用户检索意图的自由表达及三维模型特征提取方面仍存在问题，降低了三维模型检索系统人机交互的便捷性，影响了检索的性能和效果。

据此，本文研究了基于手绘草图的三维模型检索技术，主要在用户检索意图获取及三维模型特征提取方面展开研究。利用手绘草图的方式使用户自由灵活的表达检索意图，以草图特征与模型的投影轮廓特征进行相似性计算实现三维模型检索。在此基础上，给出了基于手绘草图的三维模型检索系统框架，实现了基于手绘草图的三维模型检索系统。本文主要研究内容如下：

(1) 设计了基于特征点的多边形拟合算法。在对草图进行预处理后，对草图的轮廓采用基于特征点的多边形拟合算法进行逼近，该算法弥补了传统拟合算法的不足之处，提高了后续过程中草图特征提取的准确性。

(2) 研究和改进了三维模型的特征提取方法。根据三维模型自身特性计算获取投影图像的主平面，利用改进的 SUSAN 算法对获取的三幅投影图像进行边缘提取，通过轮廓跟踪算法得到图像轮廓，进而从轮廓中提取描述三维模型的特征向量。

(3) 构建基于贝叶斯网络的推理模型和相关反馈。构建了用于三维模型检索的贝叶斯网络推理模型。根据用户绘制的草图特征通过贝叶斯网络推理模型实现三维模型检索。通过用户的反馈信息不断的对检索过程进行优化，提高三维模型检索的效率和准确性。

(4) 完成了基于手绘草图的三维模型检索原型系统的设计与主要功能开发。重点给出草图识别模块、三维模型特征提取模块等实现细节。

**关键词：** 手绘草图，草图识别，三维模型检索，贝叶斯网络，轮廓提取，特征向量

## ABSTRACT

With the development of computer graphics technology and three-dimensional model widely used in many domains, it is particularly important that how to quickly and accurately retrieve 3D models what users intent from the database in the process of design. Thus, in the field of multimedia information the research of content-based 3D model retrieval has been widespread concerned. Currently, the research of 3D model retrieval still has some problems in the freely expression of the user's search intent and the feature extraction of the model , these problems reduce the ease of human-computer interaction and affect the retrieval performance and results of 3D model retrieval system.

Accordingly, in this paper, there is a study on how to get user' retrieval intention and the feature extraction of 3D model. Users can use freehand sketch to freely express their retrieval intention.by calculating the degree of the similarity between the feature extracted from sketch and the feature of projection, users can achieve their retrieval purpose.On this basis, the framework and implementation of the retrieval system based on freehand sketch are proposed in this paper .The main researches are:

(1) Design a polygon fitting algorithm based on feature point. After pretreatment in the sketch, use the algorithm designed in this paper to fit the profile of sketch. The algorithm makes up for the traditional inadequacy of fitting algorithm and improves the accuracy of the feature extraction.

(2) Researched and improved feature extraction methods of three-dimensional model. Obtain the main plane used to project by calculating according to the properties of 3D model .uses the improved SUSAN algorithm to detect the edge of projected image, Obtain the feature used to describe 3D model from the profile which get by the profile tracking algorithm.

(3) Proposed a retrieval model based on Bayesian network and relevant feedback. Use feature data extracted from the profile to train the model, according to the characteristics of user-drawn sketches, users not only can achieve 3D model retrieval by the Bayesian network inference, but also can improved efficiency and 3D model retrieval accuracy by relevant feedback.

(4) Completed the design and the development of main function module of 3D model retrieval system based on freehand sketch, highlights the technology used in such as the pattern recognition module、 feature extraction of three-dimensional model and so on.

**Keywords:** Freehand Sketch, Sketch Recognition, 3D Model Retrieval, Bayesian Network, Profile Extraction, Feature Vector

## 目 录

第一章 绪论 .....	1
1.1 研究背景及研究意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	1
1.2 国内外研究现状.....	2
1.2.1 手绘草图研究现状.....	2
1.2.2 三维模型检索技术研究现状.....	3
1.2.3 三维模型检索系统简介.....	5
1.3 三维模型检索系统的一般框架.....	6
1.4 本文研究内容及组织结构.....	7
1.4.1 本文研究内容.....	7
1.4.2 本文组织结构.....	8
第二章 基于手绘草图的三维模型检索系统总体框架.....	9
2.1 系统的体系结构.....	9
2.2 系统处理流程 .....	10
2.2.1 草图识别数据流程.....	10
2.2.2 草图特征提取流程.....	11
2.2.3 三维模型特征提取流程.....	12
2.3 关键技术 .....	13
2.3.1 草图识别及特征提取.....	13
2.3.2 三维模型特征提取.....	14
2.3.3 相似性匹配.....	15
2.3.4 相关反馈.....	16
2.4 检索性能评估 .....	18
2.5 本章小结 .....	18
第三章 草图识别及特征提取.....	19
3.1 草图预处理 .....	19
3.1.1 冗余笔画删除.....	20
3.1.2 闭合检查.....	20

3.1.3 聚点消除.....	21
3.1.4 笔画的拟合.....	22
3.2 草图轮廓拟合 .....	23
3.2.1 轮廓去噪处理.....	24
3.2.2 基于特征点的多边形拟合.....	24
3.3 草图特征提取 .....	27
3.4 本章小结 .....	29
第四章 三维模型特征提取及检索.....	30
4.1 三维模型预处理.....	30
4.1.1 平移不变性预处理.....	31
4.1.2 旋转不变性预处理.....	31
4.1.3 缩放不变性预处理.....	32
4.2 三维模型投影图像获取.....	32
4.2.1 基本思路.....	32
4.2.2 投影平面选取.....	33
4.2.3 投影图像处理及特征提取.....	36
4.2.4 改进的 SUSAN 算法提取图像边缘.....	37
4.2.5 图像外围轮廓跟踪.....	39
4.3 基于贝叶斯网络的三维模型检索.....	40
4.3.1 贝叶斯网络.....	40
4.3.2 样本训练.....	42
4.3.3 贝叶斯网络推理模型.....	42
4.3.4 相关反馈.....	43
4.3.5 实验结果及分析.....	46
4.4 本章小结 .....	48
第五章 基于手绘草图的三维模型检索系统实现.....	49
5.1 引言 .....	49
5.2 系统功能结构设计.....	49
5.3 系统主要类设计.....	51
5.3.1 系统类图.....	51
5.3.2 草图轮廓识别及特征提取类图.....	52
5.3.3 三维模型投影图像获取及特征提取类图.....	53

5.4 数据库设计 .....	54
5.5 系统运行示例 .....	56
5.6 本章小结 .....	58
第六章 总结与展望 .....	59
6.1 总结 .....	59
6.2 展望 .....	60
参考文献 .....	61
致 谢 .....	66
在学期间的研究成果及发表的学术论文 .....	67



## 图表清单

图 2.1 基于手绘草图的三维模型检索系统框架.....	10
图 2.2 草图识别数据流程.....	11
图 2.3 草图特征提取流程.....	12
图 2.4 三维模型特征提取流程.....	13
图 3.1 不同用户的绘制习惯示意图.....	19
图 3.2 冗余笔画消除示意图.....	20
图 3.3 闭合检查与修复示意图.....	21
图 3.4 聚点消除示意图.....	21
图 3.5 直线拟合示意图.....	22
图 3.6 曲线拟合示意图.....	23
图 3.7 不同曲率笔画示意图.....	25
图 3.8 传统拟合算法的拟合效果.....	25
图 3.9 基于特征点的多边形拟合过程.....	26
图 3.10 基于特征点的拟合算法拟合结果.....	27
图 3.11 多边形上相邻的任意三个顶点.....	28
图 4.1 三维模型投影变换.....	33
图 4.2 主平面 $E_l$ 的旋转.....	35
图 4.3 三个主平面的空间关系.....	36
图 4.4 三维模型的投影图像.....	36
图 4.5 SUSAN 圆形模板.....	38
图 4.6 改进后的 SUSAN 算法与传统 SUSAN 算法边缘检测效果对比图.....	39
图 4.7 图像轮廓跟踪效果示意图.....	40
图 4.8 简单的贝叶斯网络.....	41
图 4.9 相关反馈流程.....	44
图 4.10 概率更新过程.....	46
图 4.11 相关反馈前后的查全率-查准率曲线.....	47
图 5.1 系统主要功能模块.....	50
图 5.2 系统类图.....	52
图 5.3 草图识别及特征提取.....	53

图 5.4 三维模型处理类图.....	54
图 5.5 系统 E-R 图.....	55
图 5.6 数据表设计 .....	55
图 5.7 手绘草图 .....	56
图 5.8 草图的规整 .....	57
图 5.9 检索结果 .....	57
图 5.10 反馈后的检索结果.....	58
表 4.1 不同方法的查全率与查准率比较.....	47
表 4.2 不同方法的检索时间比较.....	47

## 第一章 绪论

### 1.1 研究背景及研究意义

#### 1.1.1 研究背景

随着网络的普及、多媒体技术和计算机图形硬件的发展,三维模型以其立体表现特征及由其构成的三维场景具有更丰富的视觉感知细节等优点,成为继声音、图像和视频后的第四种多媒体数据类型。三维模型在现实生活中的应用范围也越来越广泛,并逐渐成为各领域设计的首选表现类型。其应用领域包括工业产品设计、三维游戏、影视动画、医学研究、计算机仿真、分子生物学<sup>[1][2]</sup>等。为了对三维模型进行的统一的存储和管理,各领域内相应的三维模型数据库也应运而生,如虚拟现实环境(Visual Reality)、工业实体 CAD 等等。设计师对于三维模型应用思路也在不断的进行转变,由传统的构造三维模型逐渐演变为如何对数据库中已有的三维模型进行提取、修改和重新组合,实现已有三维模型的共享和复用。这种转变既能解决大量的冗余三维模型产生问题及减少高逼真度的三维模型建模所需要的开支,也为实际中的具体应用提供了更加简便快捷的方式。设计师只需要获取个人需要的三维模型,然后进行相应逻辑上的修改或组合,便能完成实际的创意设计。比如在工业产品设计领域,如果能够利用从已构建的零件数据库中获取合适的模型作为新一轮设计的基础,经过进一步的加工和组合,形成新的产品造型,实现基于模型的设计和制造,不仅降低了产品设计的成本,也缩短了产品设计的时间。而实现这一目标的核心问题在于如何从数据库存放的大量三维模型中快速准确的检索出所需的三维模型。

手绘草图<sup>[3][4]</sup>是符号系统的一种特殊表达形式,具有语义、语法和模糊的特征,能够满足设计人员横向和纵向的思维活动。在设计过程中,手绘草图由于具有创造性思维表达的灵活性、抽象思维的外化和逻辑思想描述的自由性等优点,成为感知用户界面不可或缺的一部分。手绘草图给设计师提供了更加自然直接的交流和创意表达方式,对设计过程中产生的瞬间灵感表达具有重要的意义。手绘草图的信息连续性、信息多维性等优点减少了人机交互中的阻碍,保证了设计过程中思维的流畅性,更能对设计者所要表达的意图在内容上进行直接快速的视觉表达。目前草图识别技术研究正受到越来越多人的关注,并取得了大量的研究成果。手绘草图也逐渐成为在进行基于内容的检索系统中用户意图输入方式的首选。手绘草图识别技术在比如室内设计、工艺设计、建筑设计、图像检索等都得到了广泛的应用。

#### 1.1.2 研究意义

目前三维模型的检索技术主要包括基于关键字检索技术<sup>[5]</sup>和基于内容的检索技术<sup>[6]</sup>,尽管

基于关键词的检索技术已经取得了重大成果,但是关键字的标注往往依赖于标注者的主观因素,容易产生歧义,且关键字的描述能力比较片面,难以充分的描述复杂的三维模型,在三维模型的检索过程中已经无法满足设计和使用的需要。因此,基于内容的多媒体信息检索技术受到研究人员的广泛关注并迅速的发展起来,基于内容的三维模型检索技术主要是根据视觉特征的相似性来匹配用户检索的三维模型,这种思路也符合人们在现实生活中依据视觉感知判断和利用信息的方式。

但是,目前大多数的基于内容的三维模型检索技术在检索条件输入时并没有给用户方便表达其检索意图的交互界面。系统要求用户输入进行检索的初始条件,然后依据一定的相似性匹配算法对用户所需要的三维模型进行检索。在用户对个人的检索意图进行表达的过程中,有时面临着无法对自己所想要检索的模型有比较细致的认识,或者难以准确的表达个人检索意图,又或者表达细节方面比较繁琐等情况,由于无法清晰的对检索意图进行表达,进而会导致最后检索的结果并不能达到用户的要求,降低了用户对于系统检索性能的满意度,影响了用户实际工作中的需要。

手绘草图在表达检索条件时具有视觉空间媒体形象、直接地表达和快速的概念传递<sup>[7]</sup>等突出的优点,与传统的图形图像工具相比,更加方便、自然、随意,能够灵活的表达模糊概念,十分符合人类的使用习惯。因此,将手绘草图引入三维模型检索系统中有着明显的先天优越性和适应性。借助于手绘草图的检索接口,使用户在输入的过程中把更多的关注放在自己所要表达的意图上,通过简单的勾画将思维中呈现的三维模型的主要特征点描述出来,更加流畅的表达个人的检索意图,缩短用户意图和计算机化的检索条件之间的语义距离。因此,基于手绘草图的三维模型检索技术研究具有非常重要的理论意义及良好的实用价值。

## 1.2 国内外研究现状

### 1.2.1 手绘草图研究现状

草图识别技术的难点在于如何根据用户输入的笔画数据理解其逻辑上所要表达意图。一般识别过程可以分为两个步骤:草图识别和草图理解。草图识别主要是对用户绘制的草图进行数据采样,去除草图中影响识别效果的噪声信息,依据一定的算法识别草图中的图元信息、空间信息、约束信息等。经过识别后,用户模糊的草图表达映射为精确的图形表达,为后续理解草图的高层语义提供基础笔画信息;草图理解则是将识别后的草图符号映射为某领域内的具有特定含义的语义符号,是用户设计意图在领域范围内的体现。

目前草图识别的研究成果主要分为以下四类:

(1) 基于统计信息的识别方法<sup>[8]</sup>。该类方法根据草图中的直线段数目、弧线段数目、平行线数目、交叉线数目、封闭线数目等统计信息对用户输入的语义符号进行识别。在识别过程中,

对于草图中的笔画数据,根据其所归属的语义符号不同进行分类,在分类的基础上一般采用圆锥曲线方程的线性最小平方匹配法识别出笔画的类别,如直线、圆弧或圆等。利用统计后的数据对符号所表达的语义做出判断。

(2) 基于模糊类的识别方法<sup>[9]</sup>。该类方法采用模糊推理和模糊知识,根据草图的位置、方向、笔画速度和笔画顺序等用户习惯捕获用户的绘制意图,依据用户在绘制过程中笔画的曲率变化、速度变化等对草图的特征点进行判断,以实现草图识别。

(3) 基于几何信息的识别方法<sup>[10]</sup>。在对草图去噪、笔画规整、特征点提取等处理过程之后,将草图分解为多种类型的笔画,根据笔画中相邻三点建立角度相似函数,利用笔画的角度值与设定的阈值之间的比较实现笔画的分类。

(4) 基于过滤器和神经网类的识别方法<sup>[11]</sup>。该类方法通过构建过滤器对笔画进行分类或利用神经网络的推理和记忆功能识别草图。文献[12]提出通过提取图元几何形状的内角特征,用二进制突触的权重算法 BSW(含一个隐层的前馈网)进行识别的方法,该方法以整体方式识别三角形、椭圆和矩形三种图素,但仅适用于绘图包和掌上电脑的自动草图输入;文献[13]提出模糊样条曲线识别器 FSCI,并训练三层前向神经网实现模糊分类规则,可识别七种图元(直线、圆、圆弧、椭圆、椭圆弧、开放式自由曲线和闭式自由曲线),但各图元都被限制在水平、垂直和四分角方位,且对于不同的使用者需重新训练神经网络。

不同的草图识别技术根据识别的侧重点不同、应用领域不同等都具有各自的优缺点和具体的适用范围。随着草图识别技术研究的不断深入,将会出现更多更有效的草图识别方法,推动草图识别技术在现实生活的应用持续向前发展。

### 1.2.2 三维模型检索技术研究现状

经过近几年的不断发展,三维模型检索技术取得了较大的进展,三维模型检索技术在基于三维模型特有的性质上借鉴和融合了其他多媒体检索领域比较成熟的方法,以提高三维模型检索系统的检索效率和使用效率。根据检索时输入的方式不同,三维模型检索同样可以分为基于文本的三维模型检索(即基于关键字的检索)和基于内容的三维模型检索。

基于关键字的三维模型检索技术通过使用关键字描述数据库中的三维模型和三维场景,该类检索技术虽然在检索文本信息领域得到了广泛的应用,也取得了很好的检索效果。但是,在对于如图像、三维模型等多媒体信息检索领域内所表现出的性能却并不理想<sup>[14][15]</sup>。其中主要原因为:

(1) 三维模型具有直观、立体、多视角等符合人们视觉感知的特点,如果仅仅使用关键字对其进行描述,难以对复杂模型的特征全面细致的描述。

(2) 关键字的选取工作主要依赖于标注者个人的主观因素,受标注者知识背景的限制,往

往不具有较大范围内的代表性，对于同一个三维模型的描述，由于标注者的信息认知侧重点不同，关键字的选择也可能具有很大的差异。

(3) 关键字的描述主要通过手工添加。对于庞大的三维模型数据库而言，这无疑是一个费时费力的工作。

为了弥补基于关键字的三维模型检索技术描述能力不足的缺陷，研究人员提出了一种新的检索方法，即基于内容的三维模型检索方法。与基于关键字检索方法有所不同，基于内容的三维模型方法试图从三维模型中抽取出具有视觉特征的内容信息作为检索依据，如空间关系、表面特性、拓扑结构等。与数据库中已经建立的三维模型的特征向量进行相似性计算达到检索的目的。提取的视觉特征和建立的特征向量之间的认知差异主要依靠一定的相似性匹配算法进行计算，当两者之间的差异在用户容许的范围之内时，系统将输出用户所需要检索的三维模型。在整个基于内容的检索过程中，涉及的相关技术主要包括三维模型特征提取、相似性度量、模型检索、反馈机制、性能评价等。基于内容的检索技术的优点是人工干预少、选取的特征贴近视觉信息、检索准确率高等。

在基于内容的三维模型检索系统中，三维模型特征提取方法的好坏将直接影响到能否检索到用户理想的目标模型。因此，三维模型特征提取技术是当前研究的主要方向之一。目前已经提出的特征提取方法主要分为以下几类：

(1) 基于统计特征的特征提取方法。该方法具有思路简明、构造简单、直方图本身易于接受等优点。在三维模型特征提取时，如果直接通过参数化的方式获取三维模型的形状特征并不是一件容易的事情，然而可以利用统计学方面的知识，统计三维模型中具有独立代表意义的几何形状特征，如形状直方图、点密度、射线、约束线等，利用统计数据作为检索时的特征向量。如 Osada 等人提出的根据模型表面任意三点之间的夹角函数、任意两点之间的距离函数等对三维模型中的特征转换成形状分布<sup>[16]</sup>，但是该类方法对于模型的描述内容不够充分，并且统计过程中的计算量比较大。

(2) 基于代数矩的特征提取方法。由于代数矩的特征保证了其不会受到三维模型旋转、平移和尺度缩放改变的影响，非常适合于描述形态多样、类别不同的三维模型特征，所以被引入到三维模型的特征描述方法中。如目前常见的 Legendre 矩、三维 Zernike 矩、三维 U 系统矩、三维小波矩等<sup>[17-19]</sup>，但是基于代数矩的特征提取方法在计算量方面耗费比较大，一定程度上影响了三维模型检索系统的效率。

(3) 基于拓扑关系的特征提取方法。该类方法主要提取三维模型的拓扑结构和骨架特征信息并以此描述三维模型，比如三维模型的分支、连通性等。目前常见的提取方法主要有基于 Reeb 图<sup>[20]</sup>和基于中轴线<sup>[21]</sup>的方法。该类特征提取方法能较好的描述模型的高层结构信息，特别是基于 Reeb 图的特征提取方法，虽然同种物体的不同形状模型，但由于其拓扑结构相同，

所以提取的特征之间是相似的。基于拓扑关系提取方法的问题在于进行规范化的特征集合分割时所附加的条件过于苛刻，且对于噪声比较敏感，对于处理三维多边形模型的效果不够理想。如 Hilaga 提出的基于多分辨率 Reeb 图的检索方法和唐勇对基于骨架的三维模型检索技术的研究<sup>[22]</sup>。

(4) 基于视觉图像的特征提取方法。该类方法主要是三维模型通过正交投影或者透视投影获得二维的图像信息，将对三维模型的直接检索转换为对二维图像的检索，利用如今比较成熟的二维图像检索技术实现三维模型的检索。如 Chen 提出的光场描述方法<sup>[23]</sup>是基于三维模型的侧面影像，通过采集不同的侧面影像，利用投影图像的 Zernike 矩描述子和傅立叶描述子相结合的方法实现三维模型的检索。

(5) 基于表面属性的特征提取方法。通过三维模型的表面属性，比如颜色、纹理和反射系数等等对特征进行提取。具有典型代表性的是 Suzuki 等人提出的基于三维模型颜色的检索<sup>[24]</sup>，该方法根据用户对于三维模型外表颜色的主观认知实现模型的检索。

基于内容的三维模型检索技术仍在不断的发展和优化过程中，但如何能够实现算法检索效率和检索结果理想的状态还需要更多的探索。

### 1.2.3 三维模型检索系统简介

在三维模型检索系统的研究领域内，经过近几年的不断发展及研究人员的不断努力，许多研究机构基于现有理论和技术开发了一些通用领域或专业领域的原型检索系统<sup>[25]</sup>。其中大部分实现了基于 Web 平台的搜索，为三维模型的检索提供了更多的便捷性。在本中列举部分被广泛应用的三维模型检索系统和搜索引擎，它们主要有：

(1) 美国普林斯顿大学的形状检索与分析小组开发的三维模型搜索引擎，在三维模型的检索领域具有里程碑式的意义。该系统提供了多种交互操作方式，包括关键字输入、二维草图的绘制及建立三维模型，为用户的输入提供了便捷性。

(2) 美国卡耐基-梅隆大学 AMP 实验室开发的三维模型检索系统。该系统在检索的过程中融合了底层形状特征和语义特征，提供了相关反馈功能对系统的检索性能进行优化。

(3) 普渡大学 PRECISE 检索系统。该系统是针对工业零件模型的领域而设计，在工业零件规范化的基础上采用体素化的方法得到模型的骨架结构特征用于检索。该系统还运用了神经网络方面的相关知识理解用户的反馈，提高了模型检索的准确性。

(4) 荷兰 Utrecht 大学的 GIVE (Geometry Imaging and Virtual Environment) 实验室的三维形状搜索引擎。该引擎采用基于高斯曲率的形状特征完成模型的检索。

(5) 德国波恩大学 Klein 和 Novotni 等人开发的数字图书馆系统 RODA。该系统并没有采用通过提起模型的特征实现检索的思想，而是通过直接测定不同的模型间体积误差与分叉轮廓

实现检索功能。

(6) 中国台湾省台湾大学通讯与多媒体实验室的三维模型检索系统。该系统基于 Web 平台, 通过提取三维形状描述符和模型多视图描述符完成检索功能。

(7) 日本多媒体教育学院的 Suzuki 等人开发的多边形模型检索系统 Ogden N。该系统同时使用模型的形状和颜色特征对 VRML 模型数据实现检索, 目的在于建立三维的网络教学环境。

