

(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
2020 年 9 月 24 日 (24.09.2020)



(10) 国际公布号
WO 2020/186943 A1

- (51) 国际专利分类号:
G01C 21/00 (2006.01) **G06T 7/269** (2017.01)
- (21) 国际申请号: PCT/CN2020/075049
- (22) 国际申请日: 2020 年 2 月 13 日 (13.02.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910199169.7 2019年3月15日 (15.03.2019) CN
- (71) 申请人: 京东方科技集团股份有限公司
(**BOE TECHNOLOGY GROUP CO., LTD.**) [CN/CN];
中国北京市朝阳区酒仙桥路 10 号, Beijing
100015 (CN)。 北京大学(**PEKING UNIVERSITY**)
[CN/CN]; 中国北京市海淀区颐和园路 5
号, Beijing 100871 (CN)。
- (72) 发明人: 查红彬(**ZHA, Hongbin**); 中国北京市北京经济技术开发区地泽路 9 号, Beijing 100176 (CN)。 薛飞(**XUE, Fei**); 中国北京市北京经济技

术开发区地泽路 9 号, Beijing 100176 (CN)。 方奕庚(**FANG, Yigeng**); 中国北京市北京经济技术开发区地泽路 9 号, Beijing 100176 (CN)。 姜立(**JIANG, Li**); 中国北京市北京经济技术开发区地泽路 9 号, Beijing 100176 (CN)。

(74) 代理人: 中国国际贸易促进委员会专利商标事务所(**CCPIT PATENT AND TRADEMARK LAW OFFICE**); 中国北京市西城区阜成门外大街 2 号万通新世界广场 8 层, Beijing 100037 (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, 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, 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,

(54) **Title:** MOBILE DEVICE POSTURE DETERMINATION APPARATUS AND METHOD, AND VISUAL ODOMETER

(54) 发明名称: 移动设备的姿态确定装置、方法和视觉里程计

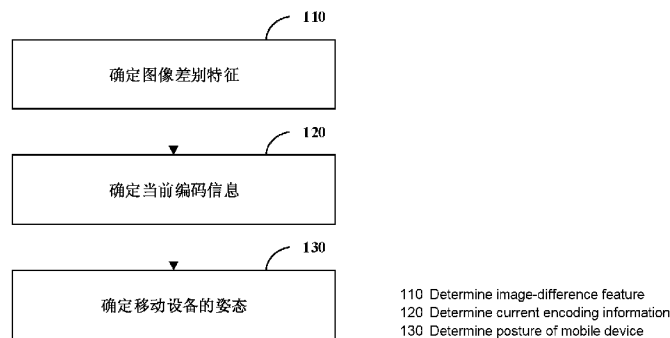


图 1

(57) **Abstract:** A mobile device posture determination apparatus and method and a visual odometer. The apparatus comprises one or more processors (81920) configured to: determine an image-difference feature with respect to a current frame and a preceding frame in a video stream acquired by a mobile device; determine, according to the image-difference feature, current encoding information by means of a first machine learning model; and determine, according to the current encoding information and at least one piece of historical encoding information, a posture of the mobile device by means of a second machine learning model.

(57) **摘要:** 一种移动设备的姿态确定装置、方法和视觉里程计。该装置包括一个或多个处理器 (81,920), 处理器 (81,920) 被配置为: 确定移动设备获取的视频流中当前帧与上一帧之间的图像差别特征; 根据图像差别特征, 利用第一机器学习模型, 确定当前编码信息; 根据当前编码信息和至少一个历史编码信息, 利用第二机器学习模型确定移动设备的姿态。

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))。

移动设备的姿态确定装置、方法和视觉里程计

相关申请的交叉引用

本申请是以 CN 申请号为 201910199169.7，申请日为 2019 年 3 月 15 日的申请为
5 基础，并主张其优先权，该 CN 申请的公开内容在此作为整体引入本申请中。

技术领域

本公开涉及计算机技术领域，特别涉及一种移动设备的姿态确定装置、移动设备的姿态方法、视觉里程计和计算机可读存储介质。

10

背景技术

视觉里程计能够通过分析处理相关图像序列，确定机器人的位置和姿态，进而记录机器人行驶的整个轨迹。

在相关技术中，视觉里程计将视频流中相邻帧的图像信息联合起来，基于图像的
15 几何特征利用局部地图优化确定对应帧的相机姿态；或者基于 IMU（Inertial measurement unit，惯性测量单元）提供的信息，确定相机姿态。

发明内容

根据本公开的一些实施例，提供了一种移动设备的姿态确定装置，包括一个或多个
20 处理器，所述处理器被配置为：确定所述移动设备获取的视频流中当前帧与上一帧之间的图像差别特征；根据所述图像差别特征，利用第一机器学习模型，获取当前编码信息；根据所述当前编码信息和至少一个历史编码信息，利用第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，所述当前帧为第 M 帧，M 为大于 1 的正整数；在从第 N 帧到
25 第 N-1 帧对应的所述移动设备的运动距离或者姿态变化中的至少一个超过阈值的情况下，存储第 N 帧的编码信息作为所述历史编码信息，N 为小于 M 的正整数。

在一些实施例中，根据所述当前编码信息的各通道分量之间的相关性，对所述当前编码信息的各通道分量进行融合，得到融合后的当前编码信息；根据所述历史编码信息的各通道分量之间的相关性，对所述历史编码信息的各通道分量进行融合，得到
30 融合后的历史编码信息；根据所述融合后的当前编码信息和所述融合后的历史编码信

息，利用第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，根据所述当前编码信息各通道分量之间的相关性，确定所述各通道分量的第一权重；根据所述第一权重，对所述各通道分量进行加权，得到所述融合后的当前编码信息。

5 在一些实施例中，根据每个历史编码信息的各通道分量之间的相关性，确定所述各通道分量的第二权重；根据所述第二权重，对所述各通道分量进行加权，得到所述融合后的历史编码信息。

