融合双尺度特征的草图化三维零件库检索方法

黄沈权, 王宗明, 张勇, 于鲁川, 陈子瑞

(温州大学机电工程学院 温州 325035)
(hshenquan@wzu.edu.cn)

摘 要: 基于草图的三维模型检索是机械零件库中三维模型检索的重要途径. 草图和三维模型之间存在着巨大的模态差异, 针对三维模型投影出的多个视图的信息冗余, 零件视图提取的特征的表征能力弱的问题, 为了提高基于草图的零件三维模型检索的效果, 通过融合零件模型的局部和全局的双尺度特征, 提出一种草图化三维机械零件库检索方法. 首先提出基于图像熵的三维模型投影视图表达方法, 减少多个视图之间的冗余性; 然后提出面向零件视图边缘轮廓提取的 mechanical sketch 模型, 深度提取零件的轮廓草图, 以逼近视图和草图的相似性; 最后构建视觉词袋模型和改进的多视图卷积神经网络模型来分别提取轮廓草图的局部特征和全局特征, 通过融合双尺度特征, 采用欧几里得距离进行基于草图的零件三维模型匹配. 在扩充的 ESB 数据集上的实验结果表明, 所提方法在准确率和召回率等系列评价指标上, 相较文中对比的 6 种方法, 有了较为明显的提高, 验证了所提方法的可行性和有效性.

关键词: 零件草图; 三维模型; 边缘轮廓; 特征提取; 特征融合; 检索系统
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2023-00732

Research on Retrieval Method for Sketching 3D Part Library by Fusing Dual Scale Features

Huang Shenquan, Wang Zongming, Zhang Yong, Yu Luchuan, and Chen Zirui

(*College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou* 325035)

Abstract: Sketch-based 3D model retrieval is an important way to retrieve 3D models in mechanical part library. There are huge modal differences between sketches and 3D models. Aiming at the information redundancy of multiple views projected by the 3D model, the weak representation ability of the features extracted from the part views, in order to improve the effect of sketch-based 3D model retrieval, a sketch-based 3D mechanical part Library retrieval method is proposed to fuse the local and global dual-scale features of the part model. Firstly, a projection view representation method of 3D model based on image entropy is proposed to reduce the redundancy among multiple views. Then, a mechanical sketch model for edge contour extraction of part view is proposed, and the contour sketch of the part is extracted in depth to approximate the similarity between the view and the sketch. Finally, a visual word bag model and an improved multi-view convolutional neural network model is constructed to extract the local and global features of the contour sketch respectively. By fusing the dual-scale features, Euclidean distance is used to match the 3D model of the part based on the sketch. The experimental results on ESB datasets show that the proposed method is significantly improved in a series of evaluation indicators such as accuracy and recall rate compared with the existing 6 common methods, which verifies the feasibility and effectiveness of the proposed method.

收稿日期: 2023-11-11; 修回日期: 2024-03-27. **基金项目**: 国家自然科学基金(52305549, 71501143); 浙江省自然科学基金(LY19G010007); 温州市基础性科研项目(20240016). **黄沈权**(1986—), 男, 博士, 副教授, 硕士生导师, 主要研究方向为大批量定制、服务型制造; **王宗明**(1997—), 男, 硕士研究生, 主要研究方向为制造业信息化; **张勇**(1997—), 男, 硕士研究生, 主要研究方向为知识图谱; **于鲁川**(1993—), 男, 博士, 讲师, 硕士生导师, 主要研究方向为工业机器人; **陈子瑞**(2001—), 男, 硕士研究生, 主要研究方向为图像识别.

Key words: part sketches; 3D models; edge contours; feature extraction; feature fusion; retrieval system

随着 CAD 系统在现代工业中的日益普及，大量的三维数字模型被生成并存储在互联网或企业数据库中，形成零件库[1-2]. 产品设计师可以通过零件库查阅和调用所需零件的 CAD 模型，加速自己的产品开发，降低设计成本，提高产品质量；对零件供应商而言，零件库是他们宣传和销售自己产品并与客户沟通协作的平台[3]. 然而，在数据巨大、种类繁多的零件库中查找到所需的三维 CAD 模型是一项耗时耗力的过程，传统的基于文本、特征描述的三维模型检索方法不能满足快速检索的需求，同时还存在特征描述不准确、三维模型名称歧义等问题. 因此，如何从海量的零件模型中快速、有效地检索出用户所需的模型，已成为产品设计领域的一个迫切需求[4-5].

近年来，基于草图的三维模型检索具有灵活表达设计者意图的优势，逐渐受到研究者的关注. 目前的研究大致分为 2 个方向：(1) 针对三维模型本身的轮廓形状、点云、体素、几何相似性等对三维模型进行处理生成特征. Shao 等[6]提出一种用于细粒度三维点云分类的弱超视网络，采用上下文感知辨别特征提取模块，分层提取不同感受野的上下文对比信息，从点云中捕获判别性局部细节. 这类方法计算量较大，主要是提取三维模型的局部特征，局部特征指三维模型局部区域所具有的特征，如轴承的滚珠. (2) 基于视图的方法. 首先利用投影的方法把三维模型投影为二维视图；然后从这些二维视图中提取深度特征，得到一个紧凑表征三维模型的特征，紧凑表征指将复杂的三维模型转换为较低维度的特征表示. 研究表明，基于视图的方法比基于模型本身的方法获得了更好的性能[7]. 刘杨圣彦等[8]提出面向三维模型视图特征提取的残差卷积网络优化算法，采用加权损失函数提高了视图特征的可分性; Yang 等[9]提出一种新颖的顺序学习框架，采用三维模型网络模块学习三维模型表示，采用二维草图网络模块学习二维草图表示.

在三维模型的视图表示方面，刘玉杰等[10]提出一种基于手绘图像融合信息熵和卷积神经网络(convolutional neural networks, CNN)的三维模型检索方法，通过计算模型投影图的一维信息熵得到模型的代表性视图；Lei 等[11]提出一种代表性视图选择模块，用于获取三维形状的最具代表性的视图；Li 等[12]提出基于视点熵的三维信息复杂度测量方法指导三维模型的自适应视图聚类，可以筛选出一组具有代表性的视图样本. 虽然一些学者已考虑多视图信息的冗余问题，但主要采用单一信息熵或复杂度测量来剔除冗余视图，未充分考虑通过多维信息熵结合的方式提高视图表达的效果.

当前，视图边缘提取普遍采用 Canny 算子、Prewitt 算子等传统边缘检测算法. Manda 等[13]提出 Canny 边缘检测和高斯模糊思想的加权组合，用于生成计算机草图; Tan 等[14]采用 Gaussians 差方法提取三维模型投影深度图像上的线图，然后采用 Bézier 曲线优化提取的线图；Yu 等[15]提出基于 Berkeley 的改进边缘提取算法，根据每个二进制点的二进制边缘图的灰度值生成边缘图. 基于 Canny, Prewitt 等算子的传统边缘提取方法虽然有着不错的效果，但不可避免地保留了大量不必要的噪声；另一方面，与普通物品的三维模型相比，零件三维模型只是整个装配体结构的一部分，具有不完整性，因此单个零件的投影轮廓往往不够清晰，独立于装配体的单独识别较为困难. 所以，如何有效地处理噪声，减少视图和草图之间的差距，增强模型对机械零件的识别和理解能力是一个关键难题.

在三维模型的特征提取方面，现有研究采用单一网络模型来提取特征. Bai 等[16]提出一个域间增强网络 Inter-DAN，采用域间特征相关学习捕获跨域互信息，以学习草图和三维形状的增强通用全局特征，取得了良好的检索效果；白静等[17]设计了基于多视图融合的三维模型特征提取网络，通过融合不同视角的特征提取更全面的三维模型特征；Zhao 等[18]提出一种模态共享特征提取网络，利用模态特定特征和模态共享信息提取特征：以上方法在三维模型视图特征提取上取得了重要进展，然而与普通三维模型不同，零件三维模型通常包含大量特殊的局部几何特征，如存在孔洞、微凸、倒角、螺纹和尖锐边缘等，现有的特征提取方法在三维模型的视图特征提取上往往忽略了视图的局部特征(指视图中小范围区域的局部信息、形状等)和全局特征的集成融合，而局部特征对于零件三维模型的特征提取非常重要.