(8) Sam3D 公司开发的商业搜索引擎。该引擎在实现检索功能的基础之上, 还为不同领域的用户交流提供了互动平台。相比于研究机构提供的免费系统, 该系统为使用者提供了更为贴切的服务, 用户检索出得模型文件也具有更高的质量和更好的设计效果。

(9) 以色列 Institute of Technology 开发的 Georgle 三维模型检索系统。该系统具备了其他在线系统所没有的用户的反馈信息学习功能。

(10) 美国布朗大学 LEMS (Laboratory for Engineering Man/MachineSystem) 实验室开发的三维模型检索系统。该系统通过用户输入的二维图形对数据库中的三维模型进行检索。

除此之外, 还有一部分用于专业领域的检索系统, 如德国慕尼黑大学 Ankerst 等人研制的用于蛋白质分子模型检索的系统以及苏格兰 Heriot-Watt 大学开发的基于 Web 平台的 CAD 模型检索系统 ShapeSifter 等。

### 1.3 三维模型检索系统的一般框架

典型的三维模型检索系统框架由在线和离线两部分构成, 按模块可以划分为数据获取、数据分析和检索匹配三个主要模块。其中数据获取模块的主要功能是获取表现用户检索意图的关键字、图像信息以及用于检索的三维模型数据库; 数据分析模块的功能主要通过特征提取算法从三维模型数据库中提取描述模型的特征信息、用户输入的检索条件中提取表现用户检索意图的初始查询数据; 而检索匹配模块则是连接数据分析与数据获取之间的桥梁, 通过对数据分析模块获取的特征信息与建立的三维模型数据库中的特征信息进行相似性计算, 查找符合用户检索意图的三维模型。查找的方法主要通过计算用户输入的特征向量与建立的三维模型特征向量之间的相似程度, 根据计算的结果给出与用户检索意图相匹配的三维模型。在系统中三个主要模块之间相互协作, 各自发挥着不同的作用。其总体框架结构如图 1.1 所示。



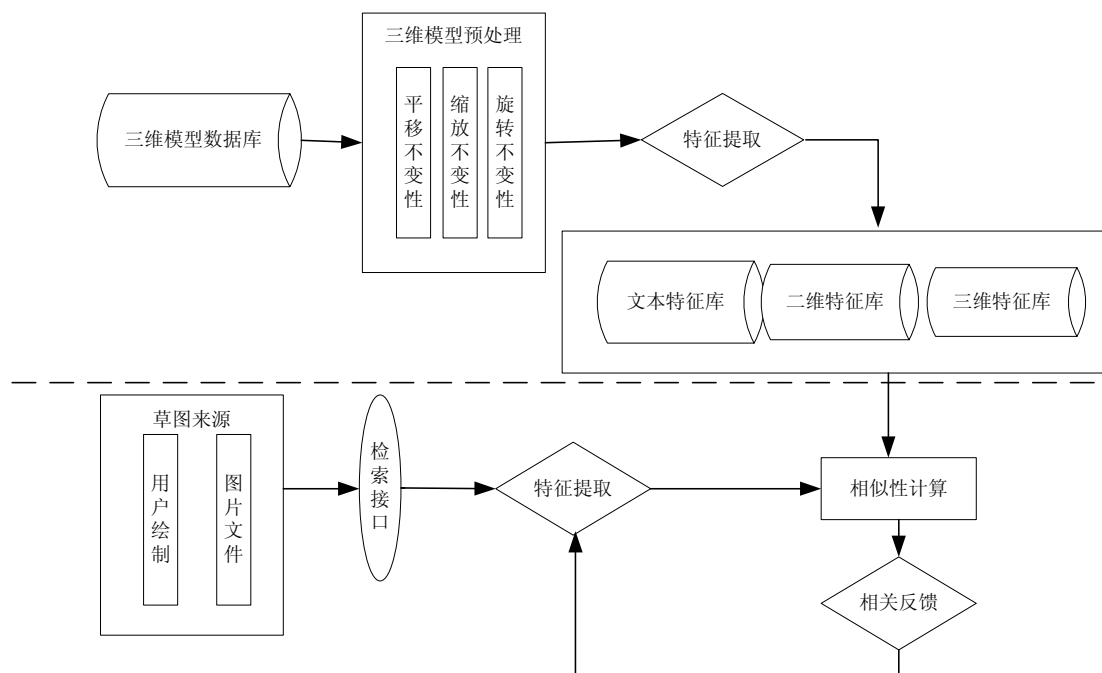


图 1.1 基于内容的三维模型检索系统一般框架

## 1.4 本文研究内容及组织结构

本文研究了手绘草图的识别技术、三维模型特征提取及检索技术。在融合草图识别和三维模型检索的关键技术基础上，提出了基于手绘草图的三维模型检索的总体方案，并依据方案初步实现了原型系统。

### 1.4.1 本文研究内容

#### （1）设计了基于特征点的多边形拟合方法

特征点是在草图中具有代表意义的点，是描述草图特征最为直接的方式。传统的多边形拟合方式并没有将草图轮廓中的特征点在拟合过程中的重要作用体现出来，且拟合算法的计算量大、执行效率较低。本文在传统多边形拟合算法基础上将草图轮廓的特征点充分利用到拟合过程中，对草图特征点附近区域使用更多的线段进行更加细致的拟合，降低了拟合过程对后续草图轮廓特征提取的影响。

#### （2）研究了基于主平面分析的三维模型投影图像获取方法

基于三维模型投影的检索系统中，大部分系统都采用了基于坐标轴的投影方式，所得到的投影图像对三维模型的描述能力有限。因此，本文采用基于主平面分析的三维模型投影，通过计算得到三维模型的三个投影主平面，获取包含三维模型自身特性的投影图像，提取投影图像的特征更好的反映三维模型的特征，以利于检索性能和准确性的提高。

(3) 改进了提取投影图像轮廓的 SUSAN 算法

传统 SUSAN 算法对图像边缘的提取主要基于中心点像素与圆形模板内的像素值之差进行比较, 通过计算所有比较结果的总和与设定阈值之间的大小判断是否为图像的边缘像素点, 这种方式对于噪声等干扰因素的抵抗性不够。本文中对 SUSAN 算法进行了改进, 将原有的中心像素与圆形模板差值的计算改进为中心像素与圆形模板中像素差值的比值, 该方法能够增强算法对于噪声的抗干扰性, 使得边缘提取更加清晰准确。

(4) 构建基于贝叶斯网络的推理模型和相关反馈

通过构建贝叶斯网络推理模型, 从数据库中根据用户输入的草图特征将用户需要的三维模型检索出来, 并呈现给用户。为了提高系统检索的准确性, 用户可以根据个人的主观判断对结果进行评价, 系统在融合用户的反馈信息之后将重新检索后的结果输出。

(5) 完成了基于手绘草图的三维模型系统的设计及开发

在前面研究和改进的关键技术基础之上, 初步实现了基于手绘草图的三维模型检索系统。在系统实现过程中改进现有三维模型检索系统中的存在一些不足之处。在提供友好的交互界面前提下, 优化了系统的检索性能和检索效果。

## 1.4.2 本文组织结构

围绕上述内容, 论文共分六章, 各章节的具体安排如下:

第一章 绪论。分析了论文研究背景及研究意义, 并对基于手绘草图的三维模型检索系统中所涉及到的草图识别技术及基于内容的三维模型检索的研究现状进行了分析与总结。给出了论文主要研究内容、研究目的与意义。

第二章 基于手绘草图的三维模型检索系统框架设计。提出了基于手绘草图的三维模型检索系统的总体框架, 对系统框架中各主要模块涉及的关键技术进行了概述, 并给出了每个模块中的处理流程。

第三章 手绘草图的识别。给出了手绘草图识别的具体流程, 并对草图预处理、特征点检测、轮廓拟合、特征提取五个环节分别进行了具体的阐述, 对其中使用的方法进行了设计或者改进。

第四章 三维模型特征提取及检索。阐述和改进了三维模型特征提取过程中的相关关键技术, 包括三维模型投影平面的选取、投影图像边缘提取及投影图像轮廓跟踪及特征提取。构建贝叶斯网络推理模型实现三维模型的检索, 并通过引入相关反馈技术提高系统检索的准确性。

第五章 系统实现。采用 UML 对系统进行了分析设计, 给出了系统功能模块划分以及主要模块的类图、顺序图。从系统的功能模块设计、系统结构模块的设计和数据库设计等方面介绍详细的实现过程, 最后给出了部分系统运行的界面。

第六章 总结与展望。总结本文研究工作, 并展望了进一步的研究工作。

## 第二章 基于手绘草图的三维模型检索系统总体框架

基于内容的三维模型检索系统的主要目标是从众多的三维模型数据库和网络中检索到理想的目标模型。虽然三维模型有多种表示方式,既能够采用多边形对其进行表示,也可以采用体素、线等表示方式。但是,由于这些表示方式的特殊性,无法对任意的三维曲面直接进行参数化,也就无法利用类似于二维图像标准的采样方式对物体空间进行定义,加上三维模型的拓扑结构没有局限性,如傅里叶变换等经典的信号处理技术也无法直接引入到三维模型领域。因此,基于内容的三维模型检索系统在技术实现方面存在不足。而基于手绘草图的三维模型检索系统,通过草图输入的方式可以使用户更加自然的表达个人检索意图,草图的特征与三维模型投影图像的特征之间可以建立直接的联系。因此,系统具有很大的灵活性和易用性。本章在给出了基于手绘草图的三维模型检索基础上,对其中涉及的数据流程及关键技术作了简要介绍。

### 2.1 系统的体系结构

本文中研究的基于手绘草图的三维模型检索系统包含的主要模块为草图输入、草图预处理模块、草图特征提取模块、三维模型投影平面及投影图像获取模块、投影图像边缘检测及特征提取模块和贝叶斯网络推理及相关反馈模块。系统的总体框架如图 2.1 所示。

(1) 草图预处理。主要对用户输入的草图根据绘制过程中产生的噪声类型进行相应的处理,包括了冗余笔画删除、聚点的消除、闭合检查及直线和曲线的拟合等,使处理后的草图能够更清晰的表达用户的检索意图,为后续的草图拟合及特征提取打下基础。

(2) 草图特征提取。主要是对预处理后的草图,利用轮廓跟踪算法获取草图的外围轮廓,再对轮廓进行平滑处理后,采用改进的基于特征点的多边形拟合算法对草图的外围轮廓进行拟合,构建基于拟合多边形的几何不变量,并以此作为草图的特征向量用于三维模型的检索。

(3) 三维模型投影平面及投影图像获取。主要根据三维模型本身固有的特性,计算能够最大程度反映三维模型特征的主平面。由于单个主平面上的投影图像特征描述能力不够,因此,本文中引入了另外两个投影主平面。并在三个主平面上对三维模型进行正交投影,获得描述三维模型特征的三幅投影图像。

(4) 图像边缘检测及特征提取。主要利用改进的 SUSAN 算法提取投影图像的边缘,并在此基础上利用轮廓跟踪算法获取投影图像的外围轮廓,通过特征提取方法提取轮廓的特征信息,以便与草图的特征之间进行相似性计算。

(5) 贝叶斯网络推理及相关反馈。通过构建基于贝叶斯网络的推理模型实现三维模型的检索。为了提高系统的检索准确性,用户可以对系统返回的结果集进行主观评价,对检索结果

是否与自己的检索意图作出相关性的判断，系统将融合用户的反馈信息后调整相关参数重新输出检索结果。

在本系统中，用户提交一幅手绘草图后，系统将对原始的笔画信息进行数据采集、草图预处理、轮廓跟踪及拟合、特征向量的生成等一系列的过程，然后根据生成的特征向量与三维模型数据库中已经建立的特征索引进行相似性计算，将计算结果呈现给用户。为了提高整个系统的交互性和系统检索效果的进一步优化，用户在检索结束后可以对系统输出的结果进行判断，对于与个人检索意图无关的模型进行标记。系统将根据用户反馈的结果采用一定的反馈算法对相应的参数进行调整和更新，在融合用户的主观评价信息后重新进行检索，并将结果呈现给用户。

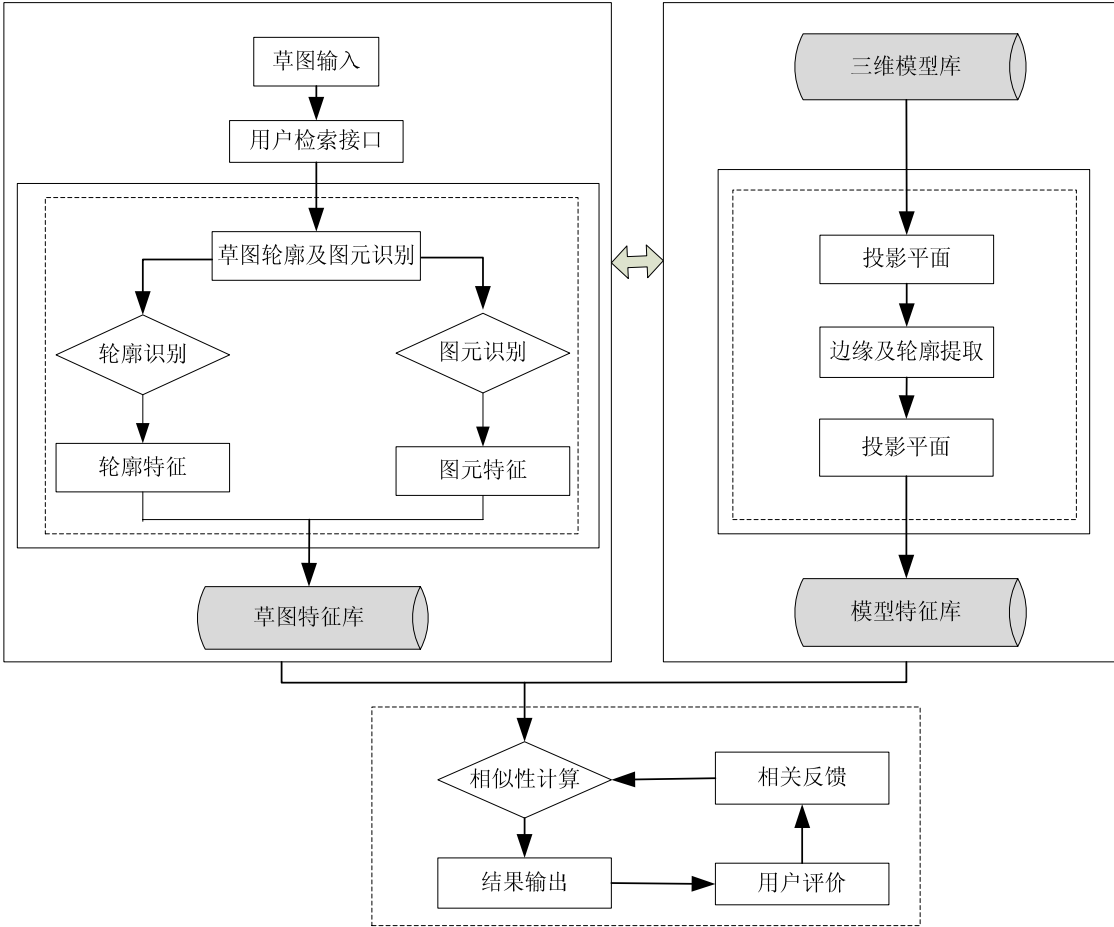


图 2.1 基于手绘草图的三维模型检索系统框架

## 2.2 系统处理流程

### 2.2.1 草图识别数据流程

草图识别<sup>[26]</sup>是利用手绘草图进行三维模型检索的前提，草图识别流程中主要包括对获取的

草图采样数据进行草图预处理、草图识别两个过程。在预处理过程中，将草图中存在的噪点、冗余笔画等从草图整体中剔除，减少对草图识别效果的影响。在草图识别过程中，主要对草图中出现的直线、曲线等信息进行规整和拟合。并根据这些信息完成对草图的识别，其主要的数据流程如图 2.2 所示。

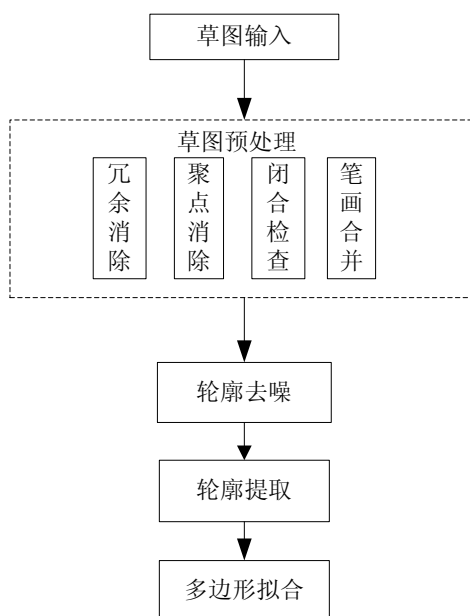


图 2.2 草图识别数据流程

### 2.2.2 草图特征提取流程

草图的特征提取是在草图识别的基础上，提取草图特征描述子构成特征向量作为检索的初始条件。由于草图中蕴藏了用户需要的三维模型的视觉感知信息，特别是草图的外围轮廓，更能体现三维模型的重要特征，用户在绘制草图的过程中，将描述三维模型特征的信息通过草图的特征点等方式体现出来。因此，在本文中主要对草图的轮廓特征进行提取，而内部的投影轮廓只是在识别过程中作为辅助信息。

草图的特征提取流程如图 2.3 所示。根据预处理后的草图数据，依据轮廓跟踪算法获取草图的外围轮廓。由于提取的草图外围轮廓往往具有比较大的噪声，而噪声的存在会影响到后续过程中提取特征信息的准确性，因此，需要对草图进行去噪。对于草图的外围轮廓，如果直接对草图轮廓进行特征提取，存在提取的特征信息准确性不高且提取过程复杂的问题，构建的特征向量也不容易直观的表达，而采用多边形对草图的轮廓进行拟合的方法在保证特征提取准确性的同时也获得计算上良好的性能。将对草图的处理转化为对拟合后的多边形进行特征提取后，由于多边形具有很好的几何特性，可以利用数学中的相关知识选取多边形的几何性质作为特征向量。在本文中，对草图轮廓拟合后，构建了多边形中的几何不变量，选取的几何不变量既满足作为草图特征描述符的要求，也具有直观和计算简单的特点，降低了后续相似度计算的复杂

度。

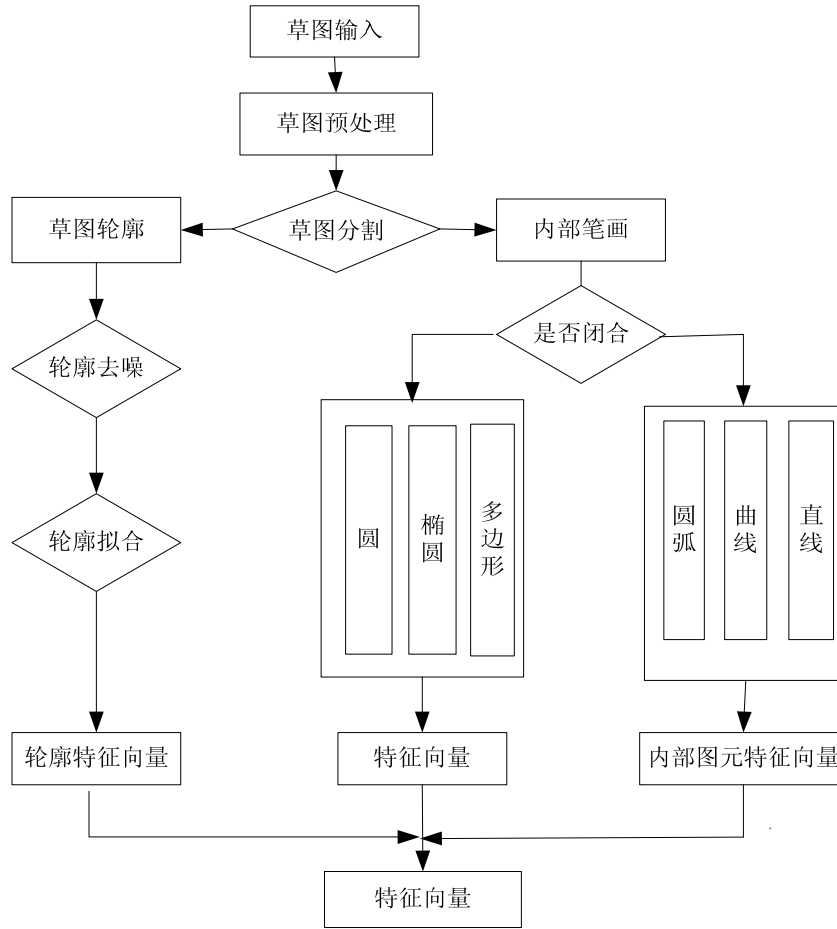


图 2.3 草图特征提取流程

### 2.2.3 三维模型特征提取流程

三维模型的特征提取如图 2.4 所示。在三维模型检索前，需要建立三维模型特征向量库，以便于相似性的计算和检索性能的优化。因此，三维模型的特征提取是三维模型检索过程一个重要的任务。但是，由于三维模型的来源、大小等不同，在提取三维模型的特征之前，同样需要对三维模型进行预处理<sup>[28]</sup>。经过预处理后的三维模型，可以利用统一的标准进行描述，而且满足平移、旋转和比例缩放的不变性。针对平移不变性，一般采用 Vranic 加权面积的方法<sup>[29]</sup>，计算公式如下：

$$\bar{P} = \sum_i P_i \cdot S_i \quad S_i = \frac{1}{3} \sum_j S_j \quad (2-1)$$

其中  $\bar{P}$  为选取的模型的中心点坐标， $S_j$  为模型中每个顶点的表面积，其由  $S_i$  的临界网格面积决定。这种方法弥补了单纯的将三维模型的坐标原点进行平移时受模型顶点采样密度的影响，提高了模型预处理的鲁棒性；针对旋转不变性可以采用主元分析法<sup>[30]</sup>、最大法线分布<sup>[31]</sup>

和最大面积法、极限投影变换方法<sup>[32]</sup>等；而针对缩放不变性，可以采用将三维模型放入一个单位球体里面，使模型的尺寸得到了统一，这样提取出来的特征向量就实现了缩放的不变性。在本文中，用户输入接口得到的草图信息主要是二维的。因此，需要得到三维草图在二维空间的信息，利用二维的特征向量描述三维模型的特征。