10 在一些实施例中，根据各历史编码信息之间的相关性，对所述各历史编码信息进行融合，得到综合历史编码信息；根据所述综合历史编码信息和所述当前编码信息，利用第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，根据各历史编码信息之间的相关性，确定所述各历史编码信息的第三权重；根据所述第三权重，对所述各历史编码信息进行加权求和，得到所述综合历史编码信息。

15 在一些实施例中，将所述当前编码信息和所述历史编码信息，按照通道维度方向拼接，生成输出编码信息；根据所述输出编码信息，利用所述第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，所述图像差别特征通过光流网络模型获取；所述第一机器学习模型和所述第二机器学习模型中的至少一个为 **ConvLSTM (Convolutional Long Short-Term Memory Network, 卷积长短期记忆网络)** 模型。

20 根据本公开的另一些实施例，提供了一种移动设备的姿态确定方法，包括：确定所述移动设备获取的视频流中当前帧与上一帧之间的图像差别特征；根据所述图像差别特征，利用第一机器学习模型，确定当前编码信息；根据所述当前编码信息和至少一个历史编码信息，利用第二机器学习模型确定所述移动设备的姿态。

25 在一些实施例中，所述当前帧为第 M 帧， M 为大于 1 的正整数；在从第 N 帧到第 $N-1$ 帧对应的所述移动设备的运动距离或者姿态变化中的至少一个超过阈值的情况下，存储第 N 帧的编码信息作为所述历史编码信息， N 为小于 M 的正整数。

30 在一些实施例中，根据所述当前编码信息的各通道分量之间的相关性，对所述当前编码信息的各通道分量进行融合，得到融合后的当前编码信息；根据所述历史编码信息的各通道分量之间的相关性，对所述历史编码信息的各通道分量进行融合，得到融合后的历史编码信息；根据所述融合后的当前编码信息和所述融合后的历史编码信

息，利用第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，根据所述当前编码信息各通道分量之间的相关性，确定所述各通道分量的第一权重；根据所述第一权重，对所述各通道分量进行加权，得到所述融合后的当前编码信息。

5 在一些实施例中，根据每个历史编码信息的各通道分量之间的相关性，确定所述各通道分量的第二权重；根据所述第二权重，对所述各通道分量进行加权，得到所述融合后的所述历史编码信息。

10 在一些实施例中，所述至少一个历史编码信息包括多个历史编码信息，根据各历史编码信息之间的相关性，对所述各历史编码信息进行融合，得到综合历史编码信息；根据所述综合历史编码信息和所述当前编码信息，利用第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，根据各历史编码信息之间的相关性，确定所述各历史编码信息的第三权重；根据所述第三权重，对所述各历史编码信息进行加权求和，得到所述综合历史编码信息。

15 在一些实施例中，将所述当前编码信息和所述历史编码信息，按照通道维度方向拼接，生成输出编码信息；根据所述输出编码信息，利用所述第二机器学习模型确定所述移动设备的姿态。

在一些实施例中，所述图像差别特征通过光流网络模型获取；所述第一机器学习模型和所述第二机器学习模型中的至少一个为 **ConvLSTM** 模型。

20 根据本公开的又一些实施例，提供了一种视觉里程计，包括：如前述任一实施例所述的姿态确定装置，用于根据移动设备拍摄的视频流确定所述移动设备的姿态。

在一些实施例中，所述的视觉里程计还包括图像传感器，用于获取所述视频流。

根据本公开的再一些实施例，提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如前述任一实施例所述的姿态确定方法。

25 通过以下参照附图对本公开的示例性实施例的详细描述，本公开的其它特征及其优点将会变得清楚。

附图说明

30 此处所说明的附图用来提供对本公开的进一步理解，构成本申请的一部分，本公开的示意性实施例及其说明用于解释本公开，并不构成对本公开的不当限定。在附图

中：

图 1 是示出根据本公开一个实施例的移动设备的姿态确定方法的流程图；

图 2a 是示出根据本公开一个实施例的移动设备的姿态确定方法的示意图；

图 2b 是示出根据本公开一个实施例的移动设备的姿态确定方法所用的

5 ConvLSTM 的示意图；

图 3 是示出图 1 中步骤 130 的一个实施例的流程图；

图 4 是示出图 3 中步骤 1320 的一个实施例的示意图；

图 5 是示出图 1 中步骤 130 的另一个实施例的流程图；

图 6 是示出图 5 中步骤 1321 的一个实施例的示意图；

10 图 7 是示出图 1 中步骤 130 的又一个实施例的流程图；

图 8 是示出根据本公开一个实施例的移动设备的姿态确定装置的框图；

图 9 是示出用于根据本公开另一个实施例的移动设备的姿态确定装置的框图；

图 10 是示出根据本公开一个实施例的视觉里程计的框图。

应当明白，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

15 此外，相同或类似的参考标号表示相同或类似的构件。

具体实施方式

下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。

20 以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本公开及其应用或使用的任何限制。基于本公开中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本公开保护的范围。

除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本公开的范围。同时，应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为授权说明书的一部分。在这里示出和讨论的所有示例中，任何具体值应被解释为仅仅是示例性的，而不是作为限制。因此，示例性实施例的其它示例可以具有不同的值。应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一
25
30 项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

图 1 是示出根据本公开一个实施例的移动设备的姿态确定方法的流程图。

如图 1 所示，该方法包括：步骤 110，确定图像差别特征；步骤 120，确定当前编码信息；和步骤 130，确定移动设备的姿态。

在步骤 110 中，确定移动设备获取的视频流中当前帧与上一帧之间的图像差别特征。

例如，移动设备可以是机器人、无人驾驶车、无人机等可以移动的平台，通过基于 CCD 或 CMOS 等图像传感器的相机拍摄图像。

例如，可以通过卷积神经网络（CNN）获取图像差别特征。

例如，可以通过光流网络（FlowNet: Learning Optical Flow with Convolutional Networks）模型获取图像差别特征。

例如，可以通过光流网络（FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks）模型获取图像差别特征。

在一些实施例中，可以将相邻两帧图像重叠起来输入光流网络模型，利用光流网络模型的特征提取部分提取图像差别特征。图像差别特征为高维特征，高维特征的通道数（如 1024 个）可以根据当前帧图像的分辨率确定。例如，光流网络模型可以对重叠后的图像进行多次卷积处理，并根据卷积处理结果提取相邻两帧图像每个像素的偏移量作为图像差别特征。