与普通三维模型相比，零件三维模型不仅反映了零件的结构，还体现了零件的参数、装配和约束等关系，但缺乏渲染等视觉效果. 考虑零件三维模型检索的特殊性，针对草图和零件三维模型间的巨大模态差异，为了消除两者之间的差距，提高

检索的精度，本文提出融合双尺度特征的草图化三维零件库检索方法．首先将三维模型投影为多个二维视图，为了消除冗余视图对检索性能的影响，提出基于联合图像熵的投影视图表达方法；为了逼近零件视图和草图的相似性，引入 Sobel 算子构建一种生成对抗网络(generative adversarial networks，GAN)，提出基于 mechanical sketch 的投影视图轮廓提取技术，增强检索模型对机械零件的识别和理解能力；对于已经完成草图化的视图，建立融合双尺度的草图特征提取模型，采用多视图 CNN(multi-view CNN，MVCNN)模型提取全局特征，针对提取到的全局特征容易漏掉一些关键的局部特征的问题，采用视觉词袋模型(bag of visual words，BOVW)提取其局部特征，进而将局部特征融入全局特征作为对应模型的聚合特征描述子，采用欧几里得距离进行基于草图的零件三维模型匹配，综合发挥检索模型的鉴别能力和泛化能力，提高检索效果．

1　本文方法概述

本文提出的融合双尺度特征的草图化三维零件库检索方法的框架如图 1 所示，包括 4 个核心模块：基于图像熵的三维模型投影表达、基于 mechanical sketch 的投影视图轮廓提取、局部特征与全局特征耦合的轮廓草图特征提取、基于欧几里得距

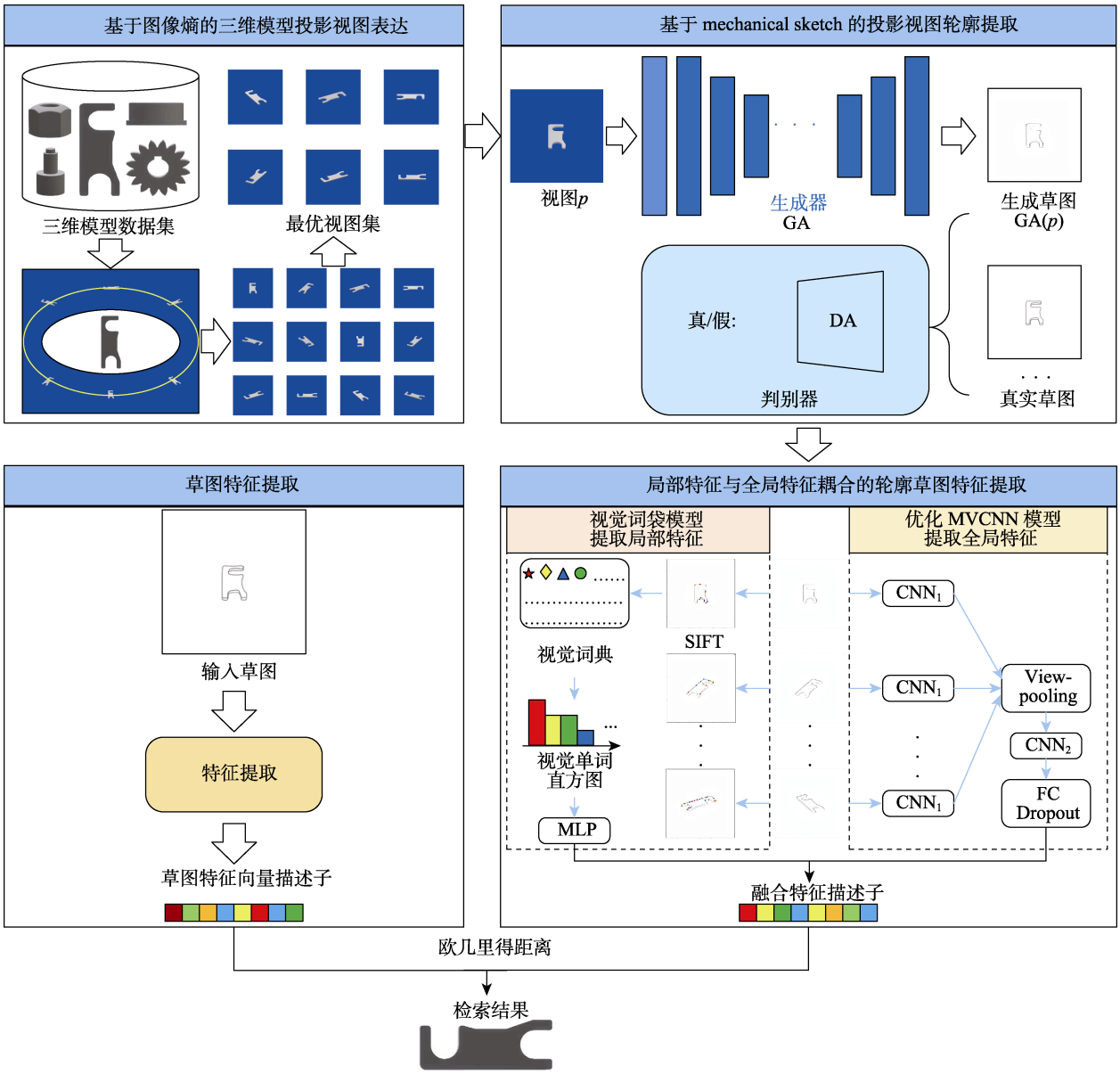


图 1　本文方法框架

离的草图和轮廓草图相似度计算.

各模块之间的执行过程与关系如下.

(1) 基于图像熵的三维模型投影视图表达. 采用固定投影的方法对三维模型进行投影形成多视图, 为了消除冗余视图的影响, 采用联合图像熵的方式从多个视图中选择最优视图集.

(2) 基于 mechanical sketch 的投影视图轮廓提取. 为了高精度地提取边缘轮廓, 缩小边缘轮廓与草图的差距, 构建面向零件视图边缘轮廓提取的 mechanical sketch 模型, 通过引入 Sobel 算子构建一种 GAN, 对边缘轮廓进行递进式的深化提取, 将其转化为轮廓草图.

(3) 局部特征与全局特征耦合的轮廓草图特征提取. 构建 BOVW 提取轮廓草图局部特征, 改进 MVCNN 提取轮廓草图全局特征, 并把局部特征和全局特征进行耦合, 得到一个更具表征力的聚合特征描述子.

(4) 基于欧几里得距离的草图和轮廓草图相似度计算. 针对输入的草图, 采用本文提出的双尺度特征提取模型提取草图特征述子. 对草图特征向量描述子和模型库中的三维模型聚合特征描述子使用欧几里得距离进行相似性度量. 草图和三维模型间的距离计算公式为

$$d(\boldsymbol{X},\boldsymbol{Y})=\sqrt{\sum_{i=1}^{n}(\boldsymbol{x}_i-\boldsymbol{y}_i)^2}\ .$$

其中, $\boldsymbol{X}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\}$ 和 $\boldsymbol{Y}=\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\}$ 分别表示任意草图和三维模型的特征描述符. $d(\boldsymbol{X},\boldsymbol{Y})$ 越小, 表示 2 个特征向量越相似, 根据距离从小到大依次排序并输出匹配结果.

2 基于图像熵的三维模型投影视图表达方法

采用固定投影的方法对三维模型进行投影得到多个视图. 例如, 假设三维模型的摆放是按照恒定的 z-轴竖直摆放的, 此时虚拟相机按照事先摆放好的位置环绕, 与摆放三维模型的水平面有 30°的夹角, 其朝向方向是径直指向三维模型的中心, 以每隔 30°的方式环绕三维模型 360°即可以产生 12 个不同的视图; 但是, 产生的多个视图存在信息冗余, 一些视图不能够有效地表达出对应三维模型的特征. 为了提高模型特征表达的精度和效率, 需要去除冗余视图, 保留信息较多的视图.