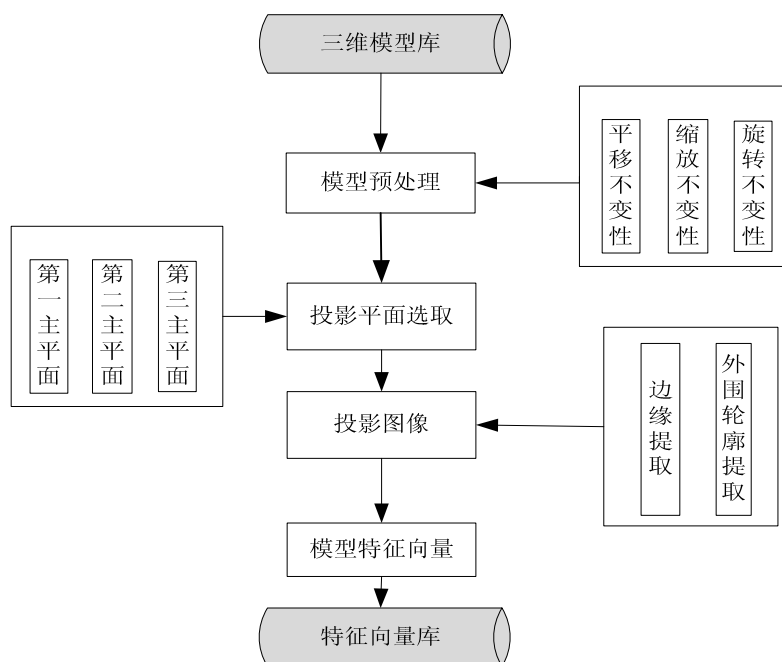


图 2.4 三维模型特征提取流程

## 2.3 关键技术

### 2.3.1 草图识别及特征提取

草图识别是基于手绘草图的三维模型检索的基础，对于用户绘制的草图，体现了用户思维中三维模型的轮廓信息及模型上突出的特征点。因此，草图的识别效果将直接决定了整个系统的检索性能。在本系统中，用户绘制的是三维模型在某一个方向上投影图像的轮廓特征，草图的结构相对比较简单，图元的个数也比较少。用户检索意图的主要由草图外围轮廓上的特征点表现出来。因此，本文中所使用的算法更侧重于外围轮廓上特征点的提取，寻找能够比较明确的区分出两个不同的三维模型的特征信息。这与一般的草图识别方法在侧重点和算法的选取上面有着很大的不同。

由于草图的模糊性和随意性等特点，用户在绘制草图时往往更侧重于思想和创意的表达，而不希望受到绘制习惯等因素的影响。在进行草图的识别之前也需要对草图进行预处理，草图的预处理主要包括冗余笔画的删除、交叉点的消除、笔画的闭合检查等，通过草图预处理过程，使得用户所要表达的检索意图在草图上有更加完整和清晰的体现，减少不必要信息的干扰，为

下一步的草图识别奠定基础。

由于草图的外围轮廓特征能反映出草图所具有的重要特征，具有较强的特征描述能力，也与人类对于物体的视觉感知相吻合。因此，用户手绘草图的轮廓提取在系统中具有重要的意义。通过对草图的笔画数据进行采样，依据一定的算法跟踪采样点数据得到草图的轮廓信息。在此基础上，通过对外围轮廓进行特征提取得到反映草图自身特性的特征向量，将特征向量作为初始的检索条件对三维模型进行检索，实现用户的检索意图和满足设计需要。

对于外围轮廓的提取首先需要对草图的轮廓进行一定的处理，用规范性或者更容易提取特征的方式近似的表达出草图的轮廓信息。目前常用于轮廓表示的主要包括基于采样点的轮廓特征提取、基于极图谱的轮廓特征提取、基于离散傅里叶变换的轮廓特征提取、基于多边形拟合的轮廓特征提取等<sup>[33-36]</sup>。不同的特征提取方式所关注的轮廓的特征点各不相同，算法的执行效率和使用范围也不相同。在本文中，采用基于特征点的多边形拟合算法对草图的外围轮廓进行拟合，并在此基础上构建了描述草图轮廓特征的几何不变量，并以此构建草图轮廓的特征向量。改进后的多边形拟合算法能对草图的轮廓进行更加准确的描述，有利于草图轮廓特征点的提取和特征向量的生成。

### 2.3.2 三维模型特征提取

目前，提取的三维模型特征需要满足以下条件：

(1) 提取的特征信息具有较强的描述能力，能够反映模型的整体信息。

(2) 提取的特征不会受到三维模型边界噪声和模型退化现象的影响，并且具有良好的鲁棒性和稳定性。

(3) 由于三维模型的来源各有不同，在空间中得自由度相对较高，三维模型可以具有任意的位位置、朝向和尺度的大小。因此，提取的三维模型特征需要满足平移、旋转和缩放的不变性。

根据用户输入的检索条件完成三维模型的检索，需要将草图的特征向量与用户将要检索的三维模型特征向量之间建立联系。但两者之间直接的相似性计算存在一定的困难，一方面三维模型基于三维的坐标空间，与草图的二维特征向量并不能直接进行相似性的判断；另一方面，三维模型的特征向量提取过程比二维图像的提取过程更加的复杂，往往需要经过长时间的处理才能得到用户查询的三维模型，造成了系统与用户之间交互性差的问题，不能满足用户对于友好的交互界面的要求。

为了避免上述两个问题的出现，在进行三维模型检索前需要提取描述三维模型的特征向量，以二维图像特征向量的特征描述方式刻画三维模型的特征，该方式既解决了计算特征向量相似性时的维度问题，也能提升计算和检索的效率。而三维模型的投影图像恰恰是完成转换的桥梁，三维模型的投影图像不仅反映了三维模型在某一视角上的特征信息，也能够运用成熟的



二维图像处理方法对图像进行处理。将投影图像的特征信息与三维模型之间建立对应关系，不仅符合人类的视觉认知习惯，也能够大幅度的减少直接提取三维模型特征过程中大量复杂的计算。

在本文中，根据三维模型自身的特性计算用户获取投影图像的三个主平面，将三维模型在三个平面上分别进行投影获得投影图像，然后利用改进的 SUSAN 算法<sup>[37]</sup>对投影图像进行边缘提取、轮廓跟踪等处理。得到对应于三维模型的特征向量，建立系统特征向量库。当特征向量库中数目巨大时，可以建立相应的索引提高检索的效率。

具体的三维模型处理及特征提取技术详见第四章。

### 2.3.3 相似性匹配

从用户输入的检索意图中提取作为检索条件的特征向量后，要检索出用户理想的模型，则需要比较输入的特征向量与数据库中的模型之间的特征向量之间的差异距离，由特征向量之间的差异程度进而得到模型的相似程度。根据距离值进行相似度排序，系统也会根据排序的结果将最符合用户检索意图的三维模型呈现出来。目前用于衡量模型之间相似程度的函数一般都选择距离函数<sup>[37]</sup>，不同距离函数虽然表达形式有所不同，但是它们之间具有统一的标准。如式 2-2 所示。

$$\begin{aligned} D(x, y) &= 0 \Leftrightarrow x = y \\ D(x, y) &= D(y, x) \\ D(x, y) &\leq D(x, z) + D(z, y) \\ D(x, y) &\geq 0 \end{aligned} \tag{2-2}$$

假设在本文中提供  $n$  维特征向量，给定输入模型和数据库中的系统特征向量分别为： $X = (x_1, x_2, \dots, x_n)$ ， $Y = (y_1, y_2, \dots, y_n)$ ，比较常见的距离函数公式可以表示为：

$$(1) \text{ Euclidean 距离: } D(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \text{ , 如果给特征向量赋予不同的权重后则}$$

$$\text{可辨识为: } D(X, Y) = \sqrt{\sum_{i=1}^n \omega_i (x_i - y_i)^2}$$

$$(2) \text{ Manhattan 距离: } D(X, Y) = \sum_{i=1}^n |x_i - y_i|$$

$$(3) \text{ Hausdorff 距离: } D(X, Y) = \max_{1 \leq i \leq n} \min_{1 \leq j \leq n} (x_i, y_j)$$

$$(4) \text{ 动态时间规整距 } d_{tw}: \text{ 定义两个长度均为 } N \text{ 的向量 } V_1, V_2, V_1 = \langle V_{1,0}, V_{1,1}, \dots, V_{1,N-1} \rangle,$$

$V_2 = \langle V_{2,0}, V_{2,1}, \dots, V_{2,N-1} \rangle$ 。  $Tail(V)$  定义为向量  $V$  去掉第一个元素的向量即

$Tail(V_1) = \langle V_{1,1}, V_{1,2}, \dots, V_{1,N-1} \rangle$  :

$$d_{tw}(\langle \rangle, \langle \rangle) = 0, \quad d_{tw}(V_1, \langle \rangle) = d_{tw}(\langle \rangle, V_2) = \infty \quad (2-3)$$

$$d_{tw} = \sqrt{(V_{1,0}, V_{2,0}) + \text{Min} \begin{cases} d_{tw}(V_1, Tail(V_2)) \\ d_{tw}(Tail(V_1), V_2) \\ d_{tw}(Tail(V_1), Tail(V_2)) \end{cases}} \quad (2-4)$$

根据特征提取算法提取的特征的不同，往往也会选择相匹配的相似度计算方法来评价最后得到的检索结果。如可以使用 Minkowski 距离<sup>[38]</sup>函数中的  $l_n$  范式进行形状特征的分布比较，公式可以表示为： $D(X, Y) = (\int |X - Y|^x)^{\frac{1}{x}}$ （分布函数）；如果对主观语义进行相似性匹配的话，可以采用多为综合度量<sup>[39]</sup>的方法。通过构造一个空间来描述用户的兴趣度，通过多元回归的分析将三维模型特征空间和用户的兴趣度空间一一对应起来，这样便实现了抽象的三维模型特征描述向量与人们的主观语义关键字联系了起来；如果只是对局部的检索来说的话，可以采用几何哈希表进行特征匹配。不同的距离度量方式也会影响到三维模型检索的效果，Euclidean 距离、Manhattan 距离主要适合与计算各维之间特征相互的特征向量；而二次型距离和 EMD 距离则是属于交叉对应类，它们的应用范围主要在与处理维度不同且具有相关性的特征向量之间。

### 2.3.4 相关反馈

相关反馈<sup>[40]</sup>是信息检索领域的一个研究焦点，它随着文本检索系统的出现而产生，它要求用户对于系统输出的检索结果进行主观上的评价，系统将通过相关反馈计算对检索过冲在那个某些参数进行调整和优化。通过这种交互的方式帮助三维模型检索系统提高检索的精度。

由于现有算法的特征描述因子并不能完全的反应出一个三维模型信息的所有特征，也无法对人类的视觉感知系统进行完整的模拟，低层的单一特征信息并不能完整的刻画出高层丰富语义的信息。往往一种特征向量只能对三维模型某一特征的信息进行体现，但是对于另一方面的特征信息却无法具有良好的体现效果。因此，系统检索出来的结果往往和设计本身的意图之间会产生一定程度的偏差，检索结果中出现的不相关的模型的存在会降低用户对于系统的满意程度。因此，相关反馈的算法被引入到检索系统中，以缩小低层特征描述与高层语义之间的语义鸿沟，从而提高系统的检索性能。

在相关反馈的过程中，首先是通过用户对系统检索出来的模型进行评价，然后采用相关反馈技术将用户的评价信息和特征向量权重之间进行融合，根据调整后的特征向量对用户的设计意图重新进行检索。在这个不断调整的过程中，系统的检索效果也得到不断的优化。

用户的判定信息是反馈算法所使用的原始数据，用户判定信息的获取方式主要是通过提示用户，让用户对系统所呈现出的检索结果进行标注。通过这种方式再次捕获到用户对于检索结果中的模型从视觉感知上的满意程度，目前原始数据的获取主要包括以下几个方面：

(1) 正负类型。正负类型的判断是指用户对于提供的结果信息给出自己的评价，但是用户对于模型的评价结果只可以分为两类：一是相关，表示用户对该结果满意；二是不相关，表示用户对该结果不满意。这种表达信息对于结果的划分过于苛刻，并不能具体的描述出用户对于一个结果模型的满意程度。后来，研究人员将这种二元选择推广到有多个相关级别的方法中，对相关的程度进行了描述上的改进，使之更加具体的反映出用户的偏好。但是这种多级的反馈方法给用户的操作便捷性带来了一定的损失。

(2) 等级排序。等级排序主要是让用户在考虑结果中的一个子集合，并根据自己的主观评判对结果进行相关性的排序，这种方法增加了用户的评判任务。

(3) 内部反馈。在内部反馈机制中，检索系统以可视化形式为用户呈现其内部查询结构，让用户交互地对这个内部结构进行操作，以改进查询表达。许多文本搜索引擎使用了内部反馈技术<sup>[41]</sup>。对于内部反馈，用户必须具备一定的领域知识，且很快就会被复杂的多媒体数据困扰。

获得了用户的评价信息之后，下一步将采用一定的相关反馈算法充分的利用用户的反馈信息对检索性能进行优化，虽然在文本检索和图像检索领域相关反馈的方法一定广泛的使用，但是是同一个算法对于三维模型领域下的性能却并不一定理想。目前常用的反馈算法主要有：

(1) 优化查询向量或距离度量的相关反馈，该方法来源于文本检索的研究过程，主要通过获取到的用户反馈信息修改查询向量向结果的中心移动对得到的检索结果进行优化，该方法实现起来较为直观简单，如 JingHuang 剔除的修改距离公式<sup>[42]</sup>的反馈方法，在采用欧式距离作为特征空间的距离度量算法的情况下，如  $D(x, q) = (x, q)^T (x, q)$ ，通过采用加权欧式距离以调整特征向量各维的权重，如  $D(x, q) = (x, q)^T A (x, q)$ ，A 为对角矩阵，则经过反馈后于查询向量等距离的点构成沿坐标轴方向的超椭球面。

(2) 基于神经网络的学习的相关反馈 神经网络学习方法的优点在于能够通过一定的训练过程，达到一个比较理想的检索效果。而获取用户的反馈信息同样可以看作是用户和系统之间的一个人机交互过程，这样，用户的反馈信息就给神经网络提供了不同的训练数据，其中包括正例样本和反例样本。基于统计学习理论的支持向量机是一种理论上比较成熟应用广泛的技术，它不仅能够有效的分离训练数据中的两类模型，而且还能利用分类器准确的将没有经过训练的模型区分出来，例如在 Chu Hong Ho 等人的工作中，利用 SVM 构造分类器<sup>[43]</sup>进行相关反馈。利用同心的超球面划分特征空间，将正理训练数据和反例训练数据分别归类到内外两个球面中。这种区分方法充分利用正例样本，但是也同时考虑反例数据的对于检索结果的影响，因此能够具有较好的性能。

## 2.4 检索性能评估

随着模型检索技术研究的发展，涌现出了大量的特征提取算法和模型检索算法。如何客观的评价出一个算法性能也成为了目前研究人员关注的焦点，而如何将不同算法的检索结果在同一个层面进行比较是解决这一问题的难点，目前常见的算法性能评价的表示主要分为以下三种：

(1) 数据曲线图。首先定义查全率和查准率，用公式表示为：

$$precision = \frac{C}{A} \quad recall = \frac{C}{RA} \quad (2-5)$$

其中  $C$ 、 $A$ 、 $RA$  分别表示正确检索的模型、检索出的模型总数、模型库中相关的模型数目。

这种表示方法主要是通过构造出检索系统的查全率和查准率的曲线示意图来直观的反映检索性能的好坏。

(2) 视觉角度。可以按照计算出的相似度的值从高到低对检索结果中的模型进行排序，并将结果在一幅图直观清晰的反应出来，在这种情况下可以直观的看出结果模型与目标模型直接的相似度值，有利于检索算法的改进。

(3) 等级匹配。设定将要检索的模型为  $C$ ，则数据库中与  $C$  相关的模型数目为  $|C|$ ，检索后返回的数目为  $K$ ，则可以定义在第一等级匹配和第二等级匹配下的  $FT$ 、 $ST$  分别为：

$$FT = \frac{K}{|C|-1} \quad ST = \frac{K}{2 \times |C|-1} \quad (2-6)$$

由上述公式可以推理出，检索系统的性能与  $FT$ 、 $ST$  的值成正比。

## 2.5 本章小结

本章在原有系统的基础上提出了基于手绘草图的三维模型检索系统总体框架，分析了系统实现中各部分的数据处理流程，并给出了具体的数据流程图。对系统中所涉及的草图处理、三维模型特征提取、相关反馈等关键技术进行了阐述。最后给出了常用于检索系统性能评估的具体策略。

### 第三章 草图识别及特征提取

在基于手绘草图的三维检索模型系统中，草图信息识别是整个系统的一个重要模块，因为草图从二维图像方面体现了用户的绘制意图，也是模型检索的初始检索条件。草图使用户的检索意图具象化，同时能够在一个流畅的环境中表达自己的创新活动。因此，草图识别过程的好坏将直接影响了后续的检索结果。而用于三维模型检索的草图信息，在识别过程中不仅要考虑草图本身所具有的特征信息，而且还要从中抽取出更能表达与其相对应的三维模型所具有的特征信息。因此，在三维模型检索系统的草图处理方式虽与普通的二维草图的识别方式大体过程相同，但是在识别的过程中其着重点不同。本节着重处理与三维模型相关的草图特征，如草图的轮廓信息、草图边缘具有较大曲率的顶点等，主要因为草图的轮廓信息能够反映三维模型在某一方向的投影，而边缘上具有较大曲率的特征点往往又能对三维模型的顶点进行表示。草图的处理流程主要包括笔画数据采集、笔画预处理、草图轮廓提取及拟合、特征提取等。

#### 3.1 草图预处理

草图的预处理过程是草图识别过程中最为基础同时也是非常重要的环节，草图识别的原则在于保持用户连贯自由的绘制表达自己意图的草图过程中对用户的设计意图进行捕捉。但是，对于不同的用户，其绘制习惯也不尽相同，一幅草图的绘制过程、笔序、布局、笔画大小、笔画数目等等也因人而异；即使同一个用户，自身的绘制习惯也会根据时间的推移而有所变化。如图 3.1 所示，两位用户以不同的笔画数目和笔序绘制矩形：用户 1 用三条笔画完成了一个矩形的绘制，但是用户 2 由于与用户 1 的习惯不同只用了两条笔画，并且他们的绘制的笔画顺序也不同。从图中也不难看出，虽然用户的绘制习惯不同，但在绘制过程中都不同程度的产生一些无关的笔画数据，如图 3.1 中圆圈标注的冗余点、闭合处的交叉等。

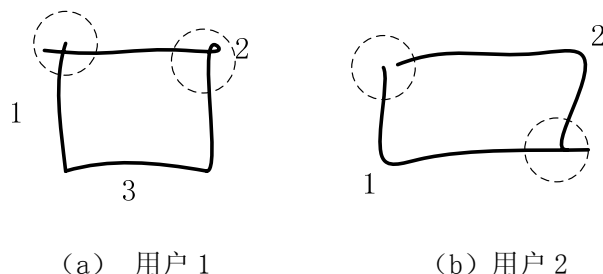


图 3.1 不同用户的绘制习惯示意图

因此，草图预处理的目的在于将不同用户绘制草图的核心部分着重提取出来，将无关的、无意识的抖动造成的笔画毛刺、聚点、闭合处的间隙等从原始草图中剔除，减少噪声对于草图

在后续的拟合和识别过程的影响，确保后续工作中笔画识别的效率和准确性。因此，草图预处理过程在草图识别过程中起着重要的作用。

草图预处理过程主要包括冗余点的消除、聚点的消除、闭合检查和笔画拟合等。针对不同的情况采用不同的预处理方法，详细的处理方法及具体步骤在 3.1.1 节、3.1.2 节、3.1.3 节 3.1.4 节阐述。

### 3.1.1 冗余笔画删除

冗余的笔画主要是用户绘制过程中由于绘制习惯或者无意中勾画的与其表达意图无关的笔画数据，这种笔画一般具有的特点是长度较小、常出现在笔画的端点处、笔画较轻或者游离于草图整体之外。由于是用户无意中绘制的，所以往往会出现一些曲率较大的采样点的前后，对后面的特征值计算造成比较大的影响，因此需要将其去除。本文采用判断曲率值大小的方法对冗余笔画进行处理，具体如下：

假定笔画  $S$  经过数据采样后可表示为  $S = \{p_i | 0 \leq i \leq n-1\}$ ，设定曲率阈值  $\varphi_0 = 0.85$ 。

(1) 判定笔画的头尾。根据笔画的采样数据计算  $S$  最小外包矩形的对角线长度  $RLength$ 。距离笔画开始或者结尾长度小于  $RLength \times 0.1$  的点则为笔画的头或尾。

(2) 计算笔画头尾之间曲率最大的点。如果曲率大于设定阈值，则从该点出去除头或尾。

冗余笔画删除后的效果如图 3.2 所示。

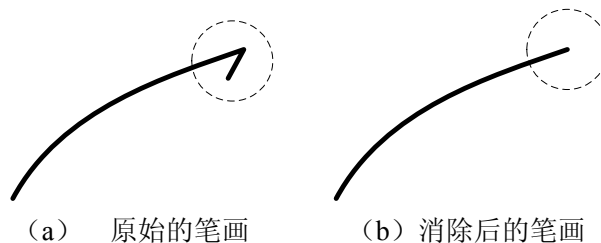


图 3.2 冗余笔画消除示意图

### 3.1.2 闭合检查

用户在输入一个闭合的曲线的时候，用户由于绘制时具有一定的随意性，所以很难做到笔画的起点和终点刚好重合。一般情况下往往会产生两种情况：

(1) 如图 3.3 (a) 所示，笔画交叉之后仍有比较小的延伸。

(2) 如图 3.3 (b) 中所示，起点和终点没有重合，两端点间存在一个间隙。

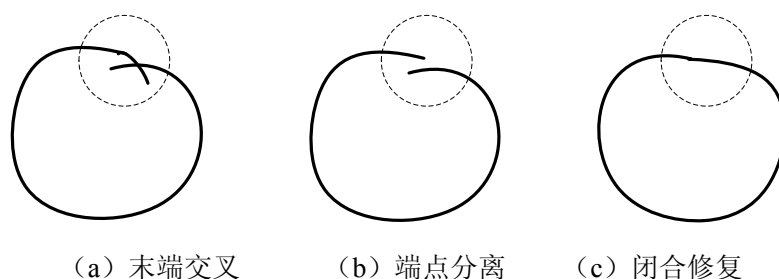


图 3.3 闭合检查与修复示意图

基于上面出现的两种情况，通过如下方式对笔画的闭合性做出检查：通过计算笔画采样点的偏移角度进行累加，当累加值在笔画的中点处达到  $2\pi \pm k$  ( $k$  为设定的阈值) 时，可以认为用户的意图是想要绘制一个闭合的图形。

对于情况 (1) 的判定及处理方法：当累加的转角值小于  $2\pi$  时，则表示为分离情况，那么闭合点取起点和终点的中点。

对于情况 (2) 的判断及处理方法：当累积的转角值大于  $2\pi$  时，则表示为交叉情况，那么闭合点为交点，将以闭合点为端点的小笔画端删除。

### 3.1.3 聚点消除

在实际绘图过程中，用户由于需要进行设计思考等原因导致笔速变缓或者笔画的抖动，会出现笔画聚集在一个较小区域内的现象，一般称之为聚点。聚点的存在既对描述笔画形状特征的作用甚微，也耗费了系统处理笔画数据时的计算时间。因此，在保持笔画形状特征的基础上，应去除聚点以提高系统处理效率。根据聚点产生的原因，多数出现在笔画的拐角处。在本文中通过滑动窗口中是否有重复采样点的方面对聚点进行判断。设定窗口大小为 8 个采样点，将采集到的笔画数据依次进入窗口中，判断窗口内是否有重复的采样点，当窗口内有重复的采样点时，则表明存在聚点，并予以消除。聚点消除后的效果如图 3.4 (b) 所示。

