

(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
2024 年 11 月 7 日 (07.11.2024)



(10) 国际公布号
WO 2024/227286 A1

(51) 国际专利分类号:
G06T 7/50 (2017.01) **G06V 10/44** (2022.01)
G06T 9/00 (2006.01)

深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(21) 国际申请号: PCT/CN2023/092055

(74) 代理人: 北京市诚辉律师事务所 (**BEIJING CHENGHUI LAW FIRM**); 中国北京市朝阳区朝阳北路99号2号楼905, Beijing 100123 (CN)。

(22) 国际申请日: 2023 年 5 月 4 日 (04.05.2023)

(25) 申请语言: 中文

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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。

(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 518055 (CN)。

(72) 发明人: 王飞 (**WANG, Fei**); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。程俊 (**CHENG, Jun**); 中国广东省

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚

(54) **Title:** LIGHTWEIGHT, RAPID, AND ACCURATE SELF-SUPERVISED DEPTH ESTIMATION METHOD AND SYSTEM

(54) 发明名称: 一种轻量快速准确的自监督深度估计方法及系统

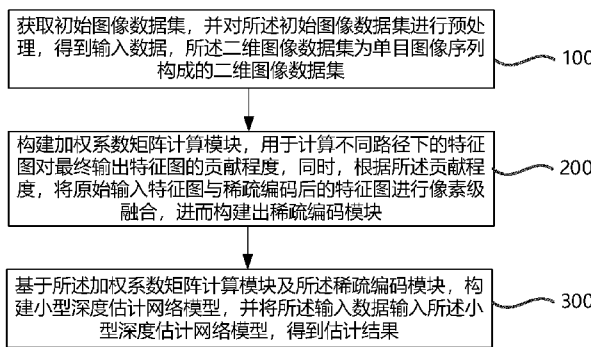


图 1

100 Acquire an initial image data set and pre-processing the initial image data set to obtain input data, the two-dimensional image data set being a two-dimensional image data set composed of monocular image sequences
200 Construct a weighting coefficient matrix calculation module, which is configured for calculating contribution of a feature map to a final output feature map in different paths, and according to the contribution, fuse an original input feature map and a sparsely coded feature map in the pixel level to construct a sparse coding module
300 Construct, on the basis of the weighting coefficient matrix calculation module and the sparse coding module, a small depth estimation network model, and input the input data into the small depth estimation network model to obtain an estimation result

(57) **Abstract:** The present application relates to the technical field of computer vision and image processing, in particular to a light-weight, rapid, and accurate self-supervised depth estimation method and system, which can solve the problem that a complex depth estimation network model is required when scene depth is recovered from a two-dimensional image to a certain extent. The method comprises: acquiring an initial image data set; constructing a weighting coefficient matrix calculation module and a sparse coding module; constructing, on the basis of the weighting coefficient matrix calculation module and the sparse coding module, a small depth estimation network model to increase the algorithm reasoning speed and reduce the number of parameters in the model; constructing and training a large depth estimation network model by taking a photometric loss as a supervision signal at the same time; training the small depth estimation network model by using the photometric loss and a function approximation loss as supervision signals to optimize the depth estimation precision of the small depth estimation network model; and finally, inputting input data into the small depth estimation network model to obtain an estimation result.

(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, ME, 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))。

(57) 摘要: 本申请涉及计算机视觉及图像处理技术领域, 具体而言, 涉及一种轻量快速准确的自监督深度估计方法及系统, 一定程度上可以解决从二维图像中恢复出场景的深度, 需要复杂的深度估计网络模型的问题。方法包括: 获取初始图像数据集; 构建加权系数矩阵计算模块及稀疏编码模块; 基于加权系数矩阵计算模块及稀疏编码模块, 构建小型深度估计网络模型, 以加快算法推理速度, 降低模型的参数量; 同时构建并以光度损失作为监督信号训练大型深度估计网络模型; 再使用光度损失及函数近似损失作为监督信号, 训练小型深度估计网络模型, 以优化小型深度估计网络模型的深度估计精度; 最后, 将输入数据输入小型深度估计网络模型, 得到估计结果。

一种轻量快速准确的自监督深度估计方法及系统

技术领域

本申请涉及计算机视觉及图像处理技术领域，具体而言，涉及一种轻量快速准确的自监督深度估计方法及系统。

背景技术

随着计算机视觉及图像处理技术的发展，能够从二维图像中准确的恢复出场景的深度信息有助于更好地理解场景的三维结构，了解当前相机与场景中物体之间的距离关系，从而协助完成避障、内窥镜手术等各种视觉任务。

但是，普通相机获取的是二维图像，丢失了场景的深度信息，通过神经网络的强大数据拟合能力，以全无监督的形式直接从单目视频中恢复的场景的深度，虽然得益于神经网络的强大数据拟合能力，其精度有了很大的提高。

但是，往往需要复杂的深度估计网络模型去拟合待输入图像中的数据分布特性，而复杂的深度估计网络模型往往会增加算法的时间复杂度和空间复杂度，致使算法的推理速度变慢，模型的参数量增加。

发明内容

为了解决现有技术通过神经网络从二维图像中恢复出场景的深度信息，需要复杂的深度估计网络模型，使得算法的推理速度变慢，模型的参数量增加的问题，本申请提供了一种轻量快速准确的自监督深度估计方法及系统。

本申请的实施例是这样实现的：

第一方面，本申请提供一种轻量快速准确的自监督深度估计方法，其特征在于，包括：

获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，所述初始图像数据集为单目图像序列构成的二维图像数据集；

构建加权系数矩阵计算模块，用于计算不同路径下的特征图对最终输出特征图的贡献程度，同时，根据所述贡献程度，将原始输入特征图与稀疏编码后的特征图进行像素级融合，进而构建出稀疏编码模块；

基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型，并将所述输入数据输入所述小型深度估计网络模型，得到估计结果。

在一种可能的实现方式中，所述获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，还包括以下步骤：

对所述初始图像数据集进行随机翻转、随机裁剪及数据归一化处理，得到初步处理数据；

对所述初步处理数据进行转换，得到所述输入数据，所述输入数据为 $C \times H \times W$ 的张量数据，所述输入数据省略了 batch 维度，其中，C 表示样本的通道维度大小，H 表示输入样本图像的高度，W 表示输入样本图像的宽度。

在一种可能的实现方式中，所述构建加权系数矩阵计算模块，还包括以下步骤：

给定第 i 条路径的特征图为 $X_i \in \mathbb{R}^{C \times H \times W}$, $i \in \{1, 2, \dots, N\}$ ；

基于通道维度，计算得到每条路径下的特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$ ，并将每条路径下的

特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$ 沿所述通道维度进行拼接, 得到混合特征图 $X_{fuse} \in \mathbb{R}^{N \times H \times W}$;

根据公式 $X_{exp} = \delta_{hs}(F_{exp}(X_{fuse}))$, 将所述混合特征图 $X_{fuse} \in \mathbb{R}^{N \times H \times W}$ 映射到高维空间, 得到高维特征图 $X_{exp} \in \mathbb{R}^{r^* N \times H \times W}$;

利用公式 $X_{mid} = \delta_{hs}(F_{mid}(X_{exp}))$, 对所述高维特征图 $X_{exp} \in \mathbb{R}^{r^* N \times H \times W}$ 进行非线性映射出处理, 得到增强高维特征图 $X_{mid} \in \mathbb{R}^{r^* N \times H \times W}$;

根据公式 $X_{squ} = F_{squ}(X_{mid})$, 将所述增强高维特征图 $X_{mid} \in \mathbb{R}^{r^* N \times H \times W}$ 映射到原始特征空间维度 $\mathbb{R}^{C \times H \times W}$, 得到最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$;

基于 Softmax 函数计算所述最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$ 在不同通道, 但位于同一空间位置处的概率值, 并根据所述概率值构成加权系数矩阵;

其中, C 为特征图的通道数, H 为特征图的高度, W 为特征图的宽度, N 表示路径的个数, r 为膨胀系数, $\delta_{hs}(\cdot)$ 表示激活函数, $F_{exp}(\cdot), F_{mid}(\cdot), F_{squ}(\cdot)$ 为特征映射函数。

在一种可能的实现方式中, 所述根据所述贡献程度, 将原始输入特征图与稀疏编码后的特征图进行像素级融合, 进而构建出稀疏编码模块, 还包括以下步骤:

在输入特征图的高维通道空间进行稀疏建模, 得到稀疏特征图;

根据所述加权系数矩阵计算模块计算出的贡献程度, 并将输入特征图与所述稀疏特征图进行像素级融合。

在一种可能的实现方式中, 所述在输入特征图的高维通道空间进行稀疏建模, 得到稀疏特征图, 还包括以下步骤:

给定输入特征图为 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$, 根据公式 $X_{exp}^{sp} = \delta_{hs}(F_{exp}^{sp}(X_{in}^{sp}))$ 将所述输入特征图映射到高维通道特征空间, 得到高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^* C \times H \times W}$;

基于公式 $X_{mid}^{sp} = \delta_{hs}(F_{mid}^{sp}(\delta_{hs}(F_{mid}^{sp}(X_{exp}^{sp}))))$, 对所述高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^* C \times H \times W}$ 的每个通道进行单独编码, 得到高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^* C \times H \times W}$ 。

在一种可能的实现方式中, 所述根据所述加权系数矩阵计算模块计算出的贡献程度, 并将输入特征图与所述稀疏特征图进行像素级融合, 还包括以下步骤:

利用公式 $X_{squ}^{sp} = F_{squ}^{sp}(X_{mid}^{sp})$, 将所述高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^* C \times H \times W}$ 压缩至与所述输入特征图 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$ 相同的特征空间维度, 得到稀疏特征图 $X_{squ}^{sp} \in \mathbb{R}^{C \times H \times W}$;

根据公式 $\Theta_{in}^{coeff}, \Theta_{sp}^{coeff} = F_{coeff}(X_{in}, X_{squ}^{sp})$, $W_1, W_2 \in \mathbb{R}^{1 \times H \times W}$, 计算出不同路径下特征图的加权系数矩阵, 并基于公式 $X_{out}^{sp} = \Theta_{in}^{coeff} * X_{in} + \Theta_{sp}^{coeff} * X_{squ}^{sp}$, 计算得到融合后的稀疏特征图 $X_{out}^{sp} \in \mathbb{R}^{C \times H \times W}$;

其中, $F_{exp}^{sp}(\cdot)$ 和 $F_{squ}^{sp}(\cdot)$ 可以采用经济快速的核大小为 1×1 的卷积核函数实现, $F_{mid}^{sp}(\cdot)$ 基于核大小为 3×3 的卷积核函数实现, 代替对通道之间信息进行稠密建模, 此处只对每一通道特征单独编码。

在一种可能的实现方式中, 所述基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型, 还包括以下步骤:

所述小型深度估计网络模型采用编码器解码器的结构形式, 根据输出特征图的分辨率将

所述编码器与所述解码器划分为 5 个阶段, 给定编码器第 $i \in \{0, 1, 2, 3, 4\}$ 阶段的输入特征图为

$$X_{enc_in}^i \in \mathbb{R}^{C_{enc_in}^i \times \frac{H}{2^i} \times \frac{W}{2^i}};$$

根据公式 $X_{enc_out}^i = F_{spblock}(F_{down}(X_{enc_in}^i))$, 计算得出第 i 阶段编码器的输出特征图

$$X_{enc_out}^i \in \mathbb{R}^{C_{enc_out}^i \times \frac{H}{2^{i+1}} \times \frac{W}{2^{i+1}}}, \text{ 其中 } F_{spblock}(\cdot) \text{ 为稀疏编码模块, } F_{down}(\cdot) \text{ 是步长为 2, 核大小为 } 3 \times 3 \text{ 的}$$

5 卷积核函数;

$$\text{根据公式 } X_{dec_out}^i = \begin{cases} F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) & , i = 0 \\ \Theta_{sp}^i * F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) + \Theta_{enc_out}^{i-1} * X_{enc_out}^{i-1} & , i > 0 \end{cases}, \text{ 计算得到第 } i$$

阶段解码器的输出特征图 $X_{dec_out}^i \in \mathbb{R}^{C_{dec_out}^i \times \frac{H}{2^i} \times \frac{W}{2^i}};$

基于公式 $X_{disp}^j = F_{disp}(X_{dec_out}^j)$, $X_{disp}^j \in \mathbb{R}^{1 \times \frac{H}{2^j} \times \frac{W}{2^j}}$ $j \in [0, 1, 2]$ 回归得到场景的视差。

10 在一种可能的实现方式中, 在所述基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型之后, 进一步包括以下步骤:

构建并以光度损失作为监督信号训练大型深度估计网络模型, 得到高精度的大型深度估计网络模型, 并冻结所述高精度的大型深度估计网络模型的模型权重;

15 使用光度损失及函数近似损失作为监督信号, 训练所述小型深度估计网络模型, 驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型各阶段的特征图的数据分布特性, 以优化所述小型深度估计网络模型的深度估计精度。

第二方面, 本申请提供一种轻量快速准确的自监督深度估计系统, 包括:

输入模块, 用于获取初始图像数据集, 并对所述初始图像数据集进行预处理, 得到输入数据, 所述初始图像数据集为单目图像序列构成的二维图像数据集;

20 计算模块, 用于构建加权系数矩阵计算模块, 用于计算不同路径下的特征图对最终输出特征图的贡献程度, 同时, 根据所述贡献程度, 将原始输入特征图与稀疏编码后的特征图进行像素级融合, 进而构建出稀疏编码模块;

处理模块, 用于基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型, 并将所述输入数据输入所述小型深度估计网络模型, 得到估计结果。

在一种可能的实现方式中, 所述轻量深度估计系统还包括:

25 标定模块, 用于构建并以光度损失作为监督信号训练大型深度估计网络模型, 得到高精度的深度估计模型, 并冻结所述高精度的深度估计模型的模型权重;

优化模块, 用于使用光度损失及函数近似损失作为监督信号, 训练所述小型深度估计网络模型, 驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型各阶段的特征图的数据分布特性, 以优化所述小型深度估计网络模型的深度估计精度。

30 本申请提供的技术方案至少可以达到以下有益效果:

本申请提供了一种轻量快速准确的自监督深度估计方法及系统, 上述轻量快速准确的自监督深度估计方法及系统, 通过加权系数矩阵计算模块来计算出不同路径下的特征图对最终输出特征图的贡献程度, 以此驱使不同路径的特征图实现更高效的像素级融合; 通过稀疏特征编码模块来加快模型的推理速度, 并降低模型的参数量; 通过函数近似损失驱使轻量快速的小型深度估计模型的各阶段特征逼近复杂耗时的大型深度估计网络模型各阶段特征图的分

布特性，以此来提高小模型的深度估计精度。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单的介绍，显而易见地，下面描述中的附图是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请一示例性实施例示出的一种轻量快速准确的自监督深度估计方法的流程示意图；

图 2 是本申请另一示例性实施例示出的获取输入数据的流程示意图；

图 3 是本申请另一示例性实施例示出的构建加权系数矩阵计算模块及稀疏编码模块的流程示意图；

图 4 是本申请另一示例性实施例示出的稀疏建模的流程示意图；

图 5 是本申请另一示例性实施例示出的输入特征图与稀疏特征图进行融合的流程示意图；

图 6 是本申请另一示例性实施例示出的构建小型深度估计网络模型的流程示意图；

图 7 是本申请另一示例性实施例示出的引入函数近似损失作为额外的监督信号来训练轻量深度估计模型的流程示意图；

图 8 是本申请一示例性实施例示出的一种轻量快速准确的自监督深度估计系统的结构示意图；

图 9 是本申请一示例性实施例示出的加权系数矩阵计算模块示意图；

图 10 是本申请一示例性实施例示出的稀疏特征编码模块示意图；

图 11 是本申请一示例性实施例示出的小型深度估计网络模型连接示意图；

图 12 是本申请一示例性实施例示出的小型深度估计网络模型整体训练示意框图。

附图标记：

700、轻量深度估计系统；710、输入模块；720、计算模块；730、处理模块；731、标定模块；732、优化模块。

具体实施方式

为了使本申请的目的、实施方式和优点更加清楚明白，下面将结合本申请示例性实施例中的附图，对本申请示例性实施方式进行清楚、完整的描述，显然，所描述的示例性实施例仅是本申请一部分实施例，而不是全部的实施例，应当理解，此处描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

需要说明的是，本申请中对于术语的简要说明，仅是为了方便理解接下来描述的实施方式，而不是意图限定本申请的实施方式。除非另有说明，这些术语应当按照其普通和通常的含义理解。

本申请中说明书和权利要求书及上述附图中的术语“第一”“第二”“第三”等是用于区别类似或同类的对象或实体，而不必然意味着限定特定的顺序或先后次序，除非另外注明。应该理解这样使用的用语在适当情况下可以互换。

术语“包括”和“具有”以及他们的任何变形，意图在于覆盖但不排他的包含，例如，包含了一系列组件的产品或设备不必限于清楚地列出的所有组件，而是可包括没有清楚地列出

的或对于这些产品或设备固有的其他组件。

在对本申请实施例提供的一种轻量快速准确的自监督深度估计方法及系统进行解释说明之前，先对本申请实施例的应用场景和实施环境进行介绍。

随着计算机视觉及图像处理技术的发展，能够从二维图像中准确的恢复出场景的深度信息有助于更好地理解场景的三维结构，了解当前相机与场景中物体之间的距离关系，从而协助完成避障、内窥镜手术等各种视觉任务。

