

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 11 月 17 日 (17.11.2022)



(10) 国际公布号
WO 2022/236802 A1

(51) 国际专利分类号:
G06T 17/00 (2006.01)

(21) 国际申请号: PCT/CN2021/093783

(22) 国际申请日: 2021 年 5 月 14 日 (14.05.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518000 (CN)。

(72) 发明人: 王磊 (WANG, Lei); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong

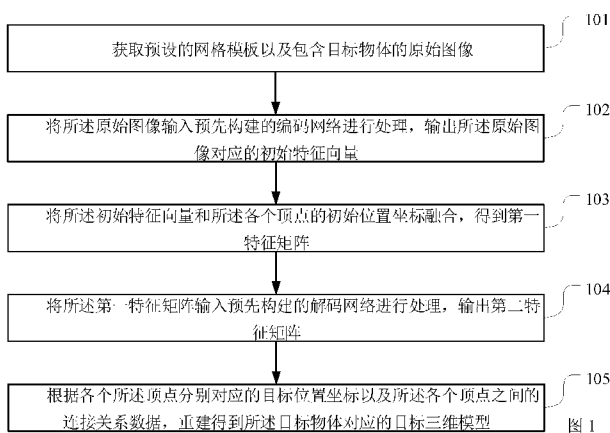
518000 (CN)。钟宏亮 (ZHONG, Hongliang); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518000 (CN)。林佩珍 (LIN, Peizhen); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518000 (CN)。程俊 (CHENG, Jun); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518000 (CN)。

(74) 代理人: 深圳中一联合知识产权代理有限公司 (SHENZHEN ZHONGYI UNION INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国广东省深圳市福田区园岭街道深南中路 1014 号报春大厦 9 楼 (5 号信箱), Guangdong 518028 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,

(54) Title: METHOD AND APPARATUS FOR RECONSTRUCTING OBJECT MODEL, AND TERMINAL DEVICE AND STORAGE MEDIUM

(54) 发明名称: 一种物体模型的重建方法、装置、终端设备和存储介质



- 101 Acquire a preset grid template, and an original image including a target object
102 Input the original image into a pre-constructed encoding network for processing, and output an initial feature vector corresponding to the original image
103 Fuse the initial feature vector and initial position coordinates of each vertex, so as to obtain a first feature matrix
104 Input the first feature matrix into a pre-constructed decoding network for processing, and output a second feature matrix
105 According to target position coordinates respectively corresponding to the vertices and connection relationship data between the vertices, perform reconstruction to obtain a target three-dimensional model corresponding to the target object

图 1

(57) Abstract: The present application relates to the technical field of image processing, and provided are a method and apparatus for reconstructing an object model, and a terminal device and a storage medium. The method comprises: firstly, acquiring an original image including a target object and a certain preset grid template, extracting a feature vector of the original image, and then fusing the feature vector and position coordinates of each vertex in the grid template, so as to obtain a feature matrix; then using a decoding network to process the feature matrix, introducing an attention mechanism during decoding, so as to take a positional correlation

CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

between the vertices of an object into consideration, so as to obtain target position coordinates of each vertex after decoding; and finally, according to the obtained target position coordinates of each vertex and previously acquired connection relationship data between the vertices, performing reconstruction to obtain a three-dimensional model corresponding to the target object. By means of the method, the occurrence of unnatural bulges or recesses on the surface of the obtained, by means of reconstruction, three-dimensional model of an object can be avoided, thereby improving the reconstruction effect of the three-dimensional model.

(57) 摘要: 本申请涉及图像处理技术领域, 提出一种物体模型的重建方法、装置、终端设备和存储介质。首先, 获取包含目标物体的原始图像和某个预设的网格模板, 提取该原始图像的特征向量, 然后将该特征向量和该网格模板中具有的各个顶点的位置坐标融合, 得到特征矩阵。接着, 采用解码网络对该特征矩阵进行处理, 在解码时引入注意力机制以考虑物体各个顶点之间的位置相关性, 得到解码后的各个顶点的目标位置坐标。最后, 根据获得的各个顶点的目标位置坐标以及之前获取的各个顶点之间的连接关系数据, 重建得到该目标物体对应的三维模型。该方法能够避免重建得到的物体三维模型的表面出现不自然的凸起或凹陷现象, 提高三维模型的重建效果。

一种物体模型的重建方法、装置、终端设备和存储介质

技术领域

本申请涉及图像处理技术领域，尤其涉及一种物体模型的重建方法、装置、终端设备和存储介质。

5 背景技术

室内场景的三维模型重建技术在虚拟现实和人机交互等领域具有巨大的应用价值。目前，通常采用基于深度学习的单目三维物体模型重建方法，该方法一般采用端到端的编码器-解码器结构的运算模型。然而，在解码器端预测物体表面的某个顶点位置分布时，通常只考量物体图像的全局信息以及该顶点的特征信息，这会使得重建得到的物体三维模型的表面
10 出现不自然的凸起或凹陷现象，导致三维模型的重建效果变差。

技术问题

有鉴于此，本申请实施例提供了一种物体模型的重建方法、装置、终端设备和存储介质，能够避免重建得到的物体三维模型的表面出现不自然的凸起或凹陷现象，提高三维模型的重建效果。

15 技术解决方案

本申请实施例的第一方面提供了一种物体模型的重建方法，包括：

获取预设的网格模板以及包含目标物体的原始图像，所述网格模板包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据；

将所述原始图像输入预先构建的编码网络进行处理，输出所述原始图像对应的初始特征
20 向量，所述编码网络为用于提取图像特征的神经网络；

将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，所述第一特征矩阵包含各个所述顶点分别对应的目标特征向量；

将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，所述第二特征矩阵包含各个所述顶点分别对应的目标位置坐标，所述解码网络为包含全连接层和注意力机制层的神经网络，所述注意力机制层用于针对所述各个顶点中的每个顶点，分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量，得到该顶点
25 对应的融合后的目标特征向量，所述融合后的目标特征向量用于确定该顶点对应的目标位置坐标；

根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据，重
30 建得到所述目标物体对应的目标三维模型。

本申请实施例首先获取包含目标物体的原始图像和某个预设的网格模板，提取该原始图像的特征向量，然后将该特征向量和该网格模板中具有的各个顶点的位置坐标融合，得到特征矩阵。接着，采用解码网络对该特征矩阵进行处理，在解码时引入注意力机制以考虑物体各个顶点之间的位置相关性，得到解码后的各个顶点的目标位置坐标。最后，根据获得的各个顶点的目标位置坐标以及之前获取的各个顶点之间的连接关系数据，重建得到该目标物体对应的三维模型。上述过程根据物体各个顶点之间的位置坐标的相关性执行特征向量的融合，能够考量物体各个顶点之间的相互影响关系，从而避免重建得到的物体三维模型的表面出现不自然的凸起或凹陷现象，提高三维模型的重建效果。

在本申请的一个实施例中，在将所述初始特征向量和所述各个顶点的初始位置坐标融合之前，还可以包括：

获取与所述目标物体对应的类别向量，所述类别向量用于表示所述目标物体所属的物体类别；

将所述类别向量和所述初始特征向量拼接，得到拼接后的特征向量；

所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，具体可以为：

将所述拼接后的特征向量和所述各个顶点的初始位置坐标融合。

在本申请的一个实施例中，所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，可以包括：

将所述各个顶点的初始位置坐标表示为维度 $N \times 3$ 的矩阵， N 为所述各个顶点的数量；

将所述初始特征向量和所述维度 $N \times 3$ 的矩阵在第二维度上拼接，得到维度 $N \times (3+X)$ 的所述第一特征矩阵， X 为所述初始特征向量的元素数量。

在本申请的一个实施例中，所述解码网络包含多个级联的解码模块，每个所述解码模块依次包含全连接层、注意力机制层和批归一化层，所述将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，可以包括：

将所述第一特征矩阵输入所述解码网络的第一个解码模块的全连接层进行处理，输出第一中间矩阵；

将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理，输出第二中间矩阵；

将所述第二中间矩阵和所述第一中间矩阵拼接，得到第三中间矩阵；

将所述第三中间矩阵输入所述第一个解码模块的批归一化层进行处理，得到第四中间矩阵；

将所述第四中间矩阵输入所述解码网络的第二个解码模块，继续采用与所述第一个解码模块相同的处理方式，直至获得由所述解码网络的最后一个解码模块输出的所述第二特征矩阵。

进一步的，所述第一中间矩阵包含各个所述顶点分别对应的目标特征向量，所述将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理，输出第二中间矩阵，可以包括：