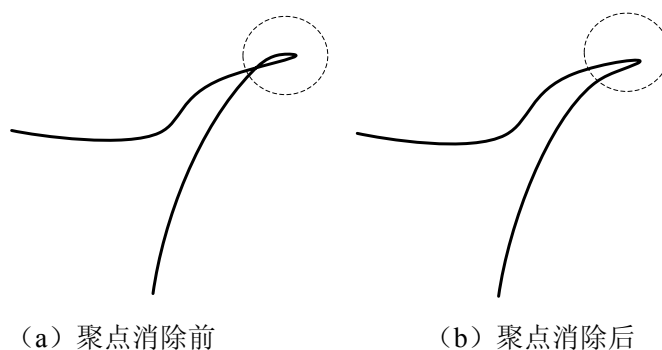


图 3.4 聚点消除示意图

### 3.1.4 笔画的拟合

笔画的拟合主要分为两种：

(1) 直线的拟合。直线的拟合过程主要是一个线性化的判断过程，通过将一个笔画的首尾距离与采样点的累计弦长的比值与设定的一个阈值相比较，用公式表示为：

$$\text{笔画首尾距离 } d \text{ 为: } d = |p_n - p_0| = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2}$$

$$\text{笔画采样点的累积弦长为: } l = \sum |p_i - p_{i-1}| = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$$

$$\text{两者之间的比值为: } \varepsilon = \frac{d}{l} (0 \leq \varepsilon \leq 1)$$

由  $\varepsilon$  值可知，当  $\varepsilon$  越接近于 1 时，说明笔画首尾的长度越大，也就说明笔画越趋近于直线；反之，则表示笔画是圆弧的可能性更大。因此，可以通过比较  $\varepsilon$  的值与设定的阈值之间的大小对笔画类型判断。拟合后的效果如图 3.5 所示。

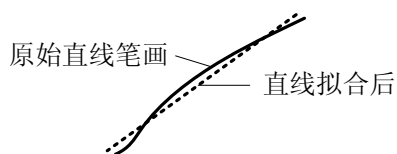


图 3.5 直线拟合示意图

(2) 曲线的拟合。曲线包括很多种情况，比如有圆、圆弧、椭圆、椭圆弧线及其他二次曲线等。可以利用二次曲线的不变量对曲线的类型进行判断，二次曲线的不变量在表示曲线性质时不会受到曲线所处坐标系的影响。根据不变量的理论，可以采用如下改进最小二乘法对曲线进行拟合：

假设曲线由  $n$  个采样点构成，即  $S = \{p_i, 1 \leq i \leq n\}$ 。设定代数距离为  $Q(x, y)$ ，则曲线的方程可以表示为：

$$Q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (3-1)$$

其中，为了避免  $a, b, c, d, e, f$  全为 0 的情况，使分  $f \neq 1$ 。则拟合后的误差为：

$$\varepsilon = \sum_{i=0}^n Q(x_i, y_i)^2 \quad (3-2)$$

为使  $\varepsilon$  的值最小，则  $\varepsilon$  对  $a, b, c, d, e$  分别进行一阶求导，并使所有的一阶导数为 0。由求导过程可以构建方程为：



$$\frac{\partial \varepsilon}{\partial a} = \sum_{i=0}^n 2x_i^2(ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + 1) = 0 \quad (3-3)$$

$$\frac{\partial \varepsilon}{\partial b} = \sum_{i=0}^n 2x_iy_i(ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + 1) = 0 \quad (3-4)$$

由构建的方程便可以计算出  $a$ 、 $b$ 、 $c$ 、 $d$ 、 $e$  的值。最后得到拟合后的曲线方程为：

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (3-5)$$

通过最小二乘法对曲线进行拟合后的效果如图 3.6 所示。

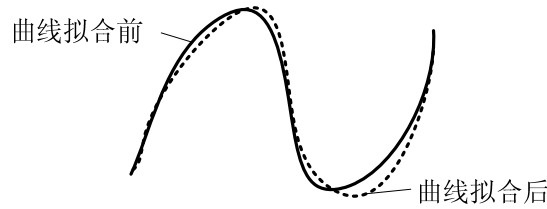


图 3.6 曲线拟合示意图

### 3.2 草图轮廓拟合

轮廓是描述三维模型一个侧面的重要信息<sup>[44]</sup>，能从视觉上对物体的类别形状进行描述，也符合人类的视觉认知习惯。因此，在三维模型的检索系统中，用户勾勒出的轮廓特征是进行三维模型检索的重要依据。但是，轮廓信息只是在二维形式上对三维物体一个视角上的描述，单一的轮廓并不能准确全面的刻画三维模型的全部特征信息，三维模型从正面、侧面等各个角度得到的轮廓图形在视觉上有可能会有很大的不同，因此，单一的二维轮廓图像在进行三维模型的检索时并不能取得良好的效果，可以通过用户绘制多个视角的轮廓草图综合起来对一个三维模型进行判断，多个视角的草图能够比较全面的描述一个具体的三维模型。

在获取用户绘制的草图轮廓信息后，并不是直接将轮廓图像与三维模型的投影图像进行相似性的判断，这种做法不仅无法取得良好的检索效果，同样也会因为大量的计算使得系统的检索效率降低。而一种替代的方法则是从外围轮廓中提取出能对外围轮廓进行描述的显著的某些特征信息，然后利用这些特征信息构建描述草图轮廓特征信息的特征向量。利用从外围轮廓中提取的特征向量与建立的三维模型特征向量索引之间进行相似性的判断，往往会节省大量不必要的运算。

提取草图外围轮廓特征的步骤主要分为以下几个部分：（1）根据 3.1 节中的草图预处理方

法对草图进行去噪、规整等处理；（2）利用轮廓跟踪算法提取出草图的外围轮廓，并采用基于特征点的多边形拟合算法对草图轮廓进行拟合；（3）根据拟合后的多边形构建几何不变量，并以此作为检索时输入的条件。

### 3.2.1 轮廓去噪处理

由于用户绘制的外围轮廓特征往往会带有不同程度的噪声，出现轮廓不够光滑、线条抖动比较大等情况。因此，在进行外围轮廓的提取之前需要对轮廓草图进行去噪处理。目前常利用卷积积分去除草图轮廓存在的噪声。表示草图轮廓  $R = (x(t), y(t))$ ， $t$  为任意的参数

通过  $x(t)$ 、 $y(t)$  与高斯函数作卷积可得到：

$$\begin{aligned} R_0 &= (x(t, \varepsilon), y(t, \varepsilon)) \\ x(t, \varepsilon) &= x(t) \bullet y(t, \varepsilon) \\ y(t, \varepsilon) &= y(t) \bullet x(t, \varepsilon) \end{aligned} \quad (3-6)$$

那么草图轮廓轮廓曲线噪声消除后的曲率可表示为：

$$K(t, \varepsilon) = \frac{x_t(t, \varepsilon) \bullet y_{tt}(t, \varepsilon) - x_{tt}(t, \varepsilon) \bullet y_t(t, \varepsilon)}{(x_t(t, \varepsilon)^2 + y_t(t, \varepsilon)^2)^{\frac{1}{2}}} \quad (3-7)$$

当标准差  $\varepsilon$  取一个最优值时，则能够在保留较小尺寸笔画的同时将噪声去除。经过处理后的轮廓线比较平滑，更有利于轮廓的拟合和特征向量提取过程。

### 3.2.2 基于特征点的多边形拟合

由于数字图像的边界信息都具有一定的冗余度，给处理过程和特征提取带来了诸多困难。利用多边形对边界进行拟合是解决这一问题常见而有效的方法。通过对边界进行拟合，图像的边界信息由与其近似的多边形表示，将对图像边界信息的特征提取转换为对拟合后多边形的特征提取。由于多边形的处理方法更加的简便和有效，因此通过转换可以将多边形处理方面成熟的技术引入并加以运用。

对于一个曲线的拟合过程，最常用的方法是逐段线性近似方法<sup>[45]</sup>。设定某一曲线  $L$  由点集  $A = \{A_0, A_1, \dots, A_n\}$  构成， $A_0$  和  $A_n$  分别代表曲线的起点和终点，那么连接  $A_0 A_n$  的直线方程可以表示为：

$$Ax + By + C = 0 \quad (3-8)$$

用  $A_0 A_n$  直线近似的替代曲线  $L$  存在一定的误差，而衡量两者之间近似程度的指标一般采用均方误差或者最大误差。误差值越小，则表明两者之间的近似程度越高，对于曲线的拟合效果也更能使人满意。对于误差的计算可以表示为：

$$d_i = \frac{|Ax_i + By_i + C|}{\sqrt{A^2 + B^2}} \quad (i = 0, 1, \dots, n-1) \quad (3-9)$$

$$E = \sum_{i=0}^{n-1} d_i^2 \quad (3-10)$$

上述方法在拟合过程中将曲线上所有的采样点到近似直线的距离都计算在内，当曲线的采样点数较少时算法的运算量比较小，能够表现出比较好的拟合效果和执行效率。但是，当采样的点集中数目变多时计算的效率就会随之下降。另外，在拟合的过程中，对于曲率较小的局部曲线段完全可以直接用比较长的线段进行代替而不需要对这一小段曲线内的所有点都进行计算。而对于曲率较大的点附近，由于体现了轮廓的特征信息，则应该在拟合过程中予以保留，所以需要使用比较多较短的线段对其进行拟合。如图 3.7 所示。

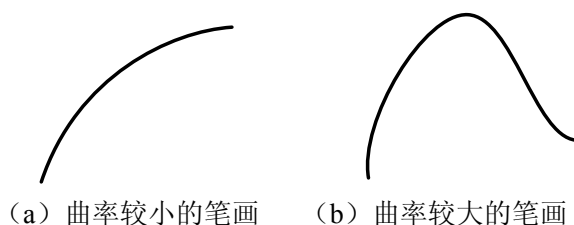


图 3.7 不同曲率笔画示意图

在对笔画 (a) 进行拟合时，由于笔画 (a) 的曲率较小，对于轮廓特征的提取贡献并不大，因此不需要对其上所有的点进行计算，用长度较大的线段对其进行拟合就能达到比较好的拟合效果。而对于笔画 (b) 的拟合，由于其笔画曲率较大，有可能是轮廓中比较重要的特征点，具有较强的描述能力。因此，在拟合过程中应尽量保留这些特征。但是传统的拟合算法并没有将上述两种情况区分对待，只是简单的对笔画 (a) 和笔画 (b) 进行了同样的处理，在拟合过程用同样长度的线段进行了拟合。传统多边形拟合算法如图 3.8 所示。

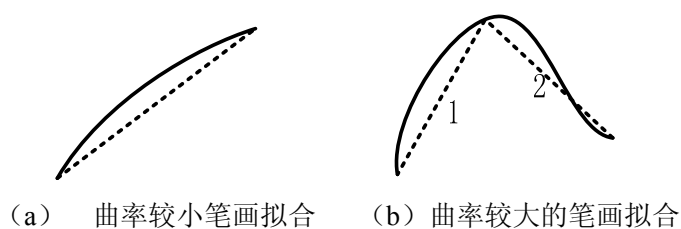


图 3.8 传统拟合算法的拟合效果

基于上述传统算法拟合过程中的不足，在本节中设计了基于特征点的多边形拟合算法。由于草图轮廓上曲率较大的点一般被当作曲线的特征点，在轮廓特征提取中起到重要的作用。因此，在特征点处的拟合需要更加的细致，尽可能的保留其特征以便于特征向量的提取。换句话说就是对于特征点附近的部分曲线需要更多的线段进行拟合。本节提出的算法通过在拟合开始

时设定一个初始的拟合线段长度，在拟合过程中，根据草图轮廓曲率的不断改变而动态的调整拟合所使用的线段长度。通过这种方式实现了在拟合过程中对不同曲率的曲线进行不同细致程度的拟合，有效的减少了拟合过程中的多余的计算。拟合算法的流程如图 3.9 所示。

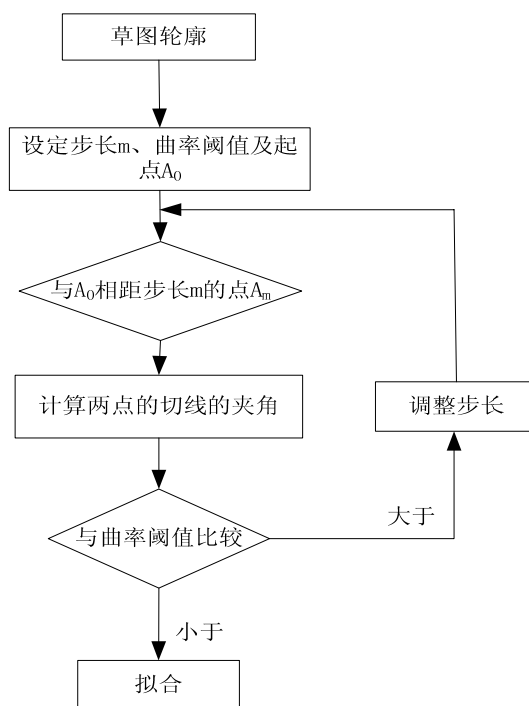


图 3.9 基于特征点的多边形拟合过程

具体的拟合过程描述如下：

假定轮廓  $L$  的点集  $A = \{A_0, A_1, \dots, A_n\}$ ，给定初始步长为  $m$ ，角度阈值为  $\theta_0$ 。则：

(1) 从轮廓的起点  $A_0$ （最左边和最下边的第一个点）出发，按逆时针方向选取与  $A_0$  间隔为  $m$  的点  $A_m$  开始计算。

(2) 分别计算过  $A_0$ 、 $A_m$  点与曲线相切的直线方程，然后计算两个方程之间的夹角值  $\theta$ ，

当  $\theta$  值小于设定的角度阈值  $\theta_0$  时，则用直线  $A_0 A_m$  代替由  $\{A_0, A_1, \dots, A_m\}$  构成的曲线段。当  $\theta$  大

于设定的阈值  $\theta_0$  时，则将重新设定步长  $m_i = \frac{m \times \theta_0}{\theta}$ （步长取整，当  $m_i = \frac{m \times \theta_0}{\theta} < 1$  时，取值为 1），转入 (3)

(3) 仍以  $A_0$  作为起点，选取与  $A_0$  间隔  $m_i$  的点  $A_{m_i}$ ，进行步骤 (2)。

(4) 直至点  $A_n$ ，拟合结束。

采用上述算法对笔画进行拟合，拟合效果如图 3.10 所示。

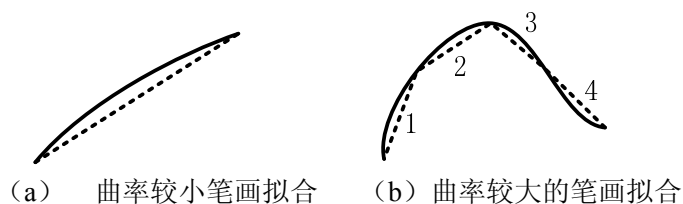


图 3.10 基于特征点的拟合算法拟合结果

从图中可以看出，传统拟合算法对于曲线段的拟合都是特定长度的线段，并没用在曲线特征点处进行特别的处理，在拟合过程中没有对特征点进行有效的保护。而基于特征点的多边形拟合算法对于曲率较大的笔画附近的拟合非常精细，拟合所使用的直线段也比较多，随着曲率值的变化，近似线段的长度也不断的变化。如图 3.10 (b) 所示，传统的拟合算法 2 条线段进行拟合，而基于特征点的多边形拟合则用了 4 条线段对该处进行拟合。充分的保证了轮廓的特征点在拟合过程中不会消失，使后续的轮廓特征提取更加的准确和完整。

### 3.3 草图特征提取

经过拟合可以得到近似表示草图轮廓的多边形，可以利用几何中相关知识对多边形处理，提取多边形的特征信息作为草图轮廓的特征信息，以便用于三维模型检索过程。

多边形的特征主要包括以下几个方面<sup>[46-48]</sup>：

(1) 多边形的顶点个数。拟合后的多边形顶点个数越多表明草图的轮廓越复杂，可以利用两个多边形顶点数的差值判断它们之间的相似程度。

(2) 多边形的内角的标准方差。内角标准方差能够从总体上对多边形的内角分布情况进行描述，其值得大小随着多边形的规则程度而改变，多边形越复杂，则其值越大。

(3) 多边形的内角直方图。将  $(0, 2\pi)$  等角度或者不等角度划分为  $k$  个区间，再统计多边形的内角在每个区间上的数目，最后得到内角直方图

(4) 紧凑度。紧凑度表示轮廓与圆的相似程度，紧凑度的值越大是，说明形状越接近于圆形。用公式表示为：

$$\theta = 4\pi S / L^2 \quad (3-11)$$

其中， $L$  为多边形的边长， $S$  为多边形面积。

在本节中，通过 3.2.2 中对草图轮廓拟合后的多边形，充分利用多边形的几何信息，构建用于检索的特征向量，采用距离和角度特征构成二维集合的特征不变量，构建的几何特征不变量不会随着草图轮廓的平移、缩放和旋转的变化而变化，这也满足对于描述子的基本要求。

设定  $P_1(x_1, y_1)$ 、 $P_2(x_2, y_2)$ 、 $P_3(x_3, y_3)$  分别为拟合后的多边形上的任意相邻的三个顶点，如图 3.11 所示。

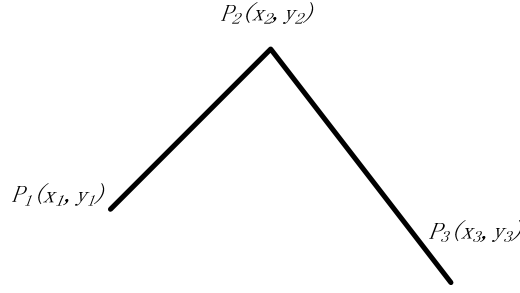


图 3.11 多边形上相邻的任意三个顶点

由  $P_1(x_1, y_1)$ 、 $P_2(x_2, y_2)$ 、 $P_3(x_3, y_3)$  三个点可以构建出两个向量和两向量的模分别为：

$$\overrightarrow{P_2P_1} = (x_1 - x_2, y_1 - y_2) \quad (3-12)$$

$$\overrightarrow{P_2P_3} = (x_3 - x_2, y_3 - y_2) \quad (3-13)$$

$$|\overrightarrow{P_2P_1}| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3-14)$$

$$|\overrightarrow{P_2P_3}| = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} \quad (3-15)$$

由于草图轮廓上相邻点的距离并不会随着草图轮廓的平移和旋转的改变而发生变化。但是，草图轮廓的缩放会直接导致两点间距离的变化。因此，采用两点间距离时要使用相对长度，而不能使用绝对的长度。为了反映轮廓点分布的距离关系，依据轮廓跟踪方向对特征量的影响，定义轮廓点  $P_2$  的距离特征量为：

$$Dis(P_2) = \frac{\max\{|\overrightarrow{P_2P_1}|, |\overrightarrow{P_2P_3}|\}}{\min\{|\overrightarrow{P_2P_1}|, |\overrightarrow{P_2P_3}|\}} \quad (3-16)$$

根据两向量的点积运算公式，向量之间的夹角可以表示为：

$$\cos \alpha = \frac{\overrightarrow{P_2P_1} \cdot \overrightarrow{P_2P_3}}{|\overrightarrow{P_2P_1}| |\overrightarrow{P_2P_3}|} \quad (3-17)$$

构建的夹角由于不会随着草图轮廓的平移、旋转和缩放的变化而改变，因此，定义轮廓点的角度特征量为：

$$\alpha(P_2) = \arccos \frac{(x_1 - x_2)(x_3 - x_2) + (y_1 - y_2)(y_3 - y_2)}{|\overrightarrow{P_2P_1}| |\overrightarrow{P_2P_3}|} \quad (3-18)$$

通过对距离和角度特征量的构建，便得到了描述草图轮廓的几何特征不变量，由不变量也就构成草图的特征向量。在相似性计算时，只要通过特征向量的比较便能从数据库中检索出用户需要的三维模型。

### 3.4 本章小结

本章主要结合在进行检索前对用户绘制的草图需要进行的各项处理过程，详细的阐述了各个过程实现的相应技术。用户绘制的表达个人检索意图的草图，需要对草图进行预处理，去除噪声对后续处理过程的干扰，然后采用改进的轮廓跟踪和拟合算法对草图轮廓进行提取。最后获取草图中用于模型检索的特征信息。

## 第四章 三维模型特征提取及检索

在三维模型检索系统中对三维模型的描述通常使用由其构造出描述自身特征的特征向量。目前提取模型特征向量的方法有很多种，但是针对不同的相似性匹配和系统检索算法而采用的提取方法也各不相同。在本文中，用户通过绘制三维模型的草图轮廓表达个人的检索意图，但由于用户绘制的草图轮廓特征存在于二维空间，无法与直接提取的三维空间内的描述符进行比较。因此，通过对三维模型进行投影，获取三维模型在二维空间内的投影图像，提取三维模型投影图像的二维特征，将二维空间的描述信息与处于三维空间的模型之间建立映射关系。在这种转换的基础上，通过三维模型投影轮廓特征与草图轮廓特征之间的相似性计算实现三维模型的检索。

由于模型的低层特征与高层语义之间的差异、特征描述符与模型本身的差异、相似性度量与人类视觉感知系统之间的差异，经常导致三维模型检索系统无法取得令用户满意的效果。因此，通过相关反馈技术，利用检索过程中的人机交互机制，将用户的主观评价信息融合到模型检索中，不断的调整相关参数以优化系统检索效果

本章给出了提取三维模型特征的相关处理技术的改进及实现细节，包括三维模型的预处理、投影平面获取、投影图像获取、投影图像的边缘检测及轮廓跟踪、图像特征提取等。

### 4.1 三维模型预处理

对于三维模型的检索，系统的准确性是评价系统性能的关键指标之一，而三维模型检索系统的检索准确性不仅取决于计算三维模型之间相似性的算法，还取决于的预处理技术。由于三维模型的来源不同、产生的技术手段不同，三维模型之间在尺度大小、坐标位置和方向也不尽相同。在检索过程中，提取的三维模型特征描述符需要满足在进行相似性计算时不会受到模型的平移、旋转和缩放影响的条件。因此，需要在进行三维模型特征提取之前进行预处理，将三维模型通过预处理变换到同一的坐标尺度下，具有相同的大小、位置和方向。这样保证了在统一的规范下提取的特征向量不依赖于任何的坐标系。

完整的三维模型的统一过程主要包括三个方面的标准化<sup>[49]</sup>：（1）坐标原点的对其预处理，保证了三维模型的评议不变性；（2）坐标方向对齐预处理，保证了模型的旋转不变性；（3）坐标尺度对齐预处理，保证了模型的缩放不变性。将三个过程统一起来可用公式表示为如下：

$$\tau(I) = s^{-1} \cdot F \cdot R \cdot (I - c) \quad (4-1)$$

其中， $s$  为缩放系数， $F$  为对角矩阵的翻转矩阵， $R$  为对应于 PCA 变换的旋转矩阵， $I$  为原



始模型坐标， $c$  为坐标系原点。

#### 4.1.1 平移不变性预处理

将不同的三维模型统一在同一坐标原点下，最基本的思想是将三维模型的中心平移到坐标原点上。由于该方法需要对三维模型进行顶点的坐标采样，所以其缺陷在于对模型的采样密度不均时方法的鲁棒性比较差。为了提高算法的鲁棒性，减小三维模型顶点的采样密度，可以采用 Vranic 等在文献[50]中提出的基于加权面积的方法提高重心平移的鲁棒性，该方法可由公式表示为：

$$\bar{P} = \sum_i P_i \cdot S_i \quad S_i = \frac{1}{3} \sum_j S_j \quad (4-2)$$

其中， $\bar{P}$  为三维模型的重心坐标点， $S_i$  为每个顶点的相邻面片平均面积， $S_j$  为每个顶点各相邻面片的面积。

#### 4.1.2 旋转不变性预处理

对于模型之间的坐标系不统一的问题，可以在对齐坐标原点之后，对三个坐标轴的方向进行对齐，得到保证三维模型具有旋转不变性的通用坐标系。大多数文献采用了主成分分析法<sup>[51]</sup>完成预处理过程。主成分分析法是普遍用于信号处理、数据统一分析的信号分析方法，其原理是通过对随机变量进行统计分析，构造用于计算特征值和特征向量的协方差矩阵，然后获得按特征值排序的特征向量构成的正交基，完成原始数据到主分量空间的映射，按照主轴的标准化将模型统一。用公式表示为：