但是，普通相机获取的是二维图像，丢失了场景的深度信息，通过神经网络的强大数据拟合能力，以全无监督的形式直接从单目视频中恢复的场景的深度，虽然得益于神经网络的强大数据拟合能力，其精度有了很大的提高。

但是，往往需要复杂的深度估计网络模型去拟合待输入图像中的数据分布特性，而复杂的深度估计网络模型往往会增加算法的时间复杂度和空间复杂度，致使算法的推理速度变慢，模型的参数量增加。

基于此，本申请提供了一种轻量快速准确的自监督深度估计方法及系统，通过构建加权系数矩阵计算模块来同时建模同一路径下特征图中像素的上下文依赖，以及不同路径特征图之间的依赖关系的方式来计算不同路径下的特征图对最终输出的特征图的贡献程度，依据其贡献程度实现不同路径获得的特征图的像素级融合；利用稀疏编码模块对通道信息进行稀疏建模，加快模型的推理速度，并降低模型的参数；基于加权系数矩阵计算模块和稀疏编码模块，构造小型深度估计网络模型；利用函数近似损失驱使小型深度估计网络模型的各阶段输出的特征图逼近大型深度估计模型对应阶段输出的特征图分布特性，来提高小型深度估计网络模型估计的场景深度的精度。

接下来，将通过实施例并结合附图具体地对本申请的技术方案，以及本申请的技术方案如何解决上述技术问题进行详细说明。各实施例之间可以相互结合，对于相同或相似的概念或过程可能在某些实施例中不再赘述。显然，所描述的实施例是本申请实施例一部分实施例，而不是全部的实施例。

图 1 是本申请一示例性实施例示出的一种轻量快速准确的自监督深度估计方法的流程示意图。

在一个示例性实施例中，如图 1 所示，提供了一种轻量快速准确的自监督深度估计方法，本实施例中，该方法可以包括以下步骤：

步骤 100：获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，所述初始图像数据集为单目图像序列构成的二维图像数据集；

步骤 200：构建加权系数矩阵计算模块，用于计算不同路径下的特征图对最终输出特征图的贡献程度，同时，根据所述贡献程度，将原始输入特征图与稀疏编码后的特征图进行像素级融合，进而构建出稀疏编码模块；

步骤 300：基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型，并将所述输入数据输入所述小型深度估计网络模型，得到估计结果。

可以看出，本实施例通过加权系数矩阵计算模块来计算出不同路径下的特征图对最终输出特征图的贡献程度，以此驱使不同路径的特征图实现更高效的像素级融合；通过稀疏特征编码模块来加快模型的推理速度，并降低模型的参数量。

图 2 是本申请另一示例性实施例示出的获取输入数据的流程示意图。

在一种可能的实现方式中，如图 2 所示，所述获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，还包括以下步骤：

步骤 110: 对所述初始图像数据集进行随机翻转、随机裁剪及数据归一化处理, 得到初步处理数据;

步骤 120: 对所述初步处理数据进行转换, 得到所述输入数据, 所述输入数据为 $C \times H \times W$ 的张量数据, 所述输入数据省略了 batch 维度, 其中, C 表示样本的通道维度大小, H 表示输入样本图像的高度, W 表示输入样本图像的宽度。

其中, 输入数据可以作为深度估计网络和相机姿态网络的输入, C 表示样本的通道维度大小, 深度估计网络中, C=3, 相机姿态网络中 C=9, H 表示输入样本图像的高度, 在本实施例中, H=128, W 表示输入样本图像的宽度, W=416。

图 3 是本申请另一示例性实施例示出的构建加权系数矩阵计算模块及稀疏编码模块的流程示意图, 图 9 是本申请一示例性实施例示出的加权系数矩阵计算模块示意图。

在一种可能的实现方式中, 如图 3 所示, 所述构建加权系数矩阵计算模块, 还包括以下步骤:

步骤 210: 给定第 i 条路径的特征图为 $X_i \in \mathbb{R}^{C \times H \times W}, i \in \{1, 2, \dots, N\}$;

步骤 220: 基于通道维度, 计算得到每条路径下的特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$, 并将每条路径下的特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$ 沿所述通道维度进行拼接, 得到混合特征图 $X_{fuse} \in \mathbb{R}^{N \times H \times W}$;

步骤 230: 根据公式 $X_{exp} = \delta_{hs}(F_{exp}(X_{fuse}))$, 将所述混合特征图 X_{fuse} 映射到高维空间, 得到高维特征图 $X_{exp} \in \mathbb{R}^{r \times N \times H \times W}$;

步骤 240: 利用公式 $X_{mid} = \delta_{hs}(F_{mid}(X_{exp}))$, 对所述高维特征图进行非线性映射出处理, 得到增强高维特征图 $X_{mid} \in \mathbb{R}^{r \times N \times H \times W}$;

步骤 250: 根据公式 $X_{squ} = F_{squ}(X_{mid})$, 将所述增强高维特征图映射到原始特征空间维度 $\mathbb{R}^{C \times H \times W}$, 得到最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$;

步骤 260: 基于 Softmax 函数计算所述最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$ 在不同通道, 但位于同一空间位置处的概率值, 并根据所述概率值构成加权系数矩阵;

其中, 为了建模不同路径之间特征图的加权系数矩阵, 将混合特征图进行高维映射, 为了增强高维特征的表达能力, 再将高维特征图进行非线性映射处理, C 为特征图的通道数, H 为特征图的高度, W 为特征图的宽度, N 表示路径的个数, r 为膨胀系数, $\delta_{hs}(\cdot)$ 表示激活函数, $F_{exp}(\cdot), F_{mid}(\cdot), F_{squ}(\cdot)$ 为特征映射函数, 在本实施例中设置 r=2, 特征映射函数 $F_{exp}(\cdot), F_{mid}(\cdot), F_{squ}(\cdot)$ 可由核大小为 3×3 步长为 1 的卷积核实现。

可以看出, 本实施例为了能够让不同路径获得的特征图在像素级别更好地融合, 构建了如图 9 所示的加权系数矩阵计算模块, 通过同时建模同一路径下特征图中的像素的上下文依赖, 以及不同路径特征图之间的依赖关系来计算不同路径下的特征图对最终输出的特征图的贡献程度, 最后, 依据计算的像素级权重对不同路径下的特征图进行像素级加权融合。

图 10 是本申请一示例性实施例示出的稀疏特征编码模块示意图。

在一种可能的实现方式中, 如图 3 所示, 所述根据所述贡献程度, 将原始输入特征图与稀疏编码后的特征图进行像素级融合, 进而构建出稀疏编码模块, 还包括以下步骤:

步骤 270: 在输入特征图的高维通道空间进行稀疏建模, 得到稀疏特征图;

步骤 280: 根据所述加权系数矩阵计算模块计算出的贡献程度, 并将输入特征图与所述稀疏特征图进行像素级融合。

可以看出, 为了加快模型的推理速度并兼顾模型的推理精度, 我们构造了如图 10 所示的稀疏特征编码模块 $F_{spblock}(\cdot)$, 该模块通过对通道信息进行稀疏建模, 达到加快模型推理速度的目的。

图 4 是本申请另一示例性实施例示出的稀疏建模的流程示意图。

在一种可能的实现方式中, 如图 4 所示, 所述在输入特征图的高维通道空间进行稀疏建模, 得到稀疏特征图, 还包括以下步骤:

步骤 271: 给定输入特征图 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$, 根据公式 $X_{exp}^{sp} = \delta_{hs}(F_{exp}^{sp}(X_{in}^{sp}))$ 将所述输入特征图映射到高维通道特征空间, 得到高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$;

步骤 272: 基于公式 $X_{mid}^{sp} = \delta_{hs}(F_{mid}^{sp}(\delta_{hs}(F_{mid}^{sp}(X_{exp}^{sp}))))$, 对所述高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 的每个通道进行单独编码, 得到高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 。

可以看出, 为了弥补因对通道信息稀疏建模而致使通道之间无法进行信息交互的缺点, 将低维通道特征映射到高维通道特征空间, 为了能与包含通道之间信息关系的原始输入特征图进行信息交流, 对所述高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 的每个通道进行单独编码。

图 5 是本申请另一示例性实施例示出的输入特征图与稀疏特征图进行融合的流程示意图。

在一种可能的实现方式中, 如图 5 所示, 所述根据所述加权系数矩阵计算模块计算出的贡献程度, 并将输入特征图与所述稀疏特征图进行像素级融合, 还包括以下步骤:

步骤 281: 利用公式 $X_{squ}^{sp} = F_{squ}^{sp}(X_{mid}^{sp})$, 将所述高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 压缩至与所述输入特征图 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$ 相同的特征空间维度, 得到稀疏特征图 $X_{squ}^{sp} \in \mathbb{R}^{C \times H \times W}$;

步骤 282: 根据公式 $\Theta_{in}^{coeff}, \Theta_{sp}^{coeff} = F_{coeff}(X_{in}, X_{squ}^{sp})$, 计算出不同路径特征图的加权系数矩阵 $W_1, W_2 \in \mathbb{R}^{1 \times H \times W}$, 并基于公式 $X_{out}^{sp} = \Theta_{in}^{coeff} * X_{in} + \Theta_{sp}^{coeff} * X_{squ}^{sp}$, 计算得到融合后的稀疏特征图 $X_{out}^{sp} \in \mathbb{R}^{C \times H \times W}$;

其中, $F_{exp}^{sp}(\cdot)$ 和 $F_{squ}^{sp}(\cdot)$ 可以采用经济快速的核大小为 1×1 的卷积核函数实现, $F_{mid}^{sp}(\cdot)$ 基于核大小为 3×3 的卷积核函数实现, 代替对通道之间信息进行稠密建模, 此处只对每一通道特征单独编码。

可以看出, 本实施例将待处理特征图进行分组, 组内特征图使用稠密建模, 切断组与组之间的信息交互, 以此来加快推理时间。

图 6 是本申请另一示例性实施例示出的构建小型深度估计网络模型的流程示意图, 图 11 是本申请一示例性实施例示出的小型深度估计网络模型连接示意图。

在一种可能的实现方式中, 如图 6 所示, 所述基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型, 还包括以下步骤:

步骤 310: 所述小型深度估计网络模型采用编码器解码器的结构形式, 根据输出特征图的分辨率将所述编码器与所述解码器划分为 5 个阶段, 给定编码器第 $i \in \{0, 1, 2, 3, 4\}$ 阶段的输入特征图

为 $X_{enc_in}^i \in \mathbb{R}^{C_{enc_in}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}$;

步骤 320: 根据公式 $X_{enc_out}^i = F_{spblock}(F_{down}(X_{enc_in}^i))$, 计算得出第 i 阶段编码器的输出特征

图 $X_{enc_out}^i \in \mathbb{R}^{C_{enc_out}^i \times \frac{H}{2^{i+1}} \times \frac{W}{2^{i+1}}}$;

步骤 330: 根据公式
$$X_{dec_out}^i = \begin{cases} F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) & , i = 0 \\ \Theta_{sp}^i * F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) + \Theta_{enc_out}^{i-1} * X_{enc_out}^{i-1} & , i > 0 \end{cases}$$
, 计算

得到第 i 阶段解码器的输出特征图 $X_{dec_out}^i \in \mathbb{R}^{C_{dec_out}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}$;

步骤 340: 基于公式 $X_{disp}^j = F_{disp}(X_{dec_out}^j)$, $X_{disp}^j \in \mathbb{R}^{1 \times \frac{H}{2^j} \times \frac{W}{2^j}}$ $j \in [0, 1, 2]$ 回归得到场景的视差。

5 其中, $F_{spblock}(\cdot)$ 为稀疏编码模块, $F_{down}(\cdot)$ 是步长为 2, 核大小为 3×3 的卷积核函数, $F_{disp}(\cdot)$ 核大小为 3×3 , 步长为 1 的卷积核函数以及 Sigmoid 激活函数组成, 连接方式如图 11 所示。

图 7 是本申请另一示例性实施例示出的引入函数近似损失作为额外的监督信号来训练轻量深度估计模型的流程示意图, 图 12 是本申请一示例性实施例示出的小型深度估计网络模型整体训练示意框图。

10 在一种可能的实现方式中, 如图 7 所示, 在所述基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型之后, 进一步包括以下步骤:

步骤 400: 构建并以光度损失作为监督信号训练大型深度估计网络模型, 得到高精度的大型深度估计网络模型, 并冻结所述高精度的大型深度估计网络模型的模型权重;

15 其中, 本申请的一些实施例, 首先使用程俊等人的发明专利 CN202111488926.6 《图像的
深度估计方法、装置、电子设备及存储介质》中发明的深度估计网络作为大型深度估计网络
模型, 以全无监督的方式去训练大型深度估计网络模型, 桥接网络和相机姿态网络, 以此来
获得一个高精度的大型深度估计模型, 再冻结已训练好的大型深度估计网络模型和桥接网络
的模型权重。

20 步骤 500: 使用光度损失及函数近似损失作为监督信号, 训练所述小型深度估计网络模
型, 驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型的各阶段的特征图的数
据分布特性, 以优化所述小型深度估计网络模型的深度估计精度。

具体的, 小型深度估计网络模型的第 i 阶段编码器、解码器的输出特征图分别为

$$X_{enc_out}^i \in \mathbb{R}^{C_{enc_out}^i \times \frac{H}{2^{i+1}} \times \frac{H}{2^{i+1}}}, X_{dec_out}^i \in \mathbb{R}^{C_{dec_out}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}, \text{第 } i \text{ 尺度的视差为 } X_{disp}^i \in \mathbb{R}^{1 \times \frac{H}{2^i} \times \frac{W}{2^i}}.$$

大型深度估计网络模型的第 i 阶段的编码器、解码器的输出特征图分别为

$$25 Z_{enc_out}^i \in \mathbb{R}^{C_{enc_out}^i \times \frac{H}{2^{i+1}} \times \frac{H}{2^{i+1}}}, Z_{dec_out}^i \in \mathbb{R}^{C_{dec_out}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}, \text{第 } i \text{ 尺度的视差为 } Z_{disp}^i \in \mathbb{R}^{1 \times \frac{H}{2^i} \times \frac{W}{2^i}}.$$

为了进一步降低每一阶段的通道数量并对齐对应阶段通道数量, 以便可以使用更轻量更
快的小型深度估计网络模型来模拟大型深度估计网络模型, 本方案利用已经训练好的桥接网
络对大型深度估计网络模型每一阶段的编码器解码器输出的特征图进行信息压缩, 让小型深

度估计网络模型的每一阶段输出去逼近信息更紧凑的特征图 $Y_{enc_hidden}^i \in \mathbb{R}^{C_{enc_small}^i \times \frac{H}{2^{i+1}} \times \frac{H}{2^{i+1}}}$,

$$30 Y_{dec_small}^i \in \mathbb{R}^{C_{dec_hidden}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}.$$

其中, C_{enc/dec_small}^i 是小型深度估计网络模型第 i 阶段编码器/解码器输出特征图的通道数
量。

将 $Y_{enc_out}^i \in \mathbb{R}^{C_{enc_x}^i \times \frac{H}{2^{i+1}} \times \frac{H}{2^{i+1}}}$ / $Y_{dec_out}^i \in \mathbb{R}^{C_{dec_x}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}$ 作为大型深度估计网络模型的编码器第 $i+1$ 阶段的输入/解码器第 $i-1$ 阶段的输入。

其中, $F_{bridge_squ}^i(\cdot)$ 由核大小为 3×3 步长为 1 的卷积核函数、Hardswish 激活函数、程俊等人的发明专利 CN202111488926.6 《图像的深度估计方法、装置、电子设备及存储介质》中的信息精调模块组成, 分别用于特征图压缩以及信息精调; $F_{bridge_exp}^i(\cdot)$ 由核大小为 3×3 步长为 1 的卷积核函数、Hardswish 激活函数组成, 用于特征图通道维度恢复。

其中, $\lambda_{enc}, \lambda_{dec}, \lambda_{disp}$ 是对应损失加权超参数, 本方案设置其为 $\lambda_{enc} = 0.05, \lambda_{dec} = 0.1, \lambda_{disp} = 10$ 。

为了进一步提高小型深度估计网络模型的精度, 本申请将光度损失及函数近似损失作为额外的监督信号, 使用王飞等人的发明专利 CN202110346713.3 《图像场景深度的估计方法、装置、终端设备和存储介质》中发明的图像重建目标函数作为本技术方案的光度损失, 同时根据:

$$L_{apx} = \lambda_{enc} \sum_i^4 \|X_{enc_out}^i - Y_{enc_hidden}^i\| + \lambda_{dec} \sum_i^4 \|X_{dec_out}^i - Y_{dec_hidden}^i\| + \lambda_{disp} \sum_i^2 \|X_{disp}^i - Z_{disp}^i\| \quad (18),$$

作为函数近似损失。

步骤 600: 将所述输入数据输入所述小型深度估计网络模型, 得到估计结果。

其中, 优化后的小型深度估计网络模型, 能够从单目图像中估计出对应的场景深度, 也可以单独使用相机姿态估计出相邻帧图像之间的相对位姿。

可以看出, 为了进一步提高小型深度估计网络模型的估计精度, 如图 12 所示, 本实施例通过光度损失及函数近似损失驱使轻量快速的小型深度估计模型的各阶段特征逼近复杂耗时的大型深度估计网络模型各阶段特征图的分布特性, 以此来提高小模型的深度估计精度。