这样，可以将高维冗余的图像信息转换为高层的、抽象的语义特征，解决了基于几何特征的相关技术易受环境因素（如遮挡、光照变化、动态物体等）影响的问题，从而提高了姿态确定的准确性。

在步骤 120 中，根据图像差别特征，利用第一机器学习模型，确定当前编码信息。例如，第一机器学习模型可以为 RNN（Recurrent Neural Network，循环神经网络）模型，如 ConvLSTM 模型。

在一些实施例中，可以从 RNN 模型的历史输出中筛选出对姿态确定具有重要影响的历史编码信息（即关键帧相应的编码信息）作为有效信息。可以将有效信息与当前编码信息融合，共同确定移动设备的当前姿态。例如，在移动设备从第 N 帧到第 N-1 帧对应的运动距离或者姿态变化中的至少一个超过阈值的情况下，确定第 N 帧为关键帧；存储 RNN 模型提取的第 N 帧的编码信息作为历史编码信息。

在步骤 130 中，根据当前编码信息和至少一个历史编码信息，利用第二机器学习模型，确定移动设备的姿态。例如，第二机器学习模型可以为 RNN 模型，如 ConvLSTM

模型。利用 RNN 模型对编码信息进行解码，可以确定移动设备的姿态。

这种基于当前编码信息和历史编码信息确定的当前姿态，是在视频流的第一帧到当前帧的全局范围内，进行全局优化确定的姿态（即绝对姿态）。相比于相关技术仅在当前帧和前一帧的局部范围内确定的局部优化姿态（即相对姿态），绝对姿态更加准确。

另外，ConvLSTM 模型不必依赖于 IMU 提供的信息，仅依赖视觉信息即可确定姿态，从而降低了姿态确定成本。

图 2a 是示出根据本公开一个实施例的移动设备的姿态确定方法的示意图。

如图 2a 所示，提取的 1 到 T 时刻的当前编码信息为 x_1 到 x_T 。各时刻存储的历史编码信息为 S^2 到 S^T 。将各时刻的当前编码信息和历史编码信息作为第一机器学习模型（如 ConvLSTM）的输入，得到各时刻的输出编码信息 O_1 到 O_T 。将 O_1 到 O_T 输入第二机器学习模型（如 ConvLSTM），得到各时刻的移动设备的姿态 P 到 P_T 。

如图 2b 所示，显示了 ConvLSTM 的一个原理性实现。 X_t, h_t, o_t 分别表示输入特征、状态变量和输出。

在一些实施例中，步骤 130 可以通过图 2a 中的步骤实现。

尽管本公开的实施例列举了 ConvLSTM 作为机器学习模型的一种实现，其它的机器学习模型也可以适用于本公开，例如 FC-LSTM（Fully Connection LSTM，全连接长短期记忆）模型等。

如本领域技术人员所理解的，为了使得机器学习模型（例如神经网络等）具有所需的功能，在使用机器学习模型前，还包括利用多个样本，如样本图像、样本数据等对机器学习模型进行训练的步骤。训练好的机器学习模型可以用于上述方法。例如，可以通过有监督的方式（样本和与样本对应的标注）训练并获得所需机器学习模型。

图 3 是示出图 1 中步骤 130 的一个实施例的流程图。

如图 3 所示，步骤 130 包括：步骤 1310，融合当前编码信息的各通道分量；步骤 1320，融合历史编码信息的各通道分量；和步骤 1330，确定移动设备的姿态。

在步骤 1310 中，根据当前编码信息的各通道分量之间的相关性，对当前编码信息的各通道分量进行融合。

在一些实施例中，根据当前编码信息各通道分量之间的相关性，确定各通道分量的第一权重；根据第一权重，对各通道分量进行加权，得到融合后的当前编码信息。

例如，当前编码信息即为第一机器学习模型在当前时刻的输出 O_t 。 O_t 具有 J 个通

道分量： O_{i1} 、 $O_{i2}\cdots O_{iJ}$ 。计算 O_{i1} 、 $O_{i2}\cdots O_{iJ}$ 之间的相关性，并根据相关性确定 O_{i1} 、 $O_{i2}\cdots O_{iJ}$ 的相应权重。利用相应权重对 O_{i1} 、 $O_{i2}\cdots O_{iJ}$ 进行加权处理得到 O'_i 。

这样，相当于根据当前编码信息的空间信息，对各通道分量进行选择。这样的技术方案增大了对姿态确定重要的通道分量，减小了不重要的通道分量，从而提高了姿态确定准确性。

在步骤 1320 中，根据历史编码信息的各通道分量之间的相关性，对历史编码信息的各通道分量进行融合。

在一些实施例中，根据每个历史编码信息的各通道分量之间的相关性，确定各通道分量的第二权重；根据第二权重，对各通道分量进行加权，得到融合后的历史编码信息。

例如，存储的历史编码信息（有效信息）的集合为 S ， S 中包含 I 个历史编码信息 S_1 、 $S_2\cdots S_I$ ， i 为小于 I 的正整数。任一个 S_i 具有 J 个通道分量： S_{i1} 、 $S_{i2}\cdots S_{iJ}$ 。计算 S_{i1} 、 $S_{i2}\cdots S_{iJ}$ 之间的相关性，并根据相关性确定 S_{i1} 、 $S_{i2}\cdots S_{iJ}$ 的相应权重。利用相应权重对 S_{i1} 、 $S_{i2}\cdots S_{iJ}$ 进行加权处理得到 S'_i 。这些 S'_i 组成了融合后的历史编码信息集合 S' 。