针对上述问题, 本文采用联合图像熵判别三维模型投影视图包含的信息量大小, 选取最优视图集. 图像熵是一种统计图像特征的方式, 目前常用的基于图像熵的计算方法有一维熵和二维熵. 一维熵表示图像中灰度分布的聚集特征所包含的信息量, 其计算公式为

$$h_1=-\sum_{i=0}^{255}P_i\times\mathrm{lb}P_i\ .$$

其中, h_1 表示一维熵, 非负值; P_i 表示图像中灰度值为 $i(0\leqslant i\leqslant 255)$ 的像素所占的比例. 一维熵计算比较简单, 用于描述像素值的分布情况, 具有一定的旋转和缩放不变性, 但其不能反映图像灰度分布的空间特征.

二维熵是对图像中每个像素及其周围像素的灰度值进行概率分布计算, 进而计算出图像的熵值, 其计算公式为

$$h_2=-\sum_{i=0}^{255}\sum_{j=0}^{255}P_{ij}\times\mathrm{lb}\,P_{ij}\ .$$

其中, h_2 表示二维熵; $P_{ij}=f(i,j)\big/N^2$ 表示图像中 $f(i,j)$ 的发生的概率, i 表示像素的灰度值 $(0\leqslant i\leqslant 255)$, j 表示邻域灰度均值 $(0\leqslant j\leqslant 255)$, (i,j) 表示特征二元组, $f(i,j)$ 表示特征二元组 (i,j) 出现的频数, N 表示图像的尺度.

二维熵能够更好地反映图像的整体空间特征, 但其对图像的分辨率、旋转和缩放等比较敏感; 而一维熵描述的是像素值的分布信息, 由于旋转和缩放不会改变图像中像素值的分布特征, 因此具有不变性. 为了兼顾图像的整体空间特征和不变性, 本文提出联合图像一维熵和二维熵的方法, 分别赋予视图一维熵和二维熵不同的权重 ω_1 和 ω_2, 对两者进行权重融合相加计算综合的熵值 T, 将熵值 T 按照从大到小的顺序作为筛选视图的标准, 选择具有较高熵值的视图作为三维模型的表达. 联合图像熵的最优视图集选取流程如图 2 所示.

3 基于 mechanical sketch 模型的投影视图轮廓提取技术

为了使得到的边缘轮廓接近实际的草图效果, 本文借鉴动漫领域的 anime2sketch 算法[19]的思想, 通过融入传统的 Sobel 边缘检测技术, 提出一种面向零件视图边缘轮廓提取的 mechanical sketch 模型, 其结构如图 3 所示. 该模型可以实现快速轮廓提取, 学习到零件视图装配、约束关系等更高级的

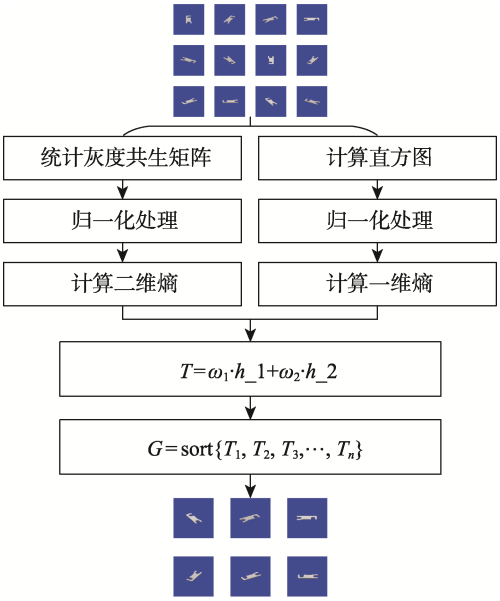


图 2　联合图像熵的最优视图集选取流程

特征，避免冗余的线条对整体轮廓的影响，使得到的轮廓草图具有更强的泛化能力.

边缘轮廓是零件视图中最富含信息的部分之一，而在边缘轮廓提取的算法中，传统的 Sobel 算法被广泛采用，具有计算简单和快速的优点；但在机械零件设计领域，该算法提取的边缘和实际的草图仍有差距，且存在不必要的线条和噪声等泛化能力弱的问题. 为了解决这一问题，本文针对机械零件的特点，基于 GAN 提出 mechanical sketch 模型，以减少视图和草图的差异性. mechanical sketch 模型的结构如图 3 所示，包含 Sobel 算子、U-Net 结构和 PatchGAN 结构. mechanical sketch 模型将 Sobel 算子作为特殊的一层卷积融入 U-Net 结构; U-Net 结构作为该模型的生成器，用来将输入的零件视图生成轮廓草图；而 PatchGAN 结构作为该模型的判别器，用来判别生成器生成轮廓草图的质量.

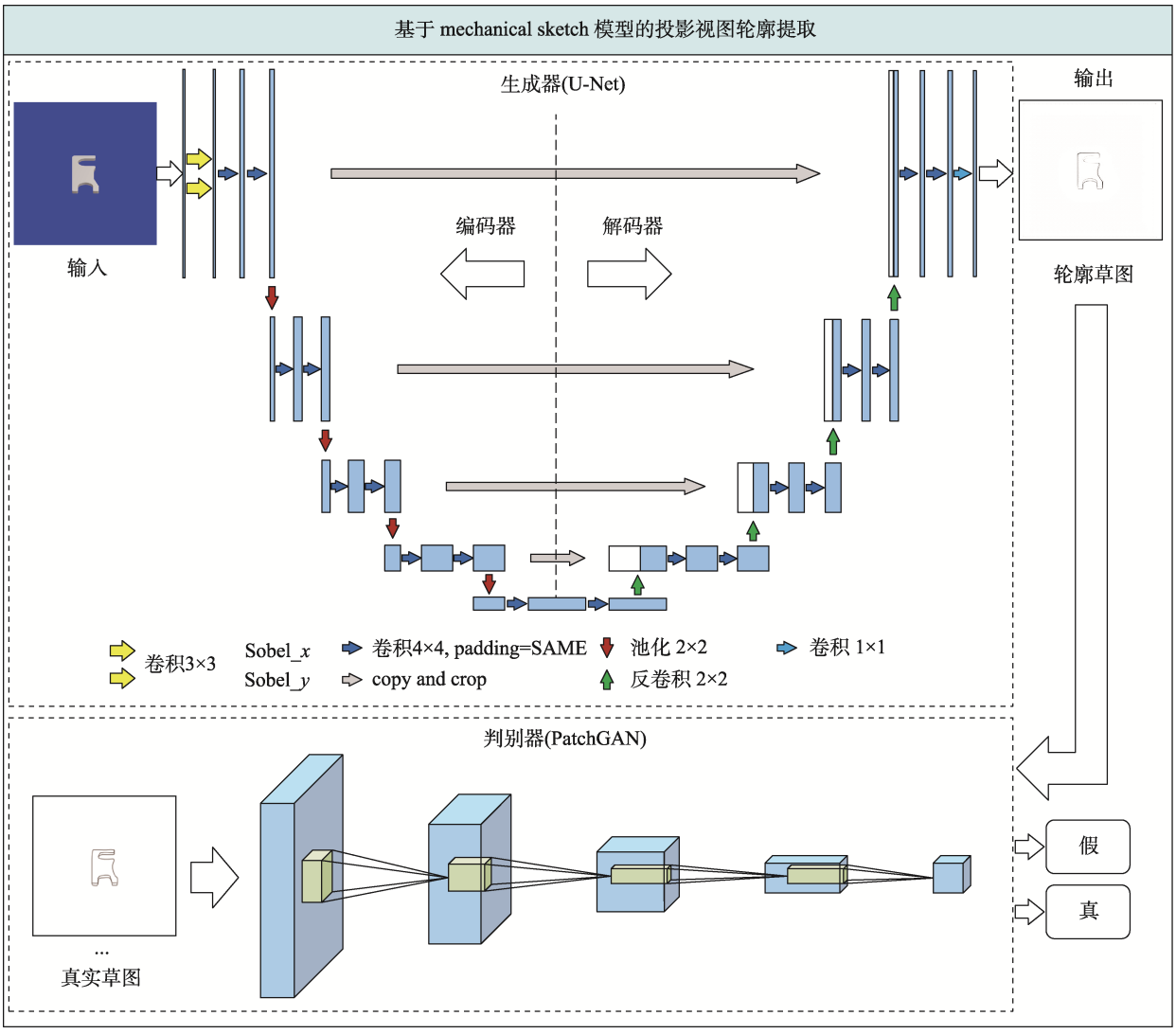


图 3　mechanical sketch 模型结构

将零件视图作为输入，经由 U-Net 结构编码器部分提取特征(即 U-Net 左半边)，由卷积和下采样(池化层)组成. 在卷积部分，引入 Sobel 算子作为第 1 层卷积，该算子包含水平和垂直 2 个方向的矩阵，分别为 $\begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix}$ 和 $\begin{bmatrix}-1&-2&-1\\0&0&0\\1&2&1\end{bmatrix}$. 将上述矩阵和零件视图作卷积，得到视图像素点横向及纵向的亮度差分近似值，通过计算亮度差分近似值较大的像素点检测到视图的边缘. 将 Sobel 算子作为第 1 层卷积，可以让模型初步学习到视图的边缘信息，简化后续卷积层的学习，提高模型的训练效率和性能. 为了适应对零件视图边缘轮廓的特征提取，提高网络的特征提取能力，本文将普通 GAN 中卷积核的大小由 3×3 改为 4×4. 下采样过程中，通道数翻倍，使图像不断缩小，提取到浅层信息. 解码器部分(U-Net 右半边)采用卷积和上采样(反卷积). 上采样过程中通道数减半，使图像不断扩大，提取深层信息. 在编码器和解码器中间使用 copy and crop 快捷连接，融合 U-Net 左右两边的特征，把浅层特征和深层特征融合起来.