应该理解的是, 虽然上述实施例所涉及的流程图中的各个步骤按照指示依次显示, 但是这些步骤并不是必然按照指示的顺序依次执行。除非本文中有明确的说明, 这些步骤的执行并没有严格的顺序限制, 这些步骤可以以其他的顺序执行。而且, 上述实施例所涉及的流程图中的至少一部分步骤可以包括多个步骤或者多个阶段, 这些步骤或者阶段并不必然是在同一时刻执行完成, 而是可以在不同的时刻执行, 这些步骤或者阶段的执行顺序也不必然是依次进行, 而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

与前述一种轻量快速准确的自监督深度估计方法的实施例相对应, 采用相同的技术构思, 本申请还提供了一种轻量快速准确的自监督深度估计系统的实施例。

图 8 是本申请一示例性实施例示出的一种轻量快速准确的自监督深度估计系统的结构示意图。

在一个示例性实施例中, 如图 8 所示, 该深度估计系统 700 包括:

输入模块 710, 用于获取初始图像数据集, 并对所述初始图像数据集进行预处理, 得到输入数据, 所述初始图像数据集为单目图像序列构成的二维图像数据集;

计算模块 720, 用于构建加权系数矩阵计算模块, 用于计算不同路径下的特征图对最终输出特征图的贡献程度, 同时, 根据所述贡献程度, 将原始输入特征图与稀疏编码后的特征图进行像素级融合, 进而构建出稀疏编码模块;

处理模块 730, 用于基于所述加权系数矩阵计算模块及所述稀疏编码模块, 构建小型深度估计网络模型, 并将所述输入数据输入所述小型深度估计网络模型, 得到估计结果。

在一种可能的实现方式中，如图 8 所示，所述轻量深度估计系统还包括：

标定模块 731，用于构建并以光度损失作为监督信号训练大型深度估计网络模型，得到高精度的深度估计模型，并冻结所述高精度的深度估计模型的模型权重；

5 优化模块 732，用于使用光度损失及函数近似损失作为监督信号，训练所述小型深度估计网络模型，驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型的各阶段的特征图的数据分布特性，以优化所述小型深度估计网络模型的深度估计精度。

10 关于轻量深度估计系统的具体限定可以参见上文中对于轻量快速准确的自监督深度估计方法的限定，在此不再赘述。上述轻量深度估计系统中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

应该理解的是，本申请实施例中的技术方案可借助软件加必需的通用硬件平台的方式来实现。因此，本申请实施例中的技术方案本质上或者对现有技术做出贡献的部分可以以软件产品的形式体现出来，该软件产品可以存储在计算机可读存储介质中。

15 以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

20 以上所述的实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权 利 要 求 书

1. 一种轻量快速准确的自监督深度估计方法，其特征在于，包括：

获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，所述初始图像数据集为单目图像序列构成的二维图像数据集；

5 构建加权系数矩阵计算模块，用于计算不同路径下的特征图对最终输出特征图的贡献程度，同时，根据所述贡献程度，将原始输入特征图与稀疏编码后的特征图进行像素级融合，进而构建出稀疏编码模块；

基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型，并将所述输入数据输入所述小型深度估计网络模型，得到估计结果。

10 2. 如权利要求 1 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，还包括以下步骤：

对所述初始图像数据集进行随机翻转、随机裁剪及数据归一化处理，得到初步处理数据；

对所述初步处理数据进行转换，得到所述输入数据，所述输入数据为 $C \times H \times W$ 的张量数据，所述输入数据省略了 batch 维度，其中，C 表示样本的通道维度大小，H 表示输入样本图像的高度，W 表示输入样本图像的宽度。

15 3. 如权利要求 1 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述构建加权系数矩阵计算模块，还包括以下步骤：

给定第 i 条路径的特征图为 $X_i \in \mathbb{R}^{C \times H \times W}$, $i \in \{1, 2, \dots, N\}$ ；

20 基于通道维度，计算得到每条路径下的特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$ ，并将每条路径下的特征图的平均值 $\bar{X}_i \in \mathbb{R}^{1 \times H \times W}$ 沿所述通道维度进行拼接，得到混合特征图 $X_{fuse} \in \mathbb{R}^{N \times H \times W}$ ；

根据公式 $X_{exp} = \delta_{hs}(F_{exp}(X_{fuse}))$ ，将所述混合特征图 $X_{fuse} \in \mathbb{R}^{N \times H \times W}$ 映射到高维空间，得到高维特征图 $X_{exp} \in \mathbb{R}^{r \times N \times H \times W}$ ；

利用公式 $X_{mid} = \delta_{hs}(F_{mid}(X_{exp}))$ ，对所述高维特征图 $X_{exp} \in \mathbb{R}^{r \times N \times H \times W}$ 进行非线性映射出处理，得到增强高维特征图 $X_{mid} \in \mathbb{R}^{r \times N \times H \times W}$ ；

25 根据公式 $X_{squ} = F_{squ}(X_{mid})$ ，将所述增强高维特征图 $X_{mid} \in \mathbb{R}^{r \times N \times H \times W}$ 映射到原始特征空间维度 $\mathbb{R}^{C \times H \times W}$ ，得到最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$ ；

基于 Softmax 函数计算所述最终输出特征图 $X_{squ} \in \mathbb{R}^{N \times H \times W}$ 在不同通道，但位于同一空间位置处的概率值，并根据所述概率值构成加权系数矩阵；

其中，C 为特征图的通道数，H 为特征图的高度，W 为特征图的宽度，N 表示路径的个数，

30 r 为膨胀系数， $\delta_{hs}(\cdot)$ 表示激活函数， $F_{exp}(\cdot)$, $F_{mid}(\cdot)$, $F_{squ}(\cdot)$ 为特征映射函数。

4. 如权利要求 1 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述根据所述贡献程度，将原始输入特征图与稀疏编码后的特征图进行像素级融合，进而构建出稀疏编码模块，还包括以下步骤：

在输入特征图的高维通道空间进行稀疏建模，得到稀疏特征图；

35 根据所述加权系数矩阵计算模块计算出的贡献程度，并将输入特征图与所述稀疏特征图进行像素级融合。

5. 如权利要求 4 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述在输入

特征图的高维通道空间进行稀疏建模，得到稀疏特征图，还包括以下步骤：

给定输入特征图为 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$ ，根据公式 $X_{exp}^{sp} = \delta_{hs}(F_{exp}^{sp}(X_{in}^{sp}))$ 将所述输入特征图映射到高维通道特征空间，得到高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ ；

5 基于公式 $X_{mid}^{sp} = \delta_{hs}(F_{mid}^{sp}(\delta_{hs}(F_{mid}^{sp}(X_{exp}^{sp}))))$ ，对所述高维输入特征图 $X_{exp}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 的每个通道进行单独编码，得到高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 。

6. 如权利要求 4 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述根据所述加权系数矩阵计算模块计算出的贡献程度，并将输入特征图与所述稀疏特征图进行像素级融合，还包括以下步骤：

10 利用公式 $X_{squ}^{sp} = F_{squ}^{sp}(X_{mid}^{sp})$ ，将所述高维稀疏特征图 $X_{mid}^{sp} \in \mathbb{R}^{r^*C \times H \times W}$ 压缩至与所述输入特征图 $X_{in}^{sp} \in \mathbb{R}^{C \times H \times W}$ 相同的特征空间维度，得到稀疏特征图 $X_{squ}^{sp} \in \mathbb{R}^{C \times H \times W}$ ；

根据公式 $\Theta_{in}^{coeff}, \Theta_{sp}^{coeff} = F_{coeff}(X_{in}, X_{squ}^{sp})$ ， $W_1, W_2 \in \mathbb{R}^{1 \times H \times W}$ ，计算出不同路径特征图的加权系数矩阵，并基于公式 $X_{out}^{sp} = \Theta_{in}^{coeff} * X_{in} + \Theta_{sp}^{coeff} * X_{squ}^{sp}$ ，计算得到融合后的稀疏特征图 $X_{out}^{sp} \in \mathbb{R}^{C \times H \times W}$ ；

15 其中， $F_{exp}^{sp}(\cdot)$ 和 $F_{squ}^{sp}(\cdot)$ 可以采用经济快速的核大小为 1×1 的卷积核函数实现， $F_{mid}^{sp}(\cdot)$ 基于核大小为 3×3 的卷积核函数实现，代替对通道之间信息进行稠密建模，此处只对每一通道特征单独编码。

7. 如权利要求 1 所述的轻量快速准确的自监督深度估计方法，其特征在于，所述基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型，还包括以下步骤：

