

基于深度学习的三维形状特征提取方法

周 燕 曾凡智 吴 臣 罗 粤 刘紫琴

(佛山科学技术学院计算机系 广东 佛山 528000)

摘 要 研究具有低维、高鉴别力的三维形状特征提取方法有助于解决三维形状数据分类和检索等问题。随着深度学习的持续发展,结合深度学习的三维形状特征提取方法已成为研究热点。将深度学习与传统的三维形状特征提取方法相结合,不仅可以突破非深度学习方法的瓶颈,而且可以提高三维形状数据分类、检索等任务的准确率,尤其是当三维形状是非刚体时。然而,深度学习尚在发展中,仍存在需要大量训练样本的问题,因此如何运用深度学习方法来高效提取三维形状特征成为了计算机视觉领域的研究重点和难点。目前,研究者大多从改进网络结构和训练方法等方面入手,着重提高神经网络提取特征的能力。文中结合深度学习和三维形状特征提取方法的发展历程,首先介绍相关深度学习模型,以及网络改进、训练方法等方面的新思路;其次重点对基于深度学习的刚体与非刚体的特征提取方法做综合的阐述,描述当前深度学习方法用于三维形状特征提取的情况;然后简述现有三维形状检索系统的现况以及相似度计算方法;最后介绍当前三维形状特征提取方法存在的问题,并探讨其未来的发展趋势。

关键词 三维形状,特征提取,深度学习,神经网络

中图分类号 TP391.4 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.09.006

3D Shape Feature Extraction Method Based on Deep Learning

ZHOU Yan ZENG Fan-zhi WU Chen LUO Yue LIU Zi-qin

(Department of Computer Science, FoShan University, Foshan, Guangdong 528000, China)

Abstract Research on extracting 3D shape features with low dimension and high discriminating ability can solve the problem such as classification, retrieval of 3D shape data. With the continuous development of deep learning, 3D shape feature extraction method combined with deep learning has become a research hotspot. Combining deep learning with traditional 3D shape feature extraction methods can not only break through the bottleneck of non-deep learning methods, but also improve the accuracy of 3D shape data classification, retrieval and other tasks, especially when 3D shape is non-rigid body. However, deep learning is still developing, and there are still problems that require a large number of training samples. Therefore, how to effectively extract 3D shape features by using deep learning methods has become the research focus and difficulty in the field of computer vision. At present, most researchers focus on improving the ability of neural network to extract features by improving network structure, training methods and other aspects. First, the relevant deep learning model are introduced, and there are some new ideas about the network improvement and training methods. Second, the feature extraction methods of rigid body and non-rigid body based on deep learning are comprehensively expounded which combined with the development of deep learning and 3D shape feature extraction methods, and the current deep learning methods for the 3D shape feature extraction are described. And then, the current situation of the existing 3D shape retrieval system and the similarity calculation method are described. Finally, the current problems of 3D shape feature extraction methods are introduced, and its future development trend are explored.

Keywords 3D shape, Feature extraction, Deep learning, Neural networks

来稿日期:2019-05-15 返修日期:2019-07-24 本文受国家自然科学基金(61602116),广东省自然科学基金(2017A030313388),广东省工程技术研究中心(G601624),佛山市工程技术研究中心(2017GA00015,2016GA10156)资助。

周 燕(1979-),女,硕士,教授,CCF 会员,主要研究方向为计算机视觉、三维模型检索,E-mail:zhouyan791266@163.com(通信作者);曾凡智(1965-),男,博士,教授,主要研究方向为计算机视觉、图像处理、数据挖掘;吴 臣(1974-),男,硕士,讲师,主要研究方向为模式识别、智能计算;罗 粤(1994-),男,硕士生,主要研究方向为计算机视觉、三维模型检索;刘紫琴(1995-),女,硕士生,主要研究方向为计算机视觉、视频分析。

1 引言

大数据时代,三维模型作为第四代数字媒体,随着软、硬件的发展呈大规模增长趋势。现阶段,有效地提取三维模型的低维、高鉴别力的形状内容特征,有利于对其进行分类与检索等,因此研究三维模型特征提取新方法是当前计算机视觉领域的重要研究内容。

三维模型是一种具有空间结构的数据表达,蕴含的内容属性较二维图像更丰富。三维模型大致可分为刚体和非刚体两大类:刚体指的是受到外力后形状和体积不发生改变的物体;相反,非刚体指的是受到外力后形状和体积发生改变的物体。图1给出了刚体和非刚体三维形状的几个示例。

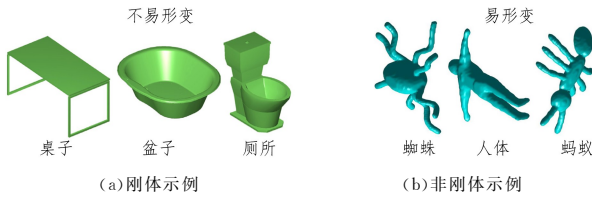


图1 三维形状示例

Fig. 1 3D shape example

三维模型数据被广泛应用于3D打印^[1]、工业产品设计^[2]、计算机辅助设计^[3]、家具设计、医学诊断^[4]、影视动画、虚拟现实、三维游戏设计^[5]、建筑物设计、分子生物研究、文物修补等领域。随着深度传感器、激光雷达、三维成像技术、三维模型渲染软件等在制造业领域的广泛使用,三维模型数据将继续呈海量增长趋势,并且将持续产生对三维模型分类与检索等任务的应用场景需求。

目前主要使用基于内容的三维模型特征^[6]进行分类与检索等任务。其中,三维模型的内容特征主要包含形状属性特征,以及颜色、纹理、材质等表面属性特征。然而,由于表面属性内容特征在三维模型特征表达中可变量易被改变,因此提取基于形状属性的三维模型内容特征是当前的主要研究方向,且主要通过文件检索、二维草图检索和三维草图检索3种方式进行三维模型分类与检索等任务。

随着人工智能技术的发展,使用深度学习工具提取模型特征已成为主流技术。深度学习领域也产生了众多模型,例如经典的卷积神经网络、深度置信网络、自动编码器、生成对抗网络等,这些模型在不同任务中取得了比传统非深度学习方法更优异的成绩。深度学习的方法由于具有强大的抽象和提取高层语义特征的能力,在语音、图像、视频等领域取得了巨大的成功,因此利用基于深度学习的方法提取三维形状特征已成为主要的研究方向。

目前在大规模刚体三维形状的分类与检索等任务上,基于深度学习的方法取得了不错的效果,因此诞生了众多基于视图^[7-9]、体素^[10-11]、点云^[12-13]、网格^[14]和八叉树^[15]的深度学习神经网络模型。不同编码方式的刚体三维形状的深度学习神经网络模型如图2所示。而依据数据组织类型,目前基于深度学习的非刚体三维形状特征提取方法大致可分为5类:基于人工特征的方法、基于原始数据的特征提取、基于投影视图的特征提取、基于三维体素的特征提取、基于多特征融合的特征提取^[16-17]。目前,如何运用深度学习提取更具表达

能力的三维模型形状内容特征,从而促进三维模型数据分类、检索、识别、匹配等任务的精确进行,是亟待解决的问题。

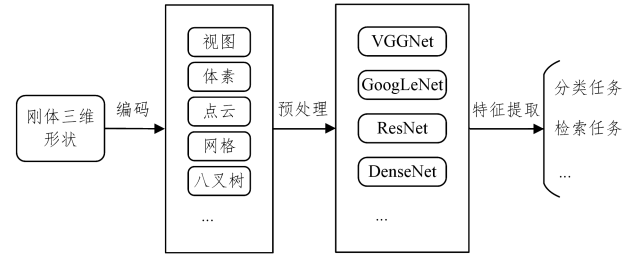


图2 不同编码方式的刚体三维形状的深度学习神经网络模型

Fig. 2 Deep learning network model of rigid 3D shape with different coding methods

本文立足于三维形状分类、检索等任务,对近年来基于深度学习方法的三维形状特征提取方法及相关工作进行了系统的梳理。本文第2节简要介绍相关的深度学习模型、网络设计与训练方法以及三维模型的预处理;第3节介绍刚体三维形状的特征提取方法和相关工作;第4节介绍非刚体三维形状的特征提取方法和相关工作;第5节介绍现有三维形状检索系统的现状、检索方式、可视化以及相似度计算方法;第6节介绍当前三维形状检索面临的挑战;最后对三维形状特征提取的研究方向进行了展望。

2 相关深度学习模型

2.1 卷积神经网络

卷积神经网络(Convolutional Neural Networks, CNN)是最经典的深度学习神经网络模型,其特点是运用了卷积操作和反向传播算法来训练神经网络,可应用于二维图像分类、检索、语义分割等相关任务。