在 U-Net 生成器进行卷积计算过程中，padding 指在输入图像的周围添加额外的像素值，这样做可以控制输出特征图的尺寸，并且有助于保留图像边缘信息. 在 padding 的过程中，有“SAME”和“VALID” 2 种方式，padding=“SAME”时，在图像的周围填“0”，达到保留图像边缘信息的效果；padding=“VALID”时则不需要在图像的周围填“0”. VALID 方式的优点是降低特征图的尺寸，减少网络参数和计算量. 然而,本文中图像边界的信息十分重要，为了尽可能地保留图像边缘信息并使输入/输出图像尺寸一致，将卷积方式设置为 SAME.

PatchGAN 判别器接收生成器生成的轮廓草图作为输入，将生成的轮廓草图分割成多个小块在局部区域上进行判别，评估每个小块的真实性. 生成器根据判别器的反馈信号调整生成的轮廓草图，逐渐接近真实的草图零件. 通过 PatchGAN 结构的判别，生成器可以受益于局部细节的评估和优化，生成更真实准确、更能体现零件模型间装配约束等高级特征的轮廓草图.

4 局部特征与全局特征耦合的轮廓草图特征提取模型

为了实现草图和轮廓草图的精确匹配，需要提取生成的轮廓草图的深层特征. 针对单个模型提取的特征表征力弱的问题，本文通过 BOVW 提取轮廓草图的局部特征，通过优化的 MVCNN 模型提取轮廓草图的全局特征，提出局部特征与全局特征耦合的轮廓草图特征提取模型，以提高轮廓草图特征的提取效果.

4.1 基于 BOVW 的局部特征提取

引入多层感知机(multilayer perceptron, MLP)到 BOVW，该模型提取的特征是视图中具有一定代表性的小区域或关键点，如零件模型中普遍存在的孔洞、微凸等不明显但重要的特征，本文称之为局部特征. 基于 BOVW 局部特征提取过程如图 4 所示.

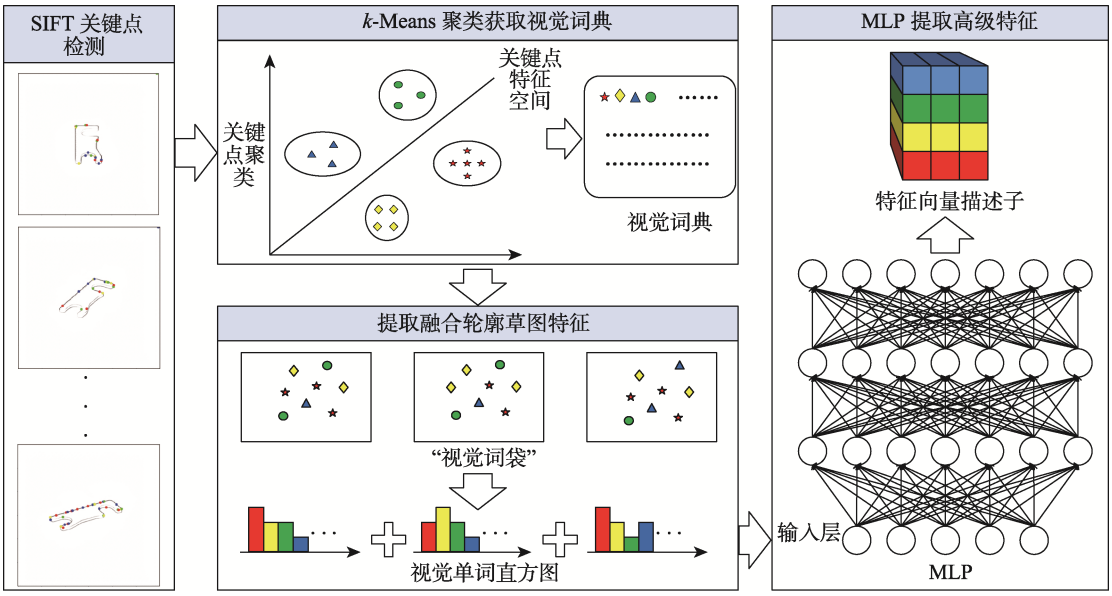


图 4 基于 BOVW 的局部特征提取过程

具体步骤如下.

Step1. 采用尺度不变特征转换(scale-invariant feature transform, SIFT)算法提取所有轮廓草图的 SIFT 关键点，并计算关键点的特征描述符. 描述符捕捉了关键点周围区域的梯度方向和强度等信息，并将其表示为一个向量. SIFT 是一种尺度不变的特征提取算法，具有很好的旋转不变性、尺度不变性、光照不变性和鲁棒性，可以检测到零件视图中不明显的特征和孔洞、尖锐边缘等关键点，并生成对应的特征描述符.

Step2. 基于 *k*-Means 聚类将每个描述符分配给其最近的聚类中心. 首先从所有样本点中随机选择 k 个作为聚类中心；然后计算每个样本点到聚类中心的欧几里得距离，并将其分配到最近的类中；再将每个类中样本点的均值定义为新中心点，作为下一个迭代的聚类中心；重复上述分配和更新，直到聚类中心不再变化，完成聚类，最终得到 k 个聚类中心(即为 k 个视觉词)，所有视觉词组成视觉词典.

Step3. 提取每个模型对应的视图轮廓草图融合特征. 首先计算每个轮廓视图的样本点到所有聚类中心的距离；然后找到距离最小的聚类中心，此样本点则归属于此聚类中心，对应的视觉单词直方图的数字加 1，每个样本点可近似用一个视觉词代替；再对所有的样本点重复以上过程，得到一个轮廓视图的特征直方图，因此每个轮廓视图都可以用特征直方图表示；最后把一个三维模型的多张视图轮廓草图的特征直方图合并为一个融合特征，即每个模型可以用多个特征直方图表示，所有模型的特征直方图构成直方图矩阵，该矩阵如表 1 所示，其中，P_{N-k}^{M} 表示三维模型 N 的第 M 个视图在视觉词 k 下的频数.

Step4. 基于 MLP 提取直方图高级特征. 特征直方图包含局部特征的频率信息，缺乏对特征本身的表征能力. 本文引入 MLP 网络，将 BOVW 提取的特征直方图矩阵作为输入，通过 MLP 网络的多层非线性变换，从原始局部特征中学习到更加丰富和抽象的特征表示，如形状信息等，提高特征直方图的表征力，最终输出轮廓草图的高级特征描述子. 本文提出的 MLP 网络是一个 3 层的 MLP，有 2 个隐藏层，且各层之间的连接是全连接的，每个节点到上一层的每个节点都有连接，使用反向传播算法训练网络参数.