针对所述各个顶点中的每个顶点，均根据可训练的权重矩阵分别计算得到各个所述顶点与该顶点之间的相关性权重，然后将各个所述顶点分别对应的目标特征向量按照各自对应的相关性权重执行加权求和处理，得到该顶点对应的融合后的目标特征向量，所述第二中间矩阵是由各个所述顶点分别对应的融合后的目标特征向量组成的矩阵。

在本申请的一个实施例中，在重建得到所述目标物体对应的目标三维模型之后，还可以包括：

根据所述目标三维模型的各个顶点的位置坐标计算得到所述目标三维模型具有的所有二面角的大小；

根据所述所有二面角的大小计算得到平滑损失；

基于所述平滑损失对所述解码网络的参数进行优化更新。

进一步的，所述根据所述所有二面角的大小计算得到平滑损失，具体可以为：

采用以下公式计算得到所述平滑损失：

$$L_{smooth} = \sum_{i,j \in F} (\cos(\theta_{ij}) + 1)^2$$

其中， L_{smooth} 表示所述平滑损失， θ_{ij} 表示所述目标三维模型具有的任意两个平面 i, j 之间的二面角， F 表示所述目标三维模型具有的所有平面。

本申请实施例的第二方面提供了一种物体模型的重建装置，包括：

数据获取模块，用于获取预设的网格模板以及包含目标物体的原始图像，所述网格模板包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据；

特征编码模块，用于将所述原始图像输入预先构建的编码网络进行处理，输出所述原始图像对应的初始特征向量，所述编码网络为用于提取图像特征的神经网络；

向量融合模块，用于将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，所述第一特征矩阵包含各个所述顶点分别对应的目标特征向量；

特征解码模块，用于将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二

特征矩阵，所述第二特征矩阵包含各个所述顶点分别对应的目标位置坐标，所述解码网络为包含全连接层和注意力机制层的神经网络，所述注意力机制层用于针对所述各个顶点中的每个顶点，分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量，得到该顶点对应的融合后的目标特征向量，所述融合后的目标特征向量用于确定该

5 顶点对应的目标位置坐标；

模型重建模块，用于根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据，重建得到所述目标物体对应的目标三维模型。

本申请实施例的第三方面提供了一种终端设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现如本

10 申请实施例的第一方面提供的物体模型的重建方法。

本申请实施例的第四方面提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时实现如本申请实施例的第一方面提供的物体模型的重建方法。

本申请实施例的第五方面提供了一种计算机程序产品，当计算机程序产品在终端设备上

15 运行时，使得终端设备执行本申请实施例的第一方面所述的物体模型的重建方法。

可以理解的是，上述第二方面至第五方面的有益效果可以参见上述第一方面中的相关描述，在此不再赘述。

附图说明

图 1 是本申请实施例提供的一种物体模型的重建方法的流程图；

20 图 2 是本申请实施例提供的一种编码网络的结构示意图；

图 3 是本申请实施例提供的一种残差模块的结构示意图；

图 4 是本申请实施例提供的一种解码网络的结构示意图；

图 5 是本申请实施例提供的注意力机制层的处理示意图；

图 6 是本申请实施例提供的物体模型的重建方法的一种操作示意图；

25 图 7 是本申请实施例提供的物体模型的重建方法的处理效果示意图；

图 8 是本申请与现有技术中的 Total3D 原模型获得的三维模型重建结果的对比图；

图 9 是本申请实施例提供的一种物体模型的重建装置的结构图；

图 10 是本申请实施例提供的一种终端设备的示意图。

本发明的实施方式

30 以下描述中，为了说明而不是为了限定，提出了诸如特定系统结构、技术之类的具体细

节，以便透彻理解本申请实施例。然而，本领域的技术人员应当清楚，在没有这些具体细节的其它实施例中也可以实现本申请。在其它情况中，省略对众所周知的系统、装置、电路以及方法的详细说明，以免不必要的细节妨碍本申请的描述。另外，在本申请说明书和所附权利要求书的描述中，术语“第一”、“第二”、“第三”等仅用于区分描述，而不能理解为指示或暗示相对重要性。

本申请提出一种物体模型的重建方法、装置、终端设备和存储介质，能够避免重建得到的物体三维模型的表面出现不自然的凸起或凹陷现象，提高三维模型的重建效果。应当理解，本申请各个方法实施例的执行主体为各种类型的终端设备或服务器，比如手机、平板电脑、笔记本电脑、台式电脑和可穿戴设备等。

请参阅图 1，示出了本申请实施例提供的一种物体模型的重建方法，包括：

101、获取预设的网格模板以及包含目标物体的原始图像；

首先，获取预设的一个网格模板。该网格模板包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据。例如，该网格模板可以是一个 Mesh 文件，该文件存储了原始三维模型的顶点位置以及顶点之间的连接关系，该原始三维模型可以是球体、立方体和长方体等各种形状的模型，而为了使各个顶点位置的分布比较均匀，一般建议采用球体形状的原始三维模型。假设该原始三维模型具有 N 个顶点，则该网格模板包含该 N 个顶点中每个顶点的三维位置坐标以及该 N 个顶点之间的连接关系数据，根据连接关系数据可以确定这 N 个顶点之间是如何连接的，从而可以获得对应的三维模型。

另外，还需要获取一幅包含目标物体的原始图像，该目标物体是需要重建出对应三维模型的任意类型的物体，例如可以是一个沙发、一个桌子或者一张床等。该原始图像具体可以是该目标物体的 RGB 图像或者灰度图像。

102、将所述原始图像输入预先构建的编码网络进行处理，输出所述原始图像对应的初始特征向量；

在获取到该原始图像之后，将该原始图像输入一个预先构建的编码网络进行处理，得到该原始图像对应的特征向量。其中，该编码网络是一个用于提取图像特征的神经网络，一般通过卷积层、池化层和全连接层等对图像进行处理以提取图像特征，获得对应的特征向量，本申请不对该神经网络的类型和结构进行限定。

本申请实施例提供的一种编码网络的结构示意图如图 2 所示，输入的维度为 224*224*3 的原始图像经过该编码网络的若干卷积层、ReLU 激活函数层、最大池化层和全连接层等网络层处理后，最终得到 1*1*1024 的特征数据，该特征数据可以视作一个 1024 个元素的向量，

也即该 $224*224*3$ 的原始图像对应的初始特征向量。另外，为了避免过深的模型结构导致的图像特征的梯度爆炸或者梯度消失现象，还可以在图 2 所示的编码网络结构中添加多个堆叠的残差模块，其中每个残差模块的结构示意图如图 3 所示，输入的特征图通过两个带边缘填充的 $3*3$ 卷积块处理，提取局部特征后，经由池化层进行特征的整合与筛选，缩小图像特征的维度。每个残差模块的输出都会与其原始输入相加，构成一条新的数据传输路径，该路径赋予残差网络进行恒等映射的能力。在实际应用中，可以使用 PyTorch 框架提供的残差网络模型 ResNet-18 及其预训练权重作为编码网络。

103、将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵；

在获得初始特征向量之后，将该特征向量和该网格模板中的各个顶点的初始位置坐标融合，得到第一特征矩阵，该第一特征矩阵包含各个所述顶点分别对应的目标特征向量。某个顶点的初始位置坐标 (x, y, z) 可以视作 3 个元素的向量，故可以将该 3 个元素的向量和初始特征向量以拼接的方式融合，得到一个新向量，即目标特征向量。各个不同顶点分别对应的各个目标特征向量可以构成一个矩阵，也即该第一特征矩阵。

在本申请的一个实施例中，所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，可以包括：

(1) 将所述各个顶点的初始位置坐标表示为维度 $N*3$ 的矩阵， N 为所述各个顶点的数量；

(2) 将所述初始特征向量和所述维度 $N*3$ 的矩阵在第二维度上拼接，得到维度 $N*(3+X)$ 的所述第一特征矩阵， X 为所述初始特征向量的元素数量。

假设总共有 N 个顶点，每个顶点的初始位置坐标表示为 3 个元素的向量，则 N 个顶点的初始位置坐标可以表示为 $N*3$ 的矩阵，再假设该初始特征向量的元素数量为 X ，则在第二维度拼接后会得到一个 $N*(X+3)$ 的矩阵，作为第一特征矩阵。