现代卷积神经网络的雏形是1998年诞生的手写字体识别网络LeNet5^[18],其初步奠定了后来出现的卷积神经网络的组成成分,基本组成成分涵盖卷积层、池化层、全连接层、输出层。如图3所示,卷积神经网络的操作层可被看作一个复杂的函数 f_{CNN} ,在反向传播阶段,由正则损失和数据损失共同驱动更新权值和偏置参数,并将误差反向传播至网络的每一层进行权重和偏置等参数更新的学习训练。

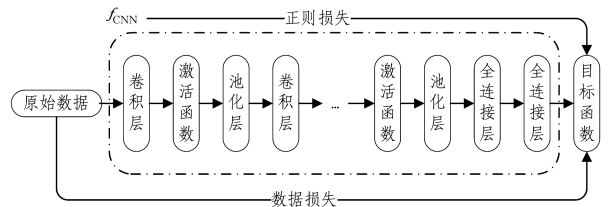


图3 卷积神经网络的基本流程图

Fig. 3 Basic flow chart of convolutional neural network

2.2 深度置信网络

深度置信网络(Deep Belief Network, DBN)由Hinton等^[19-20]于2006年提出,属于一种深度学习概率生成模型。图4给出了DBN原理结构,其由多个受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)层组成。DBN可用于数据分类、生成等任务。

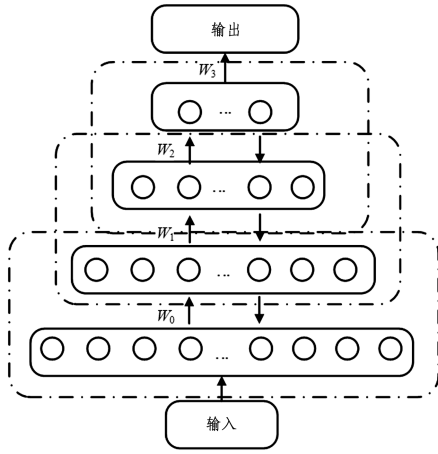


图 4 DBN 模型结构

Fig. 4 DBN model structure

2.3 自动编码器

自动编码器 (Auto-Encoder, AE) 由 Rumelhart 等^[21]于 1986 年提出,是一种神经网络模型,属于无监督学习模型,可用于数据压缩。自动编码器使用反向传播算法进行参数训练,目标是使输入尽可能等于输出。图 5 给出了自动编码器模型结构,其由编码器和解码器两部分组成。其基本结构是一个多层感知器的神经网络,从输入层到输出层含有多个中间层,特点是输入层和输出层具有相同的维数,中间的编码层维数小于输出层维数。

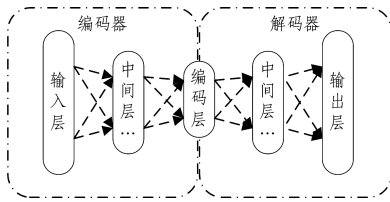


图 5 自动编码器模型结构

Fig. 5 Auto-Encoder model structure

2.4 生成对抗网络

生成对抗网络 (Generative Adversarial Networks, GAN) 是一种深度学习生成式模型,属于无监督学习模型,于 2014 年由 Goodfellow 等^[22]提出,可用于学习复杂的特征分布、图像风格转移、模型生成等任务。如图 6 所示,生成对抗网络主要包含生成模型 (Generative Model) 和判别模型 (Discriminative Model),二者互相博弈完成网络训练过程。生成器的作用是将潜在空间的随机向量转换为生成样本,来欺骗判别器,而判别器则需要分辨真实样本和生成样本的真假。

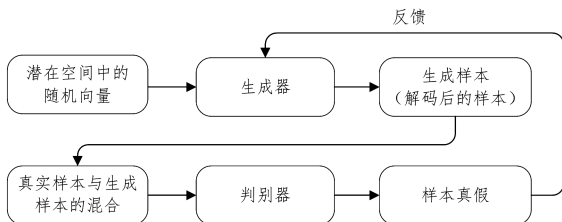


图 6 GAN 模型结构

Fig. 6 GAN model structure

2.5 深度学习网络设计与训练方法

近年来,很多经典的卷积神经网络在二维图像和刚体三

维形状的分类与检索等任务中得到了很好的应用。针对深度学习网络模型的改进或者重新设计,主要围绕激活函数^[23]与训练技巧^[24]、网络深度与宽度^[25-26]、网络连接方式^[27-30]以及生成对抗方式^[31]等方面进行优化。

2012 年, Krizhevsky 等^[24]引入了激活修正线性单元、局部规范化、数据扩增与随机失活等方式,来设计并训练卷积神经网络 AlexNet,其在大规模视觉识别挑战赛 (ImageNet Large Scale Visual Recognition Challenge, ILSVRC) 上取得了较好的效果。2014 年, Szegedy 等^[26]采用 Inception 模块拓宽网络宽度,设计的 GoogLeNet 在 ILSVRC 竞赛上将 top-5 的错误率降到了 6.7%。2015 年, He 等^[27]设计了深度达 152 层的残差网络 ResNet,提出的数据残差与跨层连接模式较好地解决了梯度消失或梯度爆炸的难题。2017 年, Huang 等^[28]在稠密网络 DenseNet 中通过网络稠密连接与特征重用,减少了网络参数,提升了网络的训练效率。2018 年, Zoph 等^[23]提出了一种新型激活函数 Swish,该函数具备不饱和、光滑、非单调性的特点,在不同的测试数据集上都表现出了比传统激活函数更优越的性能。2019 年, Sun 等^[30]提出了 HR-Net 的神经网络模型,该高分辨率网络具有并联结构,可以持续保持高分辨率表征输出。

生成对抗网络 GAN 是深度学习网络设计的热点, Chen 等^[31]提出了一个新的 InfoGAN 模型,其在生成器中除了原噪声外,通过加入互信息变量进行非监督学习,获得了一种三维形状离散型特征,在 3D 人脸检索中取得了较好的效果。

另外,自动化机器学习 (Automated Machine Learning, AutoML) 近年来发展潜力巨大,其使用了神经结构搜索 (Neural Architecture Search, NAS),它可以让计算机自动化完成神经网络的设计^[32]。2019 年, Tan 等^[33]基于 AutoML 提出了一种运用新型模型缩放方法的深度学习神经网络 EfficientNets,其在 ImageNet 数据集中实现了 84.1% 的 top-1 准确率,且模型更小,速度更快,效率更高。

此外,若将深度学习网络模型推广应用于移动互联网端,还需轻量级的神经网络,如 SqueezeNet^[34], MobileNet^[35], ShuffleNet^[36], IGCN^[37], ESPNet^[38];另外,可使用推理优化器 TensorRT 进行加速计算。

快速网络训练算法是决定三维形状深度特征提取时间效率的关键因素。经典的网络训练算法有 SGD^[39], Adam^[40] 和 Adagrad^[41]等。SGD 算法采用恒定学习率对参数进行更新,算法简单实用,但下降速度慢,且容易在最优点处产生震荡。Adam 算法和 Adagrad 算法采用自适应矩估计方式计算动态学习率,算法参数更新速度快,但可能陷入局部最优。2019 年, Luo 等^[42]提出 AdaBound 网络优化算法,该算法结合 SGD 和 Adam 的优点,既能在训练初期快速下降,又能在后期保持稳定的收敛性,但算法复杂度较高。因此,在如何构建适用于三维形状的新型深度学习网络模型、设计快速的训练方法以及提升任务时间效率等方面,有待进一步的研究。

2.6 三维模型的预处理

预处理的作用是对不同的坐标系进行归一化处理,从而保证所提取特征的平移、比例和旋转的不变性,提升了三维模型分类与检索等任务的准确率。例如,使用经典的 PCA 主成

分分析法对不同的坐标系进行归一化处理,使得三维形状特征更具有平移、比例和旋转的不变性。

在刚体三维形状的特征提取上,使用基于视图的深度神经网络的分类检索任务的效果较好。在 ModelNet40 测试集上,在分类方面,基于一致对齐的三维形状数据的精度最高的模型是 Kanazaki 等^[9]提出的一种基于多视图的无监督视图的卷积神经网络 RotationNet,其分类精度达到了 97.37%。然而,随机方向上三维形状数据的精度最高的模型是 Furuya 等^[43]提出的一种深度局部特征聚合网络(Deep Local feature Aggregation Network,DLAN),其分类精度达到了 85%。产生这种差异的原因是 RotationNet 使用了对齐的三维形状模型数据,无须对模型进行进一步的旋转对齐;而 DLAN 使用了旋转归一化,因此在随机方向上,DLAN 对三维形状数据的分类取得了比 RotationNet 更好的分类精度。