表 1　三维模型视觉单词直方图矩阵

模型序号	直方图 1				直方图 2				直方图 M			
	视觉词 1	视觉词 2	…	视觉词 k	视觉词 1	视觉词 2	…	视觉词 k	视觉词 1	视觉词 2	…	视觉词 k
模型 1	P_{1-1}^{1}	P_{1-2}^{1}	…	P_{1-k}^{1}	P_{1-1}^{2}	P_{1-2}^{2}	…	P_{1-k}^{2}	P_{1-1}^{M}	P_{1-2}^{M}	…	P_{1-k}^{M}
模型 2	P_{2-1}^{1}	P_{2-2}^{1}	…	P_{2-k}^{1}	P_{2-1}^{2}	P_{2-2}^{2}	…	P_{2-k}^{2}	P_{2-1}^{M}	P_{2-2}^{M}	…	P_{2-k}^{M}
⋮	⋮	⋮	…	⋮	⋮	⋮	…	⋮	⋮	⋮	…	⋮
模型 N	P_{N-1}^{1}	P_{N-2}^{1}	…	P_{N-k}^{1}	P_{N-1}^{2}	P_{N-2}^{2}	…	P_{N-k}^{2}	P_{N-1}^{M}	P_{N-2}^{M}	…	P_{N-k}^{M}

4.2　基于 MVCNN 优化的全局特征提取

MVCNN[20]是一种用于三维模型分类的网络架构，能够处理多个视角的输入图像. 与其他网络相比，MVCNN 提取多个视角输入图像中每幅图像的整体特征，不依赖于特定的局部区域或关键点，因此称 MVCNN 提取的特征为全局特征. 全局特征能够全面地描述图像的整体内容，提供更丰富的表达能力. 采用包括 5 个卷积层(依次记为 $\mathrm{CONV}_{1\sim5}$)、3 个全连接层(依次记为 FC6，FC7 和 FC8)和一层特殊的 View-pool 聚合层的基本网络架构，将全连接层中的 FC7 作为该模型的聚合特征描述子. 本课题组在前期实验中发现，该网络架构应用到零件视图特征提取时，在计算复杂度、参数调节和多类别分类方面存在缺点，同时存在过拟合的风险.

针对上述问题，面向零件视图全局特征提取的要求，本文构建改进的 MVCNN 来增强网络的适应性和特征提取能力，其结构如图 5 所示，包括多视角编码模块、多视角融合模块和全局特征提取模块. 首先将多个轮廓草图作为输入，经过多视角编码模块的 $\mathrm{CONV}_{1\sim5}$ 分别进行卷积操作；然后在多视角融合模块的 View-pool 层进行聚合；再经过全局特征提取模块的 3 个全连接层提取特征；最后，经过交叉损失函数进行分类输出.

MVCNN 模型的每个卷积层中包含了局部响应归一化(local response normalization, LRN)，它是一种用于控制模型复杂度并减少过拟合的常用技术. LRN 层应用于零件视图特征提取时，由于存在计算复杂度高、难以调节参数的问题，因此本文将其改为 Dropout. Dropout 舍弃 50%的神经元，减少了参数量，可以起到防止过拟合的效果，提高模型的泛化能力.

MVCNN 模型中采用的支持向量机(support vector machine, SVM)损失函数通常用于二分类，其对多分类问题处理效果差，而本文提出的三维零件库检索是一个多分类问题. 为了更好地适应多分类任务，本文将 SVM 损失函数改进为简化的交叉熵函数，直接对标签和卷积层提取出来的特

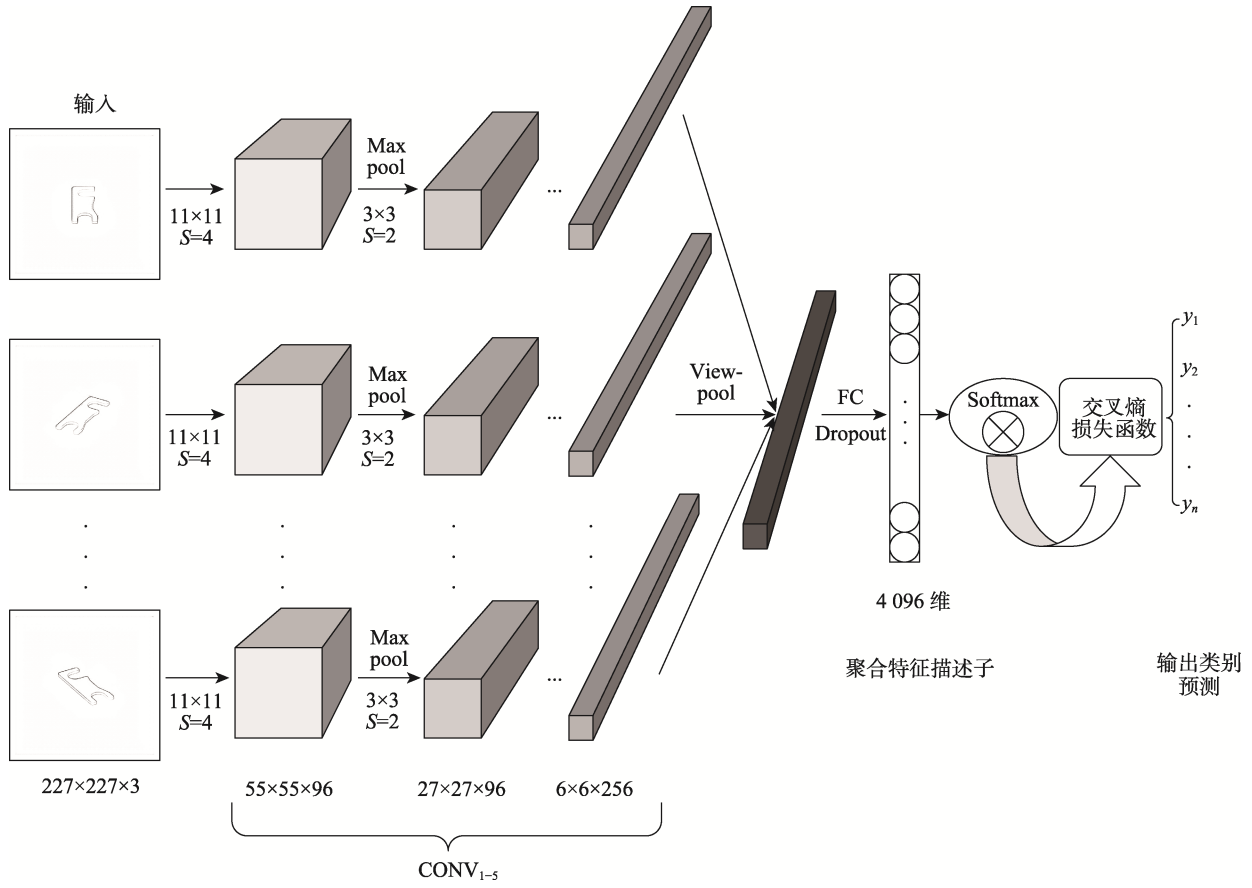


图 5 基于 MVCNN 优化的全局特征提取网络

征图进行运算. 首先通过 Softmax 函数进行分类, 然后使用交叉熵损失函数度量模型的预测结果与真实标签之间的差异, 即计算损失函数值, 以提高特征提取模型的训练效率和准确性.

4.3 基于特征融合的双尺度特征描述子提取

局部特征捕获零件多视图图像中的局部形状等细节信息, 如对孔洞、微凸和倒角等不明显的特征具有较高的辨别能力. 通过将注意力聚焦于局部特征, 可以提高模型的鉴别能力和对视角变化的鲁棒性; 全局特征则捕获零件多视图图像的整体结构和语义信息, 反映了图像的整体形状和布局等; 通过考虑全局特征, 模型可以更好地理解多视图图像的整体语义, 并具备更强的泛化能力. 为了综合发挥模型的鉴别能力和泛化能力, 本文融合局部特征和全局特征提取多视图图像的双尺度特征描述子, 其过程如图 6 所示.

特征融合方式一般有 add 和 concat[16]这 2 种, 其中, add 方式是在维度不变的情况下使视图特征在每维下的信息量增多; 而 concat 方式是通道数的合并, 即视图特征维度增加, 而每个维度特征下的信息量不变. 本文采用 concat 方式将 BOVW 提取的局部特征向量和 MVCNN 提取的全局特征进

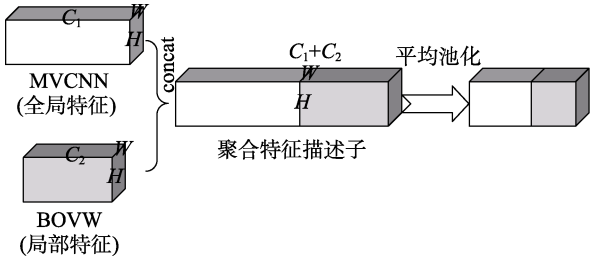


图 6 基于特征融合的双尺度特征描述子提取过程

行融合, 最大限度地保留原始信息. 为了适当减少特征维度, 提高鲁棒性, 双尺度特征融合之后构建平均池化层降低维度, 得到最终的双尺度特征描述子, 实现局部特征和整体特征的有效互补和融合.