在本申请的一个实施例中，在将所述初始特征向量和所述各个顶点的初始位置坐标融合之前，还可以包括：

(1) 获取与所述目标物体对应的类别向量，所述类别向量用于表示所述目标物体所属的物体类别；

(2) 将所述类别向量和所述初始特征向量拼接，得到拼接后的特征向量。

所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，具体可以为：

将所述拼接后的特征向量和所述各个顶点的初始位置坐标融合。

为了提高本申请的泛用性，使其能够兼容多个不同类别物体的三维模型重建处理，在将

该初始特征向量和各个顶点的初始位置坐标融合之前，还可以先将某个类别向量与该初始特征向量拼接，然后再将拼接得到的向量和该初始位置坐标融合。

具体的，每个物体类别都对应唯一的一个类别向量，故类别向量可以采用独热编码的形式。例如，若待处理的数据集中总共有 4 类物体的图像，分别为桌子、椅子、电脑和飞机，
5 则可以预先设置桌子对应的类别向量为 (0, 0, 0, 1)，椅子对应的类别向量为 (0, 0, 1, 0)，电脑对应的类别向量为 (0, 1, 0, 0)，飞机对应的类别向量为 (1, 0, 0, 0)，若当前处理的原始图像中的目标物体为桌子，则此时获取桌子对应的类别向量 (0, 0, 0, 1)，与该初始特征向量进行拼接。

举例说明具体的拼接方式如下：假设总共有 2562 个顶点，每个顶点的初始位置坐标表示为 3 个元素的向量，则 2562 个顶点的初始位置坐标可以表示为 2562×3 的矩阵。该初始特征向量的元素数量为 1024，类别向量的元素数量为 9，则先将初始特征向量和类别向量拼接，得到元素数量为 1033 的新特征向量，然后再将该新特征向量与该 2562×3 的矩阵的第二维度拼接，得到一个 2562×1036 的矩阵，作为第一特征矩阵。该第一特征矩阵中的每个 1×1036 的向量，就是每个模型顶点分别对应的语义向量。

104、将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵；

在获得第一特征矩阵之后，将其输入一个预先构建的解码网络进行处理，得到第二特征矩阵，该第二特征矩阵包含每个顶点分别对应的转换后的目标位置坐标。其中，该解码网络为包含全连接层和注意力机制层的神经网络，注意力机制层用于针对所述各个顶点中的每个顶点，分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量，得到该顶点对应的融合后的目标特征向量，所述融合后的目标特征向量用于确定该顶点对应的目标位置坐标。普通的解码网络通常采用多层堆叠的全连接网络对网格模板的顶点坐标偏移进行预测，得到转换后的目标位置坐标。然而，这种方法只能在预测时对图像的全局信息与单个目标点的信息进行考量，没有考虑目标点的相关点，尤其是局部相邻点之间的相互影响，容易导致重建的三维模型的表面出现不自然的凸起或凹陷。针对个问题，本申请
25 在解码网络中添加了注意力机制层，以捕获同一物体不同顶点之间的位置相互影响关系。

在本申请的一个实施例中，所述解码网络包含多个级联的解码模块，每个所述解码模块依次包含全连接层、注意力机制层和批归一化层，所述将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，可以包括：

(1) 将所述第一特征矩阵输入所述解码网络的第一个解码模块的全连接层进行处理，
30 输出第一中间矩阵；

(2) 将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理, 输出第二中间矩阵;

(3) 将所述第二中间矩阵和所述第一中间矩阵拼接, 得到第三中间矩阵;

(4) 将所述第三中间矩阵输入所述第一个解码模块的批归一化层进行处理, 得到第四中间矩阵;

(5) 将所述第四中间矩阵输入所述解码网络的第二个解码模块, 继续采用与所述第一个解码模块相同的处理方式, 直至获得由所述解码网络的最后一个解码模块输出的所述第二特征矩阵。

如图 4 所示, 为本申请实施例提供的一种解码网络的结构示意图。该解码网络包括多个堆叠的解码模块, 其中每个解码模块依次由全连接层、注意力机制层和批归一化层构成。全连接层可以采用 1×1 的卷积实现, 对单个顶点的坐标偏移进行预测, 之后通过注意力机制层筛选、提取与当前顶点最相关的若干个顶点 (一般是局部的相邻顶点) 的坐标信息, 与原输出进行拼接, 再经由批归一化层 (即 Batch Normalization 层, 也称作批规约层) 处理, 使数据符合高斯分布, 然后投入后续的网络中。

进一步的, 所述第一中间矩阵包含各个所述顶点分别对应的目标特征向量, 所述将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理, 输出第二中间矩阵, 可以包括:

针对所述各个顶点中的每个顶点, 均根据可训练的权重矩阵分别计算得到各个所述顶点与该顶点之间的相关性权重, 然后将各个所述顶点分别对应的目标特征向量按照各自对应的相关性权重执行加权求和处理, 得到该顶点对应的融合后的目标特征向量, 所述第二中间矩阵是由各个所述顶点分别对应的融合后的目标特征向量组成的矩阵。

如图 5 所示, 为本申请采用的注意力机制层的处理示意图。将第一特征矩阵输入第一个解码模块的全连接层处理后, 得到第一中间矩阵 $I \in \mathbb{R}^{N \times C}$, 其中, N 表示顶点数量, C 表示每个顶点对应的目标特征向量的元素个数。将第一中间矩阵 I 输入注意力机制层处理后, 得到第二中间矩阵 $A \in \mathbb{R}^{N \times C}$, 然后将两个矩阵在第二维度拼接, 得到第三中间矩阵 $O \in \mathbb{R}^{N \times 2C}$ 。接着, 将第三中间矩阵 O 输入批归一化层进行处理, 然后接入下一个解码模块执行相同方式的, 以此类推, 最终输出第二特征矩阵, 这个过程可以称作点际注意力机制处理。

在将第一中间矩阵 I 输入注意力机制层后, 具体的处理方式为: 针对某个顶点 P , 均采用一个可训练的权重矩阵分别计算得到该 $N-1$ 个顶点 (除去顶点 P) 与顶点 P 之间的相关性权重, 然后将该 $N-1$ 个顶点分别对应的目标特征向量按照各自对应的相关性权重执行加权求

和处理,得到顶点 P 对应的融合后的目标特征向量,这个过程中特征向量的维度不变(维度还是 C)。采用与顶点 P 相同的处理方式得到该 N 个顶点分别对应的融合后的目标特征向量后,就会得到 N 个融合后的目标特征向量,它们组成第二中间矩阵 $A \in \mathbb{R}^{N \times C}$ 。

在计算相关性权重时,可以采用以下的公式 (1.1) :

$$e_{ij} = \begin{cases} p_i W p_j^T, i \neq j \\ -\infty, i = j \end{cases} \quad (1.1)$$

其中, $e_{i,j}$ 表示该第 N 个顶点中的任意两个顶点 i 和 j 之间的相关性权重, p_i 表示顶点 i 对应的目标特征向量, p_j 表示顶点 j 对应的目标特征向量, W 为一个可训练的权重矩阵,该权重矩阵的初始数值可以人工设置,之后该权重矩阵的数值在该解码网络的训练过程中不断迭代更新。假设 p_i 和 p_j 都为 1×1036 的向量,则权重矩阵 W 为 1036×1036 的矩阵,这样计算得到的相关性权重就会是一个数值,表示顶点 i 和 j 之间的相关性大小。

另外,还可以采用以下的公式 (1.2) 对获得的各个顶点对应的相关性权重进行处理,以保证针对某个顶点的各个相关性权重之和为 1:

$$a_i = \text{softmax}(e_i) \quad (1.2)$$

其中, a_i 表示 softmax 规约后的 e_i , e_i 是 $e_{i,j}$ 按照第 j 维度拼接起来得到的一个向量,表示除顶点 i 外的所有其它顶点和顶点 i 之间的相关性权重。

顶点 i 对应的融合后的目标特征向量可以用以下的公式 (1.3) 表示:

$$A_i = \sum_{j \in N} a_{i,j} p_j \quad (1.3)$$

其中, A_i 表示顶点 i 对应的融合后的目标特征向量, $a_{i,j}$ 表示顶点 j 与顶点 i 之间规约后的相关性权重。

假设第一特征矩阵为 2562×1036 的矩阵,在将该矩阵输入解码网络后,解码网络中各个堆叠的解码模块会逐步对该矩阵执行降维操作(通过全连接层实现),最终得到一个 2562×3 的结果矩阵,表示 2562 个顶点分别对应的转换后的三维位置坐标。