3 刚体三维形状特征提取方法

3.1 基于深度学习方法刚体特征提取方法

数字几何模型的表示方法包括:1)实体表示,如实体几何、点云、体网络、体素;2)边界表示,如表面网格、参数曲面、细分曲面、隐式曲面。目前能应用深度学习模型的几何数据表示有视图、点云、网格、体素等几种形式^[44]。图 7 给出了深度学习模型常用的数据几何的表示。

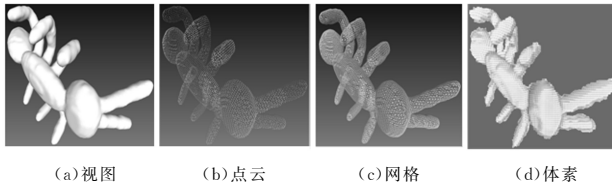


图 7 深度学习模型常用的数据几何

Fig. 7 Geometric representation of data used in deep learning models

目前,基于深度学习的方法在刚体三维形状分类与检索等任务中取得了不错的效果。过去,三维形状特征提取主要运用非深度学习算法,例如 SPH^[45]、LFD^[46]、测地距离^[47]、热核特征(Heat Kernel Signature, HKS)^[48]、波核特征(Wave Kernel Signature, WKS)^[49]等传统的基于网格、视图、点云等的非深度学习算法的特征提取方法。如今,三维形状特征提取已经发展至主流、前沿的基于深度学习算法的特征提取方法,例如运用卷积神经网络、自动编码器等深度学习神经网络模型提取三维形状特征,用于分类、检索、语义分割、三维重建、模型生成等任务。图 8 给出了使用不同类型数据的三维形状表征的发展历史,不同类型数据的数据表征如表 1 所列。

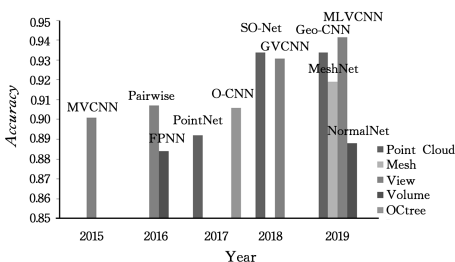


图 8 使用不同类型数据的三维形状表征的发展历史

Fig. 8 Development history of 3D shape representation using different types of data

表 1 不同类型数据表征的深度学习网络

Table 1 Deep learning networks represented by different data types

数据表征类型	深度学习网络
点云 (Point Cloud)	PointNet ^[12] , PointNet++ ^[13] , Kd-Networks ^[50] , SO-Net ^[54] , DGCCN ^[51] , Geo-CNN ^[56]
网格 (Mesh)	MeshNet ^[14]
视图 (View)	MVCNN ^[7] , Pairwise ^[52] , GVCNN ^[8] , RotationNet ^[9] , MLVCNN ^[60]
体素 (Volume)	3D ShapeNets ^[10] , VoxNet ^[62] , FPNN ^[11] , NormalNet ^[64]
八叉树 (Octree)	OctNet ^[66] , O-CNN ^[67]
组合的形式	体素+多视图 (FusionNet ^[69])、点云+多视图 (PVNet ^[70] , PVRNet ^[71])

3.1.1 基于点云的特征提取

点云是分布在三维空间中的点的集合(数学表达为 $n \times 3$ 的矩阵,其中 n 为点的数目,数学表达不唯一,具有无序性和旋转不变性),属于非结构性三维模型数据。传统的基于点云数据的方法中,点特征通常编码点的某些统计特性,并被设计为对某些变换具有不变性,例如典型的 LFD^[46]方法。点云的获取十分方便,可通过三维扫描仪扫描获取,但点云难以使用传统神经网络直接处理。目前,深度学习算法对三维模型的处理主要集中在二维图像投影映射的预处理方案,该方案较为方便,但需要将模型格式进行转换并做预处理,因此使用深度学习算法直接处理点云数据是当前研究的一大难点和热点。

2017年,Qi等^[12]首次提出了一种直接用于三维模型点云数据表示的分割和分类任务的神经网络模型 PointNet,其通过空间变换网络和 Max Pooling 解决了点云旋转和点云无序性的问题,在 ModelNet40 刚体测试数据集上的分类准确率达到 89.2%。同年,Qi等^[13]对 PointNet 做出了改进,并提出了新的网络 PointNet++。该网络在先前的基础上加入了对点云局部特征的提取,从而改善了分类和分割任务的效果,在 ModelNet40 上的分类准确率提升至 91.9%。然而,其点特征被抽象为独立和孤立的方式,忽略了相邻点的相对布局及其特征。对此,Wang等^[53]于 2018 年提出了一种新的图池化策略,即将标准的 Max Pooling 步骤替换为递归聚类 and 图池化策略。该策略在局部图上使用谱图卷积共同学习特征,在 ModelNet40 上的分类精度达到了 92.1%。

2018年,Li等^[54]提出了一种基于点云的深度学习神经网络模型 SO-Net,即自组织网络。通过构建自组织映射(SOM)建模点云的空间分布,该网络对单个点和 SOM 节点进行分层的特征提取,学习点云的特征向量表示。其在 ModelNet40 上的分类准确率达到 93.4%;同时,由于该体系结构具有并行性和简单性,训练速度也有所提升。

2019年,Zhao等^[55]提出了一种三维点胶囊网络(3D Point-Capsule Networks),其是一种自动编码器,被设计用于处理稀疏的三维点云,同时保留了输入数据的空间排列。其动态路由方案和胶囊网络部署的独特的二维潜在空间,提升了点云对象分类、对象重建和部分分割等任务的效率。

在对局部区域内点间几何结构的显式建模方面,Lan等^[56]于 2019 年提出了 Geo-CNN,该网络应用了 GeoConv 的通用类卷积操作对每个点及其局部邻域学习点之间的局部几

何关系,在 ModelNet40 上的总体分类精度达到了 93.4%。

3.1.2 基于网格的特征提取

网格是一种通过点云形成的、相对于点云而言的一种结构化的三维模型数据,包含了顶点坐标的几何信息、连接关系的拓扑信息,以及点线面上的各种属性信息,拥有比其他类型数据更强的三维形状描述能力。传统的基于网格数据的典型方法有 SPH^[45]等,然而由于网格数据具有复杂性和不规则性,近年来对针对网格数据特征提取的研究较少。

多边形网格具有良好的紧致性和几何特性。然而,使用神经网络从二维图像建模多边形网格并不容易,因为从网格到图像的转换或渲染涉及栅格化离散操作,这会阻止神经网络的反向传播。针对该问题,Kato 等^[57]于 2018 年提出了一个包含近似梯度的光栅化,使得网格渲染可以集成到神经网络中训练学习,实现了具有轮廓视图监督的单视图三维网格的重建。2018 年,Groueix 等^[58]提出了一种新形状生成框架 AtlasNet,该框架将三维形状表示为参数化表面元素的集合,可用于学习生成三维网格。

为了解决网格数据处理中的难题,Feng 等^[14]于 2019 年提出了基于网格的深度学习方法——MeshNet。其首先将面片作为单元,定义共享边的面片之间的连接关系;然后设计两个空间和结构描述块来学习空间特征和结构特征,并用一个网格卷积块来聚合相邻特征,从而从网格数据中学习三维形状表示。该模型能较好地解决网格复杂和不规则性的问题,在 ModelNet40 上的分类准确率达到 91.9%,平均检索精度达到了 81.9%。

3.1.3 基于视图的特征提取

视图是三维模型在二维平面上的投影,但是投影易受模型自身遮挡的影响,因此目前对三维模型进行表达时采用多个视角下的投影视图。二维视图在深度学习领域已经得到了广泛应用,并且在技术上已经非常成熟,故可直接移植典型的卷积神经网络模型,并基于多视图来学习和提取三维形状特征。目前,针对基于视图的三维形状特征提取方法的研究主要集中在多视图融合以及深度学习算法的使用上。