20 所述小型深度估计网络模型采用编码器解码器的结构形式，根据输出特征图的分辨率将所述编码器与所述解码器划分为 5 个阶段，给定编码器第 $i \in \{0, 1, 2, 3, 4\}$ 阶段的输入特征图为

$$X_{enc_in}^i \in \mathbb{R}^{C_{enc_in}^i \times \frac{H}{2^i} \times \frac{W}{2^i}};$$

根据公式 $X_{enc_out}^i = F_{spblock}(F_{down}(X_{enc_in}^i))$ ，计算得出第 i 阶段编码器的输出特征图 $X_{enc_out}^i \in \mathbb{R}^{C_{enc_out}^i \times \frac{H}{2^{i+1}} \times \frac{W}{2^{i+1}}}$ ，其中 $F_{spblock}(\cdot)$ 为稀疏编码模块， $F_{down}(\cdot)$ 是步长为 2，核大小为 3×3 的

25 卷积核函数；

根据公式
$$X_{dec_out}^i = \begin{cases} F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) & , i = 0 \\ \Theta_{sp}^i * F_{spblock}(F_{up}(F_{ct}(X_{dec_in}^i))) + \Theta_{enc_out}^{i-1} * X_{enc_out}^{i-1} & , i > 0 \end{cases}$$
，计算得到第 i

阶段解码器的输出特征图 $X_{dec_out}^i \in \mathbb{R}^{C_{dec_out}^i \times \frac{H}{2^i} \times \frac{W}{2^i}}$ ；

基于公式 $X_{disp}^j = F_{disp}(X_{dec_out}^j)$ ， $X_{disp}^j \in \mathbb{R}^{1 \times \frac{H}{2^j} \times \frac{W}{2^j}}$ $j \in [0, 1, 2]$ 回归得到场景的视差。

30 8. 如权利要求 1 所述的轻量快速准确的自监督深度估计方法，其特征在于，在所述基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型之后，进一步包括以下步骤：

构建并以光度损失作为监督信号训练大型深度估计网络模型，得到高精度的大型深度估

计网络模型，并冻结所述高精度的大型深度估计网络模型的模型权重；

使用光度损失及函数近似损失作为监督信号，训练所述小型深度估计网络模型，驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型各阶段的特征图的数据分布特性，以优化所述小型深度估计网络模型的深度估计精度。

5 9. 一种轻量快速准确的自监督深度估计系统，其特征在于，包括：

输入模块，用于获取初始图像数据集，并对所述初始图像数据集进行预处理，得到输入数据，所述初始图像数据集为单目图像序列构成的二维图像数据集；

10 计算模块，用于构建加权系数矩阵计算模块，用于计算不同路径下的特征图对最终输出特征图的贡献程度，同时，根据所述贡献程度，将原始输入特征图与稀疏编码后的特征图进行像素级融合，进而构建出稀疏编码模块；

处理模块，用于基于所述加权系数矩阵计算模块及所述稀疏编码模块，构建小型深度估计网络模型，并将所述输入数据输入所述小型深度估计网络模型，得到估计结果。

10. 如权利要求9所述的轻量快速准确的自监督深度估计系统，其特征在于，所述轻量深度估计系统还包括：

15 标定模块，用于构建并以光度损失作为监督信号训练大型深度估计网络模型，得到高精度的深度估计模型，并冻结所述高精度的深度估计模型的模型权重；

优化模块，用于使用光度损失及函数近似损失作为监督信号，训练所述小型深度估计网络模型，驱使所述小型深度估计网络模型模拟所述高精度的深度估计模型各阶段的特征图的数据分布特性，以优化所述小型深度估计网络模型的深度估计精度。