105、根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据,重建得到所述目标物体对应的目标三维模型。

最后,根据各个顶点分别对应的目标位置坐标,可以确定重建的三维模型中各个顶点的位置,然后再结合网格模板中包含的各个顶点之间的连接关系数据,就可以构建得到一个新的三维模型,作为该目标物体对应的目标三维模型。

在本申请的一个实施例中,在重建得到目标物体对应的目标三维模型之后,还可以包

括：

(1) 根据所述目标三维模型的各个顶点的位置坐标计算得到所述目标三维模型具有的所有二面角的大小；

(2) 根据所述所有二面角的大小计算得到平滑损失；

5 (3) 基于所述平滑损失对所述解码网络的参数进行优化更新。

在构建出目标三维模型后，由于其各个顶点坐标以及顶点之间的连接关系都是已知的，故可以方便地计算得到该目标三维模型的每个二面角的大小。然后，可以根据所有二面角的大小计算得到平滑损失，以该平滑损失作为目标函数，对该解码网络的参数进行优化更新。

进一步的，所述根据所述所有二面角的大小计算得到平滑损失，具体可以为：

10 采用以下公式 (1.4) 计算得到所述平滑损失：

$$L_{smooth} = \sum_{i,j \in F} (\cos(\theta_{i,j}) + 1)^2 \quad (1.4)$$

其中， L_{smooth} 表示所述平滑损失， $\theta_{i,j}$ 表示所述目标三维模型具有的任意两个平面之间的二面角， F 表示所述目标三维模型具有的所有平面。在采用网格模板拟合得到目标三维模型的过程中，顶点之间的连接关系是不变的，因而可以根据各个顶点坐标方便地计算出各个二面角，然后采用公式 (1.4) 计算平滑损失。

15 由于室内场景中的人造物体的表面通常是光滑的，而针对单个顶点的坐标的三维模型重建，由于神经网络的泛化性等原因，在重建物体模型的表面时往往会有不小的噪音，导致物体表面凹凸不平。为了解决这个问题，本申请实施例引入平滑损失对神经网络进行训练，对物体表面的平整性进行约束，能够使得重建得到的三维模型的表面更平整光滑，提高模型重建的效果。

本申请实施例首先获取包含目标物体的原始图像和某个预设的网格模板，提取该原始图像的特征向量，然后将该特征向量和该网格模板中具有各个顶点的位置坐标融合，得到特征矩阵。接着，采用解码网络对该特征矩阵进行处理，在解码时引入注意力机制以考虑物体各个顶点之间的位置相关性，得到解码后的各个顶点的目标位置坐标。最后，根据获得的各个顶点的目标位置坐标以及之前获取的各个顶点之间的连接关系数据，重建得到该目标物体对应的三维模型。上述过程根据物体各个顶点之间的位置坐标的相关性执行特征向量的融合，能够考量物体各个顶点之间的相互影响关系，从而避免重建得到的物体三维模型的表面出现不自然的凸起或凹陷现象，提高三维模型的重建效果。

如图 6 所示，为本申请实施例提供的物体模型的重建方法的一种操作示意图。首先，

获取一张目标物体的图片，采用编码网络对该图片进行处理，得到对应的特征向量；然后，将该特征向量与该目标物体对应的类别向量进行拼接，并且与网格模板中的顶点坐标进行拼接；接着，将拼接得到的特征矩阵输入解码网络，该解码网络由堆叠的解码模块组成，每个解码模块依次包含全连接层、注意力机制层和批归一化层，通过点际注意力机制的方式，获得每个顶点经转换后得到的目标位置坐标；最后，根据每个顶点各自对应的目标位置坐标以及各个顶点之间的连接关系数据，重建得到该目标物体对应的三维模型。另外，可以根据重建得到的三维模型中的各个二面角计算得到平滑损失，并根据平滑损失对该解码网络进行优化训练，以提高获得的三维模型的表面的平整性。

如图 7 所示，为本申请提出的物体模型的重建方法的处理效果示意图。其中，图 7 上方的 5 个三维模型为未采用点际注意力机制获得的重建的三维模型，图 7 下方的 5 个三维模型为对应的采用点际注意力机制获得的重建的三维模型。可以看出，图 7 上方的 5 个三维模型中存在很多不自然的凸起和凹陷（见图中的虚线框选处），而在图 7 下方的 5 个三维模型中则不存在这些凸起和凹陷，三维模型的重建效果更好。

为验证本申请的三维模型重建效果，现采用与现有技术中 Total3D 原模型同样的数据集进行了三维模型重建试验。模型的输入为一个 2562 个顶点的球形网格模板以及一幅 224*224 的输入图片，以下的表 1 展示了本申请的运算模型与现有技术中的 Total3D 原模型以及 AtlasNet 模型在 Pix3D 数据集上，共 9 个类别的实际场景室内物体上的三维模型重建精度对比。其中，斜切角距离反应的是重建物体模型顶点与真值之间的位置偏差，法向量距离反应的是重建物体表面与真值表面的法向量偏差。根据表 1 所示的三维模型重建指标对比可以获知，本申请提出的运算模型相较于现有技术中的 Total3D 原模型以及 AtlasNet 模型，能够获得更小的位置偏差以及法向量偏差，即有效提高了三维模型的重建效果。

表 1

	斜切角距离 (*10 ³)			法向量距离 (*10 ²)		
	本申请	Total3D	AtlasNet	本申请	Total3D	AtlasNet
杂项	35.06	31.32	84.84	0.10	0.08	0.07
床	8.94	8.41	93.20	5.04	5.64	8.54
书橱	8.11	8.22	67.55	19.03	22.75	30.26
椅子	8.72	9.60	40.50	2.70	3.43	4.03
课桌	9.22	10.75	96.29	11.90	12.84	18.80
沙发	5.38	5.31	28.98	3.34	3.35	4.85

餐桌	20.5	18.82	141.25	9.24	11.60	16.58
生活用品	3.8	5.51	103.51	0.13	0.07	0.06
衣柜	5.66	5.33	30.87	8.62	8.51	19.48
平均	10.38	10.43	37.39	5.61	6.58	9.26

图 8 是本申请与现有技术中的 Total3D 原模型获得的三维模型重建结果的对比图，其中左侧的一列是输入图片，中间的一列视采用 Total3D 原模型获得的三维模型重建结果，右侧的一列是采用本申请获得的三维模型重建结果。由此可见，采用本申请提出的运算模型能够生成更准确、平整的三维物体模型。

应理解，上述各个实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

上面主要描述了一种物体模型的重建方法，下面将对一种物体模型的重建装置进行描述。

请参阅图 9，本申请实施例中一种物体模型的重建装置的一个实施例包括：

数据获取模块 801，用于获取预设的网格模板以及包含目标物体的原始图像，所述网格模板包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据；

特征编码模块 802，用于将所述原始图像输入预先构建的编码网络进行处理，输出所述原始图像对应的初始特征向量，所述编码网络为用于提取图像特征的神经网络；

向量融合模块 803，用于将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，所述第一特征矩阵包含各个所述顶点分别对应的目标特征向量；

特征解码模块 804，用于将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，所述第二特征矩阵包含各个所述顶点分别对应的目标位置坐标，所述解码网络为包含全连接层和注意力机制层的神经网络，所述注意力机制层用于针对所述各个顶点中的每个顶点，分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量，得到该顶点对应的融合后的目标特征向量，所述融合后的目标特征向量用于确定该顶点对应的目标位置坐标；

模型重建模块 805，用于根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据，重建得到所述目标物体对应的目标三维模型。

在本申请的一个实施例中，所述物体模型的重建装置还可以包括：

类别向量获取模块，用于获取与所述目标物体对应的类别向量，所述类别向量用于表

示所述目标物体所属的物体类别；

向量拼接模块，用于将所述类别向量和所述初始特征向量拼接，得到拼接后的特征向量；

所述向量融合模块具体可以用于：

5 将所述拼接后的特征向量和所述各个顶点的初始位置坐标融合。

在本申请的一个实施例中，所述向量融合模块可以包括：

矩阵表示单元，用于将所述各个顶点的初始位置坐标表示为维度 $N \times 3$ 的矩阵， N 为所述各个顶点的数量；

10 向量拼接单元，用于将所述初始特征向量和所述维度 $N \times 3$ 的矩阵在第二维度上拼接，得到维度 $N \times (3+X)$ 的所述第一特征矩阵， X 为所述初始特征向量的元素数量。