Su 等^[7]于 2015 年提出了基于多视图的深度学习方法 (Multi-view Convolutional Neural Networks, MVCNN) 来识别对象,该模型在 ModelNet40 上的分类准确率达到 90.1%,检索精度最高达到了 80.2%。Feng 等^[8]于 2018 年,提出了群视图的卷积神经网络 (Group-View Convolutional Neural Networks, GVCNN),其采用分层框架,引入的分组模块考虑了视图之间的组内相似性和组间区分性,以便为三维形状生成更具辨别力的描述符,在 ModelNet40 上的分类准确率达到 93.1%,检索精度最高达到了 85.7%。Zhou 等^[59]于 2019 年提出了一种新的基于极坐标视图的策略,该策略通过使用 360 度投影,将三维形状的点云映射到二维平面,并通过点云的最大深度得到极坐标的视图表示;同时,提出了一种卷积神经网络,通过训练三维形状的一个单极视图,来得到极视图表示 (PVR),从而实现了三维形状的分类和检索,其在 ModelNet10 和 ModelNet40 等标准数据集上的检索精度达到了 92.54% 和 91.69%。Jiang 等^[60]于 2019 年提出了一种基于多环视卷积神经网络 (Multi-Loop-View Con-

volutional Neural Network, MLVCNN) 的三维形状检索框架模型,其在 ModelNet40 测试集上的检索精度达到了 92.84%。李荫民等^[61]于 2019 年比较了基于视图的深度特征与人工设计特征的性能,同时比较了残差网络 ResNet 和 VGG-Net。实验表明,相较于传统特征,基于视图的深度特征更为优越,且使用 ResNet 的表现更好。

3.1.4 基于体素的特征提取

三维模型的体素化表示类似于二维图像的像素表示,该表示可以直接应用卷积神经网络等经典的深度学习网络模型来学习模型特征。

Wu 等^[10]于 2015 年首次提出了一种用于三维模型体素深度表示的 3D ShapeNets 神经网络结构,其使用了深度置信网络的深度学习模型,在刚体三维模型标准库 ModelNet40 上的分类准确率达到 77%。Maturana 等^[62]于 2015 年提出了一种 VoxNet,其通过将一个体积占用网格表示 (即体素表示) 与一个监督的三维卷积神经网络相结合来处理大量的点云数据,在 ModelNet40 上的分类准确率达到 83.0%。杨军等^[63]于 2019 年提出了一种基于深度体素卷积神经网络的三维模型识别分类算法,该算法将三维多边形网格模型转化为体素,并通过深度体素卷积神经网络提取其深层语义特征,以增强特征的表达能力和差异性,在 ModelNet40 数据集上的分类准确率最高可达到 87%。Wang 等^[64]于 2019 年提出了一种基于体素的卷积神经网络 NormalNet 用于 3D 对象的分类与检索,其在 ModelNet40 数据集上实现了 88.8% 的分类精度。Kumawat 等^[65]于 2019 年提出了一种基于体素的改进三维卷积神经网络 (Unveiling Local Phase in 3D Convolutional Neural Networks, LP-3DCNN),其中的校正局部相位体积 (Rectified Local Phase Volume, ReLPV) 块有效地替代了标准三维卷积层,解决了计算量大、内存密集、容易过度拟合等问题,并且提高了特征学习能力。该网络模型在 ModelNet10 和 ModelNet40 数据集上的分类准确率分别达到了 92.1% 和 94.4%,且参数量更少。

3.1.5 基于八叉树的特征提取

八叉树是一种用于描述三维空间的树状数据结构,属于栅格数据模型。相较于体素而言,八叉树结构有助于缓解体素过多占用计算资源的问题,但其不适宜大规模点云场景的处理。

Riegler 等^[66]于 2017 年提出了基于八叉树的深度学习模型 OctNet——一种用稀疏的三维数据进行深度学习的表示,并支持高分辨率和低分辨率的三维卷积网络。同年,Wang 等^[67]提出了基于八叉树的卷积神经网络 (Octree-Based Convolutional Neural Networks, O-CNN),用于刚体和非刚体三维形状的分析,其在 ModelNet40 上的分类精度达到了 90.6%。另外,Tatarchenko 等^[15]还提出了基于八叉树的生成对抗网络,用于生成高分辨率的三维模型。

3.1.6 基于多特征融合的特征提取

传统方法中,融合多特征可以提高三维模型的表达能力,改善其在各项任务中的表现,如 Zhou 等^[68]于 2017 年提出了一种基于压缩感知与统计分析的多特征融合框架,用于刚体三维形状检索,并取得了较好的效果。但是,传统方法面临特

征融合系数大多选择经验值调配的问题。在采用深度学习网络进行多特征融合的应用研究中,目前也较多地使用将经验值进行融合的方法来提高三维形状分类与检索的精度。

Hegde等^[69]于2016年提出了一种融合体素和多视图的深度学习模型 FusionNet 用于三维模型分类,其在 ModelNet10 和 ModelNet40 数据集上的分类精度分别达到了 93.11% 和 90.8%。You等^[70]于2018年提出了一种将点云数据和多视图数据相结合的联合三维形状识别框架——点视图网络(Point-View Network, PVNet),即一种嵌入注意融合方案。该方案可以利用多视角数据中的高级特征对点云数据中不同结构特征的内在相关性和可辨别性进行建模,可将识别性描述量化并用作软注意掩模,以进一步细化 3D 形状的结构特征,在 ModelNet40 数据集上的分类精度达到了 93.2%,检索精度达到了 89.5%。You等^[71]于2019年提出了一种有效的点云关系网络(Point-View Relation Network, PVRNet),该网络将视图特征和点云特征很好地进行了融合。

4 非刚体三维形状特征提取方法

图9给出了典型非刚体三维形状的原始初态及其等距等容、拓扑、噪声、散粒噪声、孔、微孔、采样、栅格化、局部缺失、视图投影、仿射变换和尺度变换等12类非刚性变换^[72]。由于非刚体三维形状的形变具有多样性和复杂性,因此许多在刚体上成熟的检索与分类方案在非刚体上取得的效果并不理想。

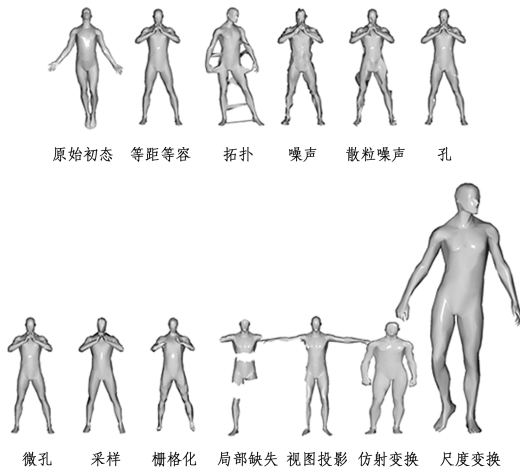


图9 典型非刚体三维形状的变换形式

Fig. 9 Transformation form of 3d shape of a typical non-rigid body

相对于刚体,非刚体三维形状的特征提取要求更高,不仅要具有平移、旋转和尺度不变性,而且也要具有等距不变性。目前对非刚体三维形状深度特征的提取和基于深度学习的大规模非刚体三维形状分类与检索系统的研究还较少,暂时没有单一特征可以对非刚体三维形状的内蕴属性做全面的描述;同时,对非刚体三维形状特征提取研究的分类检索精度和深度神经网络训练的时间效率也有待提高。

目前,基于深度学习的非刚体三维形状特征提取方法有:基于人工特征的方法、基于原始数据的特征提取、基于投影视图的特征提取、基于三维体素的特征提取、基于多特征融合的特征提取^[16-17]。

在基于人工特征的方法方面,Zeng等^[73]于2018年提出了四重卷积神经网络(Quadruplet Convolutional Neural Networks, QCNN),首先计算各顶点的波核特征 WKS 并构造多能量形状分布矩阵,然后采用 QCNN 提取深层语义特征;Luciano等^[74]于2018年提出了一种利用深度学习和测地线进行三维形状分类的综合框架,该方法继承了测地线距离的许多有用特性,采用两层叠加稀疏自编码器,以无监督的方式对隐藏层进行单独训练,然后采用 Softmax 分类器,从测地线中学习深度特征。

在基于原始数据的特征提取中,Han等^[75]于2017年提出了基于三维局部特征的网络模型,该模型能直接从原始三维体素中提取高阶、层次化的三维局部特征。

在基于投影视图的特征提取中,Sinha等^[76]于2016年将非刚体三维形状投影到二维平面上,采用卷积神经网络提取了基于投影视图的深度特征。白静等^[77]于2019年提出了一种基于投影多视图的卷积神经网络模型,并采用加权投票机制融合多视图特征,用于对三维模型进行分类与检索,其在刚体三维数据集 ModelNet10/40、非刚体三维模型库 SHREC2010/2011/2015 上的分类准确率分别为 93.18%, 93.07%, 99.5%, 99.5% 和 99.4%。

在基于三维体素的特征提取中,Li等^[11]于2016年提出了场探测神经网络(Field Probing Neural Networks, FPNN),将三维形状转换为体素化形式,并采用场探索滤波器取代卷积,提取了基于三维体素的深度特征,其在 ModelNet40 上的分类精度达到 88.4%。