这样，相当于根据历史编码信息的空间信息，对各通道分量进行选择。这样的技术方案增大了对姿态确定重要的通道分量，减小了不重要的通道分量，从而提高了姿态确定准确性。

在步骤 1330 中，根据融合后的当前编码信息和历史编码信息，利用第二机器学习模型确定移动设备的姿态。

在一些实施例中，步骤 1310 和步骤 1320 没有执行顺序，也可以并行处理；还可以仅执行步骤 1310 或步骤 1320。

图 4 是示出图 3 中步骤 1320 的一个实施例的示意图。

如图 4 所示，任一个储的历史编码信息 S_i 具有多个通道分量。根据各通道分量之间的相关系数，利用门函数计算各通道分量的权重。对通道分量进行加权处理得到融合后的 S'_i 。

在一些实施例中，130 可以通过图 3 中的步骤实现。

图 5 是示出图 1 中步骤 130 的另一个实施例的流程图。

如图 5 所示，步骤 130 包括：步骤 1321，融合各历史编码信息；和步骤 1330'，确定移动设备的姿态。

在步骤 1321 中, 根据各历史编码信息之间的相关性, 对各历史编码信息进行融合, 得到综合历史编码信息。

在一些实施例中, 根据各历史编码信息之间的相关性, 确定各历史编码信息的第三权重; 根据第三权重, 对各历史编码信息进行加权求和, 得到综合历史编码信息。

5 例如, 计算历史编码信息 $S_1, S_2 \cdots S_i \cdots S_L$ 之间的相关性, 并根据相关性确定 $S_1, S_2 \cdots S_i \cdots S_L$ 的相应权重。对 $S_1, S_2 \cdots S_i \cdots S_L$ 进行加权求和得到综合历史编码信息 $\hat{S}$ 。

这样, 利用各帧图像在时间上的连续性, 基于时间信息对历史编码信息进行融合。这样的技术方案增强了对姿态确定重要的历史编码信息, 减弱了不重要的历史编码信息, 从而提高了姿态确定准确性。

10 在一些实施例中, 可以根据图 2 中的实施例, 继续对综合历史编码信息 $\hat{S}$ 的各通道分量进行融合; 也可以先根据图 2 中的实施例对各历史编码信息的各通道分量进行融合得到 S' , 然后根据图 3 中的实施例对 S' 中的各历史编码信息进行融合。也就是说, 可以对历史编码信息先进行空间上的融合, 也可以先进行时间上的融合。

在步骤 1330' 中, 根据综合历史编码信息和当前编码信息, 利用第二机器学习模型确定移动设备的姿态。

图 6 是示出图 5 中步骤 1321 的一个实施例的示意图。

如图 6 所示, 存储的历史编码信息的集合 S 包括 $S_1, S_2 \cdots S_i \cdots S_L$ 。根据 $S_1, S_2 \cdots S_i \cdots S_L$ 之间的相关系数, 利用门函数计算 $S_1, S_2 \cdots S_i \cdots S_L$ 的相应权重。对 $S_1, S_2 \cdots S_i \cdots S_L$ 进行加权后得到 $S'_1, S'_2 \cdots S'_i \cdots S'_L$ 。对 $S'_1, S'_2 \cdots S'_i \cdots S'_L$ 求和得到综合历史编码信息 $\hat{S}$ 。

20 在一些实施例中, 步骤 130 可以通过图 7 中的步骤实现。

图 7 是示出图 1 中步骤 130 的又一个实施例的流程图。

如图 7 所示, 步骤 130 包括: 步骤 1322, 拼接当前编码信息和历史编码信息; 和步骤 1330'', 确定移动设备的姿态。

25 在步骤 1322 中, 将当前编码信息和历史编码信息, 按照通道维度方向拼接, 生成输出编码信息。也就是说, 以当前编码信息和历史编码信息为特征矩阵, 以矩阵的每一层 (即每一通道) 为一个部分进行拼接。例如, 可以通过具有两层卷积层 (如卷积核大小为 3×3 , 卷积步长为 1) 的神经网络模型进行拼接。

在一些实施例中, 可以对历史编码信息、当前编码信息进行时间上和空间上的融合后再拼接。

30 在步骤 1330'' 中, 根据输出编码信息, 利用第二机器学习模型确定移动设备的姿

态。

本公开实施例提供的姿态确定方法，在公开无人驾驶数据集 KITTI 上进行了测试，平均旋转误差不超过每 100 米 3 度，平均平移误差不超过 5%。

图 8 是示出根据本公开一个实施例的移动设备的姿态确定装置的框图。

5 如图 8 所示，移动设备的姿态确定装置 8 包括一个或多个处理器 81。

处理器 81 被配置为获取移动设备拍摄的视频流中当前帧与上一帧之间的图像差别特征。例如，图像差别特征通过光流网络模型获取。

10 处理器 81 被配置为：根据图像差别特征，利用第一机器学习模型，获取当前编码信息；根据当前编码信息和至少一个历史编码信息，利用第二机器学习模型确定移动设备的姿态。例如，第一机器学习模型和第二机器学习模型中的至少一个为 ConvLSTM 模型。

在一些实施例中，姿态确定装置还包括存储器 82。存储器 82 被配置为：在移动设备从第 N 帧到第 N-1 帧对应的运动距离或者姿态变化中的至少一个超过阈值的情况下，存储第 N 帧的编码信息作为历史编码信息。

15 在一些实施例中，处理器 81 根据当前编码信息的各通道分量之间的相关性，对当前编码信息的各通道分量进行融合。处理器 81 根据历史编码信息的各通道分量之间的相关性，对历史编码信息的各通道分量进行融合。处理器 81 根据融合后的当前编码信息和历史编码信息，利用第二机器学习模型确定移动设备的姿态。

20 例如，处理器 81 根据当前编码信息各通道分量之间的相关性，确定各通道分量的第一权重。处理器 81 根据第一权重，对各通道分量进行加权，得到融合后的当前编码信息。

例如，处理器 81 根据每个历史编码信息的各通道分量之间的相关性，确定所述各通道分量的第二权重。处理器 81 根据第二权重，对各通道分量进行加权，得到融合后的历史编码信息。

25 在一些实施例中，处理器 81 根据各历史编码信息之间的相关性，对各历史编码信息进行融合，得到综合历史编码信息。处理器 81 根据综合历史编码信息，利用第二机器学习模型确定移动设备的姿态。