在本申请的一个实施例中，所述解码网络包含多个级联的解码模块，每个所述解码模块依次包含全连接层、注意力机制层和批归一化层，所述特征解码模块可以包括：

第一处理单元，用于将所述第一特征矩阵输入所述解码网络的第一个解码模块的全连接层进行处理，输出第一中间矩阵；

15 第二处理单元，用于将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理，输出第二中间矩阵；

第三处理单元，用于将所述第二中间矩阵和所述第一中间矩阵拼接，得到第三中间矩阵；

20 第四处理单元，用于将所述第三中间矩阵输入所述第一个解码模块的批归一化层进行处理，得到第四中间矩阵；

第五处理单元，用于将所述第四中间矩阵输入所述解码网络的第二个解码模块，继续采用与所述第一个解码模块相同的处理方式，直至获得由所述解码网络的最后一个解码模块输出的所述第二特征矩阵。

25 进一步的，所述第一中间矩阵包含各个所述顶点分别对应的目标特征向量，所述第二处理单元具体可以用于：

针对所述各个顶点中的每个顶点，均根据可训练的权重矩阵分别计算得到各个所述顶点与该顶点之间的相关性权重，然后将各个所述顶点分别对应的目标特征向量按照各自对应的相关性权重执行加权求和处理，得到该顶点对应的融合后的目标特征向量，所述第二中间矩阵是由各个所述顶点分别对应的融合后的目标特征向量组成的矩阵。

30 在本申请的一个实施例中，所述物体模型的重建装置还可以包括：

二面角计算模块，用于根据所述目标三维模型的各个顶点的位置坐标计算得到所述目标三维模型具有的所有二面角的大小；

平滑损失计算模块，用于根据所述所有二面角的大小计算得到平滑损失；

网络参数优化模块，用于基于所述平滑损失对所述解码网络的参数进行优化更新。

5 进一步的，所述平滑损失计算模块具体用于：

采用以下公式计算得到所述平滑损失：

$$L_{smooth} = \sum_{i,j \in F} (\cos(\theta_{i,j}) + 1)^2$$

其中， L_{smooth} 表示所述平滑损失， $\theta_{i,j}$ 表示所述目标三维模型具有的任意两个平面 i, j 之间的二面角， F 表示所述目标三维模型具有的所有平面。

10 本申请实施例还提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时实现如图 1 表示的任意一种物体模型的重建方法。

本申请实施例还提供一种计算机程序产品，当该计算机程序产品在终端设备上运行时，使得终端设备执行实现如图 1 表示的任意一种物体模型的重建方法。

图 10 是本申请一实施例提供的终端设备的示意图。如图 10 所示，该实施例的终端设备 15 9 包括：处理器 90、存储器 91 以及存储在所述存储器 91 中并可在所述处理器 90 上运行的计算机程序 92。所述处理器 90 执行所述计算机程序 92 时实现上述各个物体模型的重建方法的实施例中的步骤，例如图 1 所示的步骤 101 至 105。或者，所述处理器 90 执行所述计算机程序 92 时实现上述各装置实施例中各模块/单元的功能，例如图 9 所示模块 801 至 805 的功能。

20 所述计算机程序 92 可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器 91 中，并由所述处理器 90 执行，以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描述所述计算机程序 92 在所述终端设备 9 中的执行过程。

25 所称处理器 90 可以是中央处理单元(Central Processing Unit, CPU)，还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现成可编程门阵列 (Field-Programmable Gate Array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

所述存储器 91 可以是所述终端设备 9 的内部存储单元，例如终端设备 9 的硬盘或内存。

所述存储器 91 也可以是所述终端设备 9 的外部存储设备，例如所述终端设备 9 上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，所述存储器 91 还可以既包括所述终端设备 9 的内部存储单元也包括外部存储设备。所述存储器 91 用于存储所述计算机程序以及所述终端设备所需的其他程序和数据。所述存储器 91 还可以用于暂时地存储已经输出或者将要输出的数据。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中，上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。另外，各功能单元、模块的具体名称也只是为了便于相互区分，并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述或记载的部分，可以参见其它实施例的相关描述。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

在本申请所提供的实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的系统实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口，装置或单元的间接耦合或通讯连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元

上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

5 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请实现上述实施例方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述各个方法实施例的步骤。。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码
10 形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如
15 在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括是电载波信号和电信信号。

 以上所述实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围，均应包含在本申请
20 的保护范围之内。

权利要求书

1. 一种物体模型的重建方法，其特征在于，包括：

获取预设的网格模板以及包含目标物体的原始图像，所述网格模板包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据；

将所述原始图像输入预先构建的编码网络进行处理，输出所述原始图像对应的初始特征向量，所述编码网络为用于提取图像特征的神经网络；

将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，所述第一特征矩阵包含各个所述顶点分别对应的目标特征向量；

将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，所述第二特征矩阵包含各个所述顶点分别对应的目标位置坐标，所述解码网络为包含全连接层和注意力机制层的神经网络，所述注意力机制层用于针对所述各个顶点中的每个顶点，分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量，得到该顶点对应的融合后的目标特征向量，所述融合后的目标特征向量用于确定该顶点对应的目标位置坐标；

根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据，重建得到所述目标物体对应的目标三维模型。

2. 如权利要求 1 所述的方法，其特征在于，在将所述初始特征向量和所述各个顶点的初始位置坐标融合之前，还包括：

获取与所述目标物体对应的类别向量，所述类别向量用于表示所述目标物体所属的物体类别；

将所述类别向量和所述初始特征向量拼接，得到拼接后的特征向量；

所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，具体为：

将所述拼接后的特征向量和所述各个顶点的初始位置坐标融合。

3. 如权利要求 1 所述的方法，其特征在于，所述将所述初始特征向量和所述各个顶点的初始位置坐标融合，得到第一特征矩阵，包括：

将所述各个顶点的初始位置坐标表示为维度 $N \times 3$ 的矩阵， N 为所述各个顶点的数量；

将所述初始特征向量和所述维度 $N \times 3$ 的矩阵在第二维度上拼接，得到维度 $N \times (3+X)$ 的所述第一特征矩阵， X 为所述初始特征向量的元素数量。

4. 如权利要求 1 所述的方法，其特征在于，所述解码网络包含多个级联的解码模块，

每个所述解码模块依次包含全连接层、注意力机制层和批归一化层，所述将所述第一特征矩阵输入预先构建的解码网络进行处理，输出第二特征矩阵，包括：

将所述第一特征矩阵输入所述解码网络的第一个解码模块的全连接层进行处理，输出第一中间矩阵；

5 将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进行处理，输出第二中间矩阵；

将所述第二中间矩阵和所述第一中间矩阵拼接，得到第三中间矩阵；

将所述第三中间矩阵输入所述第一个解码模块的批归一化层进行处理，得到第四中间矩阵；

10 将所述第四中间矩阵输入所述解码网络的第二个解码模块，继续采用与所述第一个解码模块相同的处理方式，直至获得由所述解码网络的最后一个解码模块输出的所述第二特征矩阵。

5. 如权利要求 4 所述的方法，其特征在于，所述第一中间矩阵包含各个所述顶点分别对应的目标特征向量，所述将所述第一中间矩阵输入所述第一个解码模块的注意力机制层进
15 行处理，输出第二中间矩阵，包括：

针对所述各个顶点中的每个顶点，均根据可训练的权重矩阵分别计算得到各个所述顶点与该顶点之间的相关性权重，然后将各个所述顶点分别对应的目标特征向量按照各自对应的相关性权重执行加权求和处理，得到该顶点对应的融合后的目标特征向量，所述第二中间矩阵是由各个所述顶点分别对应的融合后的目标特征向量组成的矩阵。

20 6. 如权利要求 1 至 5 中任一项所述的方法，其特征在于，在重建得到所述目标物体对应的目标三维模型之后，还包括：

根据所述目标三维模型的各个顶点的位置坐标计算得到所述目标三维模型具有的所有二面角的大小；

根据所述所有二面角的大小计算得到平滑损失；

25 基于所述平滑损失对所述解码网络的参数进行优化更新。

7. 如权利要求 6 所述的方法，其特征在于，所述根据所述所有二面角的大小计算得到平滑损失，具体为：

采用以下公式计算得到所述平滑损失：

$$L_{smooth} = \sum_{i,j \in F} (\cos(\theta_{ij}) + 1)^2$$

其中, L_{smooth} 表示所述平滑损失, θ_{ij} 表示所述目标三维模型具有的任意两个平面 i, j 之间的二面角, F 表示所述目标三维模型具有的所有平面。

8. 一种物体模型的重建装置, 其特征在于, 包括:

数据获取模块, 用于获取预设的网格模板以及包含目标物体的原始图像, 所述网格模板
5 包含原始三维模型的各个顶点的初始位置坐标以及所述各个顶点之间的连接关系数据;

