



(12)发明专利申请

(10)申请公布号 CN 108960288 A

(43)申请公布日 2018.12.07

(21)申请号 201810580064.1

(22)申请日 2018.06.07

(71)申请人 山东师范大学

地址 250014 山东省济南市文化东路88号

(72)发明人 王吉华 李梦鸽

(74)专利代理机构 济南圣达知识产权代理有限公司 37221

代理人 李琳

(51)Int.Cl.

G06K 9/62(2006.01)

G06N 3/04(2006.01)

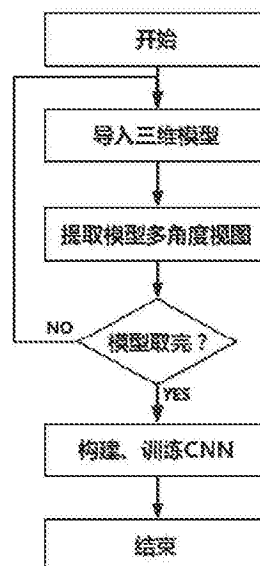
权利要求书2页 说明书6页 附图1页

(54)发明名称

基于卷积神经网络的三维模型分类方法及系统

(57)摘要

本发明公开了一种基于卷积神经网络的三维模型分类方法及系统,包括以下步骤:获取待分类的三维模型,对三维模型多角度投影得到视图,作为构建的卷积神经网络的输入,训练卷积神经网络,经过网络层次地提取三维模型形状特征,得到三维模型分类结果,本发明避免了模型数据转化等操作,有效提高三维模型分类效率,推动计算机视觉和计算机辅助设计研究的发展。



1. 一种基于卷积神经网络的三维模型分类方法,其特征是:包括以下步骤:

获取待分类的三维模型,对三维模型多角度投影得到视图,作为构建的卷积神经网络的输入,训练卷积神经网络,经过网络层次地提取三维模型形状特征,得到三维模型分类结果。

2. 如权利要求1所述的一种基于卷积神经网络的三维模型分类方法,其特征是:对三维模型多角度投影得到视图的过程中,具体包括:

(1) 获取三维模型文件,显示三维模型;

(2) 调整三维模型按照ZYX坐标轴放置,Y轴向上;

(3) 设定模型绕Y轴旋转一周,放置虚拟相机渲染,得到多角度的若干幅视图,循环导入三维模型文件,直到所有的三维模型视图渲染完毕;

(4) 将提取的多幅视图进行Picasa图片拼贴制作,按照旋转顺序整理提取到的多幅视图,从中选择需要的视图,使其排列成规则的行和列,呈现为正方形页面格式,创建拼贴,形成一幅汇总视图,导出为任意像素的图像,作为卷积神经网络的输入。

3. 如权利要求1所述的一种基于卷积神经网络的三维模型分类方法,其特征是:构建的卷积神经网络包括输入层、隐含层及输出层,且隐含层由多个卷积层和池化层交替构成。

4. 如权利要求3所述的一种基于卷积神经网络的三维模型分类方法,其特征是:在卷积层,使用若干过滤器在输入的汇总视图上滑动进行卷积运算,提取局部特征;每个过滤器共享卷积上一层的特征图的参数。

5. 如权利要求3所述的一种基于卷积神经网络的三维模型分类方法,其特征是:在卷积层,利用卷积区域内的各个单元与其相对应的权值乘积之和,与每个过滤器的固定偏置参数的加法运算,最终形成单元特征值。

6. 如权利要求3或4或5所述的一种基于卷积神经网络的三维模型分类方法,其特征是:卷积层的激活函数选择ReLU激活函数,对卷积后得到的单元特征值进行非线性映射归一化。

7. 如权利要求1所述的一种基于卷积神经网络的三维模型分类方法,其特征是:池化层构建在卷积层后,通过平均池化或最大池化操作对特征压缩映射。

8. 如权利要求1所述的一种基于卷积神经网络的三维模型分类方法,其特征是:根据提取到的特征对三维模型进行分类的基本步骤如下:

(1) 池化层输出特征值至全连接层,全连接层连接所有的特征,将低层特征抽象为高层特征。

(2) 在第一个全连接层后添加一个池化层,进一步降低特征的维度。

(3) 第二个全连接层将降维的表征模型的抽象特征输出为特征向量。

(4) 最后一个全连接层为Softmax分类层,使用Softmax回归将输入的特征向量输出为模型属于每个类别的概率,选取概率最大的类作为模型分类。

9. 一种基于卷积神经网络的三维模型分类系统,其特征是:运行于处理器或存储器上,被配置为执行以下指令:

获取待分类的三维模型,对三维模型多角度投影得到视图,作为构建的卷积神经网络的输入,训练卷积神经网络,经过网络层次地提取三维模型形状特征,输入三维模型分类结果。

10. 一种基于卷积神经网络的三维模型分类系统,其特征是:包括:
- 导入模块,被配置为导入待分类的三维模型;
 - 提取模块,并配置为提取三维模型多角度视图;
 - 学习单元构建模块,被配置为构建卷积神经网络模型;
 - 特征提取模块,被配置为使用视图表征三维模型作为网络的输入,训练卷积神经网络模型,完成特征的提取;
 - 分类模块,被配置为根据卷积神经网络模型提取的三维模型抽象特征,对三维模型进行分类。

基于卷积神经网络的三维模型分类方法及系统

技术领域

[0001] 本发明涉及一种基于卷积神经网络的三维模型分类方法及系统。

背景技术

[0002] 在日常生活中,大量物体通常以三维模型形式存在,能够被人类视觉神经系统所辨识。随着计算机硬件与软件技术的迅速发展,三维模型被广泛应用于工业生产、虚拟现实、医学影像等领域。三维模型数量大规模的增长,对三维模型的管理工作及生产设计造成了困难。

[0003] 因此,为了解决人们对已有三维模型的查找和复用需求,三维模型分类技术应运而生。如何在尽可能完美的表达三维模型形状的情况下对三维模型识别是当前的研究热点。

[0004] 传统的三维模型分类方法过于依赖工程师的经验知识,需要消耗大量时间和精力;并且鉴于三维数据高维性和复杂性特点,直接用于表达三维形状,计算代价大,特征提取复杂。

发明内容

[0005] 本发明为了解决上述问题,提出了一种基于卷积神经网络的三维模型分类方法及系统,本发明通过对三维模型多角度投影得到视图,作为卷积神经网络的输入,经过网络层次地提取三维模型形状特征,最终得到一个三维模型分类结果。避免了模型数据转化等操作,有效提高三维模型分类效率,推动计算机视觉和计算机辅助设计研究的发展。

[0006] 为了实现上述目的,本发明采用如下技术方案:

[0007] 一种基于卷积神经网络的三维模型分类方法,包括以下步骤:

[0008] 获取待分类的三维模型,对三维模型多角度投影得到视图,作为构建的卷积神经网络的输入,训练卷积神经网络,经过网络层次地提取三维模型形状特征,得到三维模型分类结果。

[0009] 进一步的,对三维模型多角度投影得到视图的过程中,具体包括:

[0010] (1) 获取三维模型文件,显示三维模型;

[0011] (2) 调整三维模型按照ZYX坐标轴放置,Y轴向上;

[0012] (3) 设定模型绕Y轴旋转一周,放置虚拟相机渲染,得到多角度的若干幅视图,循环导入三维模型文件,直到所有的三维模型视图渲染完毕;

[0013] (4) 将提取的多幅视图进行Picasa图片拼贴制作,按照旋转顺序整理提取到的多幅视图,从中选择需要的视图,使其排列成规则的行和列,呈现为正方形页面格式,创建拼贴,形成一幅汇总视图,导出为任意像素的图像,作为卷积神经网络的输入。

[0014] 可以循环制作图片拼贴,直至所有模型的汇总视图制作完毕。将所有模型的汇总视图整理为训练集和测试集,作为卷积神经网络的输入。

[0015] 进一步的,构建的卷积神经网络包括输入层、隐含层及输出层,且隐含层由多个卷

积层和池化层交替构成。

[0016] 更进一步的,在卷积层,使用若干过滤器在输入的汇总视图上滑动进行卷积运算,提取局部特征;每个过滤器共享卷积上一层的特征图的参数。

[0017] 更进一步的,在卷积层,利用卷积区域内的各个单元与其相对应的权值乘积之和,与每个过滤器的固定偏置参数的加法运算,最终形成单元特征值。

[0018] 更进一步的,卷积层的激活函数选择ReLU激活函数,对卷积后得到的单元特征值进行非线性映射归一化。

[0019] 更进一步的,池化层构建在卷积层后,通过平均池化或最大池化操作对特征压缩映射。

[0020] 进一步的,根据提取到的特征对三维模型进行分类的基本步骤如下:

[0021] (1)池化层输出特征值至全连接层,全连接层连接所有的特征,将底层特征抽象为高层特征。

[0022] (2)在第一个全连接层后添加一个池化层,进一步降低特征的维度。

[0023] (3)第二个全连接层将降维的表征模型的抽象特征输出为特征向量。

[0024] (4)最后一个全连接层为Softmax分类层,使用Softmax回归将输入的特征向量输出为模型属于每个类别的概率,选取概率最大的类作为模型的分类。

[0025] 一种基于卷积神经网络的三维模型分类系统,运行于处理器或存储器上,被配置为执行以下指令:

[0026] 获取待分类的三维模型,对三维模型多角度投影得到视图,作为构建的卷积神经网络的输入,训练卷积神经网络,经过网络层次地提取三维模型形状特征,输入三维模型分类结果。

[0027] 一种基于卷积神经网络的三维模型分类系统,包括:

[0028] 导入模块,被配置为导入待分类的三维模型;

[0029] 提取模块,并配置为提取三维模型多角度视图;

[0030] 学习单元构建模块,被配置为构建卷积神经网络模型;

[0031] 特征提取模块,被配置为使用视图表征三维模型作为网络的输入,训练卷积神经网络模型,完成特征的提取;

[0032] 分类模块,被配置为根据卷积神经网络模型提取的三维模型抽象特征,对三维模型进行分类。

[0033] 与现有技术相比,本发明的有益效果为:

[0034] 本发明利用多角度视图表征三维模型,通过构建卷积神经网络自动提取模型特征,实现对三维模型分类。该方法使用多角度视图描述三维模型外形,能够很好的保持孔洞等几何特征。与传统分类方法相比,计算代价小,避免了显式特征的提取。在大规模的分类应用中过程高效。

附图说明

[0035] 构成本申请的一部分的说明书附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。

[0036] 图1为本发明模型分类的流程图;

具体实施方式：

[0037] 下面结合附图与实施例对本发明作进一步说明。

[0038] 应该指出,以下详细说明都是例示性的,旨在对本申请提供进一步的说明。除非另有指明,本文使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解相同含义。

[0039] 需要注意的是,这里所使用的术语仅是为了描述具体实施方式,而非意图限制根据本申请的示例性实施方式。如在这里所使用的,除非上下文另外明确指出,否则单数形式也意图包括复数形式,此外,还应当理解的是,当在本说明书中使用术语“包含”和/或“包括”时,其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0040] 在本发明中,术语如“上”、“下”、“左”、“右”、“前”、“后”、“竖直”、“水平”、“侧”、“底”等指示的方位或位置关系为基于附图所示的方位或位置关系,只是为了便于叙述本发明各部件或元件结构关系而确定的关系词,并非特指本发明中任一部件或元件,不能理解为对本发明的限制。

[0041] 本发明中,术语如“固接”、“相连”、“连接”等应做广义理解,表示可以是固定连接,也可以是一体地连接或可拆卸连接;可以是直接相连,也可以通过中间媒介间接相连。对于本领域的相关科研或技术人员,可以根据具体情况确定上述术语在本发明中的具体含义,不能理解为对本发明的限制。

[0042] 正如背景技术中所述的,传统的三维模型分类方法过于依赖工程师的经验知识,需要消耗大量时间和精力,且计算代价大,特征提取复杂,本发明为解决上述问题,提供了一种基于卷积神经网络的三维模型分类方法,包括以下步骤:

[0043] (1) 导入待分类的三维模型。

[0044] (2) 提取三维模型多角度视图。视图作为三维模型的原始表征,符合人类的视觉系统,可用作深度学习模型的直接输入数据。且多角度视图可以丰富表达三维模型形状。

[0045] (3) 构建卷积神经网络(CNN)模型。

[0046] (4) 使用视图表征三维模型作为网络的输入,训练CNN,完成特征的提取。

[0047] (5) 根据CNN提取的三维模型抽象特征,对三维模型进行分类。

[0048] 步骤(2)中提取三维模型多角度视图的基本步骤如下:

[0049] (1) 利用Solidworks引擎打开三维模型step文件,显示三维模型。

[0050] (2) 调整三维模型按照ZYX坐标轴放置,Y轴向上。

[0051] (3) 设定模型绕Y轴旋转一周,放置虚拟相机渲染,得到多角度的若干幅视图。循环导入step文件,直到所有的三维模型视图渲染完毕。

[0052] (4) 对获取的视图按序整理,拼贴为一幅视图。汇总视图,作为卷积神经网络的输入。

[0053] 当然,步骤(1)中的应用软件或文件类型可以在其他实施例中进行替换。

[0054] 步骤(3)中卷积神经网络的基本构成要素为:卷积层、池化层和全连接层。构建的CNN模型由输入层、隐含层及输出层组成。隐含层即由多个卷积层和池化层交替构成。

[0055] 步骤(4)中在卷积层,使用过滤器在输入的视图上滑动进行卷积运算,提取局部特征。每个过滤器能够参数共享地卷积上一层的特征图,使用多个过滤器可学习到输入数据

的丰富特征。其中，卷积层的运算公式为：

$$[0056] \quad y_j = b_j + \sum_i w_{ij} * x_i \quad (1)$$

[0057] 其中，卷积区域内的单元为 x_i ，相对应的权值为 w_{ij} ，二者做乘积运算，之后将得到的乘积值与每个过滤器的固定偏置参数 b_j 做加法运算，最终形成一个单元值 y_j 。卷积层的激活函数选择ReLU激活函数，可以有效减轻网络训练时的梯度消失问题，从而提高学习速度。

[0058] ReLU激活函数为：

$$[0059] \quad \text{ReLU}(X) = \max(0, x) \quad (2)$$

[0060] 按照一定的方式使卷积后得到的特征值 y_j 非线性映射到归一化，与传统的sigmoid函数相比计算速度明显加快。卷积层通过若干过滤器提取到的特征若直接用于分类，网络将会产生过拟合，所以需要将调整后的特征进行降采样处理，即池化。池化层一般构建在卷积层后，通过平均池化或最大池化操作对特征压缩映射，简化网络计算复杂度。

[0061] 池化的运算公式为：

$$[0062] \quad y_j = \begin{cases} \frac{1}{9} \sum_{i=0}^8 x_i \\ \max \{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\} \end{cases} \quad (3)$$

[0063] 其中， x_i ($x=0, 1, 2, 3, 4, 5, 6, 7, 8$) 为网络上层输出特征图的输入像素值， \max 为最大池化操作， $\frac{1}{9} \sum_{i=0}^8 x_i$ 为平均池化操作。两种池化操作均在 $3*3$ 的区域内进行。 y_j 表示池化后的输出值。经过卷积层和池化层提取的特征为模型的主要特征。

[0064] 使用步骤(5)中提取到的特征对三维模型进行分类的基本步骤如下：

[0065] (1) 池化层输出特征值至全连接层，全连接层连接所有的特征，将低层特征抽象为高层特征。

[0066] (2) 在第一个全连接层后添加一个池化层，进一步降低特征的维度。

[0067] (3) 第二个全连接层将降维的表征模型的抽象特征输出为特征向量。

[0068] (4) 最后一个全连接层为Softmax分类层。使用Softmax回归将输入的特征向量输出为模型属于每个类别的概率，选取概率最大的类作为模型的分类。

[0069] 更为具体的，如图1所示，基于卷积神经网络的三维模型分类方法主要包括以下几个过程：

[0070] 过程1：三维模型的导入。

[0071] 过程2：提取三维模型多角度视图

[0072] 第一步：调整导入的三维模型位置，使坐标轴Y轴向上。

[0073] 第二步：设置运动算例。选择模型绕Y轴旋转一周。放置虚拟相机渲染旋转过程，输出指定的16幅视图。

[0074] 第三步：按照时间顺序，使用拼贴工具将16幅视图整理为1幅视图，归化视图为 $256*256$ 的RGB图像。

[0075] 过程3：构建卷积神经网络(CNN)模型。

[0076] 第一步：确定CNN网络层数。网络基本构成要素为：卷积层、池化层和全连接层。

[0077] 第二步：定义CNN网络结构。5层卷积(Conv1, ..., 5)，3层全连接(Fc6, ..., 8)。Conv1

和Conv2、Conv5、Fc6层后连接池化层,即Pool1、Pool2、Pool5、Pool6。卷积层的过滤器在输入数据视图或特征图上滑动,即与其局部区域连接,故整个视图或特征图权值共享。