定义具有  $N$  个点的三维模型点集合  $P$ ， $P_i$  表示处于三维空间坐标系中的一个点，那么点集合  $P$  均值的计算公式则可以表示为：

$$\mu_p = E(P) = \frac{1}{N} = \sum_{i=1}^N P_i \quad (4-3)$$

由点集合  $P$  的均值可以计算  $P$  的协方差为：

$$\sigma_p = \frac{1}{N} = \sum_{i=1}^N (p_i - \mu_p)(p_i - \mu_p) \quad (4-4)$$

其中， $\sigma_p$  是一个对称矩阵，存在 3 个特征值： $\lambda_1 > \lambda_2 > \lambda_3$ ，三个特征值依次对应的单位特征向量为： $e_1, e_2, e_3, (|e|=1)$ ，三者之间满足：

$$C_r e_i = \lambda_i e_i \quad (4-5)$$

由上式便可得到旋转矩阵为：

$$R = \{e_1, e_2, e_3\}^T \quad (4-6)$$

但是，实验证明 PCA 方法具有一定的缺陷，因此可以采用最大化法线分布和最大面积法完成坐标系的归一化，该方法避免了由于顶点采样不均带来的误差。其基本的思想是计算网格模型中每个面片的法线或者面积，然后将法线或者面积在不同的方向上进行投影，再从中选取出具有投影发现或者投影面积最大的方向并以此作为一个坐标轴的方向。在通过对其余的法线或者面积计算后得到第二个坐标轴的方向。当两个坐标轴确定之后，则可以通过右手规则去的第三个坐标轴方向。

#### 4.1.3 缩放不变性预处理

对三维模型实现旋转不变性的预处理，则需要寻求一个变换因子实现三维模型在三个坐标轴上的缩放变换，设定三个方向的变换因子分别为  $s_x, s_y, s_z$ ，则三维模型缩放变换因子可表示为：

$$S = \sqrt{\frac{(s_x^2 + s_y^2 + s_z^2)}{3}} \quad (4-7)$$

$$\text{其中, } s_x = \frac{\sum_{i=1}^K |p_{ix}| w_i}{S}, \quad s_y = \frac{\sum_{i=1}^K |p_{iy}| w_i}{S}, \quad s_z = \frac{\sum_{i=1}^K |p_{iz}| w_i}{S}, \quad p \text{ 为三维模型上任意一点,}$$

$w_i$  为三角面片的面积。

通过上面的计算，保证了三维模型的缩放不变性，得到了规范化的三维模型，有利于后续的处理过程。

### 4.2 三维模型投影图像获取

#### 4.2.1 基本思路

目前基于二维图像的检索算法众多，无论在检索效果还是检索性能上，都要比三维模型的检索技术更加成熟。主要归结于两个方面原因：（1）由于三维模型本身固有的特性，使三维模型的特征提取方法相对较少且都提取过程具有较高的复杂度和较大的计算量；（2）提取出的三维模型特征并不能完整的刻画出三维模型的全部特征，往往具有相似特征向量的三维模型之间在视觉上却迥然不同。因此，利用二维图像检索的相关技术并加以改进引用到三维模型的检索领域是一个便捷性的方法。而三维模型的投影轮廓从视觉上能够具有较强的描述模型特征的能力，可以作为利用二维图像检索技术实现三维模型检索的中间桥梁。获取三维模型的二维轮廓

图像的最简便的方法便是对三维模型进行投影，在本文中，采用一种改进的基于主平面分析的三维模型投影方法，在原有单个投影主平面的基础上，通过计算获取另外两个投影主平面，得到三维模型的在主平面上的三个投影图像，然后提取出投影图像的轮廓信息，从二维图像轮廓信息中提取具有描述三维模型能力的投影图像特征，将用户手绘的草图轮廓与系统建立的三维模型的投影轮廓之间进行特征向量相似性计算，最后将得到的结果显示给用户。

目前，获取三维模型投影的方法主要有两种：平行投影变化（正交投影）和中心投影变换<sup>[52]</sup>（透视投影）。如图 4.1 所示，两者之间的区别在于正交投影采用的投影射线之间是相互平行的，而透视投影变换采用的投影射线则是来源于选取的一个公共点，射线之间以锥体的方式投向三维模型。

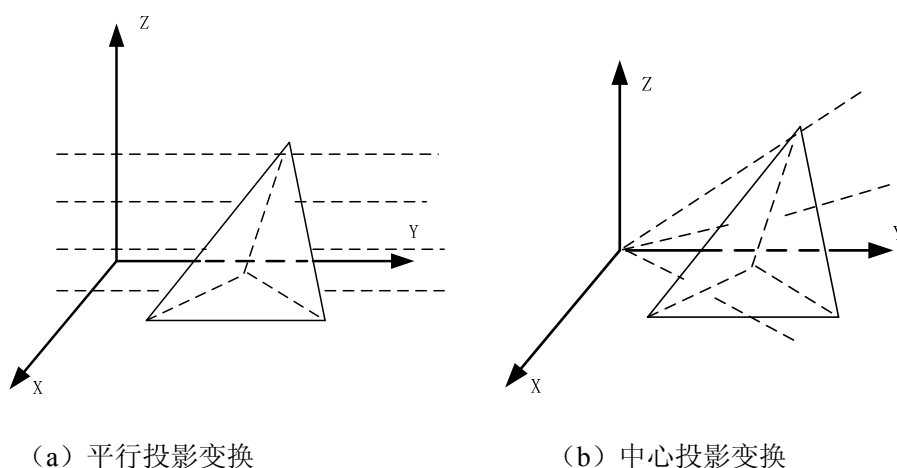


图 4.1 三维模型投影变换

本文中利用了三维模型的平行投影变换，并采用了一种基于主平面分析的三维模型投影方法<sup>[53]</sup>，该方法避免了传统的基于二维投影的三维模型检索方法<sup>[54]</sup>的缺点。基于主平面分析的投影方法，根据三维模型本身的特性选取最佳的投影平面，减少了投影平面选择的随机性，获得的投影图像能更好的反应出三维模型的性质，有利于投影图像特征的提取，在描述三维模型上具有良好的效果。

基于主平面分析的三维模型投影过程主要分为三个步骤：

- (1) 根据三维模型的性质选取最佳的投影平面。
- (2) 将三维模型投影到选取的投影平面上获得蕴含三维模型特征的二维图像。

(3) 根据得到的投影图像进行平移、缩放等不变性处理后提取出其中的特征信息，然后计算草图的特征信息与三维模型投影图像的特征信息之间的相似程度，最后得出检索结果。

#### 4.2.2 投影平面选取

三维模型的主平面指的是获取三维模型空间中的一个平面，而这个平面需要满足三维模型

上所有的点到这个主平面的距离和的值最小的条件，三维模型的重心也正好在获取的主平面上面。由于单个的主平面上得到的投影图像描述能力有限，并不能完整的刻画三维模型的特征。因此，在得到第一个主平面基础上通过计算引入了另外两个主平面，在三个正交的主平面上进行三维模型投影，得到用于特征提取的三幅投影图像。

根据主平面具有上述的性质，对于第一个主平面  $E_1$ ，因为三维模型的重心在这个平面上，因此，平面的方程可以表示为：

$$A_1x + B_1y + C_1z = 0 \quad (4-8)$$

定义  $P(x, y, z)$  是三维模型上任意的一点，则  $P(x, y, z)$  到平面  $E_1$  的距离  $d$  可以表示为：

$$d = |A_1x + B_1y + C_1z| \quad (4-9)$$

其中  $z = (A, B, C)$  为选取的平面  $E_1$  的单位法向量，那么三维模型上所有的点到平面  $E_1$  的距离和的最小值可以表示为：

$$A_1^2 + B_1^2 + C_1^2 = 1 \quad (4-10)$$

$$E_1 = \min \sum_{i=1}^n |A_1x_i + B_1y_i + C_1z_i|^2 \quad (4-11)$$

只需要求得式 4-11 的最小值，然后联合式 4-9、式 4-10 便可以计算出  $A$ 、 $B$ 、 $C$  的具体值。对式 4-12 进行微分可得：

$$\begin{aligned} 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n x(A_1x + B_1y + C_1z) &= 0 \\ 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n y(A_1x + B_1y + C_1z) &= 0 \\ 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n z(A_1x + B_1y + C_1z) &= 0 \end{aligned} \quad (4-12)$$

则可以得到第一个主平面  $E_1$  的单位法向量：

$$(A_1, B_1, C_1) = \left( \frac{k_1}{\sqrt{1+k_1^2+h_1^2}}, \frac{1}{\sqrt{1+k_1^2+h_1^2}}, \frac{h_1}{\sqrt{1+k_1^2+h_1^2}} \right) \quad (4-13)$$

由上式可得到选取的第一个主平面的单位法向量，即可得到平面  $E_1$ 。

根据上面所得的第一主平面  $E_1$  的计算过程，对三维模型进行旋转后使得  $E_1$  与坐标系的  $XOY$  平面重合，然后结合  $Z$  轴就可以计算出一个过  $Z$  轴的第二主平面  $E_2$ 。首先根据第一主平面的法向量  $(A, B, C)$  与  $Z$  轴的单位向量  $(0,1,1)$  求得三维模型的旋转角度  $\theta$ ，由两者之间的向量积

便能够计算出三维模型进行旋转时的旋转轴为  $(p_1, p_2, p_3)$ ，那么三维模型进行旋转时所使用的矩阵  $R$  则可以表示为：

$$R = \begin{bmatrix} \cos \theta + \lambda p_1^2 & \lambda p_2 p_1 - p_3 \sin \theta & \lambda p_3 p_1 - p_2 \sin \theta \\ \lambda p_2 p_1 - p_3 \sin \theta & \cos \theta + \lambda p_2^2 & \lambda p_3 p_2 - p_1 \sin \theta \\ \lambda p_3 p_1 - p_2 \sin \theta & \lambda p_2 p_3 - p_1 \sin \theta & \cos \theta + \lambda p_3^2 \end{bmatrix} \quad (4-14)$$

其中， $\lambda = 1 - \cos \theta$ 。

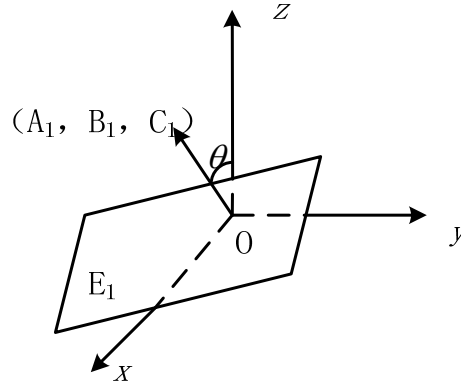


图 4.2 主平面  $E_1$  的旋转

同样的，三维模型上任意点到平面  $E_2$  的距离和的最小值表示为：

$$E_2 = \min \sum_{i=1}^n (A_2 x + B_2 y)^2 \quad (4-15)$$

其中， $A_2$  和  $B_2$  分别是平面  $E_2$  的  $x, y, z$  的系数， $(x', y', z')$  是第一主平面与  $XOY$  平面重合后点  $(x, y, z)$  的坐标。由上式进行微分可得：

$$\begin{aligned} 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n x' (A_2 x' + B_2 y') &= 0 \\ 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n y' (A_2 x' + B_2 y') &= 0 \end{aligned} \quad (4-16)$$

那么根据式 4-16 并联合可以得到平面  $E_2$  的法向量为

$$(A_2, B_2, 0) = \left( \frac{k}{\sqrt{1+k^2}}, \frac{1}{\sqrt{1+k^2}}, 0 \right) \quad (4-17)$$

根据已经计算出的两个主平面，结合主平面之间两两相交的原则，那么可以容易的得到第三个主平面的法向量：

$$(A_3, B_3, C_3) = \left( \frac{-h_1}{\sqrt{1+k_1^2+h_1^2}} \cdot \frac{1}{\sqrt{1+k_2^2}}, \frac{h_1}{\sqrt{1+k_1^2+h_1^2}} \cdot \frac{k_2}{\sqrt{1+k_2^2}}, \right. \\ \left. \frac{k_1}{\sqrt{1+k_1^2+h_1^2}} \cdot \frac{k_2}{\sqrt{1+k_2^2}} - \frac{1}{\sqrt{1+k_1^2+h_1^2}} \cdot \frac{k_2}{\sqrt{1+k_2^2}} \right) \quad (4-18)$$

通过上面的计算，得到了进行三维模型投影的三个主平面法向量，而三个正交的主平面的方程也随之确定。由整个计算过程可以得知，主平面并不依赖于选取的坐标系，而是与三维模型自身的固有特性密切相关的。

#### 4.2.3 投影图像处理及特征提取

根据 4.2.2 中得到的三个主平面，三个主平面之间的关系如图 4.3 所示。

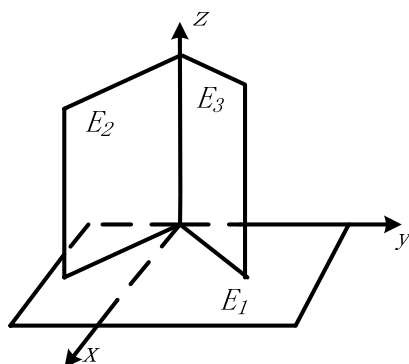


图 4.3 三个主平面的空间关系

将三维模型在三个主平面上进行正交投影，得到三张投影图像，如图 4.4 所示。

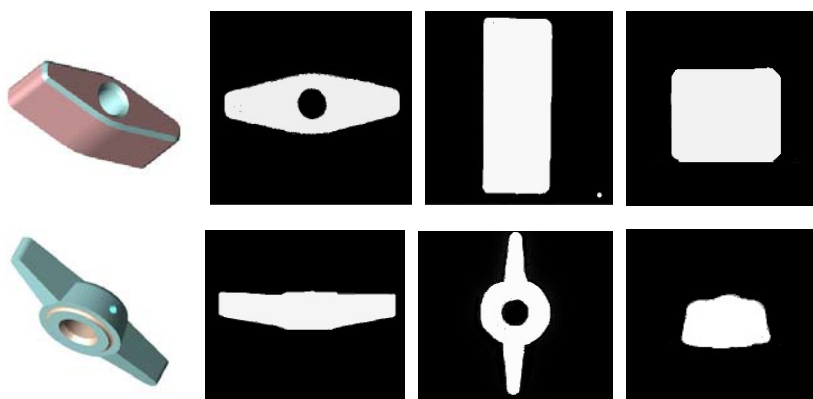


图 4.4 三维模型的投影图像

从图 4.4 可以看出，选择根据三维模型本身的性质选取投影主平面，获取的投影图像能够对三维模型的特征进行全面的描述。每个投影图像都能从某一个视觉角度对三维模型进行描述，而主平面的选取规则也同样符合人类的视觉习惯。因此，从投影图像中提取的特征向量更有利

于三维模型的检索。

#### 4.2.4 改进的 SUSAN 算法提取图像边缘

获取模型的投影图像后，需要对投影图像进行边缘检测，对图像进行边缘检测是得到图像形状特征的重要基础。边缘检测的方法主要依赖于图像的亮度或者是图像亮度的导数是否连续进行计算。根据图像亮度的变化规则不同，其不连续性则可以分为以下两种：

(1) 阶跃的不连续性。主要是指分布在非连续处两侧的图像像素灰度有比较大的差异。

(2) 线条的不连续性。主要是指图像的亮度从一个值到另外一个值具有比较剧烈的变换，但是这个变化制造一个相对较小的范围之内。

而图像的边缘并不是只是具有单一的上述类型，往往同时具有阶跃的不连续和线条的不连续性两种边缘特性。根据图像的边缘特征对图像的边缘提取时一般可以通过以下步骤：滤波、增强、提取、定位四个步骤。前两个步骤主要是对得到的图像进行预处理，通过滤波器可以改善边缘与噪声之间的交叉性，而增强步骤则可以突出像素值具有明显变化的邻域内的某个点。检测则是利用图像中某些点像素值的变化幅度较大，采用梯度的最大值或者根据二阶导数为零的点提取图像的边缘信息。目前常用于边缘检测的方法主要有基于灰度直方图的边缘检测、一阶导数边缘检测、二阶导数边缘检测、SUSAN (Smallest Univalued Segment Assimilating Nucleus, 简称 SUSAN) 边缘检测等<sup>[55]</sup>。

由于本文中是对三维模型的投影图像进行边缘检测，投影图像的内部像素与边缘像素之间的差别相对较大，图像中的焦点区域比较少，噪声主要存在于投影图像的边缘，再加之基于草图的模型检索系统对于系统的运行效率要求高等方面的考虑，采用改进的 SUSAN 边缘检测算法提取获得的投影图像。选取 SUSAN 算法主要于它直接根据图像的灰度值对图像进行处理，在实现上面不需要进行像素的梯度运算，并且能够在不使用积分算子的情况下体现出积分的特性，方法简单且具有良好的算法执行效率。

传统的 SUSAN 边缘检测方法<sup>[56]</sup>主要基于定义的具有方向性的圆形模板，根据不同的情况给模板选取大小不同的半径，为了计算上的简便，在这里选取大小为  $3 \times 3$  的模板。如图 4.5 所示。

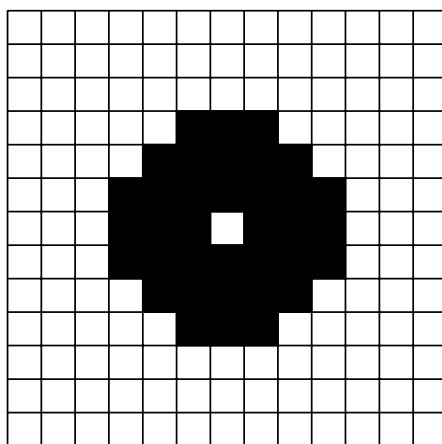


图 4.5 SUSAN 圆形模板

在图像边缘检测时，将模板的中心像素与图像中的每一个像素进行灰度值的比较，构建相似比较函数。用公式表示为：

$$c(r, r_0) = \begin{cases} 1, & |I(r) - I(r_0)| \leq t \\ 0, & \text{其他} \end{cases} \quad (4-19)$$

其中， $I(r)$  表示像素  $r$  的灰度值， $t$  为灰度差的阈值， $c(r, r_0)$  为相似比较函数，取值为 0 或者 1，分别表示  $r_0$  周围的像素与他的相似程度。 $r_0$  为模板中心的像素， $r$  为其他像素。计算  $c(r, r_0)$  的和值为：

$$n(r_0) = \sum_r c(r, r_0) \quad (4-20)$$

和值  $n(r_0)$  就是 USAN (Univalue Segment Assimilating Nucleus) 区域的像素个数，即得到了 USAN 区域的面积。最后根据选定的阈值和定义响应函数对该像素是否为边缘像素进行判断：

$$R(r_0) = \begin{cases} g - n(r_0), & n(r_0) < g \\ 0, & \text{其他} \end{cases} \quad (4-21)$$

其中， $R(r_0)$  为响应函数， $g$  为选取的阈值， $g$  的选取根据不同的情况而不同。

设定的阈值  $g$  能够决定提取边缘的 USAN 区域的最大值，即从图像中提取边缘的多少和所检测到的边缘的程度。

在传统的 SUSAN 方法中，采用绝对灰度差值的比较方法，利用灰度差值决定相似度的取值为 0 或者 1。所以，该方法并没有考虑图像像素直接相似度的影响。在本文中提出了改进的 SUSAN 边缘检测算法，在原理上与传统的方法基本相同，同样是将圆形模板中得像素与中心像素进行比较。但是，改进的算法通过从邻域像素的灰度差值的相对值决定它们之前的相似度。这样选取目的是降低传统的 SUSAN 算法阈值选取对于检测结果的影响，同时也提高了算法的对噪声的抗干扰性。用公式表示如下：



$$c(r, r_0) = \begin{cases} 1, & \frac{|I(r) - I(r_0)|}{I(r_0)} \leq t \\ 0, & \text{其他} \end{cases} \quad (4-22)$$

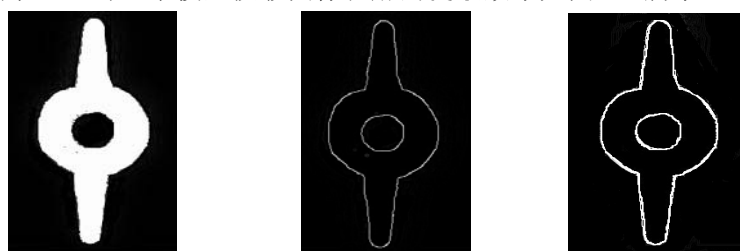
同样计算其和值为：

$$n(r_0) = \sum_{r_i} c(r_i, r_0) \quad (4-23)$$

最后，如传统的表示方法相同，通过比较选定的阈值与所计算出的变化程度值之间的关系而获得进行边缘判断的判断函数。

对于改进算法中阈值的选取决定了检测出边缘是否更接近实际的效果。由于图像的多样性，无法使用一个固定的阈值进行界定，但是通过大量实验得出，当阈值取 0.15~0.45 时能够取得较好的检测结果。

在选取阈值为 0.27 时三维模型投影图像轮廓的提取效果如图 4.6 所示。



(a) 投影轮廓 (b) 传统边缘检测 (c) 改进后的 SUSAN 算法

图 4.6 改进后的 SUSAN 算法与传统 SUSAN 算法边缘检测效果对比图

从图 4.6 中不难看出，改进后的 SUSAN 算法对于阈值的依赖性不强，能够减少内部纹理噪声对于边缘检测过程的干扰，提取的图像边缘只包含了很少的纹理，着重提取了图像的外围轮廓边缘，算法在边缘检测时效果更加明显。这说明改进后的算法非常适合基于外围轮廓特征提取的三维模型检索。

#### 4.2.5 图像外围轮廓跟踪

在对投影图像的边缘进行检测之后，则需要对图像的外围轮廓进行跟踪，去除内部与检索关系不大的冗余的轮廓信息，只得到图像的外围轮廓，为下一步进行特征提取和草图的特征进行相似性计算做好铺垫。

由于进行边缘提取之后的图像为二值图像，那么只需要按照一定的探测准则按顺序对图像进行搜索，便能够得到图像的外围轮廓特。常见的传统边界跟踪算法包括光栅扫描法、T 算法、爬虫法等<sup>[57-59]</sup>，但这些算法都存在跟踪效率不高的问题。爬虫法在进行边界跟踪时往往要进行多次的重复跟踪，甚至出现在一个封闭的局部区域内重复循环的情况；光栅扫描法需要对图像进行多次的扫描，不同的扫描过程中不断的更改阈值，因此得到的图像的轮廓不够准确；T 算

法是一种基于 4 领域搜索的算法，根据进入方向的不同而采用不同的搜索规则，容易受到内孔洞存在的影响而且算法的搜索效率较低。

本文中采用的轮廓跟踪算法，能较大程度上避免了内部轮廓对于外围轮廓提取的影响，同时对于具有断裂的轮廓线也具有好的提取效果。对于二值图像，算法的跟踪方法如下：

(1) 搜索第一个边界点，按照从左到右自上而下的顺序扫描图像，当发现像素值为 1 的点时则将该点作为起点，并记下起点的坐标。

(2) 搜索下一个边界点，假定边界的断裂点为  $M_0$ ，在某一个领域内进行搜索，则搜索当前边界点  $(2 \times M_0 + 3) \times (2 \times M_0 + 3)$  领域内的目标点，将与起点最近的像素点为后续的边界点，并将该点的像素置为 0，仍然将该点坐标记下。