特征编码模块, 用于将所述原始图像输入预先构建的编码网络进行处理, 输出所述原始图像对应的初始特征向量, 所述编码网络为用于提取图像特征的神经网络;

向量融合模块, 用于将所述初始特征向量和所述各个顶点的初始位置坐标融合, 得到第一特征矩阵, 所述第一特征矩阵包含各个所述顶点分别对应的目标特征向量;

10 特征解码模块, 用于将所述第一特征矩阵输入预先构建的解码网络进行处理, 输出第二特征矩阵, 所述第二特征矩阵包含各个所述顶点分别对应的目标位置坐标, 所述解码网络为包含全连接层和注意力机制层的神经网络, 所述注意力机制层用于针对所述各个顶点中的每个顶点, 分别根据各个所述顶点与该顶点之间的相关性融合各个所述顶点分别对应的目标特征向量, 得到该顶点对应的融合后的目标特征向量, 所述融合后的目标特征向量用于确定该
15 顶点对应的目标位置坐标;

模型重建模块, 用于根据各个所述顶点分别对应的目标位置坐标以及所述各个顶点之间的连接关系数据, 重建得到所述目标物体对应的目标三维模型。

9. 一种终端设备, 包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序, 其特征在于, 所述处理器执行所述计算机程序时实现如权利要求 1 至 7
20 中任一项所述的物体模型的重建方法。

10. 一种计算机可读存储介质, 所述计算机可读存储介质存储有计算机程序, 其特征在于, 所述计算机程序被处理器执行时实现如权利要求 1 至 7 中任一项所述的物体模型的重建方法。