30 例如，处理器 81 根据各历史编码信息之间的相关性，确定各历史编码信息的第三权重。处理器 81 根据第三权重，对各历史编码信息进行加权求和，得到综合历史编码信息。

在一些实施例中，处理器 81 将当前编码信息和历史编码信息，按照通道维度方向拼接，生成输出编码信息。处理器 81 根据输出编码信息，利用第二机器学习模型确定移动设备的姿态。

图 9 是示出用于根据本公开另一个实施例的移动设备的姿态确定装置的框图。

如图 9 所示，姿态确定装置可以通用计算设备的形式表现。计算机系统包括存储器 910、处理器 920 和连接不同系统组件的总线 900。

存储器 910 例如可以包括系统存储器、非易失性存储介质等。系统存储器例如存储有操作系统、应用程序、引导装载程序 (Boot Loader) 以及其他程序等。系统存储器可以包括易失性存储介质，例如随机存取存储器 (RAM) 和/或高速缓存存储器。

非易失性存储介质例如存储有执行显示方法的对应实施例的指令。非易失性存储介质包括但不限于磁盘存储器、光学存储器、闪存等。

处理器 920 可以用通用处理器、数字信号处理器 (DSP)、应用专用集成电路 (ASIC)、现场可编程门阵列 (FPGA) 或其它可编程逻辑设备、分立门或晶体管等分立硬件组件方式来实现。相应地，诸如判断模块和确定模块的每个模块，可以通过中央处理器 (CPU) 运行存储器中执行相应步骤的指令来实现，也可以通过执行相应步骤的专用电路来实现。

总线 900 可以使用多种总线结构中的任意总线结构。例如，总线结构包括但不限于工业标准体系结构 (ISA) 总线、微通道体系结构 (MCA) 总线、外围组件互连 (PCI) 总线。

计算机系统还可以包括输入输出接口 930、网络接口 940、存储接口 950 等。这些接口 930、940、950 以及存储器 910 和处理器 920 之间可以通过总线 900 连接。输入输出接口 930 可以为显示器、鼠标、键盘等输入输出设备提供连接接口。网络接口 940 为各种联网设备提供连接接口。存储接口 940 为软盘、U 盘、SD 卡等外部存储设备提供连接接口。

图 10 是示出根据本公开一个实施例的视觉里程计的框图。

如图 10 所示，视觉里程计 10 包括上述任一个实施例中的姿态确定装置 11，用于根据移动设备拍摄的视频流确定所述移动设备的姿态。

在一些实施例中，视觉里程计 10 还包括成像器件，例如图像传感器 12，用于获取视频流。

在一些实施例中，成像器件可以通过无线，例如蓝牙、Wi-Fi 等方式与姿态确定

装置 11 中的处理器通讯连接；也可以通过有线，例如网线、线缆、走线等与姿态确定装置 11 中的处理器通讯连接。

本领域内的技术人员应当明白，本公开的实施例可提供为方法、系统、或计算机程序产品。因此，本公开可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用非瞬时性存储介质（包括但不限于磁盘存储器、**CD-ROM**、光学存储器等）上实施的计算机程序产品的形式。

至此，已经详细描述了根据本公开的移动设备的姿态确定装置、移动设备的姿态方法、视觉里程计和计算机可读存储介质。为了避免遮蔽本公开的构思，没有描述本领域所公知的一些细节。本领域技术人员根据上面的描述，完全可以明白如何实施这里公开的技术方案。

可能以许多方式来实现本公开的方法和系统。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法和系统。用于所述方法的步骤的上述顺序仅是为了进行说明，本公开的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施例中，还可将本公开实施为记录在记录介质中的程序，这些程序包括用于实现根据本公开的方法的机器可读指令。因而，本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

虽然已经通过示例对本公开的一些特定实施例进行了详细说明，但是本领域的技术人员应该理解，以上示例仅是为了进行说明，而不是为了限制本公开的范围。本领域的技术人员应该理解，可在不脱离本公开的范围和精神的情况下，对以上实施例进行修改。本公开的范围由所附权利要求来限定。

权 利 要 求

1. 一种移动设备的姿态确定装置,包括一个或多个处理器,所述处理器被配置为:
确定所述移动设备获取的视频流中当前帧与上一帧之间的图像差别特征;
根据所述图像差别特征,利用第一机器学习模型,确定当前编码信息;
根据所述当前编码信息和至少一个历史编码信息,利用第二机器学习模型确定所述移动设备的姿态。

2. 根据权利要求 1 所述的姿态确定装置,其中,所述当前帧为第 M 帧, M 为大于 1 的正整数;

所述姿态确定装置还包括存储器,所述存储器被配置为:

在从第 N 帧到第 $N-1$ 帧对应的所述移动设备的运动距离或者姿态变化中的至少一个超过阈值的情况下,存储第 N 帧的编码信息作为所述历史编码信息, N 为小于 M 的正整数。

3、根据权利要求 1 所述的姿态确定装置,其中,根据所述当前编码信息和至少一个历史编码信息,利用第二机器学习模型确定所述移动设备的姿态包括:

根据所述当前编码信息的各通道分量之间的相关性,对所述当前编码信息的各通道分量进行融合,得到融合后的当前编码信息;

根据所述历史编码信息的各通道分量之间的相关性,对所述历史编码信息的各通道分量进行融合,得到融合后的历史编码信息;

根据所述融合后的当前编码信息和所述融合后的历史编码信息,利用第二机器学习模型确定所述移动设备的姿态。

4. 根据权利要求 3 所述的姿态确定装置,其中,对所述当前编码信息的各通道分量进行融合包括:

根据所述当前编码信息各通道分量之间的相关性,确定所述各通道分量的第一权重;

根据所述第一权重,对所述各通道分量进行加权,得到所述融合后的当前编码信息。

5. 根据权利要求 3 所述的姿态确定装置, 其中, 对所述历史编码信息的各通道分量进行融合包括:

根据每个历史编码信息的各通道分量之间的相关性, 确定所述各通道分量的第二权重;