在基于多特征融合的特征提取中,Bu等^[78]于2017年结合投影视图和三维体素的优点,提出了一种基于多模态的三维特征学习框架,该框架取得了比单一特征更好的效果,发挥了融合特征的优势。2018年,Zeng等^[79]提出了一种基于卷积神经网络的多特征融合学习方法,用于非刚性三维模型检索。其首先计算并使用热核(HKS)和波核(WKS)两种特征描述符的二维形状分布来表示三维模型;然后将其作为网络的输入,并分别构造 HKS 和 WKS 两种分布的卷积神经网络,利用多特征融合层将其连接起来;接着构建交叉连接层,将低层特征与高层语义特征相结合,进一步提高特征表达能力。2019年,He等^[80]提出了 GeoNet 的深度学习体系结构,用于对用点云表示的表面的内在结构进行建模;并提出了融合方案,该方案将 GeoNet 与其他网络(如 PU-Net^[81] 和 PointNet++^[13])相结合,用于点云分析,并改善了非刚性形状分类效果。

然而,在这几类基于深度学习的特征提取方法中,学习数据源的不同导致各方法均存在一定的局限:基于人工特征的方法依赖人工选择参数;基于投影图像的方法在变换过程中忽略了形状的局部细节和全局结构的相互关系,降低了特征鉴别力;基于三维体素的方法中,体素化的二值表达所蕴含的形状信息有限,而且对高分辨率的计算复杂度大,对低分辨率则限制了特征的辨别力;基于原始数据的方法设计能适应原始三维数据特点的深度神经网络的难度较大;基于多特征融合的方法的自适应的融合系数计算和融合理论科学设计方案

选取困难。此外,基于深度学习的非刚体三维形状的分类与检索方法较少,仍然有许多难题需要进一步深入研究。

5 现有三维形状检索系统与相似度计算

5.1 现有三维模型检索系统与检索方式

现有三维模型检索系统中,针对刚体三维形状的检索系统的应用已较为成熟。第一个基于内容的通用三维形状检索系统 Nefertiti 于 1997 年由加拿大国家研究院的 Paquet 等^[82]提出并开发。2003 年,美国普林斯顿大学团队研发了三维形状搜索引擎(3D Model Search Engine)^[83],该系统实现了关键字、二维草图、三维草图以及三维形状的检索功能。Sagara 等^[84]提出了基于 KAZE 图像特征的三维形状检索系统,该系统以二维图像为例进行检索。日本仙台东北大学的 Sintunata 等^[85]提出了一种基于单个查询图像的三维目标检索系统,该方法不仅不需要将数据库中的每一幅图像与查询图像进行比较,而且可以立即推断出三维对象的水平和垂直角度。2016 年,中国台湾台南大学的 Kao^[86]则提出了一种可以保存模型局部几何信息的三维形状检索系统。然而,由于非刚体三维形状的特殊性,针对非刚体三维形状检索系统的研究还相对较少,并且在采用深度学习网络进行多特征融合方面也有待突破。而且,大多数三维模型检索系统不提供开放接口,并且没有相关算法进行详细说明,故无法被有效整合到机构知识库系统中^[87]。图 10 给出了一个典型的三维模型的检索过程。

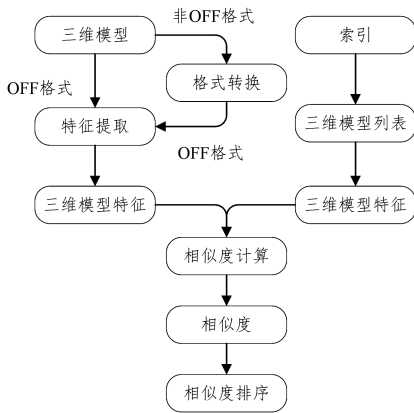


图 10 三维模型的检索过程

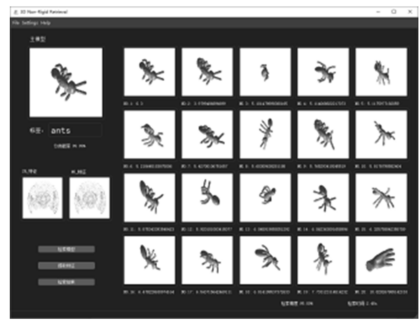
Fig. 10 3D model retrieval process

目前,大多数基于内容的三维模型检索系统主要采用文件检索或草图检索的方式,即选择要查找的三维模型,在三维形状模型库中根据相似度计算结果和设定的阈值返回查询结果。

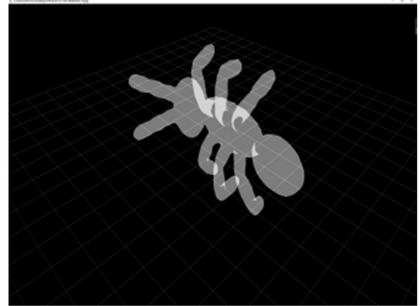
5.2 三维模型的可视化

目前,三维模型大多使用 OBJ 和 OFF 等格式存储,主要通过调用 3D 绘图库,利用模型的顶点 V 、三角面片 F 等信息绘制并渲染成三维形状。针对三维模型开发使用的可视化库有基于 Javascript 的 Threejs 库和基于 Python 的 Pyqtgraph 库,前者主要用于 B/S 端,且需要浏览器支持 WebGL,而后者用于 C/S 端。展示方式主要有 Web 在线浏览(Web3D)^[87]、客户端浏览;展示涉及范围包括模型展示、用户界面交互(模型放

大、缩小、旋转)。图 11 为三维模型检索系统界面的展示图。



(a) 检索界面



(b) 三维形状模型(模型可缩放、旋转)

图 11 三维模型检索系统界面的展示图

Fig. 11 3D model retrieval system interface display diagram

5.3 三维模型的存储管理

传统的针对三维模型的存储管理都是使用传统的单点文件系统或数据库系统进行处理,在面对海量增长的三维模型文件时处理效能会降低,故采用分布式存储管理三维形状模型文件是未来的发展趋势。

然而,采用分布式存储管理三维模型会面临文件大小不一,以及分布式存储对三维模型文件数据结构产生影响的问题。为了解决三维模型文件大小差异的问题,李海生等^[88]提出了基于 HDFS 的三维存储与管理策略,该策略考虑了三维模型文件大小差异对集群性能的影响,并结合三维模型文件结构的特点,设计和实现了使用 MapReduce 对具有严格要求的三维模型文件处理的方法。

5.4 三维模型的相似度计算

对三维模型进行相似度计算的常用方法有欧氏距离、马氏距离、余弦距离、Hausdorff 距离、曼哈顿距离、汉明距离、推土机距离、层次匹配距离等。

Biasotti 等^[89]提出了基于函数变换相似度的计算方法,包括基于 Riemannian 流形分析的 Gromov-Hausdorff 距离、基于热核扩散理论的扩散几何距离,以及基于代数同构变换与 Morse 理论的拟欧氏距离。Tabia 等^[90]提出基于顶点多特征融合的协方差矩阵作为三维模型的特征描述,采用 Riemannian 流形分析方法定义相似度距离,实现三维模型的检索与匹配。Li 等^[91]构建了一种融合 PSB 标准库的基准模型库,用于对特征提取、相似度计算的性能进行分析。Wang 等^[92]采用 PCA 分析方法,采用训练集稀疏系数作为特征向量,通过训练集定义了一种 Mahalanobis 距离,在人体三维模型上实现整体与部分的检索。

在相似度计算方面,这些算法的检索相似度计算基本都

使用单一特征,在多特征融合模型与相似度计算方法方面仍有待进行进一步的研究。

5.5 三维模型分类与检索性能评价标准

三维模型分类与检索任务中常见的评价性能指标有:查准率 P (Precision)、查全率 R (Recall)、 PR 曲线 (Precision-Recall curves)、最近邻方法 NN 、第 1 层级 FT 和第 2 层级 ST 、 E 度量、折扣的累积结果 DCG ^[93-94]、平均检索精度 mAP (mean Average Precision)。其中常用的有:查准率 P ,指检索返回结果中同类模型的占比;查全率 R ,指检索返回结果中同类模型在数据集中该类模型总量的占比;平均检索精度 mAP ,即同一类模型检索精度的均值,而检索精度 AP 定义为检索返回结果中同类模型的占比。查准率 P 和查全率 R 通过函数关系生成 PR 曲线,一般而言, PR 曲线所围面积越大,检索性能就越好。