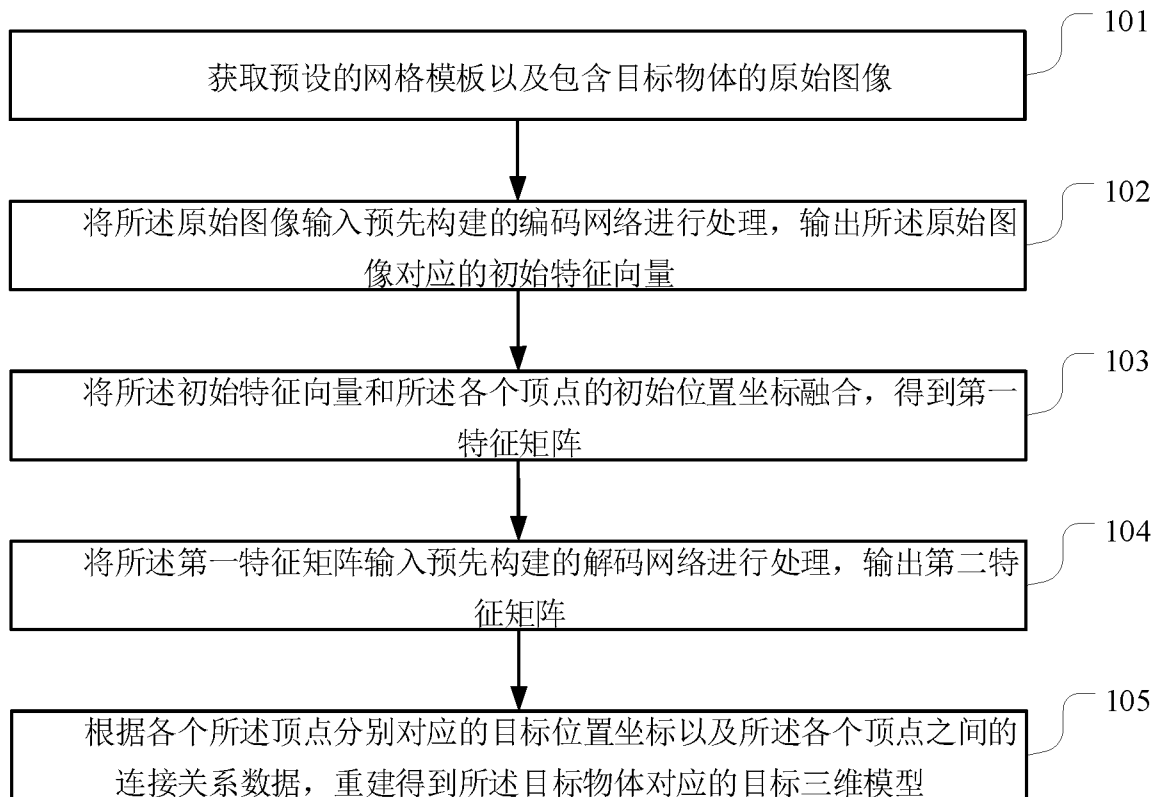


图 1

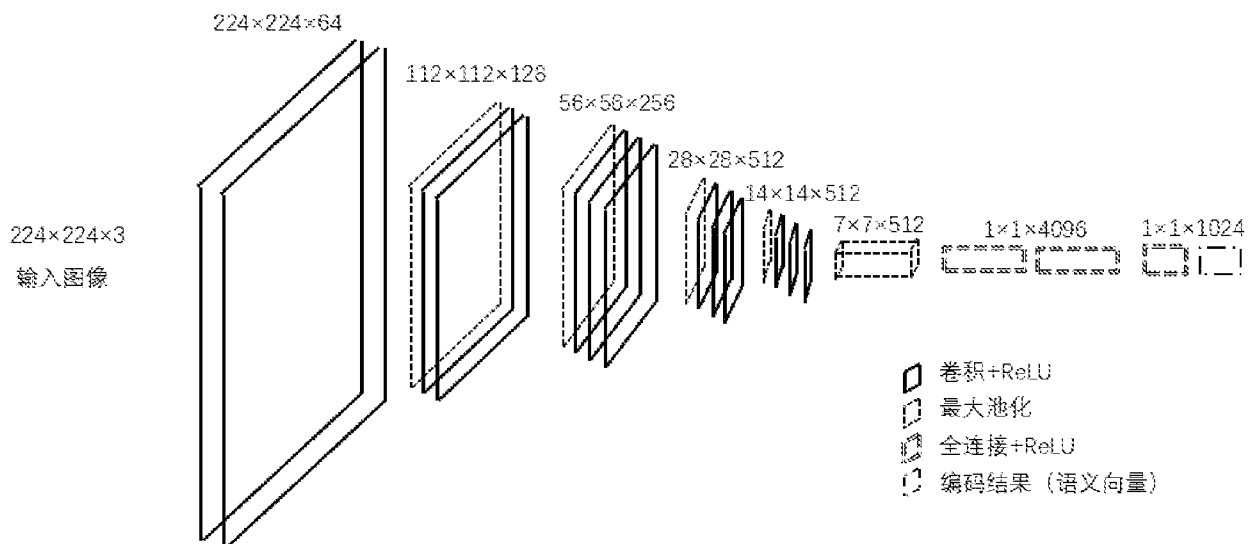


图 2

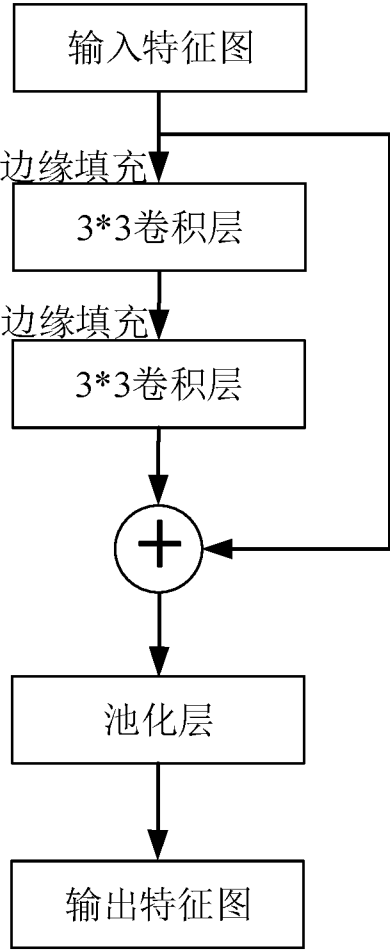


图 3

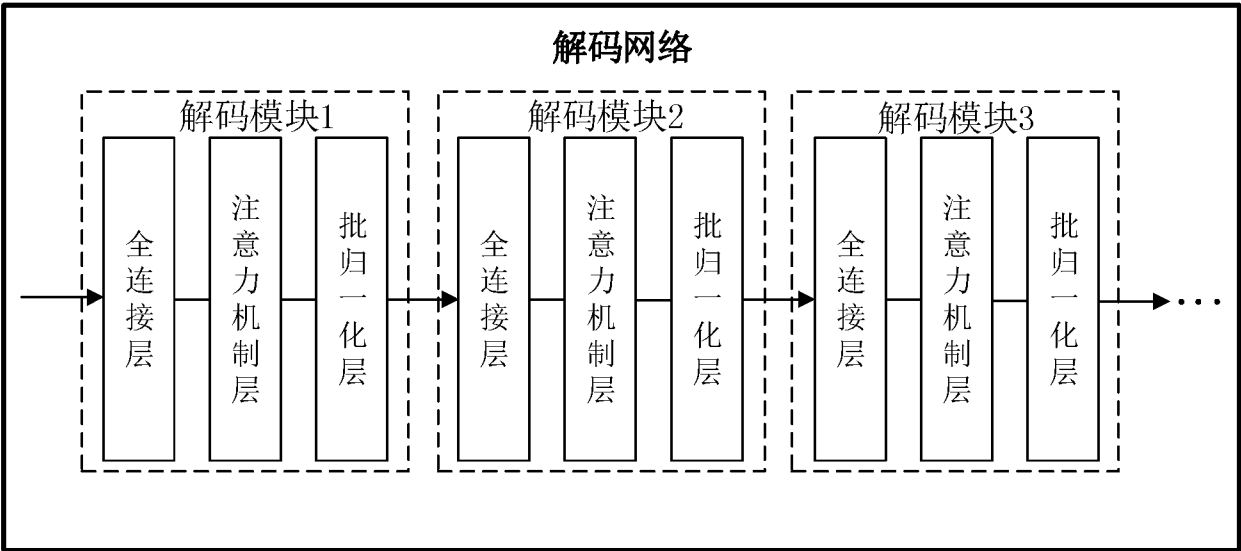


图 4

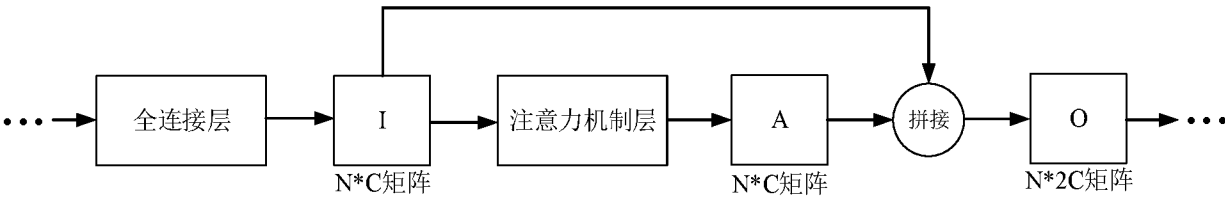


图 5

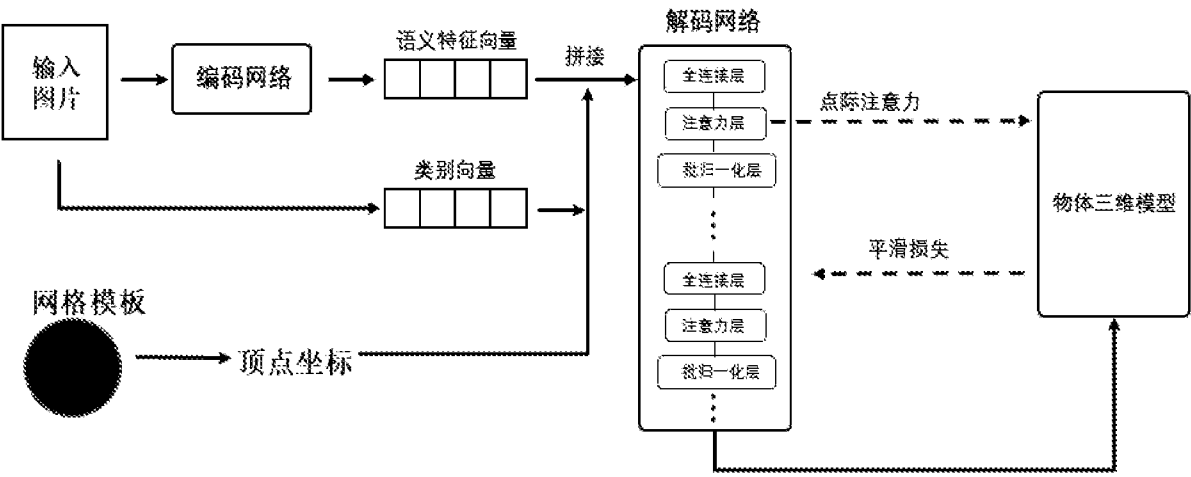


图 6

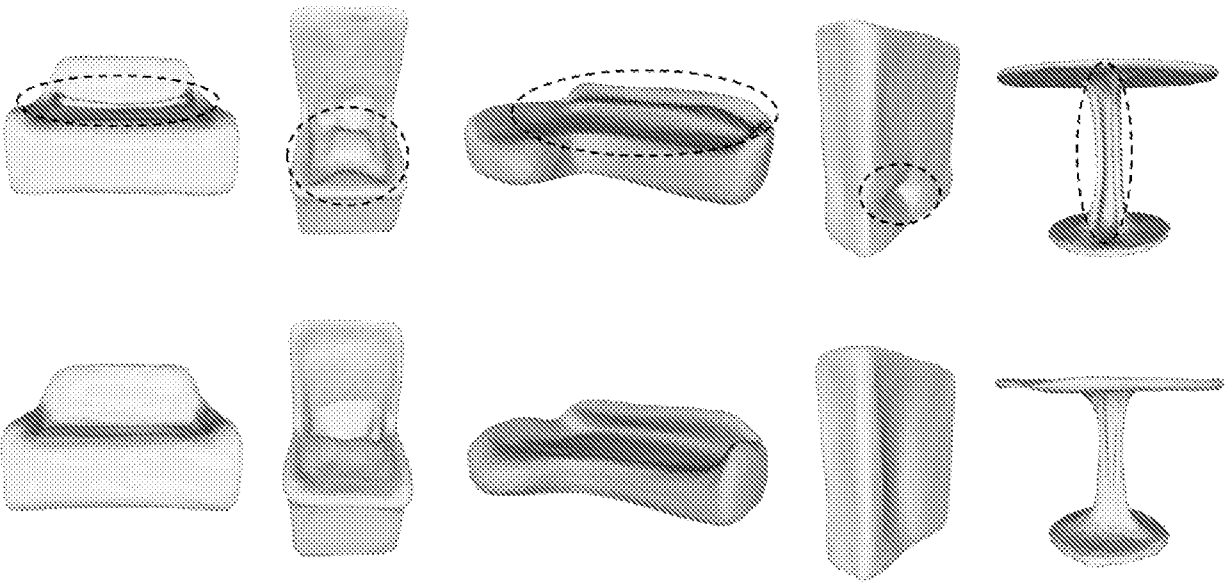


图 7

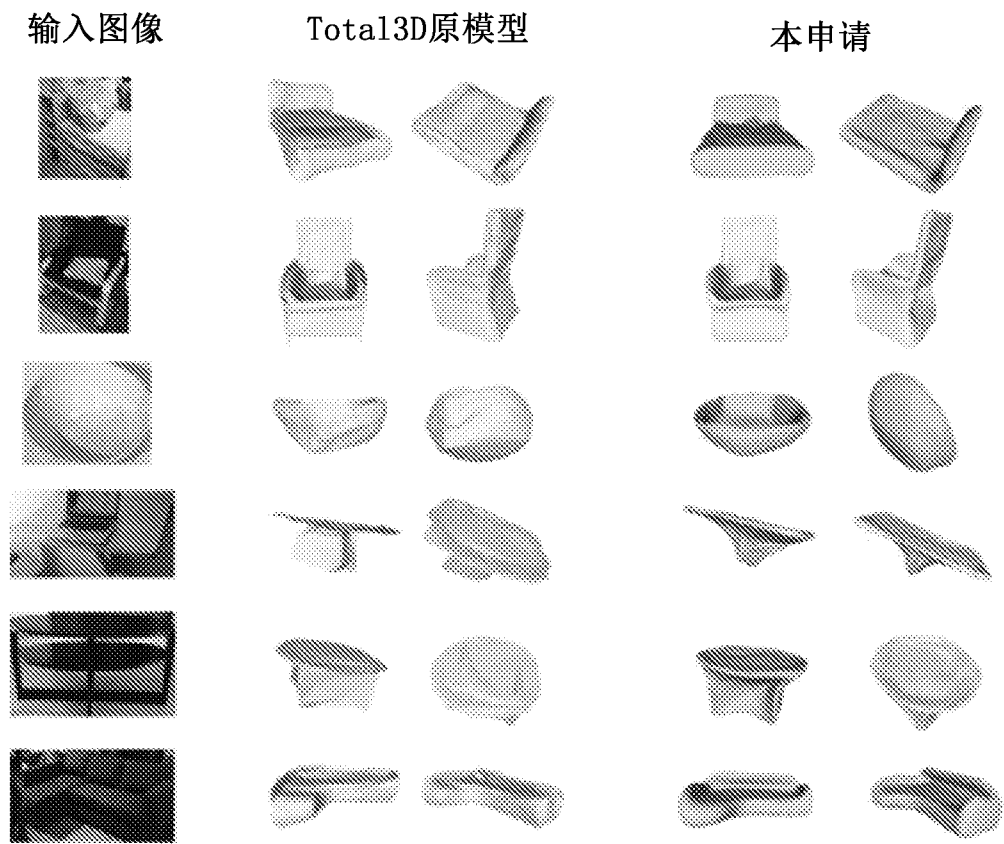


图 8

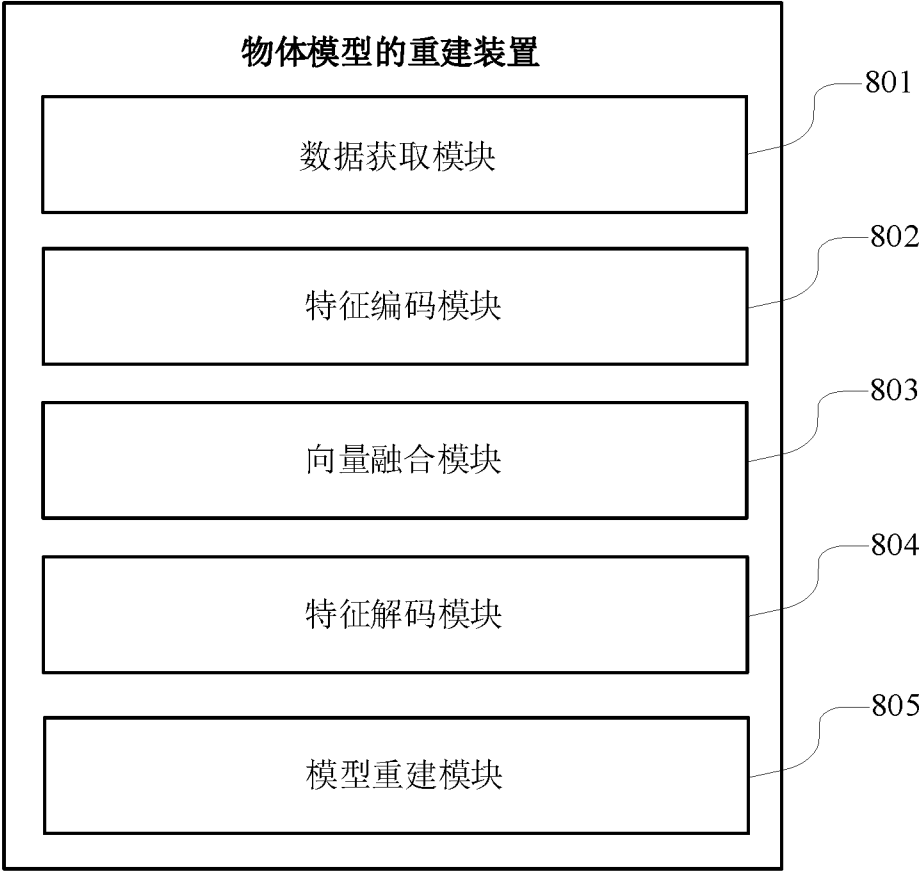


图 9

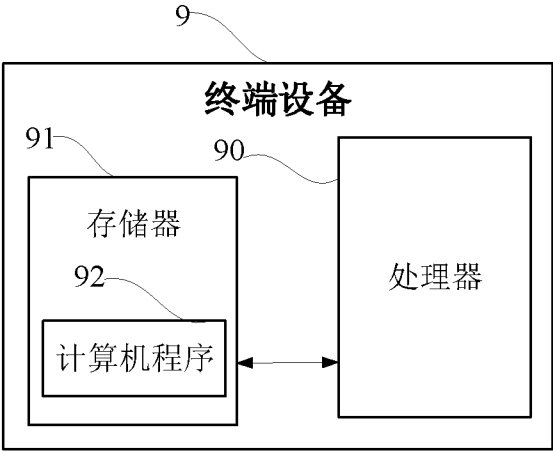


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/093783

A. CLASSIFICATION OF SUBJECT MATTER

G06T 17/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT; CNKI; WPI; EPODOC: 物体, 重建, 模型, 模板, 图像, 三维, 顶点, 位置, 坐标, 编码, 解码, 特征, 提取, 向量, 融合, 矩阵, 注意力, 全连接, 凹凸, object, reconstruction, model, template, image, three-dimensional, vertex, position, coordinate, encoding, decoding, feature, extraction, vector, fusion, matrix, attention, full join, concave, convex

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 110728219 A (TIANJIN UNIVERSITY) 24 January 2020 (2020-01-24) description, paragraphs [0030]-[0046]	1-10
A	CN 110458957 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 15 November 2019 (2019-11-15) entire document	1-10
A	CN 112784782 A (UNIVERSITY OF SHANGHAI FOR SCIENCE AND TECHNOLOGY) 11 May 2021 (2021-05-11) entire document	1-10
A	CN 110544297 A (BEIJING UNIVERSITY OF TECHNOLOGY) 06 December 2019 (2019-12-06) entire document	1-10
A	US 2021012558 A1 (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 14 January 2021 (2021-01-14) entire document	1-10
A	US 2021042557 A1 (HERE GLOBAL B.V.) 11 February 2021 (2021-02-11) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 “&” document member of the same patent family

Date of the actual completion of the international search

18 January 2022

Date of mailing of the international search report

11 February 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/093783