[0078] 第三步:定义网络各层参数。

[0079] Conv1层使用96个11*11大小的过滤器,过滤器在视图上滑动步幅为4;Pool1层使用3*3大小的池化核进行最大池化,步长为2;Conv2层使用256个5*5的过滤器,像素填充为2;Pool2层使用3*3池化核进行最大池化,步长为2;Conv3层使用384个3*3的过滤器,像素填充1;Conv4层使用384个3*3过滤器,像素填充1;Conv5使用256个3*3过滤器,像素填充1;Pool5层使用3*3池化核进行最大池化,步长为2;Fc6层使用4096个神经元;Pool6层使用6*6池化核进行平均池化,步长为1;Fc7层使用4096个神经元;Fc8层使用n个神经元,表示类别。

[0080] 过程4:将获取的视图预整理完毕,作为卷积神经网络的直接输入数据。开始训练网络,学习输入数据的特征。

[0081] 第一步:输入256*256的视图,经过Conv1层后,输出96个55*55的特征图。其中,计算任意卷积层输出的公式由公式(1)解为:

$$[0082] \quad O = (W - K + 2 * P) / S + 1 \quad (4)$$

[0083] 其中,O为输出特征图大小,W为输入图像大小,K为卷积层过滤器大小,P表示像素填充,S为过滤器滑动步幅。

[0084] 第二步:Conv1层输出的特征图经过Pool1层最大池化,压缩映射为96个27*27的特征图。

[0085] 第三步:Conv2层将96个27*27的特征图经过256个5*5的过滤器,输出256个27*27的特征图。

[0086] 第四步:Conv2层将256个27*27的特征图经过Pool2层最大池化,输出256个13*13的特征图。

[0087] 第五步:Conv3层将256个13*13的特征图经过384个3*3的过滤器,输出384个13*13的特征图。

[0088] 第六步:Conv4层将384个13*13的特征图经过384个3*3的过滤器,依旧输出384个13*13的特征图。

[0089] 第七步:Conv5层将384个13*13的特征图经过256个3*3的过滤器,输出256个13*13的特征图。

[0090] 第八步:Conv5层输出的特征图经过Pool5层最大池化,压缩为256个6*6的特征图。

[0091] 过程5:全连接层将卷积层和池化层提取的局部特征与所有的神经元相连,从而综合为全局特征。根据全连接层综合到的三维模型抽象特征,进行分类。

[0092] 第一步:Fc6层4096个神经元与输入的特征图中每个像素相连接。通过Dropout操作,随机舍弃一些神经元的权重,从而加速训练,进一步防止过拟合。最后输出一个4096维向量。

[0093] 第二步:Fc6层输出的4096向量经过Pool6层平均池化,输出4096维向量。

[0094] 第三步:Fc7层使用4096个神经元与上层输出进行全连接,依旧输出一个向量。

[0095] 第四步:Fc8层n个神经元与上层输出向量相连。使用Softmax回归函数,得到每个类别的概率,最终选取概率最大的类别作为三维模型的类。

[0096] 在本具体实施方式中,还提供一种基于卷积神经网络的三维模型分类系统,运行

于处理器或存储器上,被配置为执行以上处理过程。

[0097] 以上所述仅为本申请的优选实施例而已,并不用于限制本申请,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

[0098] 上述虽然结合附图对本发明的具体实施方式进行了描述,但并非对本发明保护范围的限制,所属领域技术人员应该明白,在本发明的技术方案的基础上,本领域技术人员不需要付出创造性劳动即可做出的各种修改或变形仍在本发明的保护范围以内。

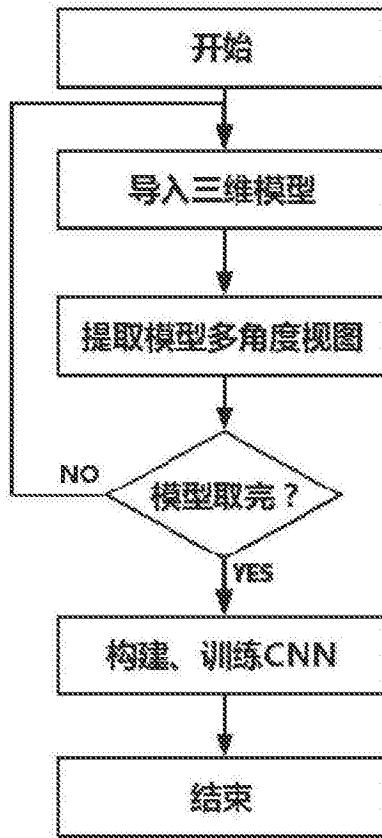


图1