(3) 当 (2) 中的方法无法再搜索到下一个边界点时，则该方向上的边界跟踪结束。

(4) 重新将起点置为当前边界点，在相同的邻域范围内搜索目标点，重复 (2) 中的搜索过程。

(5) 当 (4) 中的方法也同样无法搜索到下一个边界点时，则该曲线的跟踪结束，根据搜索的曲线长度与设定的阈值大小决定该曲线的保留或者删除。

(6) 搜索其他的边界。

利用上面的跟踪算法，在实验中对经过边缘检测后的外围轮廓进行跟踪，结果如图 4.7 所示。

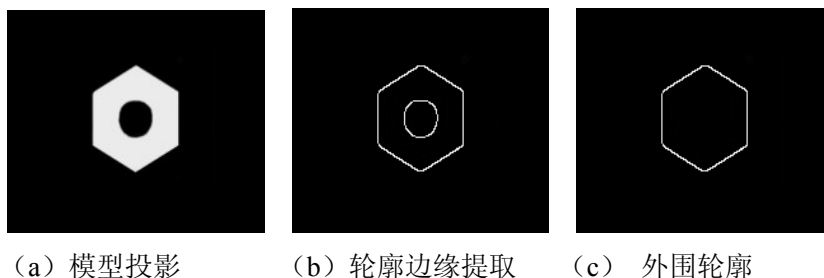


图 4.7 图像轮廓跟踪效果示意图

获取投影图像的外围轮廓后，则可以使用前面所阐述的特征提取方法对图像的外围轮廓进行提取，生成投影图像的特征向量。这样就可以通过计算草图轮廓的特征向量与投影图像的特征向量的相似性实现三维模型的检索。

## 4.3 基于贝叶斯网络的三维模型检索

### 4.3.1 贝叶斯网络

贝叶斯网络<sup>[60]</sup>最初在 1981 年由 R.Howard 和 J.Matheson 两人共同提出，主要被运用在人工智能的领域对不确定的知识进行推理，是建立在概率论和图论相结合的一个重要方法。它通过由概率标注的有向无环图自然的表达出事物之间的因果关系，利用概率计算的知识描述数据之

间的关联程度，其优点在于清晰的语义表示和较强的预测分析能力。

采用贝叶斯决策就是利用概率的不同分类决策与相应的决策代价之间的定量折中。通过概率的形式对决策的问题进行描述，在描述的过程中假定了所有有关的概率的结构均为已知。而将这种决策方式以图的形式描述出来，就是“贝叶斯置信网”，或者称之为贝叶斯网络。贝叶斯网络通过拓扑连续的方式表现出模型中变量间的相互依赖和相互独立的关系。当对变量的子集给予某些确定的值之后，则可以利用贝叶斯网络的推理功能进行相应的计算，贝叶斯网络中的每一个节点将获得一个概率值。简单的贝叶斯网络如图 4.8 所示。

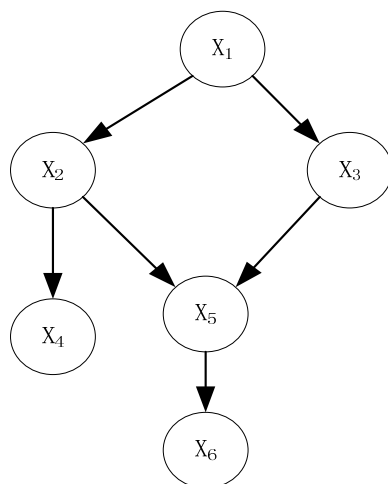


图 4.8 简单的贝叶斯网络

给定随机变量  $X = \{X_1, X_2, \dots, X_n\}$ ，则可定义基于  $X$  的贝叶斯网络  $B = \{G, \varphi\}$ ，其中  $G$  表示有向无环图， $\varphi$  表示图中所有节点的条件概率集合。一个完整的贝叶斯网络主要由两个部分构成：

(1) 有向无环图  $G$ ， $G$  中的每一个节点代表了一个随机变量，节点间的边连接关系则体现了节点之间的相互依赖关系，图  $G$  中的每个节点  $N_i$  则独立于由  $N_i$  的父节点给点的非  $N_i$  后代节点构成的任何节点子集，用公式可以表示为：

$$P(X_i | A(X_i), Pa(X_i)) = P(X_i | Pa(X_i)) \quad (4-24)$$

其中， $A(X_i)$  表示非  $X_i$  后代节点构成的子集， $Pa(X_i)$  表示  $X_i$  的双亲节点。

(2) 条件概率表  $CPT$ 。节点之间的关系构成了一个条件概率表，根据条件概率表及节点间的相互关系，就可以通过概率知识的计算得到贝叶斯网络所有节点的联合概率，用公式表示为：

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i)) \quad (4-25)$$

$$P(X_1, X_2, X_3, X_4, X_5, X_6) = P(X_6 | X_5)P(X_5 | X_3, X_2)$$

$$\begin{aligned} & \cdot P(X_4 | X_2)P(X_3 | X_1) \\ & \cdot P(X_2 | X_1)P(X_1) \end{aligned} \quad (4-26)$$

由上式可见，如果计算某一节点的概率，则可以通过对贝叶斯网络进行正向或反向的概率计算，对节点的概率计算过程也就反映了贝叶斯网络的推理过程。

#### 4.3.2 样本训练

首先，通过构建高斯分布来表示  $X$  的分布情况，那么根据贝叶斯网络推理模型，如果想要得到网络中任意一个节点的概率，则要得到相应的先验概率和条件概率，即需要得到如下的概率分布情况：

$$P(M=i) = p_i \quad (i=1 \dots k) \quad (4-27)$$

$$P(X=\bar{x} | M=i) = \frac{1}{(2\pi)^{N/2} |\sum_i|} \exp\left\{-\frac{1}{2}(x-u_i)\right\}, \quad (i=1 \dots k) \quad (4-28)$$

其中的未知参数  $\varepsilon = \{P_i, P_{ij}\}$ ， $P_i$  为先验概率集合， $P_{ij}$  为条件概率的集合。

通过训练贝叶斯网络模型确定未知集合中的值，选取一定数目的训练样本，根据前面的处理方法得到三维模型投影图像的特征向量。对于某一个三维模型  $i$ ，根据其投影图像获得的特征向量为：

$$F = \{f_1, f_2, \dots, f_n\} \quad (4-29)$$

对于样本中的  $m$  个三维模型的特征向量，则可以构建一个训练样本的数据集合  $D_j^i = \{i, \text{Vector}(i, j)\}$ ，其中  $j$  表示训练样本的个数，利用得到的数据集  $D_j^i$ ，则可以采用  $EM$  算法估计出  $\varepsilon$  集合的未知参数的值：

$$L(\varepsilon) = \ln \prod_{i=1}^n P(D_i | \varepsilon) = \sum_{i=1}^n \ln P(D_i | \varepsilon) \quad (4-30)$$

这也便得到了进行贝叶斯网络推理的所有必备条件。

#### 4.3.3 贝叶斯网络推理模型

当用户输入草图后，首先通过对草图的预处理、识别、轮廓的提取等一系列的过程，得到草图轮廓的特征向量  $f_i$ ，那么根据  $f_i$  便可以采用贝叶斯网络模型来推理用户的检索意图，推理出在特征向量  $f_i$  的情况下三维模型为  $i$  的后验概率为：

$$P(M=i|X=f_i) = \frac{\sum_j P(M=i, X=f_i)}{\sum_l \sum_j P(M=l, X=f_i)} \quad (4-31)$$

当给定一个草图轮廓的特征向量集合  $V_i = \sum_{j=0}^n \text{Vector}(i, j)$  时，用每个特征向量通过贝叶斯

网络推理模型推理出语义符号为  $i$  的概率则可以表示为  $P(S=i|X=\text{Vector}(i, j))$ ，对于语义

符号  $i$  全部的特征向量则得到语义符号  $i$  的后验概率为：

$$\varphi(i) = \frac{1}{n} \sum_j^n P(M=i|X=\text{Vector}(i, j)) \quad (4-32)$$

在进行三维模型检索时，只需要计算上式求得  $\varphi(i)$  的值来确定用户所要检索的三维模型  $i$ ，将检索到的三维模型按照值从大到小的顺序进行排列呈现给用户。

#### 4.3.4 相关反馈

相关反馈通常是人机不断交互的循环过程，在用户得到系统输出的检索结果后，根据个人的主观信息对结果进行评价或标注，是优化系统检索性能的有效手段。系统将根据用户的标注重新调整检索过程中的相关参数或相似性计算过程，以人机不断交互的方式将高层语义特征融入到检索过程中，检索获得的结果集更能符合人们的主观意图，从而优化了系统的检索效果。

建立在贝叶斯理论基础上的相关反馈算法是利用用户在检索过程中的反馈信息，通过进行统计和概率推算来实现检索性能的提高。该相关反馈算法在基于内容的信息检索中表现出很好的效果。

本节在已构建的贝叶斯网络推理模型的基础上，充分利用用户对检索结果的反馈对模型中的概率进行更新，在尽量少反馈次数基础上使得检索效果得到充分的优化，具体的反馈流程如图 4.9 所示。

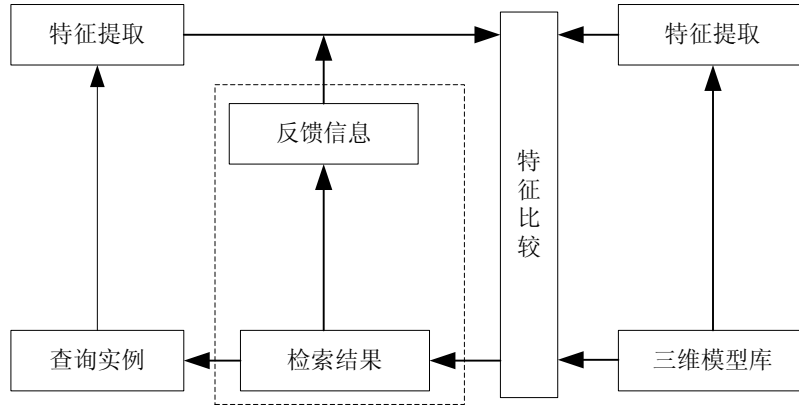


图 4.9 相关反馈流程

下面对基于贝叶斯的相关反馈算法进行总体的描述：

设定三维模型的特征向量库为  $D = \{D_1, D_2, \dots, D_n\}$ ， $D_r$  为用户绘制的草图中提取的某一特征向量，用户检索的目标三维模型为  $M$ ，假设在经过了  $t$  轮的用户反馈后，对用户的反馈记录可以表示为：

$$H_t = \{R_0, A_1, R_1, A_2, R_2, \dots, A_t, R_t\} \quad (4-33)$$

其中： $R_i$  为第  $i$  轮检索系统显示的检索结果； $A_i$  为用户在第  $i$  轮对于结果中的三维模型的评价信息，即检索显示的结果模型是否符合用户的检索意图。

根据用户的反馈信息集合，计算在反馈信息  $H_t$  集合的基础上，利用贝叶斯推理公式推理出检索目标  $M$  的概率可以表示为：

$$\begin{aligned} P(D_r = M | H_t) &= P(D_r = M | R_t, A_t, H_{t-1}) \\ &= P(A_t | D_r = M, R_t, H_{t-1}) P(D_r = M | H_{t-1}) / \\ &\quad \left[ \sum_{j=1}^n P(H_t | D_j = M, R_t, H_{t-1}) P(D_j = M | H_{t-1}) \right] \end{aligned} \quad (4-34)$$

其中： $P(D_r = M | H_t)$  为目标模型的概率，它反映了当前检索经过  $t$  轮反馈后得到目标模型  $M$  的概率。系统将根据目标概率的分布，从三维模型库中选择一个概率最大特征向量所对应的三维模型，再次把检索后的结果呈现给用户。通过这种方式，系统能够将用户的评价信息引入到建立的贝叶斯网络推理模型中，并根据用户的评价调整得到目标模型的概率，使检索出的三维模型更符合用户的主观意图。模型的推理准确性也在整个交互过程中不断的提高和优化。

通过每一轮的迭代过程，根据得到的概率大小对结果中的三维模型进行排序，最后将结果呈现给用户。

在相关反馈的过程中，根据用户评价信息不同，所得到目标三维模型的概率的也存在一个不断更新的过程。由式 4-34 可知，在进行第  $t$  轮反馈的基础上，目标模型的概率

$P(D_r = M | H_t)$  可以利用第  $t-1$  次的反馈信息表示为  $P(A_t | D_r = M, R_t, H_{t-1})$  的计算。它根据给定前一轮反馈历史  $H_{t-1}$ 、第  $t$  轮显示给用户的检索结果  $R_t$  及假定  $D_r$  为检索目标  $M$  时预测用户的行为。据此可以看出，第  $t$  轮的检索结果中融入了用户的评价信息，检索到的目标模型更加准确和贴近用户的检索意图。

在得到用户的反馈信息后，则需要根据该信息对检索过程中的相关概率进行更新。在所有的检索结果集合  $R$  集合中，包含了两种类别：一种是已经被用户评价过的三维模型，表示为  $Y = \{M_{y1}, M_{y2}, \dots, M_{yn}\}$ ；另一种则是未进行评价的三维模型，表示为  $N = \{M_{n1}, M_{n2}, \dots, M_{nm}\}$ 。两者与集合  $R$  的关系可以表示为  $R = Y + N$ 。对于用户输入的草图特征  $D_r$ ，如果其与  $Y$  集合更接近，那么说明  $D_r$  是目标模型的特征的概率比较大；反之，如果与  $N$  相似，则是目标模型的概率比较小。设定  $Dist(D_r, M_{yi})$ 、 $Dist(D_r, M_{nj})$  分别表示  $D_r$  与集合  $Y$  和  $N$  中的模型的相似性度量值。那么计算  $P(A_t | D_r = M, R_t, H_{t-1})$  的程序伪代码可以表示为：

$S$  表示  $P(A_t | D_r = M, R_t, H_{t-1})$  的最后计算结果。

$R$  则表示系统经过检索后显示给用户的三维模型。

$U(i)$  表示用户的标记情况，如果  $U(i)$  是 1，则用户选择了  $R(i)$ ；如果为 0，则没有选中  $R(i)$ 。

$\mu$  和  $\sigma$  表示用户参数。

If 没有模型被选中

$S = 1$

Else

For  $R$  中的每个三维模型  $R(i)$

Score = 0

For  $R$  中的每个  $R(j)$ ,  $R(j) \neq R(i)$

If  $Dist(D_r, M_{yi}) < Dist(D_r, M_{nj})$

Score = Score + 1

Elseif  $Dist(D_r, M_{yi}) > Dist(D_r, M_{nj})$

Score = Score + 0.5

$P = 1 / (1 + \exp(\mu - \text{score}) / \sigma)$

If  $U(i) == 1$

$S = S * P$

Else

$S = S * (1 - P)$

通过上面的分析，可以按照以下步骤对目标模型的概率进行更新：

(1) 对于给定的检索模型  $M$ ，计算数据库中每个三维模型与给定的  $M$  之间的相似度，并以此作为初始的目标模型概率  $P_t$ 。

(2) 根据  $P_t$  的大小对  $R$  集合中的三维模型进行排序，取概率最大的前  $N$  个三维模型构成  $R_t$  呈现给用户。

(3) 如果用户对检索所得到的结果满意，那么本次的检索过程结束；否则，用户可以根据个人检索意图对输出的结果重新进行标记。

(4) 根据用户的反馈信息重新计算  $P(A_t | D_r = M, R_t, H_{t-1})$ 。

(5) 根据公式计算  $P_r^{t+1}$ ，然后转 (2)。

对于上面概率的更新过程，其流程图如图 4.10 所示。

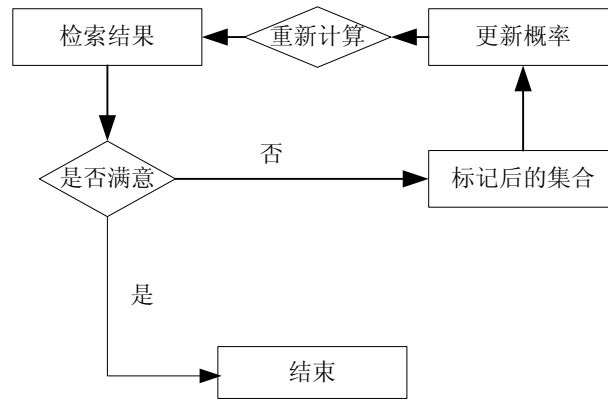


图 4.10 概率更新过程

#### 4.3.5 实验结果及分析

在多媒体检索领域，经常使用查全率 (Recall) 和查准率 (Precision) 衡量检索系统性能的好坏。在三维模型检索系统中，查全率代表返回的正确的模型占整个相关模型的比。查全率越高，则表明系统返回正确结果的能力越强；反之，则越弱；查准率代表在所有检索的返回结果中，正确检索的三维模型的比例。查准率越高，表明系统返回结果的精确性越高。

实验中，采用如下方法评价系统的检索系统：用户通过交互界面绘制三幅表现个人检索意图的草图，系统将提取草图的外围轮廓特征，通过贝叶斯网络推理模型计算与数据库中每个三维模型之间的匹配度；然后根据模型的匹配程度进行排序，取前面最相近的  $S$  个匹配三维模型作为检索结果。

将本文中的算法与基于单个主平面分析的方法、文献[34]中的球面调和方法和文献[28]中的光场描述符方法进行比较。如表 4.1 所示。从表中可以看出，本文中的算法与其他的算法相比查全率和查准率有了明显的提高。



表 4.1 不同方法的查全率与查准率比较

方法	查全率/%	查准率/%
本文方法	92.5	91.6
单个主平面方法	89.9	88.2
球面调和方法	91.2	90.6
光场描述符方法	90.3	87.9

在评价三维模型检索系统性能时，还应考虑到检索效率的问题。如果检索时间太长就会导致检索的实时性和交互性下降。由于本中选取主平面的方法是对三维模型自身性质的数学反映，受模型的位置和姿态的影响较小。因此，有效减少了三维模型预处理时需要的计算量。如表 4.2 所示。本文的检索算法与其他算法的检索时间相差不大，保证了系统良好的交互性。

表 4.2 不同方法的检索时间比较

检索方法	检索时间/ms	平均检索时间/ms
本文方法	153	145
单个主平面方法	138	126
球面调和方法	142	132
光场描述符方法	168	156

对于系统输出的检索结果，经过相关反馈后，基于手绘草图的三维模型检索系统的查全率与查准率比较曲线如图 4.11 所示。

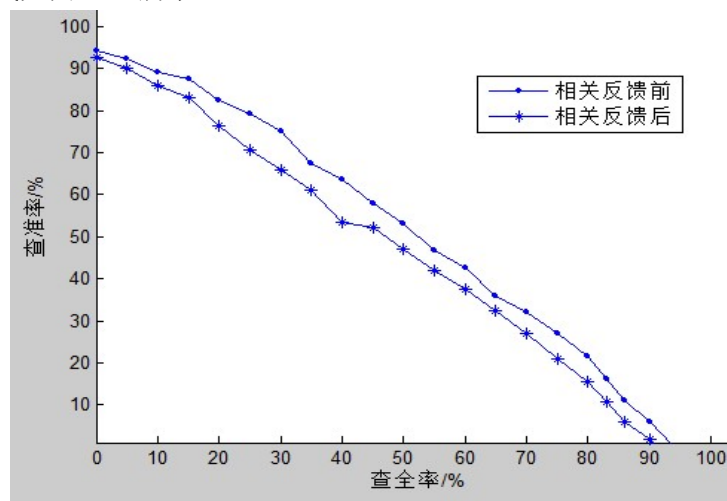


图 4.11 相关反馈前后的查全率-查准率曲线

从图中可以看出，在通过相关反馈后，三维模型检索系统的查全率和查准率都得到了相应的改进，系统的检索性能得到了优化。

## 4.4 本章小结

本章是基于手绘草图的三维模型检索系统实现技术的核心章节，主要对三维模型的特征提取过程和所运用的相关技术作了阐述，对三维模型投影图像获取方法、图像边缘提取和轮廓跟踪、基于贝叶斯网络检索过程等系统实现的关键部分作了深入的研究和改进。通过相关反馈技术，提高了系统检索的准确性。最后通过实验验证了本章相关技术的研究能有效的提高三维模型检索系统的性能。

## 第五章 基于手绘草图的三维模型检索系统实现

### 5.1 引言

前面章节分别阐述和研究建立一个基于手绘草图的三维模型检索系统需要完成的任务和各阶段中使用的关键技术。首先,提供便捷的草图绘制界面让用户通过绘制草图的方式勾画自己在设计过程中需要的三维模型轮廓,通过第三章中的草图预处理与识别技术对用户输入的草图进行去噪、拟合、规整和识别等,利用跟踪算法获取草图的轮廓,在对草图的外围轮廓进行基于特征点的多变形拟合后,依据本文中的几何不变量的构造方法进行草图特征提取,构建表达用户模型检索条件的特征向量;其次,根据三维模型自身的特性计算三维模型的投影平面,对三维模型在不同的主平面方向进行投影,得到二维投影图像,再利用改进的边缘提取算法提取出投影图像中的轮廓,构建描述三维模型的特征向量;最后,通过训练贝叶斯网络模型,计算草图的特征向量与三维模型的特征向量之间的相似度完成三维模型检索过程,将最后的检索结果呈现给用户。如果用户对系统输出的检索结果中某些模型不满意,则可以对结果进行主观评价。系统将融合用户的反馈信息再次将检索结果输出。

本章在前面的设计的技术和研究基础上,开发实现了基于手绘草图的三维模型检索系统 Retrieval3D。该系统提供给用户进行手写或者鼠标绘制的用户接口,让用户能够自由灵活的绘制草图,并运用本文中研究和设计的关键技术完成用户检索三维模型的过程。从实验的角度对系统中使用的关键技术作进一步的阐述,并对获得的实验结果进行了分析。在介绍了系统的设计方案后,给出了系统主要模块的实现细节及系统运行示例。

### 5.2 系统功能结构设计

系统设计的目的是在保证实现逻辑功能的基础上,尽可能提高系统的简单性、一致性、完整性、可靠性、系统的运行效率 and 安全性。设计实现过程中体现了如下原则:

- (1) 具有良好的扩展性。
- (2) 具有友好的人机交互接口。
- (3) 便于对算法进行性能评价。

在系统的设计部分,本节主要对系统的体系结构及主要功能模块进行了详细的描述。在借鉴原有三维模型检索系统的体系结构基础上,对基于手绘草图的三维模型检索系统进行了部分模块功能的优化,主要体现在用户的检索条件的输入方式和三维模型的检索过程的性能和效率两个方面。目的在于使用户既能够自由表达个人检索意图,同时也能够检索到满意的结果。系统的主要功能模块结构如图 5.1 所示。