根据所述第二权重, 对所述各通道分量进行加权, 得到所述融合后的历史编码信息。

6. 根据权利要求 1 所述的姿态确定装置, 其中, 所述至少一个历史编码信息包括多个历史编码信息, 利用第二机器学习模型确定所述移动设备的姿态包括:

根据各历史编码信息之间的相关性, 对所述各历史编码信息进行融合, 得到综合历史编码信息;

根据所述综合历史编码信息和所述当前编码信息, 利用所述第二机器学习模型确定所述移动设备的姿态。

7. 根据权利要求 6 所述的姿态确定装置, 其中, 对所述各历史编码信息进行融合包括:

根据各历史编码信息之间的相关性, 确定所述各历史编码信息的第三权重;

根据所述第三权重, 对所述各历史编码信息进行加权求和, 得到所述综合历史编码信息。

8. 根据权利要求 1 所述的姿态确定装置, 其中, 所述利用第二机器学习模型确定所述移动设备的姿态包括:

将所述当前编码信息和所述历史编码信息, 按照通道维度方向拼接, 生成输出编码信息;

根据所述输出编码信息, 利用所述第二机器学习模型确定所述移动设备的姿态。

9. 根据权利要求 1-8 任一项所述的姿态确定装置, 其中,

所述图像差别特征通过光流网络模型获取;

所述第一机器学习模型和所述第二机器学习模型中的至少一个为卷积长短期记忆 ConvLSTM 模型。

10. 一种移动设备的姿态确定方法，包括：

确定所述移动设备获取的视频流中当前帧与上一帧之间的图像差别特征；

根据所述图像差别特征，利用第一机器学习模型，确定当前编码信息；

根据所述当前编码信息和至少一个历史编码信息，利用第二机器学习模型，确定所述移动设备的姿态。

11. 根据权利要求 10 所述的姿态确定方法，其中，所述当前帧为第 M 帧， M 为大于 1 的正整数，

所述的姿态确定方法还包括：

在从第 N 帧到第 $N-1$ 帧对应的所述移动设备的运动距离或者姿态变化中的至少一个超过阈值的情况下，存储第 N 帧的编码信息作为所述历史编码信息， N 为小于 M 的正整数。

12. 根据权利要求 10 所述的姿态确定方法，其中，根据所述当前编码信息和至少一个历史编码信息，利用第二机器学习模型确定所述移动设备的姿态包括：

根据所述当前编码信息的各通道分量之间的相关性，对所述当前编码信息的各通道分量进行融合，得到融合后的当前编码信息；

根据所述历史编码信息的各通道分量之间的相关性，对所述历史编码信息的各通道分量进行融合，得到融合后的历史编码信息；

根据所述融合后的当前编码信息和所述融合后的历史编码信息，利用所述第二机器学习模型确定所述移动设备的姿态。

13. 根据权利要求 12 所述的姿态确定方法，其中，对所述当前编码信息的各通道分量进行融合包括：

根据所述当前编码信息各通道分量之间的相关性，确定所述各通道分量的第一权重；

根据所述第一权重，对所述各通道分量进行加权，得到所述融合后的当前编码信息。

14. 根据权利要求 12 所述的姿态确定方法，其中，对所述历史编码信息的各通道分量进行融合包括：

根据每个历史编码信息的各通道分量之间的相关性，确定所述各通道分量的第二权重；

根据所述第二权重，对所述各通道分量进行加权，得到所述融合后的所述历史编码信息。

15. 根据权利要求 10 所述的姿态确定方法，其中，所述至少一个历史编码信息包括多个历史编码信息，

利用第二机器学习模型确定所述移动设备的姿态包括：

根据各历史编码信息之间的相关性，对所述各历史编码信息进行融合，得到综合历史编码信息；

根据所述综合历史编码信息和所述当前编码信息，利用第二机器学习模型确定所述移动设备的姿态。

16. 根据权利要求 15 所述的姿态确定方法，其中，对所述各历史编码信息进行融合包括：

根据各历史编码信息之间的相关性，确定所述各历史编码信息的第三权重；

根据所述第三权重，对所述各历史编码信息进行加权求和，得到所述综合历史编码信息。

17. 根据权利要求 10 所述的姿态确定方法，其中，所述利用第二机器学习模型确定所述移动设备的姿态包括：

将所述当前编码信息和所述历史编码信息，按照通道维度方向拼接，生成输出编码信息；

根据所述输出编码信息，利用所述第二机器学习模型确定所述移动设备的姿态。

18. 根据权利要求 10-17 任一项所述的姿态确定方法，其中，

所述图像差别特征通过光流网络模型获取；

所述第一机器学习模型和所述第二机器学习模型中的至少一个为卷积长短期记忆 ConvLSTM 模型。

19. 一种视觉里程计，包括：

权利要求 1-9 任一项所述的姿态确定装置，用于根据移动设备拍摄的视频流确定所述移动设备的姿态。

20. 根据权利要求 19 所述的视觉里程计，还包括：

图像传感器，用于获取所述视频流。

21. 一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如权利要求 10-18 中任一项所述的姿态确定方法。