利用上述双尺度特征提取模型, 分别提取输入草图的特征描述子和零件库中三维模型的特征描述子, 将两者基于欧几里得距离进行相似性度量, 根据相似性大小输出匹配结果, 最终实现基于草图的机械零件三维模型检索.

5 实验及结果分析

5.1 数据集

为了验证本文方法性能, 使用 ESB(engineering

shape benchmark)数据集[21]进行评估. 该数据集包含 867 个模型，共有 42 个类别的零件，其中有 66 个模型被分在 3 个杂项类中. 由于杂项模型是无法进一步细分的模型，会干扰模型的训练，因此实验中予以忽略. 另外，从 TraceParts 和 3DSource 零件库中下载了 254 个零件模型，作为数据扩充. 最终，本文实验数据集中共有 1 055 个零件模型.

5.2　实施细节

本文实验在 Intel i7 CPU，24 GB 内存和一张 NVIDIA GeForce GTX 1650 GPU 的服务器上运行；软件环境为 python3.6.13 和 tensorflow1.14.0.

实验中，将数据集按照 6∶2∶2 的比例划分为训练集、验证集和测试集. 在联合图像熵的视图表达实验中，一维熵权重 $\omega_1=0.4$，二维熵权重 $\omega_2=0.6$. 在基于 mechanical sketch 的投影视图轮廓提取实验中，分别设置输入图像的大小为 256×256×3 像素，批次大小为 1，学习率为 2×10^{-4}，并使用 Adam 算法进行参数优化，epoch 数为 200. 在改进的 MVCNN 模型上进行训练时，每个模型选取 6 个视图作为三维模型的表达，图像数据分辨率调整为 227×227×3 像素，批次大小设置为 4，设置学习率为 10^{-6}，交叉熵损失采用 sparse_softmax_cross_entropy_with_logits 函数，并使用 Adam 算法进行参数优化，对模型文件每经过 1 000 次迭代保存一次到检查点文件，epoch 数为 500. 在对测试集进行评估时，首先将测试集中的所有草图依次输入到本文的检索模型中，对每个查询草图根据相似度对检索出的模型进行排序；然后统计测试集中所有草图检索结果的评价指标的平均值，将其作为最终的评价结果.

5.3　评价指标

为了充分地评估模型检索性能的多个方面，采用三维模型检索中常用的 6 个评价指标[22]评估实验结果.

(1) 最近邻(nearest neighbor, NN). 表示查询到的最准确的第 1 个返回结果的检索精度.

(2) 第 1 批次(first tier, FT). 表示查询到的前 k 个检索结果的召回率，k 表示数据集中相关模型的个数.

(3) 第 2 批次(second tier, ST). 表示查询到的前 $2k$ 个检索结果的召回率.

(4) E 度量(E-measure, E). 是一种对查全率和查准率进行调和平均数求解后得到的综合度量值，可以更全面地评估检索系统的性能.

(5) 折扣累计增益(discounted cumulative gain, DCG). 是一种综合考虑检索结果相关性和排名质量的指标，其以相关性作为正比例系数为检索结果分配权重，并对权重进行折扣，得到累计增益值.

(6) 平均精度均值(mean average precision, MAP). 是一种用于多个查询的检索评估指标，对于每个查询都有一个平均精度分数，将所有查询的平均精度分数求平均得到结果. MAP 指标是衡量检索系统性能的重要指标之一，可以反映出系统对于不同查询的准确性和稳定性.

5.4　实验结果

5.4.1　对比实验

采用扩充的 ESB 数据集，对本文方法进行实验，并与当前常用的 6 种基于草图的三维模型检索方法进行对比，包括跨模态流形评价方法(cross-domain manifold ranking, CDMR)[23]、基于视图聚类的三维模型草图检索方法(sketch-based 3D model retrieval based on view clustering, SBR-VC)[24]、基于二元度量学习的 Siamese 方法[25]、三维形状的重心表征方法(learned Wasserstein barycentric representation, LWBR)[26]、深度点到子空间度量学习方法(deep point-to-subspace metric learning, DPSML)[11]和不确定性的草图特征抗噪学习方法(sketch feature learning with uncertainty, SUL)[27].

图 7 所示为 7 种方法在扩充的 ESB 数据集上的查全率-查准率(precision-recall, PR)曲线. 可以看出，本文方法在 PR 曲线上优于 6 种对比方法.

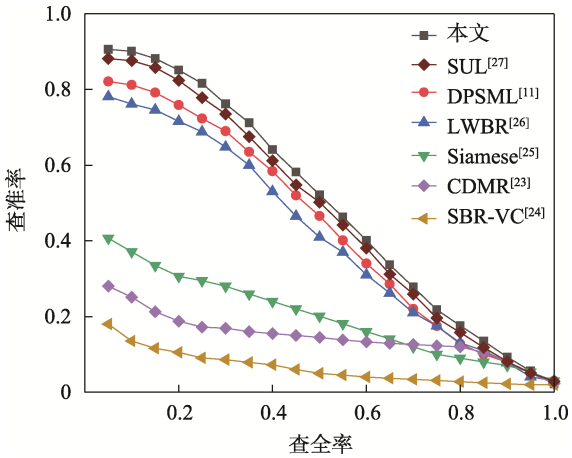


图 7　不同方法的 PR 曲线对比

表 2 所示为 7 种方法的 6 个检索指标对比. 可以看出：(1) 由于 CDMR 方法采用传统方式提取描述符，受到来自草图和三维模型不同域间的差异性、三维模型复杂性等影响，使得人工提取特征异常困难. (2) LWBR 方法虽然利用深度度量学习解

决了特征提取的问题，但在三维模型投影产生的过多视图造成了特征冗余，对特征匹配的准确性产生消极影响.

表 2 与 6 种典型方法评价指标对比

方法	NN	FT	ST	*E*	DCG	MAP
CDMR[23]	0.269	0.211	0.302	0.199	0.472	0.248
SBR-VC[24]	0.171	0.092	0.152	0.079	0.340	0.118
Siamese[25]	0.407	0.411	0.553	0.292	0.601	0.471
LWBR[26]	0.707	0.713	0.772	0.350	0.809	0.748
DPSML[11]	0.677	0.732	0.795	0.379	0.830	0.751
SUL[27]	**0.738**	0.771	0.823	0.397	**0.855**	0.790
本文	0.735	**0.794**	**0.846**	**0.448**	0.851	**0.813**

注. 粗体表示最优值.

(3) 本文方法在各个评价指标上都有一定优势，比 Siamese[25]方法的 ST 指标提升 0.293，比 LWBR[26]方法的 MAP 指标提升 0.065，在 NN, FT, *E* 和 DCG 指标上，本文方法也具备明显的优势；对比 SUL[27]方法，在 FT, ST, *E* 和 MAP 这 4 个指标上，本文方法最优，在 NN 和 DCG 指标上, SUL 方法的性能最佳，本文方法与之十分接近. 当前常用的方法或利用双网络模型分别提取视图和草图特征，或者寻找最佳视角来提高跨域学习能力，相比之下，本文方法不仅解决了多个视图的信息冗余问题，减少了边缘提取轮廓和草图之间的差距，同时关注全局特征和局部特征的融合，表现出更好的性能.

5.4.2 草图轮廓提取的实验

为了验证轮廓草图的有效性，本文设计了 2 组实验，并使用 PR 曲线进行验证，实验结果如图 8 所示. PR 曲线通过改变查询返回的模型数量生成不同的点 PR 点，以区分模型检索的相关性和不相关性，全面展示检索性能. 2 组实验如下: (1) 手绘草图. 由于 ESB 数据集未提供对应的手绘草图，因此本文收集并手绘了 42 个类别下的 186 个零件模型的不同视角下的手绘草图，共计 1 116 个手绘草图，将其作为测试集. (2) 计算机生成的草图. 使用与之对应的计算机生成草图作为测试集. 为了评估这 2 组实验的一致性，生成它们的 PR 曲线. 可以看出，2 组实验得到的 PR 曲线整体趋势接近，因此可以认为，计算机生成草图和手绘草图具有较高的相似性和一致性.