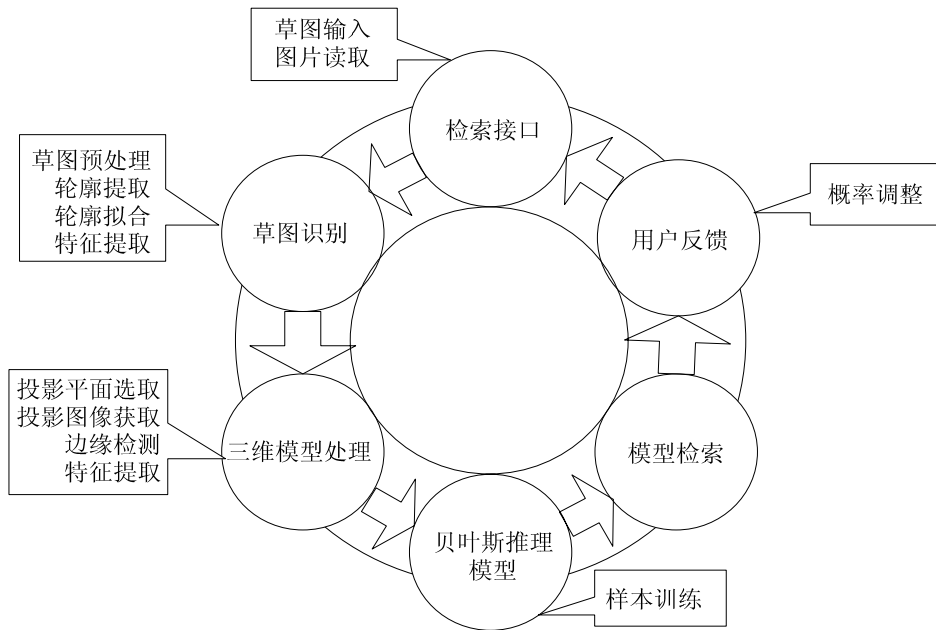


图 5.1 系统主要功能模块

系统包括以下主要功能模块：

（1）检索接口设计模块。良好的人机交互界面往往能够得到用户对于系统性能更多的认可，能否提供方便快捷的交互接口帮助用户完成个人检索意图的表达是实现这一目标的重要基础。在 Retrieval3D 系统中，为了方便用户的绘制，为用户提供了以下检索意图的表达方式：

- 绘制完整的草图。该方式为用户提供了表达个人检索意图的自由空间，用户在绘制过程中，可以根据设计的需要对草图进行修改，如笔画的撤销、删除、调整等。

- 混合方式绘图。系统中预定义了一部分常见的几何图形（如圆形、椭圆、三角形等）供用户根据自己绘制过程中的需要进行选择，用户在选取的特定的几何图形之后，可以通过鼠标的拉拽等动作改变几何图形的形状、大小、线条粗细等；也可以将给定的几何图形某一部分擦除，然后将修改后的图形根据自己的意图连接起来构成一个完整的逻辑表达，直到达到个人的设计意图。该绘图方式为用户的绘制提供了设计表达上的便捷性，用户不再需要对常用图形重复的进行绘制，只需要将部分图形的逻辑关系用少量的笔画建立起来。在这种绘图方式中，用户将最能体现个人检索意图的草图部分进行绘制，在节省工作量的同时也能使特征点更明显。

（2）草图识别模块。草图识别模块分为草图预处理和草图轮廓特征提取两个子模块：

- 草图的预处理模块。主要是对用户已经绘制满意的草图进行去噪、规整等处理，将用户绘制的模糊的、不连续的、具有歧义的草图经过处理之后变为平滑的、规范的笔画信息，排除了冗余笔画等无关信息对于草图轮廓的拟合和特征提取的干扰。通过预处理模块后的草图信

息，能够更清晰的反映用户的检索意图。

- 草图外围轮廓拟合及特征提取模块。通过本文第三章中设计的基于特征点的多边形拟合算法对草图的外围轮廓进行拟合，利用拟合后的多边形几何特性构建作为特征向量的几何不变量信息及特征点的边及曲率信息中提取出描述该草图轮廓的特征向量，建立进行三维模型检索的初始化输入数据。该模块是进行检索过程前的一个重要阶段，拟合效果的好坏和特征向量的选取好坏将直接决定了整个检索系统的检索效果。该模块也是本文中关键技术体现的一个重要部分。

(3) 三维模型处理模块。三维模型处理模块并不是在用户进行检索时才会运行的模块，而是主要用于向数据库中添加三维模型时对三维模型进行处理和特征提取的模块，该模块是 Retrieval3D 系统的一个核心模块，包含了模型检索系统中的关键技术部分，其由以下子模块构成：

- 主平面的选取子模块。该子模块的功能主要是对某个具体的三维模型根据其自身的特性计算出最佳投影的三个主平面，主平面能够最大化的反映出三维模型的某些显著特征。

- 投影图像的边缘及特征提取子模块。该子模块主要根据选取的主平面对三维模型进行三个主平面方向的投影，得到三张二维空间的投影图像，然后利用本文中改进的 SUSAN 算法对投影图像的边缘进行提取，得到图像边缘后即进行外围轮廓的提取，最后通过多边形的拟合提取出投影图像的特征向量。

- 三维模型特征向量库建立模块。根据前两个子模块所进行的处理，将得到的投影图像特征向量按一定的权重构建描述三维模型唯一的特征向量，建立该三维模型的特征向量库，便于同用户绘制的草图中提取的特征向量进行相似性的判断。

(4) 基于贝叶斯网络模型的检索及相关反馈模块。该模块是检索系统的核心模块，主要是基于贝叶斯网络模型的三维模型检索，通过构建基于贝叶斯的推理模型，利用贝叶斯网络的推理能力来完成用户意图的检索过程。对于检索结果，用户可以根据个人的满意度对三维模型进行标注，系统将结合用户的反馈信息重新将检索结果输出。

## 5.3 系统主要类设计

### 5.3.1 系统核心类图

本系统是面向对象设计的，对于所有的功能都是封装在相应的类库中。类图是 UML 设计时最常用于描述系统的结构化设计，图 5.2 为本系统中涉及的所有类。其中，SketchBeauty 类用于对手绘草图进行规整，而 SketchPretreatment 类、StrokeAnalyzer 类主要完成草图的预处理及笔画特征点的计算，Model 类继承 IModel 接口中的方法完成三维模型的预处理等功能，BayesianModel 类则主要封装了利用贝叶斯网络推理时所需要的操作，另外还包含了相应完成

系统检索结果反馈的 `Feedback()` 方法等。而 `SketchUI`、`RetrievalUI` 等主要负责处理用户输入表现个人检索意图的界面维护、笔画绘制及撤销等。

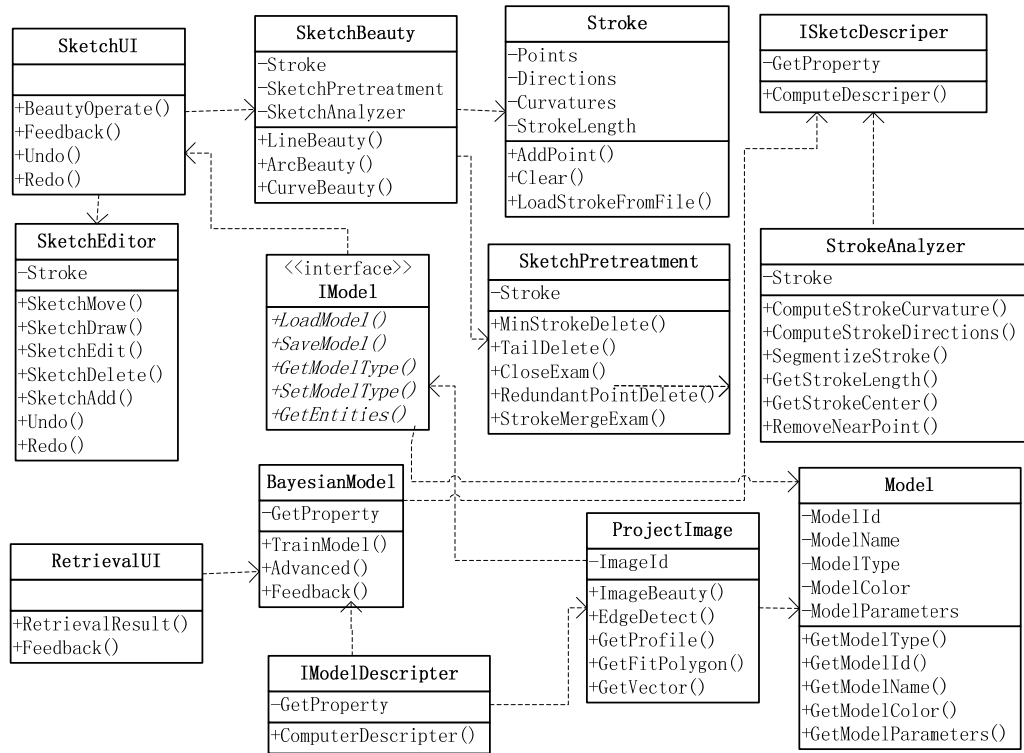


图 5.2 系统类图

### 5.3.2 草图轮廓识别及特征提取类图

手绘草图轮廓及特征提取模块主要完成用户绘制草图的预处理。目的是消除因输入条件及不同用户的输入习惯的差别而产生在草图中的噪声。对草图进行不同的预处理过程及处理采用的方法在第三章中已经作了详细的介绍。在本节中主要对该模块中的类设计进行描述，根据预处理后的草图信息，采用基于特征点的多边形拟合的算法对草图的轮廓特征进行提取，最后得到草图的特征向量。草图预处理模块的类图如图 5.3 所示。

(1) 用户绘制的草图定义为一个 `Sketch` 对象，对草图采集的数据被存放在 `Point[]` 数组中，`Sketch` 对象拥有自己本身的属性 `Length`、`Area`、`Curvature` 等。

(2) `Sketch` 通过调用接口 `PreProcess` 中的方法实现对于草图笔画的预处理，包括冗余点的去除、交点的判断、图形闭合性的检查等。预处理结束得到 `SketchPretreatment` 对象。

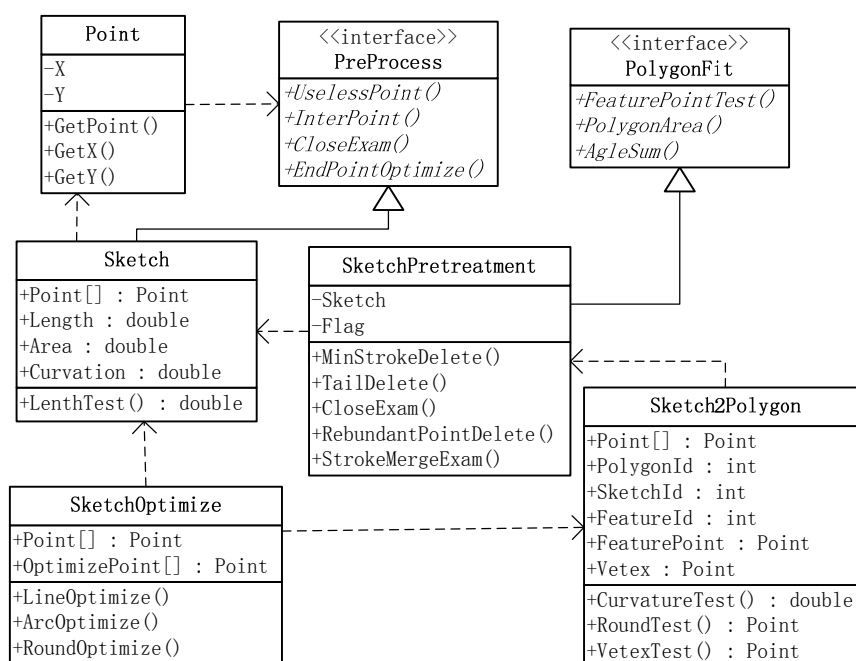


图 5.3 草图识别及特征提取

(3) OptimizeSketch 对象调用了接口 PolygonFit 中的方法进行多边形的拟合，建立 Sketc2Polygon 对象。

(4) Sketch2Polygon 对象调用其自身的方法计算出草图轮廓的特征信息，构建基于拟合多边形的几何不变量，将几何不变量作为检索的初始条件。

### 5.3.3 三维模型投影图像获取及特征提取类图

三维模型的特征提取主要是在三维模型入库时对三维模型进行处理，获取与草图特征向量进行相似性计算的特征向量。其主要的处理过程在第四章做了详述，在设计的三维模型处理类图中，主要包含了用于三维模型预处理、投影面选取、投影图像处理、特征提取等主要部分的功能类，具体如图 5.4 所示。

(1) Model 对象继承了 PreModel 接口中的方法完成三维模型的预处理过程，包括三维模型的平移、旋转和缩放等处理，处理结束后输出 PostModel 对象。

(2) PostModel 对象利用 Plane 类中的方法进行获取投影图像的三个主平面，而其中 Plane 类则实现了 SelectPlane 接口中的方法完成这一过程。

(3) 通过在三个主平面方向作正交投影，获取到三张三维模型的投影图像，用 Image 对象对其表示。Image 对象调用接口 ImageProcess 中的方法完成对投影图像的边缘检测和轮廓的提取，最后计算出投影图像的特征向量。

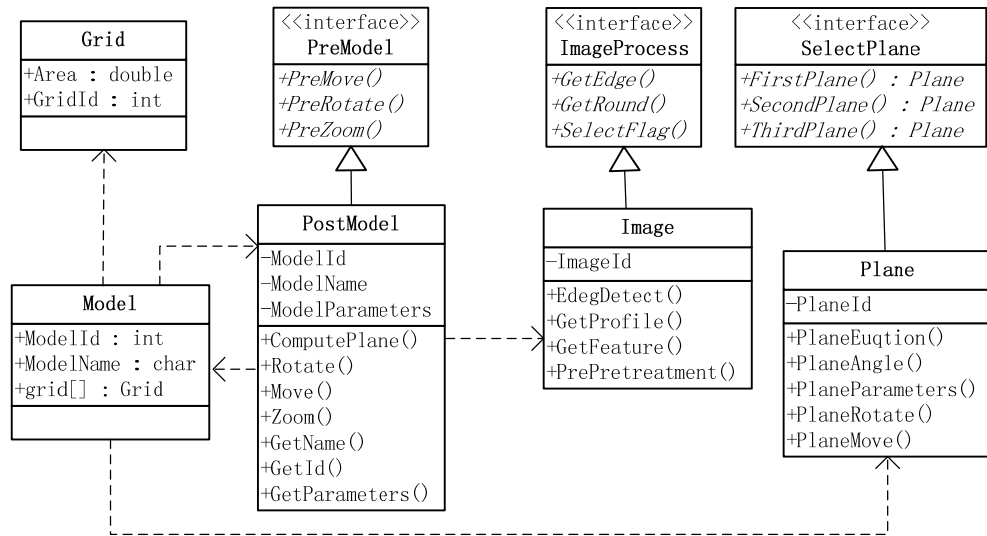


图 5.4 三维模型处理类图

## 5.4 数据库设计

基于手绘草图的三维模型检索系统的数据库设计主要考虑以下几个问题：

(1) 三维模型的文件格式。由于目前三维模型的类型不同，不同的检索方法和系统决定了数据库中的三维模型也具有不同的格式，在本系统中，并没将三维模型的格式进行严格的限制，主要是基于系统在多领域的应用性问题，在系统中采用的三维模型格式主要为 3DS、TDO 和 CTS 的数据格式，除此之外的格式可以通过转换的处理存入数据库中。

(2) 表的设计。基于手绘草图的三维模型检索系统将数据库分为两个部分，即特征向量库与模型库。通过使用第四章中三维模型特征向量的提取方法得到三维模型的特征向量，为这些向量集合单独的建立了一张特征向量表，这张表主要是对三维模型特征建立索引，将特征向量与其所代表的三维模型之间建立对应的关系，在检索时并不是直接将草图中的特征向量与数据库中的三维模型进行比较，而是与建立的特征向量表中的值进行比较，然后通过特征向量表与三维模型的对应关系得到用户所要检索的模型。

在数据库设计过程中，现实世界到数据世界是通过概念模型来转换的。E-R 模型是最常见的一种概念模型，它既能真实反映用户需求，又与计算机无关且容易被用户理解接受。基于手绘草图的三维模型检索系统的 E-R 图如图 5.5 所示。



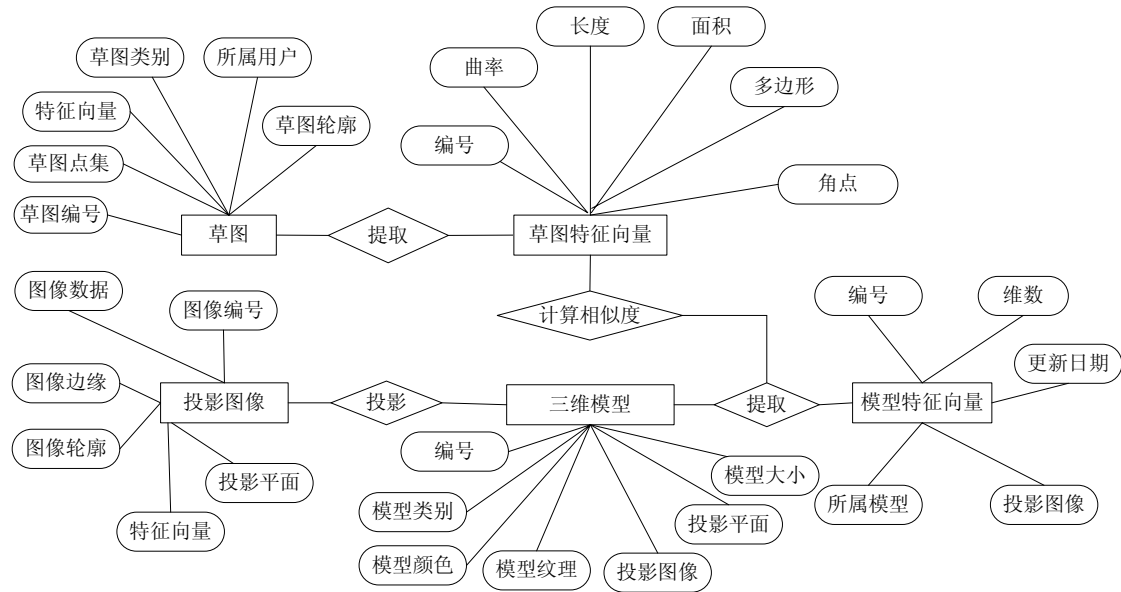


图 5.5 系统 E-R 图

基于给出的 E-R 图，给出了在本系统中建立的数据库表结构，如图 5.6 所示。

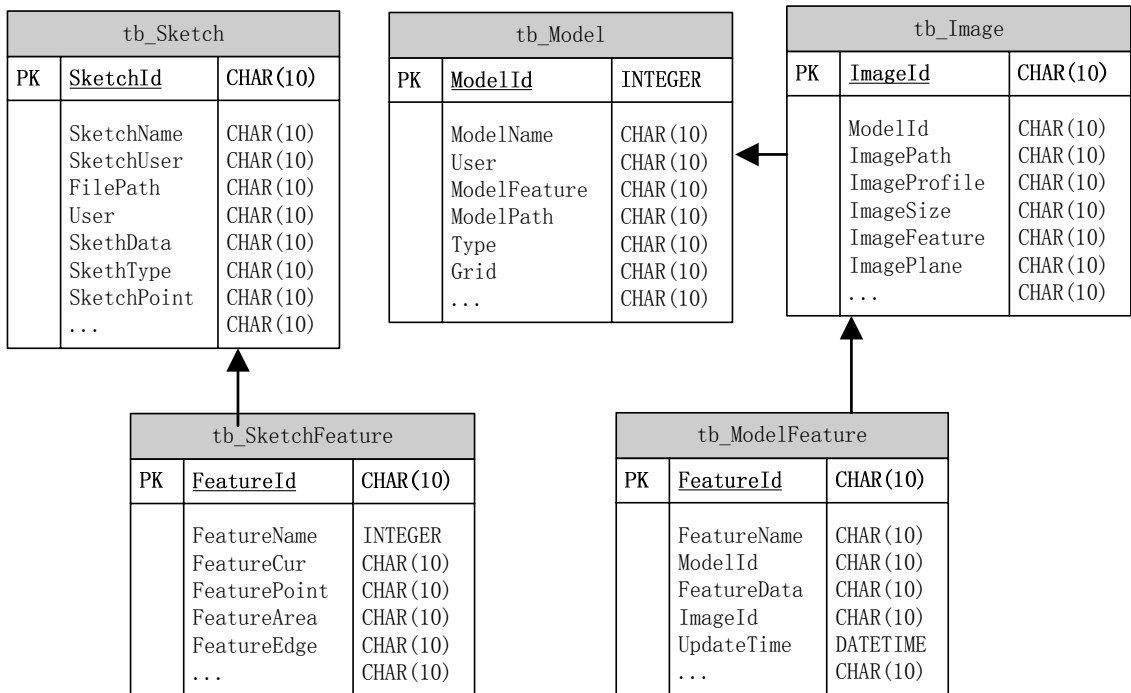


图 5.6 数据表设计

(1) 表 **tb\_Sketch** 表示了用户绘制的草图信息，将其归类为一个表主要是为了方便用户下次检索时，可以直接调用原来已经绘制的草图进行修改或直接用于三维模型检索。

(2) **tb\_Model** 对系统中的三维模型进行了描述，不同的用户之间可能会对不同领域的三维模型感兴趣。其中，**ModelPath** 表明了三维模型存放的具体路径，比较特征向量结束后才对三维模型根据路径进行提取，这也保证了系统的检索效率。

(3) `tb_SkechFeature`、`tb_ModelFeature` 分别表示了提取的草图特征向量和三维模型的特征向量。其中, `tb_ModelFeature` 的属性中包含了该特征向量对应的三维模型及其投影后得到的投影图像。

## 5.5 系统运行示例

系统的草图绘制界面如图 5.7 所示。用户可以通过系统提供的绘制界面自由的绘制表现自己检索意图的草图, 或者选取原来已经绘制的草图加载到系统中, 在此基础上进行修改。修改主要包括选取不符合个人意图的笔画信息。

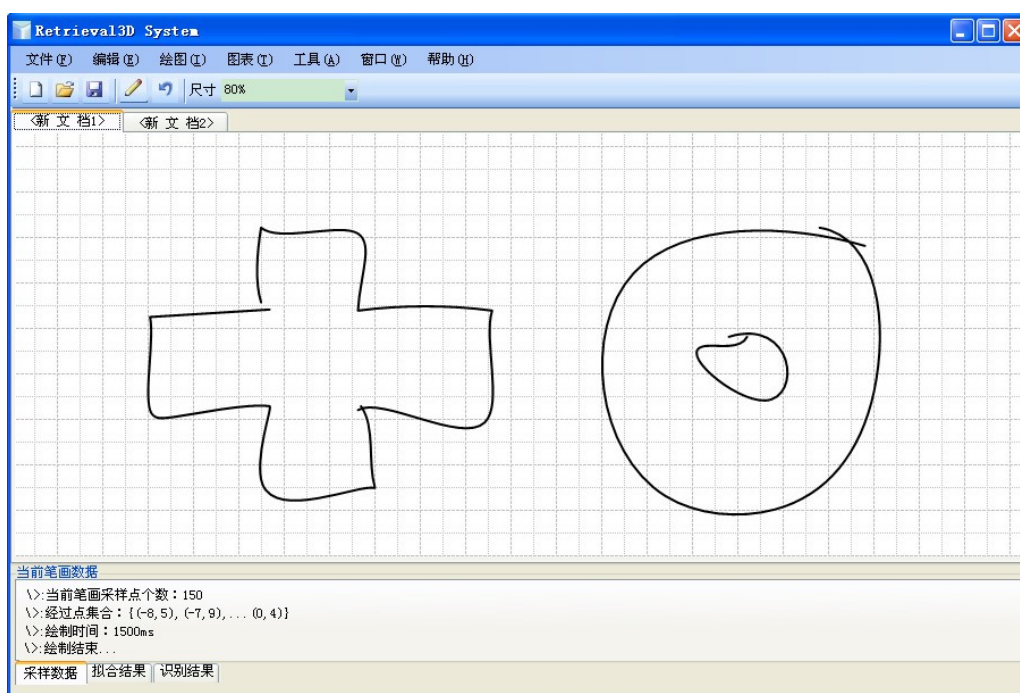


图 5.7 手绘草图

草图的规整结果如图 5.8 所示。对于草图的笔画信息进行了拟合, 使之更加的准确和平滑, 消除了绘制过程中产生的冗余笔画、聚点、及交叉笔画等。经过规整后的草图去除了原始草图中的噪声信息, 能更加清晰的反映了用户的检索意图, 减少了对后续的特征提取等过程的干扰。