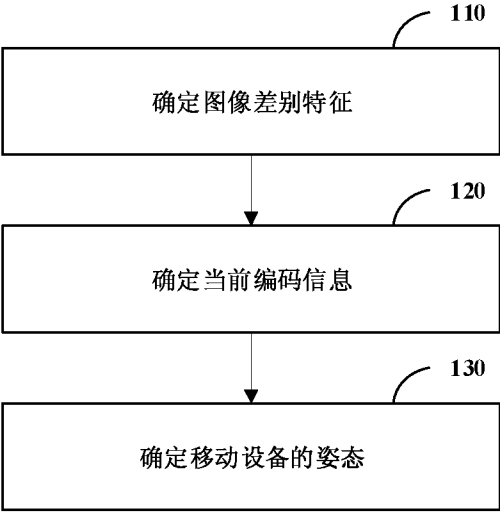


图 1

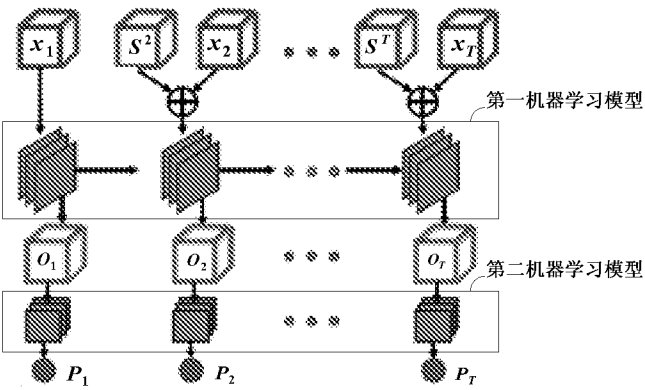


图 2a

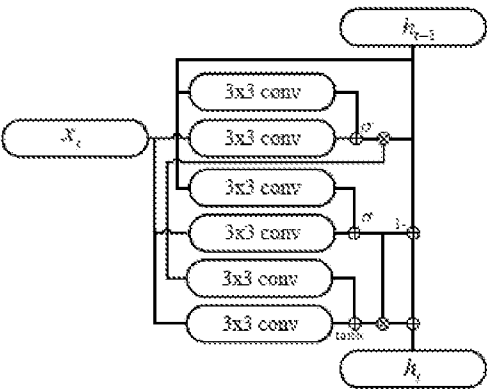


图 2b

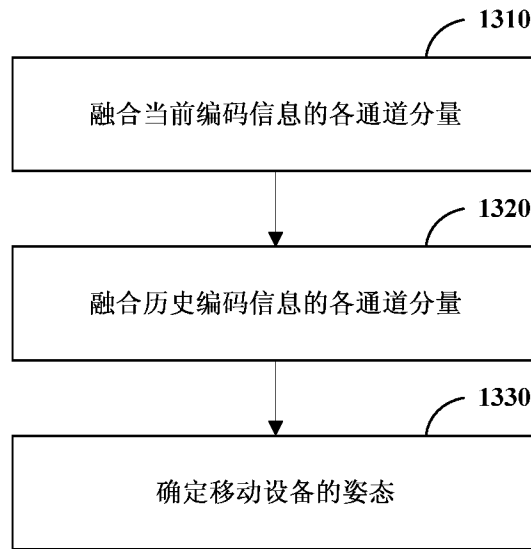


图 3

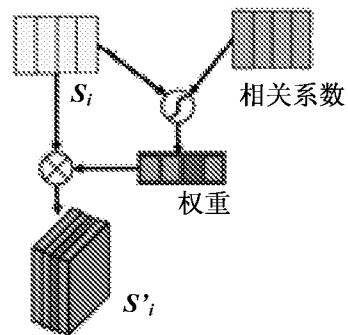


图 4

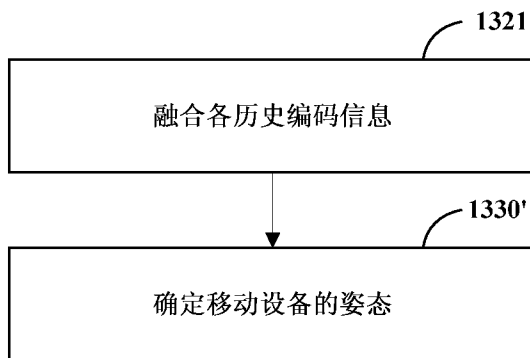


图 5

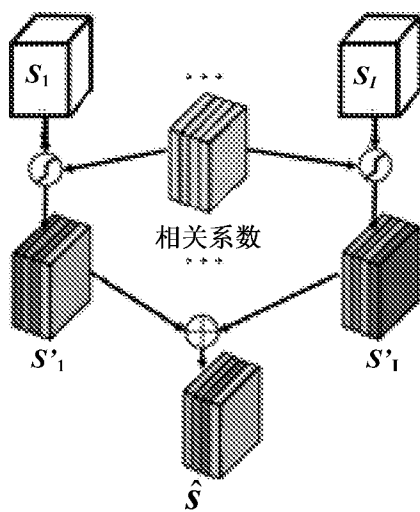


图 6

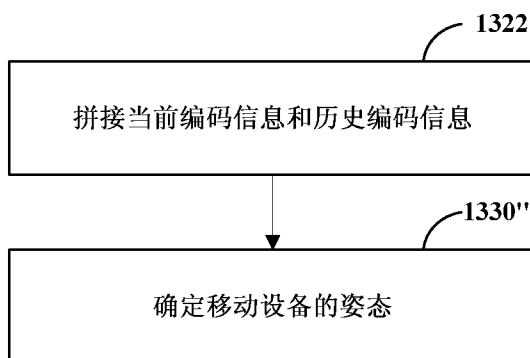


图 7

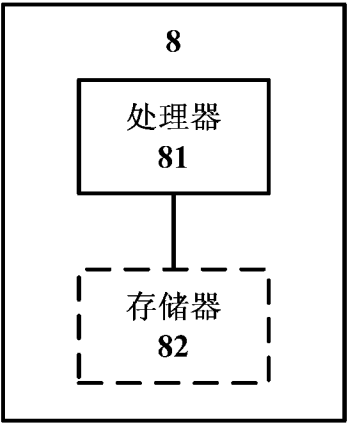


图 8

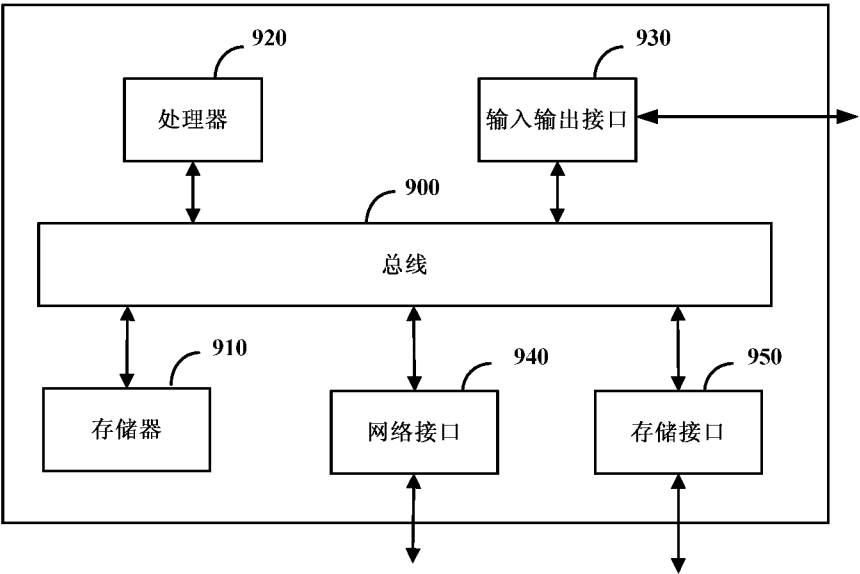


图 9

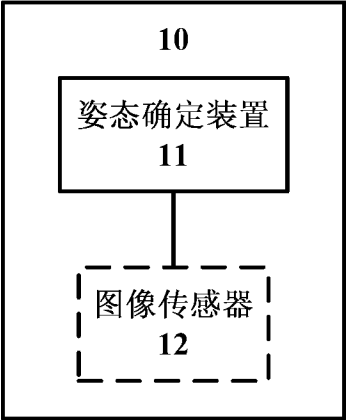


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/075049

A. CLASSIFICATION OF SUBJECT MATTER

G01C 21/00(2006.01)i; G06T 7/269(2017.01)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)

G01C,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, WPI, EPODOC, CNKI: 学习, 相邻帧, 上一帧, 下一帧, 视觉里程表, 视觉历程计, 姿态, 编码, 解码, 学习模型, 神经网络, 权重, 加权, ConvLSTM, VO, visual odometry, CNN, RNN, neural net, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109798888 A (BOE TECHNOLOGY GROUP CO., LTD. et al.) 24 May 2019 (2019-05-24) claims 1-21	1-21
Y	CN 108648216 A (CHANGSHA UNIVERSITY) 12 October 2018 (2018-10-12) description, paragraphs [0048]-[0077]	1-2, 9-11, 18-21
Y	CN 102519481 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 27 June 2012 (2012-06-27) description, paragraphs [0042]-[0053]	1-2, 9-11, 18-21
A	CN 108491763 A (BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) 04 September 2018 (2018-09-04) entire document	1-21
A	CN 108332750 A (SHENZHEN KUNGFU ROBOTICS CO., LTD.) 27 July 2018 (2018-07-27) entire document	1-21
A	CN 106504265 A (RICOH CO., LTD.) 15 March 2017 (2017-03-15) entire document	1-21



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

23 March 2020

Date of mailing of the international search report

22 April 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/075049**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2019066326 A1 (NEC LABORATORIES AMERICA, INC.) 28 February 2019 (2019-02-28) entire document	1-21
A	US 2019079533 A1 (TUSIMPLE) 14 March 2019 (2019-03-14) entire document	1-21

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/075049

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109798888	A	24 May 2019	None			
CN	108648216	A	12 October 2018	None			