5.4.3 消融实验

为了验证本文提出的基于图像熵的三维模型投影视图表达方法的效果，根据本文提出的固定投影方法，分别对每隔 30°环绕三维模型 360°得到 12 个视图、每隔 60°环绕得到 6 个视图，以及采用联合图像熵得到的视图集进行实验，结果如表 3 所示. 可以看出，基于图像熵的三维模型投影视图表达技术能有效地提高检索精度.

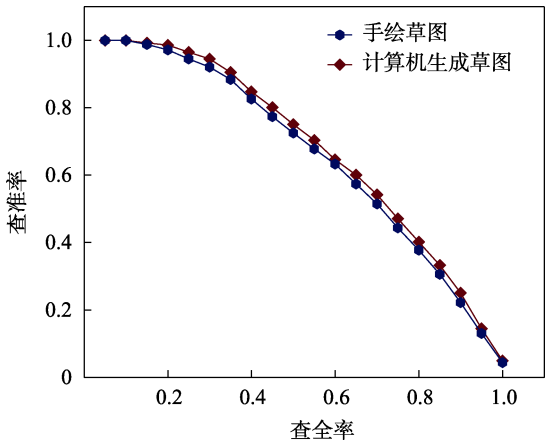


图 8 草图轮廓提取实验的 PR 曲线

本文设计了 2 组消融实验，进一步验证本文提出的双尺度特征融合方法的作用. 2 组实验如下: (1) 仅采用优化的 MVCNN 模型提取全局特征; (2) 仅采用 BOVW 提取局部特征，在 ESB 数据集上进行 PR 比较，结果如图 9 所示. 可以看出，融合 BOVW 和 MVCNN 模型的 PR 曲线高于仅使用单一的 MVCNN 模型或 BOVW 模型的情况. 实验结果表明，本文提出的双尺度特征融合方法是有效的.

表 3 选择不同视图集的检索结果对比

方法	MAP
12 个视图(每隔 30°)	0.775
6 个视图(每隔 30°)	0.756
6 个视图(基于图像熵的视图集)	0.813

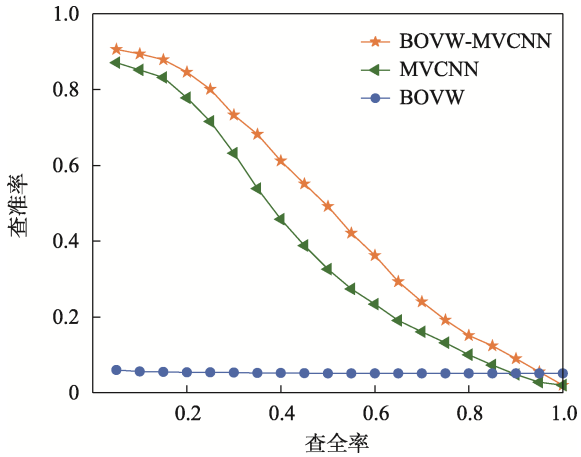


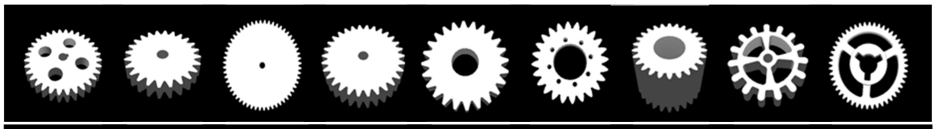
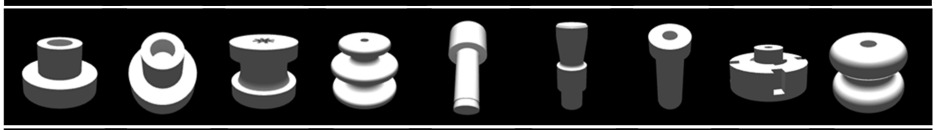
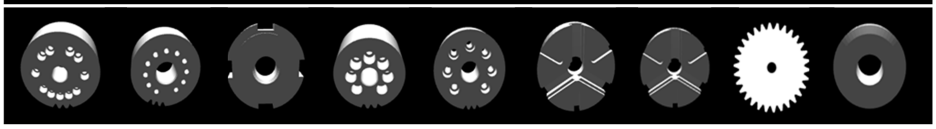
图 9 消融实验的 PR 曲线

第 2 组实验中, 仅采用局部特征的查准率很低, 这是由于 BOVW 仅关注视图的局部特征点, 无法捕捉特征之间的位置关系. 因此, 对于具有相似内容但不同布局的图像, 单独使用 BOVW 无法捕捉到它们之间的相似性, 极大地影响相似图像的检索准确性. BOVW 忽略了特征的空间排列和上下文信息, 导致图像语义表达不准确, 从而检索准确率很低, 而 MVCNN 可以学习到更高级的语义特征, 通过将两者融合, 可以弥补彼此的不足, 提高视图整体的特征表示能力, 对于机械零件处理中相似内容但不同布局的视图很重要, 因为全局特征可以帮助捕捉到不同部分之间的关系. 此外, 特别是在存在视点变化等情况下, BOVW 提取的局部特征点易受到噪声的影响, 而 MVCNN 提取的全局特征有助于降低局部特征的噪声影响, 提高视图整体的鲁棒性和准确性.

5.4.4　实验结果的可视化

表 4 所示为对 ESB 数据集的检索结果示例, 其中, 左侧为随机选取的 6 个草图, 右侧为采用本文方法获得的前 9 个检索结果. 可以看出, 对于简单机械部件(如螺栓类零件), 本文方法能够得到高度准确的检索结果, 同时, 由于局部特征的筛选能力, 还实现了易混淆模型的准确检索. 如在对齿轮和轴承检索时, 本文提出的 mechanical sketch 和双尺度特征模型由于学习和捕捉了轴承和齿轮之间的装配约束关系和局部特征, 能够理解和识别其中键的特征, 对于提高零件模型的检索效果非常重要. 因为在普通的三维模型中键是不显眼的结构, 其特征可能不重要, 但是在零件模型中, 键嵌入于轴上的切槽和贯穿齿轮轴的孔上的切槽中来固定轴和齿轮使之不产生旋转错位, 在传递动力中发挥着重要的作用.

表 4　ESB 数据集中的检索结果示例

零件类别	草图	检索的零件模型结果
齿轮类		
轴承类		
圆柱类		
螺栓类		
圆盘类		
法兰类		

为了验证本文方法的有效性, 对得到的融合特征采用 *t*-SNE 特征降维方法进行可视化, ESB 数据集上 20 个类别零件模型特征的部分可视化结果如图 10 所示, 其中, 颜色和形状共同表示类别, 每个点表示一个零件模型. 可以看出, 来自不同类别的零件模型被分散开, 而具有相同类别信息的零件模型则相对接近, 表明本文方法具有有效的表示能力.

6　结　语

零件三维模型的结构信息与草图的像素信息存在天然的模态语义鸿沟, 基于草图的零件三维模型有效检索是一项具有挑战性的任务. 针对零件三维模型投影出的多个视图的信息冗余, 表征零件视图提取的特征能力弱的问题, 为了提高基

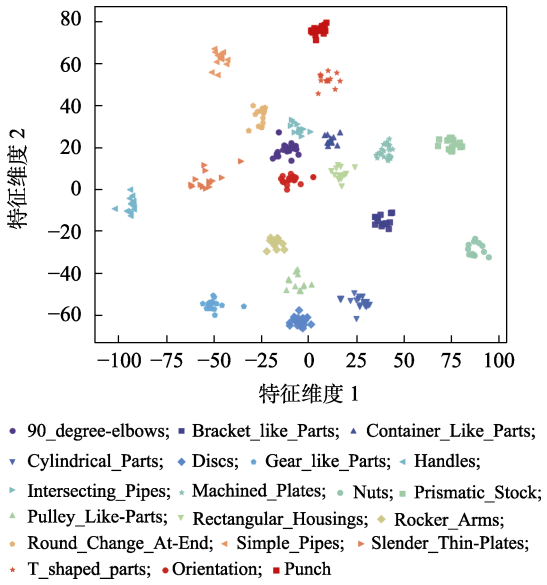


图 10 基于 *t*-SNE 的融合特征可视化结果

于草图的零件三维模型检索的准确率，本文提出一种融合双尺度特征的草图化机械零件库检索方法，并通过实验验证了其优势.