系统的检索结果如图 5.9 所示。从图中可以看出, 在所有的检索结果中, 既包含了符合用户检索意图的三维模型, 也检索出了一小部分不相关的模型。该情况主要由于特征向量并不能完全的将三维模型的所有特征描述出来所造成的, 使得检索结果中特征相近的轮廓检索出的模型之间类型可能迥然不同。图中显示的模型大小、方向也不尽相同, 这说明本文中选取的特征向量可以满足三维模型的旋转、平移和缩放的不变性, 通过手绘草图检索三维模型也取得了良好的效果。

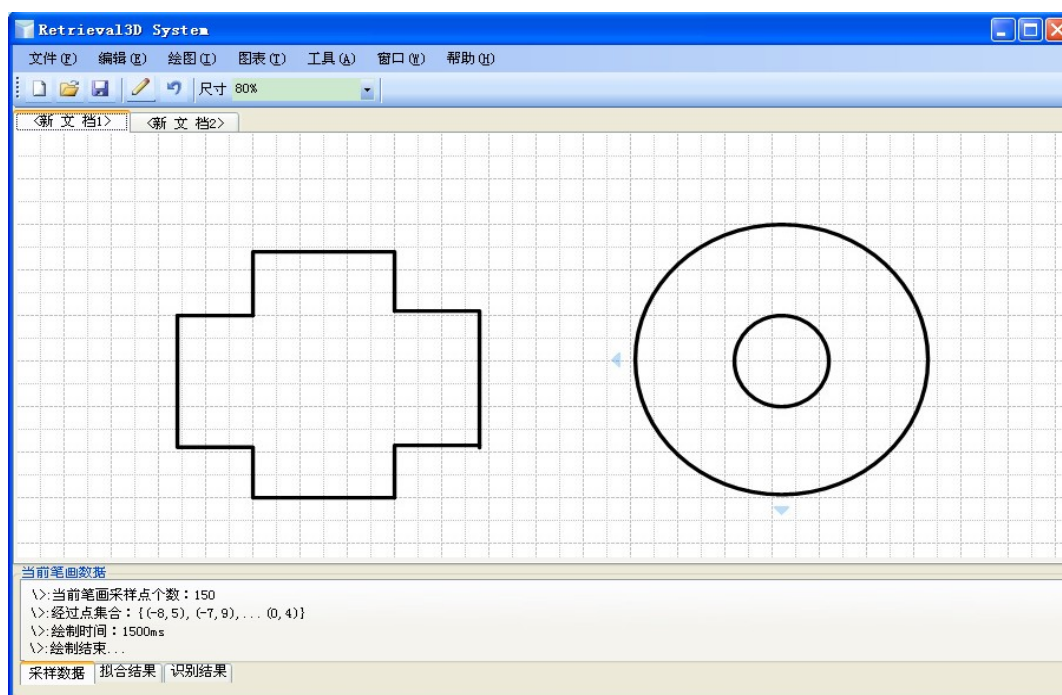


图 5.8 草图的规整



图 5.9 检索结果

在检索出的三维模型中，系统输出了用户当前选择的模型的编号、路径、类别及各个详细参数。用户如果想进一步对查看当前的三维模型，可以点击预览按钮，系统将展现三维模型的

材料设置、光线设置、颜色及缩放等属性。



图 5.10 反馈后的检索结果

用户通过对检索结果的标注对系统的检索性能进行优化。在这一过程中主要通过本文中的相关反馈实现。用户标记后，系统将自动对特征向量的权重进行更新，将用户的主观评价信息融入到模型的检索过程中，使检索结果更能表示用户的检索意图。图表示了在经过用户的反馈后重新检索的结果。从图 5.10 中可以看出，符合用户检索意图的模型增加，而与用户检索意图无关的模型则相应的减少。

## 5.6 本章小结

本章根据前面章节中关键技术研究 and 阐述，采用 .Net 框架实现了基于手绘草图的三维模型检索系统 Retrieval3D。对系统的总体结构、界面设计、系统功能模块划分、模块类结构设计及系统数据库设计等方面作了详尽的说明。最后，给出了系统的运行界面和检索结果。

## 第六章 总结与展望

### 6.1 总结

手绘草图是进行概念设计和思维表达的便捷方式,给设计师们提供了充分自由的创意空间,将手绘草图应用于三维模型的检索领域,使用户更加准确的表达个人的检索意图,有效提高了三维模型检索系统的交互便捷性,缩短了用户检索意图与表达方式之间的语义鸿沟。

本文阐述和研究了当前在草图识别领域和三维模型检索领域所应用的部分关键技术,并在此基础上将两个领域内的技术融合和改进,给出了基于手绘草图的三维模型检索系统的实现技术,通过构建基于手绘草图的三维模型检索系统对关键的技术进行实现。系统旨在提高用户检索过程中的人机交互性、三维模型检索性能和检索效果。本文主要研究工作总结如下:

(1) 给出了基于手绘草图的三维模型检索系统总体框架。系统主要包括草图笔画预处理模块、草图轮廓及特征提取模块、三维模型特征提取模块、基于贝叶斯网络模型的检索实现模块以及检索结果输出模块。各模块在系统中各自起着不同的作用:提取草图预处理模块主要完成用户绘制草图中的噪声去除,减少噪声对于后续过程的影响;草图轮廓及特征提取模块主要采用设计的基于特征点的多边形拟合算法对草图轮廓进行近似拟合,并通过拟合后的多边形提取草图轮廓的特征向量;三维模型特征提取模块主要完成三维模型投影图像获取、图像特征提取等;检索模块主要依据贝叶斯网络的推理特性通过低层的特征向量推理用户的检索意图,并通过相关反馈的机制实现系统检索性能的优化。

(2) 草图识别。草图识别是基于手绘草图的三维模型检索系统的基础环节,目的在于以草图的方式捕获用户的检索意图。本文中,在草图识别的基本原理、草图识别流程和草图识别关键方法等方面作了比较深入的研究,通过对草图的识别完成对用户检索意图的获取。

(3) 三维模型特征提取。完成了三维模型的预处理过程,根据三维模型自身特性通过计算选取投影平面后,获取模型在三个平面上的投影图像。利用改进的 SUSAN 算法对投影图像进行边缘提取,并利用选取的轮廓跟踪算法从获得的投影图像边缘中获取图像的轮廓信息,最后得到表示三维模型特征的特征向量。

(4) 构建了基于贝叶斯网络的推理模型。通过得到的三维模型特征向量与草图特征向量,构建了基于贝叶斯网络推理的模型,并利用样本数据对模型进行训练,使模型能够完成从低层特征向量到高层语义信息的推理过程。

(5) 给出了基于手绘草图的三维模型检索系统的设计与实现。根据本文中研究和改进的关键技术,实现了基于手绘草图的三维模型检索系统。对系统的主要模块划分、系统类结构设计、系统数据库设计等都作了详细的阐述,优化了检索系统人机交互界面和检索性能。通过实

践完成了对草图识别和三维模型检索领域内的关键技术运用。

## 6.2 展望

目前，基于手绘草图的三维模型检索系统虽然取得了一定的研究成果，但是仍存在许多问题，有待于进一步的探索和研究。本文中构建的基于手绘草图的三维模型检索系统，虽然在总体上对系统实现的关键技术进行了研究和实现，但仍处于一个探索的阶段，其中尚有不足之处，系统具体的细节还有待于进一步的优化。

以本文为基础，根据在系统构建过程所遇到的困难和挑战，需要进一步研究的工作主要有以下几个方面：

（1）草图识别及特征提取技术。基于手绘草图的三维模型检索系统中，草图识别及特征提取是系统能否拥有一个良好检索效果的前提条件。草图识别是对用户检索意图的抽象表示，只有正确理解了用户的检索意图，才能保证正确的检索结果。因此，寻找更为准确的草图特征提取方法来获得用户的检索意图是进一步研究的方向之一。

（2）特征向量索引的建立和组织。本文中并未对三维模型提取后的特征向量组织方法作具体深入的研究，而如何构建有效的特征索引能够在很大程度上改进系统的检索效率。因此，引入多维索引机制是今后的一个重点研究方向。

（3）系统功能扩展。可以从两个方面对基于手绘草图的三维模型检索系统进行功能扩展：一方面，优化系统交互界面，提供更为方便的交互方式；另一方面，将网页搜索引擎融入到模型检索系统中，以网站后台程序的方式管理从网络上搜索到的数据，并实现对这些数据进行检索的功能。

## 参考文献

- [1] Guetat G, Maitre M, Joly L, et al. Automatic 3d grayscale volume matching and shape analysis. *IEEE Transactions on information technology in biomedicine*, 2006, 10(2): 362~376.
- [2] Ankerst M, Kastenmuller G, Kriegel H P, et al. Nearest neighbor classification in 3D protein databases. *Proceedings of the Seventh International Conference on Intelligent Systems for Moducular Biology*, Heidelberg Germany, 2004: 34~43.
- [3] Alvarado C, Davis R. SketchREAD: a multi-domain sketch recognition engine. *Proceedings of User Interface and Software Technology(UIST04)*, 2004: 23~32 .
- [4] Casella M, Deufemia V, Martelli M. An agent-based framework for sketched symbol interpretation. *Journal of Visual Languages and Computing*, 2008, 19(2): 225~257.
- [5] Jia J, Z Qin, Q C Zhang, et al. An Overview of Content-Based Three-dimensional Model Retrieval Methods. *IEEE International Conference on System of Systems Engineering*. 2008: 217~222.
- [6] Vranie D, Saupel V. 3Dshape descriptor based on 3D fourier transform. In *Proceedings of the EURASIP Conference Digital Signal Processing for Multimedia Communications and Services(ECMCS2001)*, Budast Hungary, 2001: 271~274.
- [7] Alvarado C, Oltmans M, Davis R. A framework for multi-domain sketch recognition.. *Proc of AAAI Spring Symposium on Sketch Understanding*, 2002: 1~8.
- [8] Mankoff J, Abowd G, Hudson S. OOPS: A toolkit Supporting Mediation Techniques for Resolving Ambiguity in Recognition-based Interfaces. *Computers and Graphics*, 2000, 24(6): 819~834.
- [9] Gross M D, Do E Y. Drawing on the Back of an Envelope: A Framework for Interacting with Application Programsy Freehand Drawing. *Computers and Graphics*, 2000, 24(6): 835~849.
- [10] Lin J, Newman M, Hong J, et al. Finding a Tighter Fit Between Tools and Practice for Web Site Design, *Human Factors in Computing Systems*, 2000, 2(1): 510~517.
- [11] Fonseca M J, Pimentel C, Jorge J A. An Online Scribble Recognizer for Calligraphic Interfaces. *Proc of the AAAI Spring Symposium Sketch Understanding*, California USA, 2002: 51~58.
- [12] Tevfik Metin, Sezgin, Randall Davis. Early Sketch Processing with Application in HMM Based Sketch Recognition. *Massachusetts institute of technology, computer science and artificial intelligence laboratory*, 2004: 126~132.

- [13] Keshari B, Watt S M. Hybrid mathematical symbol recognition using support vector machines. Proceedings of the 9th International Conference on Document Analysis and Recognition, Curitiba Brazil, 2007: 859~863.
- [14] C.Zhang, T Chen. An Active Learning Framework for Content-Based Information Retrieval. IEEE Transactions on Multimedia Special Issue on Multimedia Database, June 2002, 4(2): 260~268.
- [15] M Kazhdan, T Funkhouser, S Rusinkiewicz. Shape Matching And Anisotropy AcmTrans Graph. 2004(23): 623 - 629.
- [16] R Osada, T Funkhouser, B Chazelle, D Dobkin. Shape Distributions.Acm Transactions On Graphics, 2002, 21(4): 807 - 832.
- [17] Kazhdan M, Funkhouser T, Rusinkiewicz S. Rotation invariant spherical harmonic representation of 3D shape descriptors. Proceedings of Symposium on Geometry Processing.Aachen, Eurographics Association Germany, 2003: 156 - 164.
- [18] Marcin Novotni, Reinhard Klein. 3D Zernike Descriptors for Content Based Shape Retrieval. Proceedings of the eighth ACM symposium on Solid modeling and application, Washington USA, 2003: 67~72.
- [19] 王勇睿, 李华, 崔丽等. 基于形状的三维模型特征提取和识别. 中国图形学大会, 2004: 58~63.
- [20] Yang Jun, Jiang Yugang. Evaluating Bag of Visual Words Representations in Scene Classification. Proceedings of International Workshop on Multimedia Information Retrieval, ACM Press, 2007: 197~206.
- [21] Hessellink H, Visser Menno, Roerdink Jos B T M. Euclidean skeletons of 3D data sets in linear time by the integer medial axis transform. ISMM2005, Paris France, 2005: 259~268.
- [22] Hilaga M, Shinagawa Y, Komura T, et al. Topologymatching for fully automatic similarity estimation of 3D shapes. Computer Graphics Proceedings Annual Conference Series, ACM SIG-GRAPH Los Angeles California, 2001: 203~212.
- [23] Zhang C, Chen T. Efficient feature extraction for 2d and 3d objects in mesh representation. Proceedings of IEEE International Conference on Image Processing, Thessaloniki, Greece, 2001: 935~938.
- [24] Suzuki, M. T, Yaginuma, Y. Kodama. A similarity retrieval technique for textured 3D models. Information Retrieval & Knowledge Management(CAMP), International Conference, 2010: 75~80.



- [25] Yaras F. Hoonjong Kang. Circularly configured multi-SLM holographic display system Transmission and Display of 3D Video (3DTV-CON), 2011: 86~92.
- [26] BoYu. Recognition of Freehand Sketches Using Mean Shift. Proceeding of the IUI0' 3, 2003: 204~210.
- [27] Chan-Yet Lai , Zakaria N. As sketchy as possible: Application Programming Interface (API) for sketch based user interface , Information Technology (ITSIM), 2010 : 1~6.
- [28] Wei Yongjun, Ji Guojie, Hou Suyuan, Chen Yingdong. Research on 3D symbolic representation of geographical information. 18th International Conference on Geoinformatics, 2010: 1~4.
- [29] Varanic D V, Saupe D. Description of 3D-shape using a complex function on the sphere. In: IEEE International Conference onMultimedia and Expo (ICME 2002), Lausanne Switzerland, 2002: 177~180.
- [30] Fan Yachun, Zhou Mingquan, Geng Guohua. Semantic-oriented 3D Model Retrieval Using Visual Vocabulary Labelling. International Joint Conference on Artificial Intelligence, 2009: 642~647.
- [31] Qianwei Bian, Yuanjun He, Weimin Luo, Hongming Cai. 3D model retrieval from multiple photographic images. IEEE International Conference on Progress in Informatics and Computing (PIC), 2010: 954~958.
- [32] Jun Hui Wang, Zhang Zhi yong, 3D Presentation and 3D Model Retrieval for industry. Third International Symposium on Electronic Commerce and Security (ISECS), 2010: 140~143.
- [33] Suzuki, M T , Yaginuma Y , Kodama. A similarity retrieval technique for textured 3D models. International Conference on Information Retrieval & Knowledge Management; (CAMP) , 2010: 131~137.
- [34] Cui, Chen yang, Liu Zhen, 3d Model Similarity Matching Based on Multiple Characteristic Views. Congress on Image and Signal Processing(CISP), 2008: 595~599.
- [35] Jean Philippe Valois , Mohamed Cheriet . Online recognition of sketched electrical diagrams. ICDAR 2001: 460~464.
- [36] Lowe D G. Distinctive image features from scale invariant key point. International Journal of Computer Vision, 2004, 60(2): 91~110.
- [37] 孙建勇, 金翔宇, 彭彬彬等. 一种快速在线图形识别与规整化方法, 计算机科学, 2003, 30(2): 102~106.
- [38] You Shin Park, Yong In Yun, Jong Soo Choi, A new shape descriptor using sliced image histogram for 3D model retrieval. IEEE Transactions on Consumer Electronics, 2009, 2:

240~247.

- [39] Atmosukarto I, Y Huang. Feature combination and relevance Feedback for 3D model retrieval. 11th International Multimedia Modeling Conference, 2005: 334~339.
- [40] Fan Yachun, Zhou Mingquan, Geng Guohua. Semantic-oriented 3D Model Retrieval Using Visual Vocabulary Labelling, International Joint Conference on Artificial Intelligence(JCAI '09), 2009: 642~647.
- [41] Zhi yong Zhang, Jian-qiu Jin, Fuzzy relevance feedback in content-based 3D model retrieval. Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) , 2010: 565~568.
- [42] Kim-Hui Yap, Kui Wu. Fuzzy relevance feedback in content-based image retrieval systems using radial basis function network. IEEE International Conference on Multimedia and Expo(ICME 2005), 2005: 4~8.
- [43] Kui Wul, Kim-Hui Yap. Content-based image retrieval using fuzzy perceptual feedback, Multimedia Tools and Applications, 2007, 32(3): 235~251.
- [44] Hema R, Omid M, Rosie I. Active Learning with Feedback on Both Features and Instances Journal of Machine Learning Research, 2006: 1655~1686.
- [45] Rayashi T, Kiyono T. Retrieval of 2D vector images by matching Weighted Feature Points. 15th IEEE International Conference on Image Processing(ICIP 2008), 2008: 961~964.
- [46] Xiwen Zhang, Jiqiang Song, Guozhong Dai. Extraction of line segments and circular arcs from freehand strokes based on segmental homogeneity features. IEEE Transactions on Systems, Man and Cybernetics, 2006: 300~311.
- [47] Foster I, Carl K, Tuecke S. The Anatomy of The Grid: Enabling scalable virtual organizations. International Supercomputer Applications, 2001, 15 (3): 200~222.
- [48] 王淑侠. 支持概念设计的手绘图在线识别研究. 西北工业大学研究生院, 2006: 135~139.
- [49] W B Heinzelman, A P Chandrakasan, H Balakrishnan. An Application-Specific Protocol Architecture for Wireless Microsensor Networks . IEEE Transactions on Wireless Communications (S1536-1276), 2002, 1(4): 660~670.
- [50] Hiroyasulehid ichida , Yuiehi Itoh, Yoshifumi Kitamura Fumio Kishino. Interactive Retrieval of 3D ShaPe Models using Physical Objects. 2004: 692~699 .
- [51] Tabbone S, Wendling L, Salmon J P. A new shape descriptor defined on the Radon transform. Computer Vision and Image Understanding, 2006, 102 (1): 42~51.
- [52] Russo F. New method for performance evaluation of grayscale image denoising filters. IEEE

- Signal Processing Letters, 2010, 17(5): 417–420.
- [53] Kinser J M. Image Signatures: Classification and Ontology. the 4th International Conference on Computer Graphics and Imaging. Hawaii, USA, 2001: 63~68.
- [54] Wei Liu, Zhengming Ma. Image Threshold Denoising Based on Edge Detection, IMACS Multiconference on Computational Engineering in Systems Applications, 2006: 72~78.
- [55] Chunhua Li, Kun He, Jiliu Zhou. Edge Detection of Image on the Local Feature, Second International Symposium on Intelligent Information Technology Application, 2008: 326~330.
- [56] 季虎, 孙即祥, 邵晓芳. 图像边缘提取方法及展望. 计算机工程用, 2004(14): 70~73.
- [57] Li Gang, Ju Yongfeng. Novel approach to pavement cracking detection based on morphology and mutual information. 2010 Chinese Control and Decision Conference, 2010: 3219 ~3- 223.
- [58] Xiwen Zhang, Jiqiang Song, Guozhong Dai. Extraction of line segments and circular arcs from freehand strokes based on segmental homogeneity features. IEEE Transactions on Systems, Man and Cybernetics, 2006: 300~311.
- [59] K Q Zhao, Y Wang. New Strategy to Improve Electromagnetic Torque at Starting in Thyristor Controlled InductionMotors. The 29thAnnualConference of the IEEE, 2003: 2555~2560.
- [60] M J Fonseca. Sketch-Based Retrieval in Large Sets of Drawings. Technical University of Lisbon, 2004: 128~132.

## 致 谢

回首南京航空航天大学两年半的研究生生涯，期间收获了很多，有欢乐也有泪水，自己也不断的成长起来。在毕业之际，怀念这一段美好的校园时光，留恋美丽的校园和良好的学习氛围。

感谢我的导师谢强副教授，衷心地感谢他在生活上给予我的帮助和学习上的指导。谢老师，其为人则温文敦厚、和善热心，其为师则循循善诱、细致有心，其为学则专注执着、严谨入微。在读研期间，谢老师赋予我大量的学习指导和提高综合能力的机会，不仅使我树立远大的学术目标、掌握了先进的科研方法，而且让我从实践中锻炼自己全局规划的能力。本论文从课题的选取、大纲的编排、论文的撰写无不受到谢老师教诲指点。在此，谨向谢老师表示崇高的敬意和衷心的感谢。

感谢英国皇家航空学会高级会员、南京航空航天大学信息科学与技术学院博士生导师丁秋林教授。丁秋林教授富有哲学的思想、渊博的知识、积极乐观的生活态度、崇高的人生观价值观、悉心耐心的教学方针以及高尚的人格魅力使我终身难忘。他教会了我以哲理思考人生洞察人生，以乐观活跃的态度处事处人。丰富的知识让自己的生活多姿多彩，健康活跃让自己拥有持久的年轻。

感谢周良老师和郑洪源老师在我研究生学习期间所给予的耐心和细心地指导。

感谢实验室同窗好友在我学习期间的帮助和指导，感谢师弟师妹的在论文写作过程中给我提出的宝贵意见。短短的两年半，我们一起学习，你们不同的优秀品质影响着我，我非常珍惜共同度过的这段宝贵时光，将终生难忘。

感谢我的室友在生活中对我的关心和帮助。因为你们，我勇敢地面对挫折，在困难面前永不言败。将永远记住与你们一起度过的欢乐时刻，是我记忆中最珍贵的回忆。

最后，谨向百忙中抽出宝贵时间来为我评审本文而付出辛勤劳动的各位专家教授和参与我论文答辩的老师表示衷心的感谢！

## 在学期间的研究成果及发表的学术论文

### 攻读硕士学位期间发表（录用）论文情况

- [1] 张海龙, 谢强, 丁秋林. 基于贝叶斯网络的草图符号识别研究. 中国制造业信息化(已发表).
- [2] 张海龙, 谢强, 丁秋林. 基于对象关系数据库的患者病情图像存储研究. 第 25 届南京地区研究生通信年会（已发表）.

### 攻读硕士学位期间参加科研项目情况

- [1] 2009 年 10 月到 2010 年 4 月 负责“某公司货车快速设计系统”三维模型检索及构建模块。
- [2] 2010 年 5 月到 2011 年 9 月 负责“无线激光打靶系统”系统设计、模块开发及系统维护。