CN	102519481	A	27 June 2012	CN	102519481	B	04 September 2013
CN	108491763	A	04 September 2018	None			
CN	108332750	A	27 July 2018	None			
CN	106504265	A	15 March 2017	JP	2017054501	A	16 March 2017
				CN	106504265	B	16 August 2019
				JP	6194995	B2	13 September 2017
US	2019066326	A1	28 February 2019	US	2019063932	A1	28 February 2019
				US	2019064851	A1	28 February 2019
US	2019079533	A1	14 March 2019	None			

国际检索报告

国际申请号

PCT/CN2020/075049

A. 主题的分类

G01C 21/00(2006.01)i; G06T 7/269(2017.01)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)

G01C, 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, WPI, EPDOC, CNKI:学习, 相邻帧, 上一帧, 下一帧, 视觉里程表, 视觉里程计, 姿态, 编码, 解码, 学习模型, 神经网络, 权重, 加权, ConvLSTM, VO, visual odometry, CNN, RNN, neural net, weight

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
PX	CN 109798888 A (京东方科技集团股份有限公司 等) 2019年 5月 24日 (2019 - 05 - 24) 权利要求1-21	1-21
Y	CN 108648216 A (长沙学院) 2018年 10月 12日 (2018 - 10 - 12) 说明书第[0048]-[0077]段	1-2、9-11、18-21
Y	CN 102519481 A (中国科学院自动化研究所) 2012年 6月 27日 (2012 - 06 - 27) 说明书第[0042]-[0053]段	1-2、9-11、18-21
A	CN 108491763 A (北京市商汤科技开发有限公司) 2018年 9月 4日 (2018 - 09 - 04) 全文	1-21
A	CN 108332750 A (深圳市功夫机器人有限公司) 2018年 7月 27日 (2018 - 07 - 27) 全文	1-21
A	CN 106504265 A (株式会社理光) 2017年 3月 15日 (2017 - 03 - 15) 全文	1-21
A	US 2019066326 A1 (NEC LABORATORIES AMERICA, INC.) 2019年 2月 28日 (2019 - 02 - 28) 全文	1-21

☒ 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

2020年 3月 23日

国际检索报告邮寄日期

2020年 4月 22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

授权官员

何孟珂

电话号码 86-(10)-53962472

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	US 2019079533 A1 (TUSIMPLE) 2019年 3月 14日 (2019 - 03 - 14) 全文	1-21

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/075049

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	109798888	A	2019年 5月 24日	无			
CN	108648216	A	2018年 10月 12日	无			
CN	102519481	A	2012年 6月 27日	CN	102519481	B	2013年 9月 4日
CN	108491763	A	2018年 9月 4日	无			
CN	108332750	A	2018年 7月 27日	无			
CN	106504265	A	2017年 3月 15日	JP	2017054501	A	2017年 3月 16日
				CN	106504265	B	2019年 8月 16日
				JP	6194995	B2	2017年 9月 13日
US	2019066326	A1	2019年 2月 28日	US	2019063932	A1	2019年 2月 28日
				US	2019064851	A1	2019年 2月 28日
US	2019079533	A1	2019年 3月 14日	无			