20

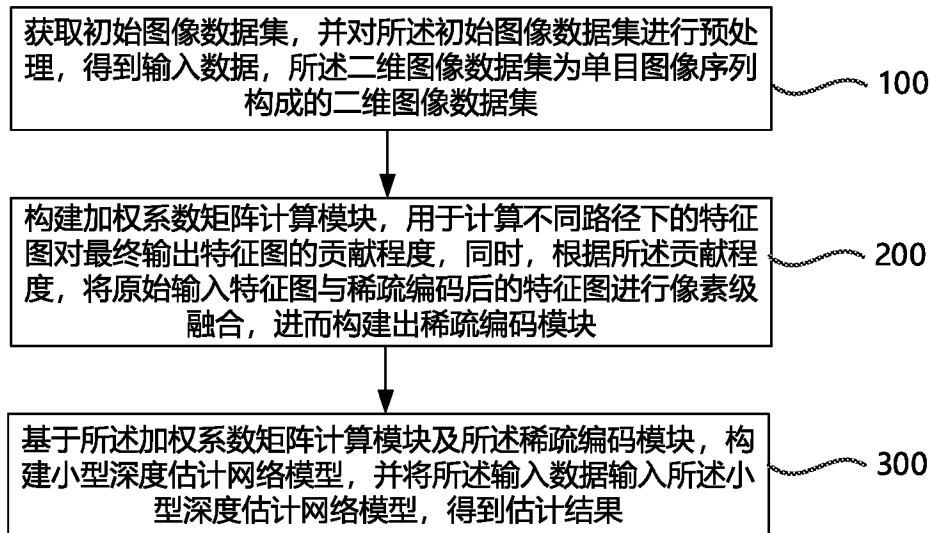


图 1

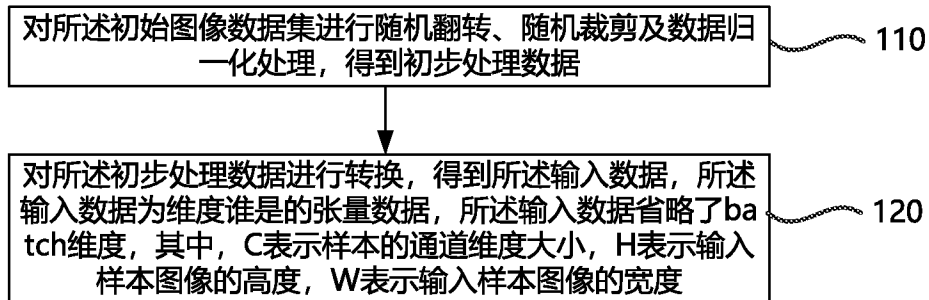


图 2

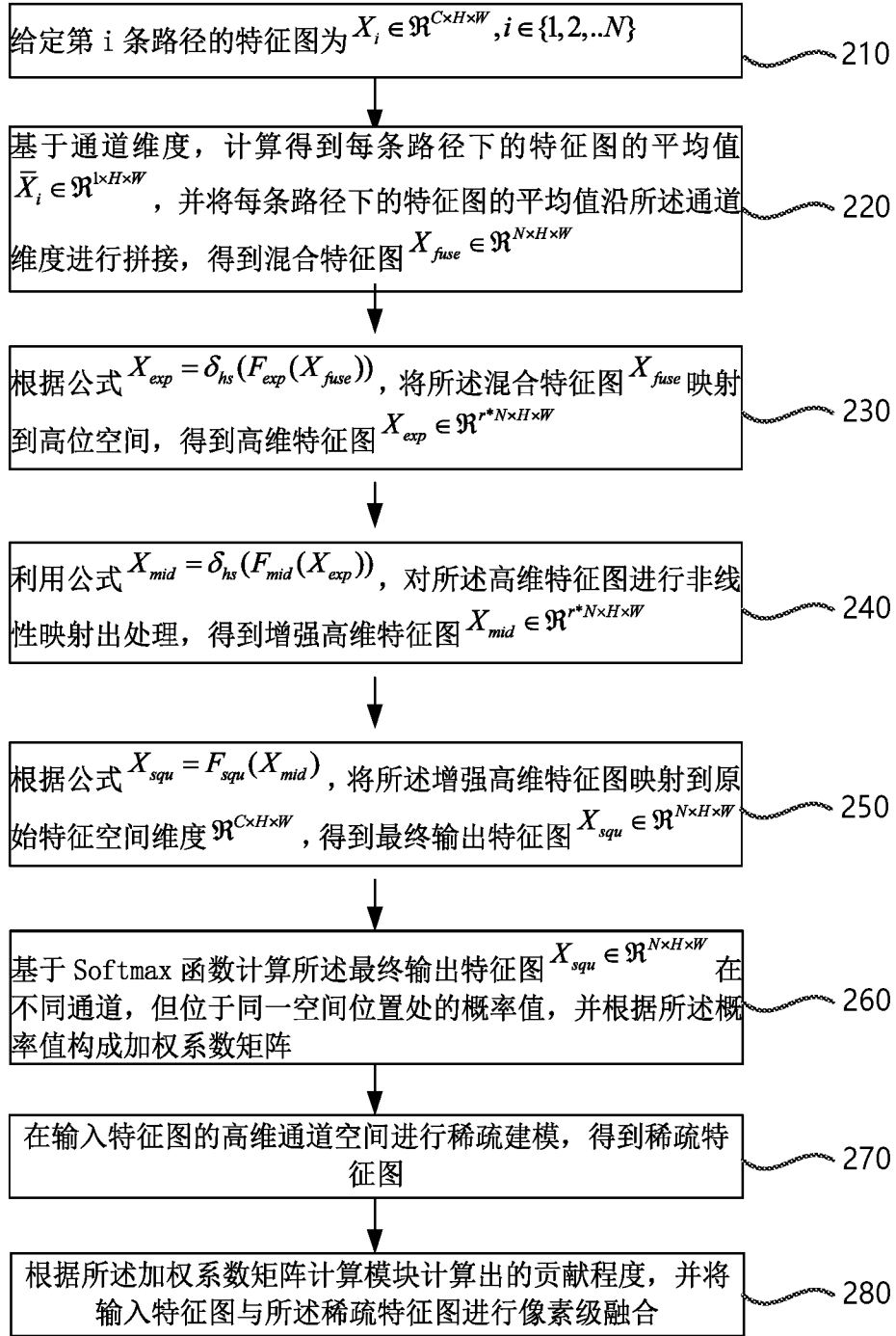


图 3

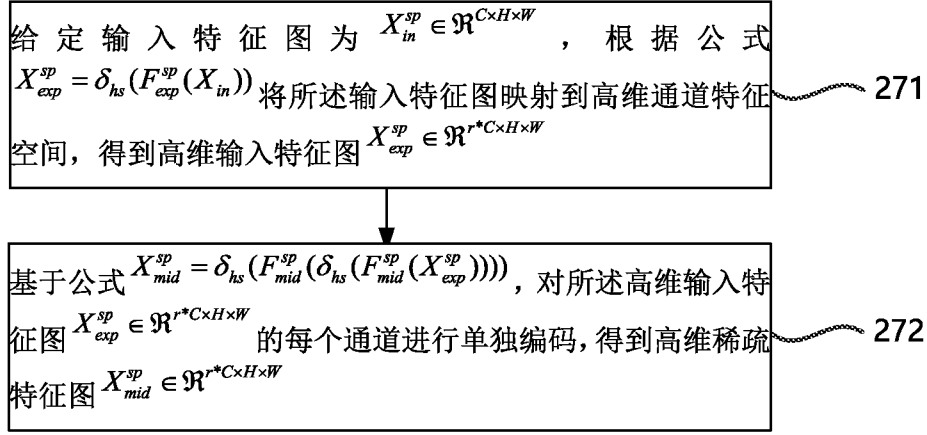


图 4

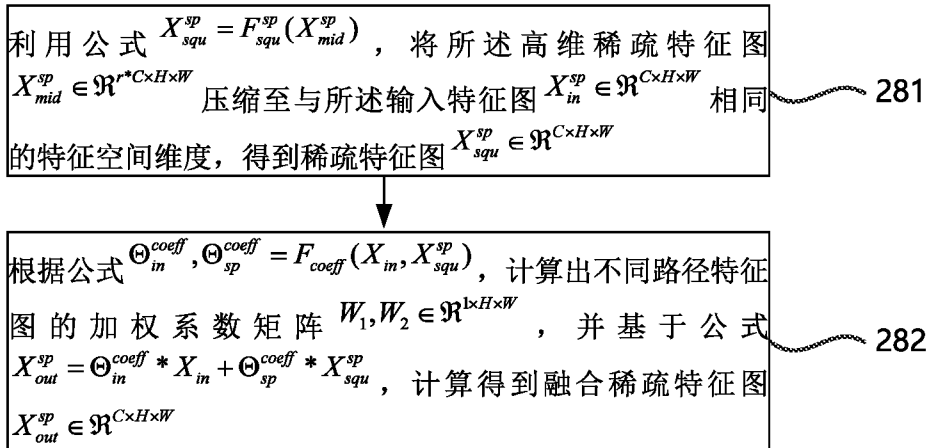


图 5

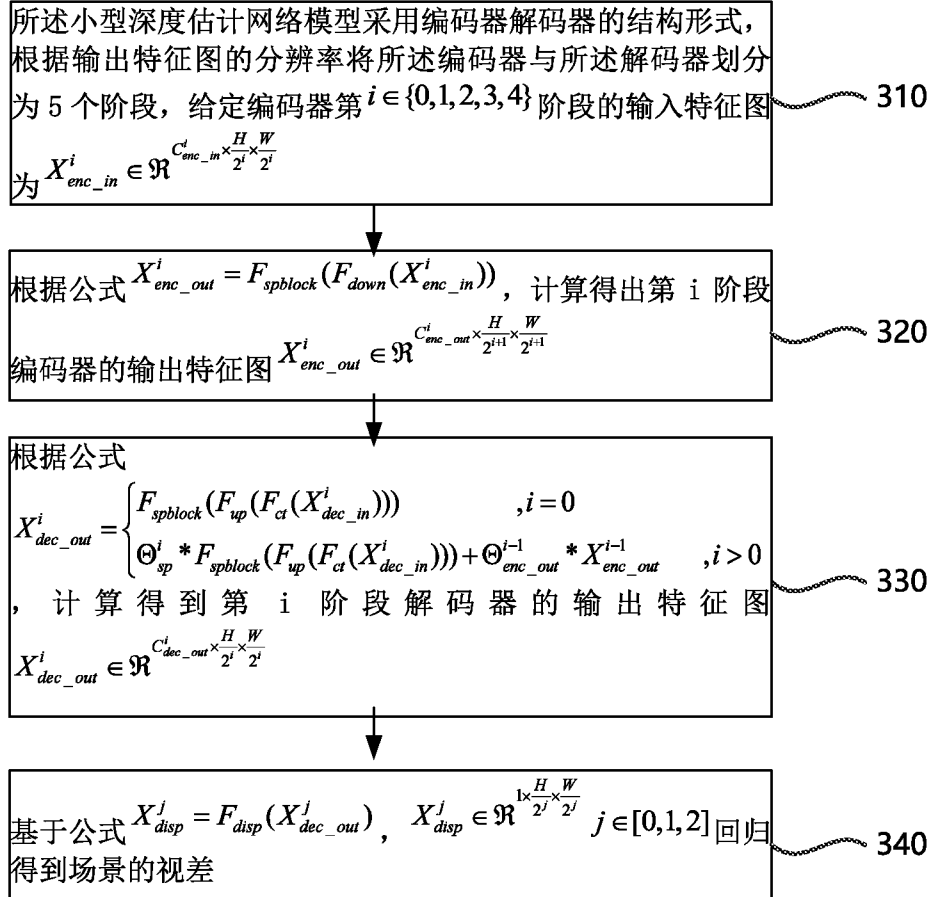


图 6

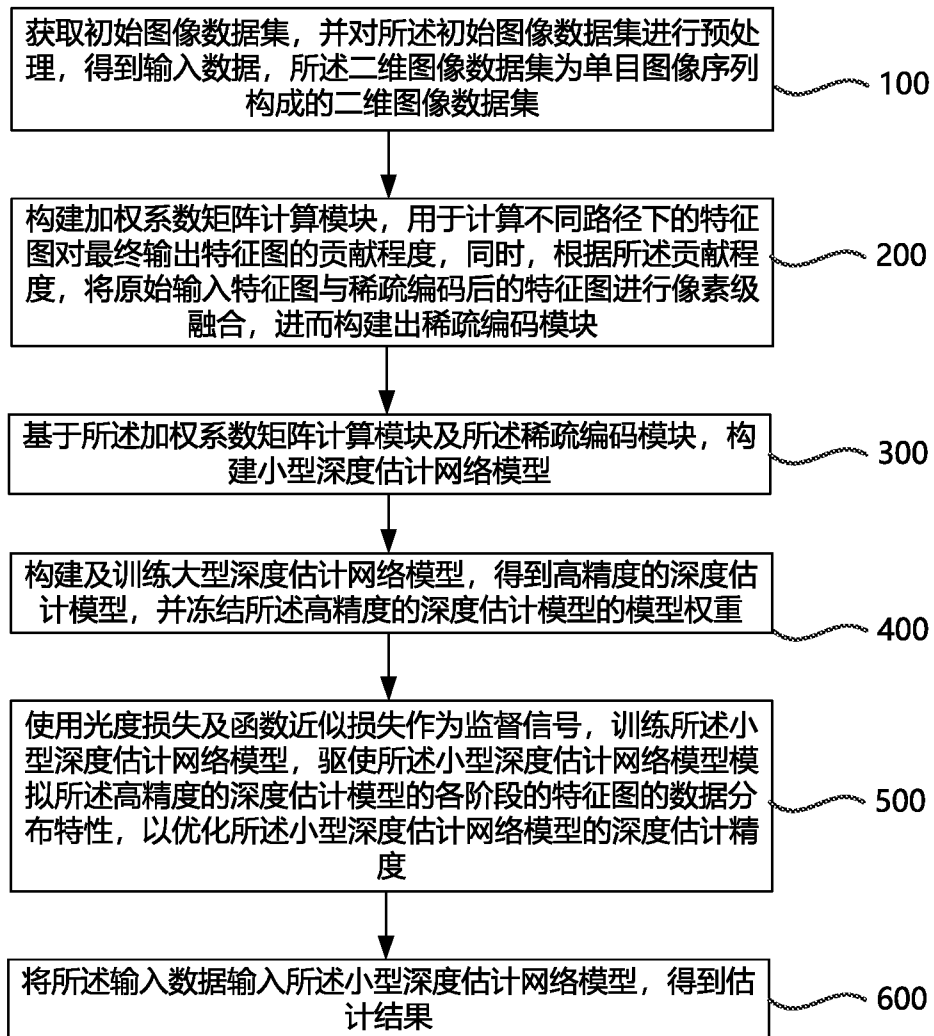


图 7

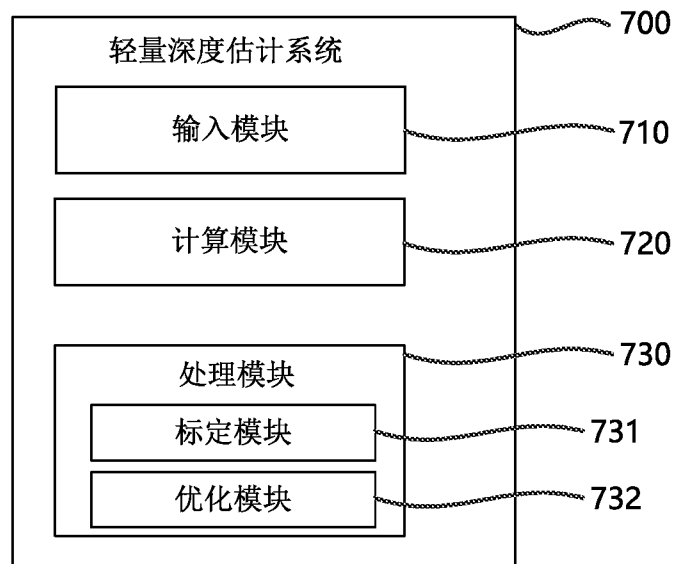


图 8

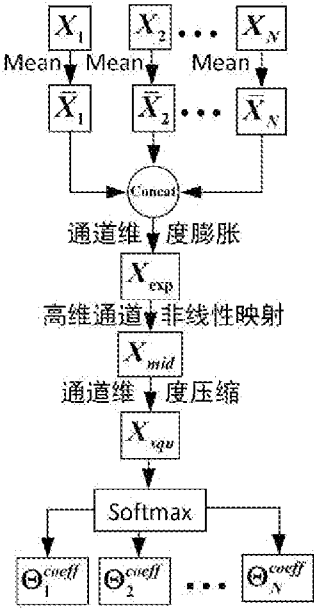


图 9

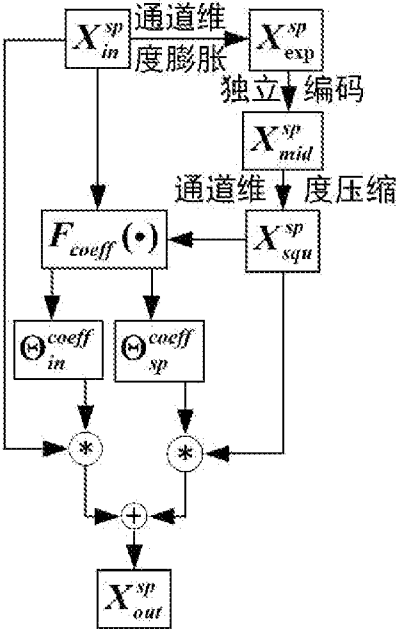


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/092055

A. CLASSIFICATION OF SUBJECT MATTER

G06T7/50(2017.01)i; G06T9/00(2006.01)i; G06V10/44(2022.01)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)

IPC: G06T, G06V

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)

CNABS, WPABSC, CNTXT, ENTXTC, DWPI, VEN, CNKI: 单目, 二维, 深度估计, 深度预测, 深度重建, 三维重建, 稀疏, 编码, 解码, 加权, 自监督, 压缩, 轻量, monocular, two dimension, 2D, depth estimate, depth predict, depth reconstruct, 3D reconstruct, sparse, encode, decode, weight, self supervised, compress, light weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2021004646 A1 (TOYOTA RESEARCH INSTITUTE, INC.) 07 January 2021 (2021-01-07) description, paragraphs 24-60, and figures 1-7	1-3, 8-10
Y	CN 115861723 A (BEIJING AEROSPACE 706 INFORMATION TECHNOLOGY CO., LTD.) 28 March 2023 (2023-03-28) description, paragraphs 23-73	1-3, 8-10
Y	CN 113192149 A (XI'AN JIAOTONG UNIVERSITY) 30 July 2021 (2021-07-30) description, paragraphs 58-71	1-3, 8-10
A	CN 114022799 A (NATIONAL DEFENSE SCIENCE AND TECHNOLOGY INNOVATION INSTITUTE ACADEMY OF MILITARY SCIENCES PLA CHINA) 08 February 2022 (2022-02-08) entire document	1-10
A	CN 115375844 A (NORTHEASTERN UNIVERSITY AT QINHUANGDAO) 22 November 2022 (2022-11-22) 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

“D” document cited by the applicant in the international application

“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

16 January 2024

Date of mailing of the international search report

17 January 2024

Name and mailing address of the ISA/CN

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

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/092055