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110728219	A	24 January 2020	None			
CN	110458957	A	15 November 2019	None			
CN	112784782	A	11 May 2021	None			
CN	110544297	A	06 December 2019	None			
US	2021012558	A1	14 January 2021	CN	109285215	A	29 January 2019
				EP	3745358	A1	02 December 2020
				WO	2020042720	A1	05 March 2020
US	2021042557	A1	11 February 2021	None			

国际检索报告

国际申请号

PCT/CN2021/093783

A. 主题的分类

G06T 17/00 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06T

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNPAT; CNKI; WPI; EPDOC: 物体, 重建, 模型, 模板, 图像, 三维, 顶点, 位置, 坐标, 编码, 解码, 特征, 提取, 向量, 融合, 矩阵, 注意力, 全连接, 凹凸, object, reconstruction, model, template, image, three-dimensional, vertex, position, coordinate, encoding, decoding, feature, extraction, vector, fusion, matrix, attention, full join, concave, convex

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 110728219 A (天津大学) 2020年1月24日 (2020 - 01 - 24) 说明书第[0030]-[0046]段	1-10
A	CN 110458957 A (浙江工业大学) 2019年11月15日 (2019 - 11 - 15) 全文	1-10
A	CN 112784782 A (上海理工大学) 2021年5月11日 (2021 - 05 - 11) 全文	1-10
A	CN 110544297 A (北京工业大学) 2019年12月6日 (2019 - 12 - 06) 全文	1-10
A	US 2021012558 A1 (TENCENT TECHNOLOGY SHENZHEN COMPANY LIMITED) 2021年1月14日 (2021 - 01 - 14) 全文	1-10
A	US 2021042557 A1 (HERE GLOBAL B. V.) 2021年2月11日 (2021 - 02 - 11) 全文	1-10

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2022年1月18日

国际检索报告邮寄日期

2022年2月11日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

王璐

电话号码 86-(10)-53961303

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/093783

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	110728219	A	2020年1月24日	无	
CN	110458957	A	2019年11月15日	无	
CN	112784782	A	2021年5月11日	无	
CN	110544297	A	2019年12月6日	无	
US	2021012558	A1	2021年1月14日	CN 109285215 A	2019年1月29日
				EP 3745358 A1	2020年12月2日
				WO 2020042720 A1	2020年3月5日
US	2021042557	A1	2021年2月11日	无	