本文的创新点有：提出基于联合图像熵的三维模型投影视图表达方法，有效地降低了多视图表达的信息冗余；构建面向零件视图的 mechanical sketch 模型，实现投影视图的边缘轮廓的深度提取；构建 BOVW 来提取轮廓草图的局部特征，改进 MVCNN 模型来提取轮廓草图的全局特征，通过局部和全局 2 个尺度的特征融合，解决了单尺度特征的表征能力弱的问题；在 ESB 数据集上的实验结果表明，与其他一些常用的草图检索方法相比，本文方法具有更好的检索效果，能有效地提高零件库的检索效果.

本文方法已在 ESB 数据集上取得了良好效果，未来将在 Shrec2013 和 Shrec2014 等通用数据集上进行实验和可能的方法改进，以进一步提高研究成果的普适性和泛化能力.

此外，本文提出局部特征和全局特征融合的双尺度特征模型，可以提高检索模型的识别能力和鲁棒性，但尚未考虑一些关键特征的提取对检索效果的提升作用，后续将引入物体显著性检测有关的特征提取网络，提取零件多视图的独特外观、独特形状等关键区域信息特征，进一步提升检索效果. 另一方面，拟结合超图技术对局部特征和全局特征使用不同感受野的卷积操作，以更好地利用数据的拓扑结构和局部特征信息，提高零件模型特征的表现力和提取性能.

参考文献(References):

[1] Chen X, Gao S M, Guo S, *et al*. A flexible assembly retrieval approach for model reuse[J]. Computer-Aided Design, 2012, 44(6): 554-574

[2] Xu Cijun, Tian Liu, Li Weilong, *et al*. 3D CAD model retrieval based on mechanical structure morphology[J]. Computer Integrated Manufacturing Systems, 2021, 27(6): 1650-1661(in Chinese)
(徐赐军, 田柳, 李炜龙, 等. 基于机械结构形态的三维CAD模型检索[J]. 计算机集成制造系统, 2021, 27(6): 1650-1661)

[3] Qin F W, Qiu S, Gao S M, *et al*. 3D CAD model retrieval based on sketch and unsupervised variational autoencoder[J]. Advanced Engineering Informatics, 2022, 51: Article No.101427

[4] Huang Shenquan, Gu Xinjian, Qi Guoning, *et al*. Web 2.0-based parts library management system[J]. Computer Integrated Manufacturing Systems, 2009, 15(10): 2057-2062(in Chinese)
(黄沈权, 顾新建, 祁国宁, 等. 基于 Web 2.0 的零件库管理系统[J]. 计算机集成制造系统, 2009, 15(10): 2057-2062)

[5] Bai Jing, Zhou Wenhui, Tuo Jiwen, *et al*. End-to-end sketch-3D model retrieval with spatiotemporal information joint embedding[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(6): 826-836(in Chinese)
(白静, 周文惠, 拖继文, 等. 时空信息联合嵌入的端到端三维模型草图检索[J]. 计算机辅助设计与图形学学报, 2021, 33(6): 826-836)

[6] Shao H H, Bai J, Wu R S, *et al*. FGPNet: a weakly supervised fine-grained 3D point clouds classification network[J]. Pattern Recognition, 2023, 139: Article No.109509

[7] Xu Y, Hu J, Wattanachote K, *et al*. Sketch-based shape retrieval via best view selection and a cross-domain similarity measure[J]. IEEE Transactions on Multimedia, 2020, 22(11): 2950-2962

[8] Liu Yangshengyan, Pan Xiang, Liu Fuchang, *et al*. Residual convolution network optimization for view features extraction of 3D model[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(6): 936-942(in Chinese)
(刘杨圣彦, 潘翔, 刘复昌, 等. 面向三维模型视图特征提取的残差卷积网络优化[J]. 计算机辅助设计与图形学学报, 2019, 31(6): 936-942)

[9] Yang H R, Tian Y, Yang C F, *et al*. Sequential learning for sketch-based 3D model retrieval[J]. Multimedia Systems, 2022, 28(3): 761-778

[10] Liu Yujie, Song Yang, Li Zongmin, *et al*. Sketch-based 3D shape retrieval with representative view and convolutional neural network[J]. Journal of Graphics, 2018, 39(4): 735- 741(in Chinese)
(刘玉杰, 宋阳, 李宗民, 等. 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(4): 735-741)

[11] Lei Y J, Zhou Z Q, Zhang P P, *et al*. Deep point-to-subspace metric learning for sketch-based 3D shape retrieval[J]. Pattern Recognition, 2019, 96: Article No.106981

[12] Li B, Lu Y J, Johan H, *et al*. Sketch-based 3D model retrieval utilizing adaptive view clustering and semantic information[J]. Multimedia Tools and Applications, 2017, 76(24): 26603- 26631

[13] Manda B, Dhayarkar S, Mitheran S, *et al*. ‘CADSketchNet’-An

annotated sketch dataset for 3D CAD model retrieval with deep neural networks[J]. Computers & Graphics, 2021, 99: 100-113

[14] Tan X H, Fan Y C, Guo R L. Local features and manifold ranking coupled method for sketch-based 3D model retrieval[J]. Frontiers of Computer Science, 2018, 12(5): 1000-1012

[15] Yu D, Liu Y J, Pang Y P, *et al*. A multi-layer deep fusion convolutional neural network for sketch based image retrieval[J]. Neurocomputing, 2018, 296: 23-32

[16] Bai S J, Bai J. HDA^{2}L: hierarchical domain-augmented adaptive learning for sketch-based 3D shape retrieval[J]. Knowledge-Based Systems, 2023, 264: Article No.110302

[17] Bai Jing, Tuo Jiwen, Bai Shaojin, *et al*. Adaptive multi-class centers and semi-heterogeneous network for sketch-based 3D model retrieval[J]. Journal of Graphics, 2022, 43(1): 36-43(in Chinese)
(白静, 拖继文, 白少进, 等. 基于自适应多类中心和半异构网络的三维模型草图检索[J]. 图学学报, 2022, 43(1): 36-43)

[18] Zhao Y, Liang Q, Ma R X, *et al*. JFLN: joint feature learning network for 2D sketch based 3D shape retrieval[J]. Journal of Visual Communication and Image Representation, 2022, 89: Article No.103668

[19] Wang N, Niu M Y, Dou Z, *et al*. Coloring anime line art videos with transformation region enhancement network[J]. Pattern Recognition, 2023, 141: Article No.109562

[20] Su H, Maji S, Kalogerakis E, *et al*. Multi-view convolutional neural networks for 3D shape recognition[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2015: 945-953

[21] Jayanti S, Kalyanaraman Y, Iyer N, *et al*. Developing an engineering shape benchmark for CAD models[J]. Computer-Aided Design, 2006, 38(9): 939-953

[22] Liu A N, Li W H, Nie W Z, *et al*. 3D models retrieval algorithm based on multimodal data[J]. Neurocomputing, 2017, 259: 176-182

[23] Furuya T, Ohbuchi R. Ranking on cross-domain manifold for sketch-based 3D model retrieval[C] //Proceedings of International Conference on Cyberworlds. Los Alamitos: IEEE Computer Society Press, 2013: 274-281

[24] Li B, Lu Y J, Johan H. Sketch-based 3D model retrieval by viewpoint entropy-based adaptive view clustering[C] //Proceedings of the Sixth Eurographics Workshop on 3D Object Retrieval. Goslar: Eurographics Association, 2013: 49-56

[25] Wang F, Kang L, Li Y. Sketch-based 3D shape retrieval using convolutional neural networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 1875-1883

[26] Xie J, Dai G X, Zhu F, *et al*. Learning barycentric representations of 3D shapes for sketch-based 3D shape retrieval[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 3615-3623

[27] Liang S, Dai W D, Wei Y C. Uncertainty learning for noise resistant sketch-based 3D shape retrieval[J]. IEEE Transactions on Image Processing, 2021, 30: 8632-8643