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2021183083 A1 (ROBERT BOSCH GMBH) 17 June 2021 (2021-06-17) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/092055

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2021004646	A1	07 January 2021	US	2021004976	A1	07 January 2021
				US	11138751	B2	05 October 2021
				US	11386567	B2	12 July 2022
CN	115861723	A	28 March 2023	None			
CN	113192149	A	30 July 2021	None			
CN	114022799	A	08 February 2022	None			
CN	115375844	A	22 November 2022	None			
US	2021183083	A1	17 June 2021	US	11315266	B2	26 April 2022
				DE	102020214863	A1	17 June 2021
				CN	112991413	A	18 June 2021

A. 主题的分类

G06T7/50(2017.01)i; G06T9/00(2006.01)i; G06V10/44(2022.01)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)

IPC: G06T, G06V

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))

CNABS,WPABSC,CNXTX,ENTXTC, DWPI,VEN, CNKI; 单目, 二维, 深度估计, 深度预测, 深度重建, 三维重建, 稀疏, 编码, 解码, 加权, 自监督, 压缩, 轻量, monocular, two dimension, 2D, depth estimate, depth predict, depth reconstruct, 3D reconstruct, sparse,encode,decode,weight,self supervised, compress,light 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
Y	US 2021004646 A1 (TOYOTA RES INST INC) 2021年1月7日 (2021 - 01 - 07) 说明书第24-60段, 图1-7	1-3, 8-10
Y	CN 115861723 A (北京航天七零六信息科技有限公司) 2023年3月28日 (2023 - 03 - 28) 说明书第23-73段	1-3, 8-10
Y	CN 113192149 A (西安交通大学) 2021年7月30日 (2021 - 07 - 30) 说明书第58-71段	1-3, 8-10
A	CN 114022799 A (中国人民解放军军事科学院国防科技创新研究院) 2022年2月8日 (2022 - 02 - 08) 全文	1-10
A	CN 115375844 A (东北大学秦皇岛分校) 2022年11月22日 (2022 - 11 - 22) 全文	1-10
A	US 2021183083 A1 (BOSCH GMBH ROBERT) 2021年6月17日 (2021 - 06 - 17) 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

“D” 申请人在国际申请中引证的文件

“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

2024年1月16日

国际检索报告邮寄日期

2024年1月17日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

许菲菲

电话号码 (+86) 62411752

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/092055

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)
US	2021004646	A1	2021年1月7日	US 2021004976 A1	2021年1月7日
				US 11138751 B2	2021年10月5日
				US 11386567 B2	2022年7月12日
CN	115861723	A	2023年3月28日	无	
CN	113192149	A	2021年7月30日	无	
CN	114022799	A	2022年2月8日	无	
CN	115375844	A	2022年11月22日	无	
US	2021183083	A1	2021年6月17日	US 11315266 B2	2022年4月26日
				DE 102020214863 A1	2021年6月17日
				CN 112991413 A	2021年6月18日