6 面临的挑战

6.1 没有统一标准格式的三维形状数据表示

三维形状数据有多样化格式,如 OBJ, OFF, STL, PLY, 3DS, DXF, AMF 和 3MF 等,目前应用得较多的格式是 OBJ 和 OFF,但仍然没有一个统一标准格式的三维形状数据表示,这其实是由大数据时代的特点所决定的,即数据量爆炸性增长、数据源的多样性以及异构性所带来的影响。这决定了针对不同数据源以及不同任务都有特定的处理方式,从而使得难以移植一些成熟的深度学习神经网络模型,也难以移植迁移学习 (Transfer Learning)^[95]、强化学习 (Reinforcement Learning)^[96] 等一些机器学习方法模型。

6.2 深度学习网络模型参数多、运算量大、训练时间长

相比二维视图,三维模型储存的信息更加丰富,因此在相同分辨率情况下,深度学习神经网络模型的运算量比视图大得多。例如,用于三维模型体素化表示的 3D ShapeNets^[10] 神经网络结构,其在高分辨率模型下的计算量相比二维视图方法呈现出几何级数的增长,从而加重计算负载和存储压力。另外,使用深度学习神经网络模型时,需要大量的训练样本和训练时间去训练模型,因此高效的网络设计和训练方法以及平衡计算性能和三维形状特征表达的精细度是未来研究的重要方向。

6.3 三维形状深度学习神经网络结构缺乏针对性

目前,能应用深度学习模型的数据类型表征有视图、点云、网格、体素,这些模型都需要将三维模型转化为相应的数据类型进行处理,并且在转换过程中可能损失部分信息。然而,PointNet^[12] 和 MeshNet^[14] 等直接基于原始数据的深度学习神经网络模型^[75,97] 还较少,这限制了深度学习算法在三维形状特征提取领域的应用范围。

6.4 非刚体三维形状的姿态变化复杂、变换多样

非刚体形变的复杂性和多样性,导致针对非刚体的特征提取难度大且易受 12 种非刚性变换的影响。由于非刚体三维物体的姿态是任意的,尽管非刚体在高层语义表达上是同一类,但在姿态表达上却不同。因此,在设计深度学习神经网络模型时,需要考虑对姿态变化的变换不变性。这方面可以借鉴 PointNet^[12] 引入 Max-Pooling 步骤消除点云旋转的思

路,或者考虑使用姿态归一化^[98] 的操作。

6.5 针对三维形状的公共数据集较少

相比 ImageNet^[99] 等千万数量级的二维图像数据集,传统的刚体或非刚体三维形状数据公共集仍然很小。典型的常用三维形状公共数据集包括 ShapeNet^[100], ModelNet10/40^[10], SHREC2010/2011/2014/2015/2016/2017^[93] 等,但目前仍旧缺乏针对三维形状的大规模公开数据集,尤其是缺乏非刚体三维形状的大规模公开数据集。而深度学习算法仍需要一定规模数量的数据集样本来去训练使用。因此,目前更加需要高效的深度学习模型的设计与训练。

针对数据集样本少的问题,深度学习模型常用数据扩充方法来解决数据样本少的问题,从而提高深度学习模型的泛化能力。然而,目前针对三维形状的数据扩充方法仍较缺乏,虽然也有针对三维模型的插值^[101] 和生成^[102] 的研究,但仍然需要将其结合到深度学习方法中。

7 展望

未来,针对基于深度学习的三维形状特征提取技术以及分类与检索等任务的应用仍有待挖掘:1) 跨域检索^[103],优化任务界面,开发语义检索方式;2) 设计并运用新型深度学习神经网络模型,提取具有低维、高鉴别力的三维形状特征,如尝试使用 GAN 进行数据分类与检索^[31,104];3) 开发和更高效的方式,如使用二进制存储,或提出新的更高效的数据编码方式、分布式处理系统;4) 提高深度学习网络训练时间效率,研究高效的深层网络训练方法,开发和软件和应用硬件进行加速;5) 直接对原始三维模型数据进行处理,使用新型卷积核子吞吐式处理三维模型数据;6) 克服三维模型数据结构的异构性,实现模型格式转换的统一;7) 设计高效的多特征融合方法,生成具有自适应的融合系数;8) 持续构建大规模的三维形状数据集。

结束语 在三维形状分类、检索、识别、检测、重建、生成领域,可以利用深度学习算法的场景很多。特别地,三维形状蕴含的信息更加丰富,表达更加充分,使用深度学习算法可以更加充分地分析理解和挖掘三维形状内蕴的深层语义信息^[105],帮助现实场景中针对三维模型的应用设计生产。深度学习算法作为近几年人工智能领域最热门的研究应用之一,在教育、医疗、金融、制造业等领域取得了卓越的成就,更深地吸引着学者进行更深入的研究。本文对近年来基于深度学习的三维形状特征提取方法进行总结和阐述分析,以深度学习为技术手段,着重介绍深度学习算法和模型在刚体和非刚体三维形状特征提取方面的应用与研究,以及近年来三维形状检索系统的发展、三维模型预处理和相似度计算方法,最后探讨了深度学习与三维形状特征提取未来的发展趋势与面临的挑战。

参考文献

- [1] WEI X Z, QIU S Q, ZHU L, et al. Toward Support-Free 3D Printing: A Skeletal Approach for Partitioning Models[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(10):2799-2812.

- [2] CHU C H, WANG I J, WANG J B, et al. 3D parametric human face modeling for personalized product design[J]. *Advanced Engineering Informatics*, 2017, 32: 202-223.
- [3] FRADI A, LOUHICHI B, MAHJOUB M A, et al. 3D Object Retrieval Based on Similarity Calculation in 3D Computer Aided Design Systems[C]// *Proceedings of 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications*. Hammamet: IEEE, 2017: 160-165.
- [4] SHAN H, ZHANG Y, YANG Q, et al. 3-D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning From a 2-D Trained Network[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1522-1534.
- [5] HU W F, ZHAO S, REN Y, et al. 3D model dynamic cutting technology based on game engine[C]// *Proceedings of Annual Acis International Conference on Computer and Information Science*. Wuahn: IEEE, 2017: 1-6.
- [6] YANG Y B, LIN H, ZHU Q. Content-Based 3D Model Retrieval: A Survey[J]. *CHINESE JOURNAL OF COMPUTERS*, 2004, 27(10): 1297-1310. (in Chinese)
杨育彬, 林辉, 朱庆. 基于内容的三维模型检索综述[J]. *计算机学报*, 2004, 27(10): 1297-1310.
- [7] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view Convolutional Neural Networks for 3D Shape Recognition[C]// *Proceedings of International Conference on Computer Vision*. Santiago: IEEE, 2015: 945-953.
- [8] FENG Y, ZHANG Z, ZHAO X, et al. GVCNN: Group-View Convolutional Neural Networks for 3D Shape Recognition[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 264-272.
- [9] KANEZAKI A, MATSUSHITA Y, NISHIDA Y, et al. RotationNet: Joint Object Categorization and Pose Estimation Using Multiviews from Unsupervised Viewpoints[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 5010-5019.
- [10] WU Z, SONG S, KHOSLA A, et al. 3D ShapeNets: A deep representation for volumetric shapes[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1912-1920.
- [11] LI Y, PIRK S, SU H, et al. FPNN: Field Probing Neural Networks for 3D Data[C]// *Proceedings of Neural Information Processing Systems*. Barcelona: Curran Associates, 2016: 307-315.
- [12] QI C R, SU H, MO K, et al. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 77-85.
- [13] QI C R, YI L, SU H, et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space[C]// *Proceedings of Neural Information Processing Systems*. Long Beach: Curran Associates, 2017: 5099-5108.
- [14] FENG Y, FENG Y, YOU H, et al. MeshNet: Mesh Neural Network for 3D Shape Representation[C]// *Proceedings of AAAI Conference on Artificial Intelligence*. Honolulu: AAAI, 2019.
- [15] TATARCHENKO M, DOSOVITSKIY A, BROX T, et al. Occlusion Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs[C]// *Proceedings of International Conference on Computer Vision*. Venice: IEEE, 2017: 2107-2115.
- [16] LI H S, SUN L, WU Y J, et al. Survey on Feature Extraction Techniques for Non-Rigid 3D Shape Retrieval[J]. *Journal of Software*, 2018, 29(2): 483-505. (in Chinese)
李海生, 孙莉, 武玉娟, 等. 非刚性三维模型检索特征提取技术研究[J]. *软件学报*, 2018, 29(2): 483-505.
- [17] GUO Y L. Depth feature representation of 3d shape data[J]. *CCF Computer Vision Newsletter*, 2017(2): 8-11. (in Chinese)
郭裕兰. 三维形状数据的深度特征表示[J]. *CCF 计算机视觉专委会简报*, 2017(2): 8-11.
- [18] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [19] HINTON G E, OSINDERO S, TEH Y W, et al. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [20] HINTON G E, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [21] RUMELHART D E, HINTON G E, WILLIAMS R J, et al. Learning representations by back-propagating errors[J]. *Nature*, 1988, 323(6088): 696-699.
- [22] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]// *Proceedings of Neural Information Processing Systems*. Montreal: Curran Associates, 2014: 2672-2680.
- [23] ZOPH B, LE Q V. Searching for Activation Functions[C]// *Proceedings of International Conference on Learning Representations*. Vancouver, 2018.
- [24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E, et al. ImageNet Classification with Deep Convolutional Neural Networks[J]. *Neural Information Processing Systems*, 2012, 141(5): 1097-1105.
- [25] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]// *Proceedings of International Conference on Learning Representations*. San Diego, 2015.
- [26] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1-9.
- [27] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.
- [28] HUANG G, LIU Z, LAURENS V D M, et al. Densely Connected Convolutional Networks[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 2261-2269.
- [29] YANG Y, ZHONG Z, SHEN T, et al. Convolutional Neural Networks with Alternately Updated Clique[C]// *Proceedings of*

- IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018; 2413-2422.
- [30] SUN K, XIAO B, LIU D, et al. HRNet: Deep High-Resolution Representation Learning for Human Pose Estimation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019; 5693-5703.
- [31] CHEN X, DUAN Y, HOUTHOOFT R, et al. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets[C]//Proceedings of Neural Information Processing Systems. Barcelona: Curran Associates, 2016; 2180-2188.
- [32] LIU C, CHEN L C, SCHROFF F, et al. Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019; 82-92.
- [33] TAN M, LE Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[C]//Proceedings of International Conference on Machine Learning. Long Beach: ACM, 2019.
- [34] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5MB model size [C]//Proceedings of International Conference on Learning Representations. Toulon, 2017.
- [35] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018; 4510-4520.
- [36] ZHANG X, ZHOU X, LIU M, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018; 6848-6856.
- [37] XIE G, WANG J, ZHANG T, et al. Interleaved Structured Sparse Convolutional Neural Networks [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018; 8847-8856.
- [38] MEHTA S, RASTEGARI M, CASPI A, et al. ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation[C]//Proceedings of The European Conference on Computer Vision. Munich; Springer, 2018; 552-568.
- [39] GOYAL P, DOLLAR P, GIRSHICK R B, et al. Accurate, Large Minibatch SGD; Training ImageNet in 1 Hour[J]. arXiv; Computer Vision and Pattern Recognition, 2017.
- [40] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[C]//Proceedings of International Conference on Learning Representations. San Diego, 2015.
- [41] DUCHI J C, HAZAN E, SINGER Y, et al. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization [J]. Journal of Machine Learning Research, 2011, 12(7): 2121-2159.
- [42] LUO L, XIONG Y, LIU Y, et al. Adaptive Gradient Methods with Dynamic Bound of Learning Rate[C]//Proceedings of International Conference on Learning Representations. New Orleans, 2019.
- [43] FURUYA T, OHBUCHI R. Deep Aggregation of Local 3D Geometric Features for 3D Model Retrieval[C]//Proceedings of British Machine Vision Conference. York; BMVA Press, 2016.
- [44] XIA Q, LI S, HAO A M, et al. Deep Learning for Digital Geometry Processing and Analysis: A Review[J]. Journal of Computer Research and Development, 2019, 56(1): 155-182. (in Chinese)
夏清, 李帅, 郝爱民, 等. 基于深度学习的数字几何处理与分析技术研究进展[J]. 计算机研究与发展, 2019, 56(1): 155-182.
- [45] KAZHDAN M M, FUNKHOUSER T A, RUSINKIEWICZ S, et al. Rotation invariant spherical harmonic representation of 3D shape descriptors[C]//Proceedings of Eurographics Symposium on Geometry Processing. Aachen: The Eurographics Association, 2003; 156-164.
- [46] CHEN D Y, TIAN X P, SHEN Y T, et al. On Visual Similarity Based 3D Model Retrieval[J]. Computer Graphics Forum, 2003, 22(3): 223-232.
- [47] MAHMOUDI M, SAPIRO G. Three-dimensional point cloud recognition via distributions of geometric distances[J]. Graphical Models, 2009, 71(1): 22-31.
- [48] SUN J, OVSEJANIKOV M, GUIBAS L. A Concise and Provably Informative Multi-Scale Signature Based on Heat Diffusion[J]. Computer Graphics Forum, 2009, 28(5): 1383-1392.
- [49] AUBRY M, SCHLICKWEI U, CREMERS D. The wave kernel signature: A quantum mechanical approach to shape analysis [C]//Proceedings of International Conference on Computer Vision Workshops. Barcelona; IEEE, 2011; 1626-1633.
- [50] KLOKOV R, LEMPITSKY V. Escape from Cells; Deep Kd-Networks for the Recognition of 3D Point Cloud Models[C]//Proceedings of International Conference on Computer Vision. Venice; IEEE, 2017; 863-872.
- [51] WANG Y, SUN Y, LIU Z, et al. Dynamic Graph CNN for Learning on Point Clouds[J]. arXiv; Computer Vision and Pattern Recognition, 2018.
- [52] JOHNS E, LEUTENEGGER S, DAVISON A J. Pairwise Decomposition of Image Sequences for Active Multi-View Recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas; IEEE, 2016; 3813-3822.
- [53] WANG C, SAMARI B, SIDDIQI K, et al. Local Spectral Graph Convolution for Point Set Feature Learning[C]//Proceedings of European Conference on Computer Vision. Munich; Springer, 2018; 56-71.
- [54] LI J, CHEN B M, LEE G H, et al. SO-Net: Self-Organizing Network for Point Cloud Analysis[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018; 9397-9406.
- [55] ZHAO Y, BIRDAL T, DENG H, et al. 3D Point-Capsule Networks[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019; 1009-1018.
- [56] LAN S, YU R, YU G, et al. Modeling Local Geometric Structure of 3D Point Clouds using Geo-CNN[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long

- Beach; IEEE, 2019: 998-1008.
- [57] KATO H, USHIKU Y, HARADA T. Neural 3D Mesh Renderer [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018: 3907-3916.
- [58] GROUEIX T, FISHER M, KIM V G, et al. AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018: 216-224.
- [59] ZHOU Y, ZENG F, QIAN J, et al. 3D shape classification and retrieval based on polar view [J]. Information Sciences, 2019, 474: 205-220.
- [60] JIANG J, BAO D, CHEN Z, et al. MLVCNN: Multi-Loop-View Convolutional Neural Network for 3D Shape Retrieval [C]//Proceedings of AAAI Conference on Artificial Intelligence. Honolulu; AAAI, 2019.
- [61] LI Y M, XUE K X, GAO Z, et al. 3-D Model Retrieval Algorithm Based on Residual Network [J]. COMPUTER SCIENCE, 2019, 46(3): 148-153. (in Chinese)
李荫民, 薛凯心, 高赞, 等. 基于残差网络的三维模型检索算法 [J]. 计算机科学, 2019, 46(3): 148-153.
- [62] MATURANA D, SCHERER S. VoxNet: A 3D Convolutional Neural Network for real-time object recognition [C]//Proceedings of Intelligent Robots and Systems. Hamburg; IEEE, 2015: 922-928.
- [63] YANG J, WANG S, ZHOU P. Recognition and Classification for Three-Dimensional Model Based on Deep Voxel Convolution Neural Network [J]. Acta Optica Sinica, 2019, 39(4): 314-324. (in Chinese)
杨军, 王顺, 周鹏. 基于深度体素卷积神经网络的三维模型识别分类 [J]. 光学学报, 2019, 39(4): 314-324.
- [64] WANG C, CHENG M, SOHEL F, et al. NormalNet: A voxel-based CNN for 3D object classification and retrieval [J]. Neurocomputing, 2019, 323: 139-147.
- [65] KUMAWAT S, RAMAN S. LP-3DCNN: Unveiling Local Phase in 3D Convolutional Neural Networks [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019: 4903-4912.
- [66] RIEGLER G, ULUSOY A O, GEIGER A, et al. OctNet: Learning Deep 3D Representations at High Resolutions [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu; IEEE, 2017: 6620-6629.
- [67] WANG P S, LIU Y, GUO Y X, et al. O-CNN: octree-based convolutional neural networks for 3D shape analysis [J]. ACM Transactions on Graphics, 2017, 36(4): 72:1-72:11.
- [68] ZHOU Y, ZENG F. 2D compressive sensing and multi-feature fusion for effective 3D shape retrieval [J]. Information Sciences, 2017, 409-410: 101-120.
- [69] HEGDE V, ZADEH R. Fusionnet: 3D Object Classification Using Multiple Data Representations [J]. arXiv: Computer Vision and Pattern Recognition, 2016.
- [70] YOU H, FENG Y, JI R, et al. PVNet: A Joint Convolutional Network of Point Cloud and Multi-View for 3D Shape Recognition [J]. Acm Multimedia, 2018: 1310-1318.
- [71] YOU H, FENG Y, ZHAO X, et al. PVRNet: Point-View Relation Neural Network for 3D Shape Recognition [C]//Proceedings of AAAI Conference on Artificial Intelligence. Honolulu; AAAI, 2019.
- [72] BOYER E, BRONSTEIN A M, BRONSTEIN M M, et al. SHREC 2011: robust feature detection and description benchmark [C]//Proceedings of Eurographics Conference on 3D Object Retrieval. Llandudno; The Eurographics Association, 2011: 71-78.
- [73] ZENG H, LIU Y, LIU J, et al. Non-Rigid 3D Model Retrieval Based on Quadruplet Convolutional Neural Networks [J]. IEEE Access, 2018, 6: 76087-76097.
- [74] LUCIANO L, HAMZA A B. Deep learning with geodesic moments for 3D shape classification [J]. Pattern Recognition Letters, 2018: 182-190.
- [75] HAN Z, LIU Z, HAN J, et al. Unsupervised Learning of 3-D Local Features From Raw Voxels Based on a Novel Permutation Voxelization Strategy [J]. IEEE Transactions on Systems, Man, and Cybernetics, 2019, 49(2): 481-494.
- [76] SINHA A, BAI J, RAMANI K, et al. Deep Learning 3D Shape Surfaces Using Geometry Images [C]//Proceedings of European Conference on Computer Vision. Amsterdam; Springer, 2016: 223-240.
- [77] BAI J, SI Q L, QIN F W. 3D Model Classification and Retrieval Based on CNN and Voting Scheme [J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(2): 123-134. (in Chinese)
白静, 司庆龙, 秦飞巍. 基于卷积神经网络和投票机制的三维模型分类与检索 [J]. 计算机辅助设计与图形学学报, 2019, 31(2): 123-134.
- [78] BU S, WANG L, HAN P, et al. 3D shape recognition and retrieval based on multi-modality deep learning [J]. Neurocomputing, 2017, 259(2017): 183-193.
- [79] ZENG H, LIU Y, LI S, et al. Convolutional Neural Network Based Multi-feature Fusion for Non-rigid 3D Model Retrieval [J]. Journal of Information Processing Systems, 2018, 14(1): 176-190.
- [80] HE T, HUANG H, YI L, et al. GeoNet: Deep Geodesic Networks for Point Cloud Analysis [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019: 6888-6897.
- [81] YU L, LI X, FU C, et al. PU-Net: Point Cloud Upsampling Network [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City; IEEE, 2018: 2790-2799.
- [82] PAQUET E, RIOUX M. Nefertiti: A Query by Content Software for Three-Dimensional Models Databases Management [C]//International Conference on Recent Advances in 3-D Digital Imaging and Modeling (3DIM '97). Ottawa, Ontario, Canada; DBLP, 1997.
- [83] MIN P, HALDERMAN J A, KAZHDAN M M, et al. Early experiences with a 3D model search engine [C]//Proceedings of International Conference on 3D Web Technology. New York; ACM, 2003: 7-18.
- [84] SAGARA H, TAKANO S, OKADA Y, et al. 3D Model Data

- Retrieval System Using KAZE Feature for Accepting 2D Image as Query[C]//Proceedings of International Conference on Complex, Intelligent, and Software Intensive Systems. Fukuoka: IEEE, 2016: 617-622.
- [85] SINTUNATA V, AOKI T. 3D object retrieval system using skewness database[C]//Proceedings of International Conference on Control, Decision and Information Technologies. St. Julian's: IEEE, 2016: 123-128.
- [86] KAO C. Shape-based 3D model retrieval system[J/OL]. International Journal of Computers and Applications. <https://www.tandfonline.com/doi/abs/10.1080/1206212X.2016.1226541>.
- [87] WU Z Q, ZHU Z M, LIU W, et al. Retrieving 3D Models from Institutional Repository[J]. Data Analysis and Knowledge Discovery, 2017, 1(1): 73-80. (in Chinese)
吴志强, 祝忠明, 刘巍, 等. 机构知识库三维模型检索与展示技术研究与实践[J]. 数据分析与知识发现, 2017, 1(1): 73-80.
- [88] LI H S, LAI L, CAI Q, et al. 3D Model Storage and Shape Distribution in Hadoop Environment[J]. Journal of Computer Research and Development, 2014, 51(S2): 18-29. (in Chinese)
李海生, 赖龙, 蔡强, 等. Hadoop 环境下三维模型的存储及形状分布特征提取[J]. 计算机研究与发展, 2014, 51(S2): 18-29.
- [89] BIASOTTI S, CERRI A, BRONSTEIN A M, et al. Quantifying 3D Shape Similarity Using Maps; Recent Trends, Applications and Perspectives[C]//Proceedings of eurographics. Strasbourg: The Eurographics Association, 2014: 135-159.
- [90] TABIA H, LAGA H, PICARD D, et al. Covariance Descriptors for 3D Shape Matching and Retrieval[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 4185-4192.
- [91] LI B, LU Y, LI C, et al. A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries[J]. Computer Vision and Image Understanding, 2015, 131: 1-27.
- [92] WANG J, MA K, SINGH V K, et al. BodyPrint: Pose Invariant 3D Shape Matching of Human Bodies[C]//Proceedings of International Conference on Computer Vision. Santiago: IEEE, 2015: 1591-1599.
- [93] Princeton University. The Princeton shape benchmark [EB/OL]. (2005-03-15) [2019-07-05]. <http://shape.cs.princeton.edu/benchmark/>.
- [94] SHILANE P, MIN P, KZAHDAN M, et al. The Princeton shape benchmark[C]//Proceedings of Conference on the Shape Modeling Applications. Genova: IEEE, 2004.
- [95] PAN S J, YANG Q. A Survey on Transfer Learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [96] SUTTON R S, BARTO A G. Reinforcement Learning: An Introduction[J]. IEEE Transactions on Neural Networks, 1998, 9(5): 1054-1054.
- [97] HAN Z, LIU Z, HAN J, et al. Mesh Convolutional Restricted Boltzmann Machines for Unsupervised Learning of Features With Structure Preservation on 3-D Meshes[J]. IEEE Transactions on Neural Networks, 2017, 28(10): 2268-2281.
- [98] WANG J T, ZHAO L, QI X B. Face Recognition Method Based on Adaptive 3D Morphable Model and Multiple Manifold Discriminant Analysis[J]. Computer Science, 2017, 44(S1): 232-235, 239. (in Chinese)
王渐韬, 赵丽, 齐兴斌. 自适应三维形变模型结合流形分析的人脸识别方法[J]. 计算机科学, 2017, 44(S1): 232-235, 239.
- [99] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [100] CHANG A X, FUNKHOUSER T A, GUIBAS L J, et al. ShapeNet: An Information-Rich 3D Model Repository [J]. arXiv: 1512.03012, 2015.
- [101] DUAN P. Anisotropy Radial Basis Function Spatial Interpolation Model Research for 3D Spatial Field[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(12): 1696. (in Chinese)
段平. 三维空间场各向异性径向基函数空间插值模型研究[J]. 测绘学报, 2018, 47(12): 1696.
- [102] WU J, ZHANG C, XUE T, et al. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling [C]//Proceedings of Neural Information Processing Systems. Barcelona: Curran Associates, 2016: 82-90.
- [103] PENG Y X, QI J W, HUANG X. Current Research Status and Prospects on Multimedia Content Understanding[J]. Journal of Computer Research and Development, 2019, 56(1): 183-208. (in Chinese)
彭宇新, 綦金玮, 黄鑫. 多媒体内容理解的研究现状与展望[J]. 计算机研究与发展, 2019, 56(1): 183-208.
- [104] ODENA A. Semi-Supervised Learning with Generative Adversarial Networks[J]. arXiv: 1606.01583, 2016.
- [105] LI H S, WU Y J, ZHENG Y P, et al. A Survey of 3D Data Analysis and Understanding Based on Deep Learning[J/OL]. Chinese Journal of Computers, 2019: 1-25. [2019-07-25]. <http://kns.cnki.net/kcms/detail/11.1826.TP.20190709.1509.002.html>. (in Chinese)
李海生, 武玉娟, 郑艳萍, 等. 基于深度学习的三维数据分析理解方法研究综述[J/OL]. 计算机学报, 2019: 1-25. [2019-07-25]. <http://kns.cnki.net/kcms/detail/11.1826.TP.20190709.1509.002.html>.