

(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
2025 年 6 月 19 日 (19.06.2025)



(10) 国际公布号
WO 2025/123370 A1

(51) 国际专利分类号:

G06T 7/73 (2017.01)

(21) 国际申请号:

PCT/CN2023/139298

(22) 国际申请日:

2023 年 12 月 15 日 (15.12.2023)

(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西丽深圳大学城学苑大道1068号 518000 (CN)。

(72) 发明人: 陈思凯 (CHEN, Sikai); 中国广东省深圳市南山区西丽深圳大学城学苑大道1068号 518000 (CN)。刘畅 (LIU, Chang); 中国广东省深圳市南

山区西丽深圳大学城学苑大道1068号 518000 (CN)。刘嘉 (LIU, Jia); 中国广东省深圳市南山区西丽深圳大学城学苑大道1068号 518000 (CN)。

(74) 代理人: 深圳五邻知识产权代理事务所 (普通合伙) (SHENZHEN WULIN INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市龙华区民治街道民新社区深国际华南数字谷1#12F 518000 (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, 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,

(54) Title: TARGET SIX-DEGREE-OF-FREEDOM POSE ESTIMATION METHOD BASED ON BINOCULAR VISION AND PARTICLE SWARM OPTIMIZATION

(54) 发明名称: 基于双目视觉与粒子群优化的目标六自由度位姿估计方法

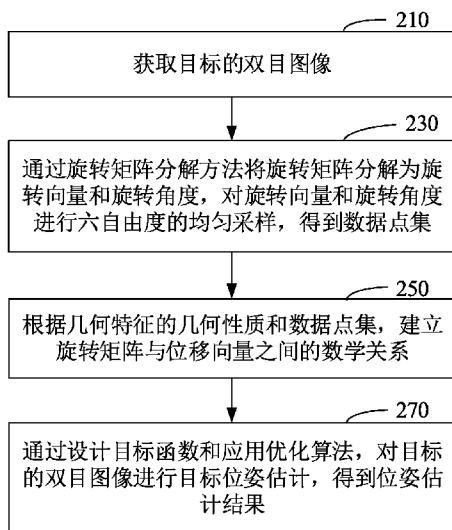


图 2

- 210 Obtain a binocular image of a target
230 Decompose a rotation matrix into a rotation vector and a rotation angle by means of a rotation matrix decomposition method, and perform six-degree-of-freedom uniform sampling on the rotation vector and the rotation angle to obtain a data point set
250 Establish a mathematical relationship between the rotation matrix and a displacement vector on the basis of the geometric property of geometric features and the data point set
270 Perform target pose estimation on the binocular image of the target by designing an objective function and an application optimization algorithm to obtain a pose estimation result

(57) Abstract: Embodiments of the present application relate to the field of computer vision. Disclosed are a target six-degree-of-freedom pose estimation method and apparatus based on binocular vision and particle swarm optimization. The method comprises: obtaining a binocular image of a target, wherein the binocular image at least comprises a rotation matrix, a displacement vector and geometric features of the target; decomposing the rotation matrix into a rotation vector and a rotation angle by means of a rotation matrix decomposition method, and performing six-degree-of-freedom uniform sampling on the rotation vector and the rotation angle to obtain a data point set; establishing a mathematical relationship between the rotation matrix and the displacement vector on the basis of the geometric property of the geometric features and the data point set; and performing target pose estimation on the binocular image of the target by designing an objective function and an application optimization algorithm to obtain a pose estimation result. The present application solves the problems in the related art of low accuracy, poor universality and poor robustness.

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。

(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 (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: 本申请实施例公开了一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法和装置, 涉及计算机视觉领域, 其中, 所述方法包括: 获取目标的双目图像; 所述双目图像中至少包括目标的旋转矩阵、位移向量和几何特征; 通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度, 对所述旋转向量和旋转角度进行六自由度的均匀采样, 得到数据点集; 根据所述几何特征的几何性质和所述数据点集, 建立所述旋转矩阵与位移向量之间的数学关系; 通过设计目标函数和应用优化算法, 对所述目标的双目图像进行目标位姿估计, 得到位姿估计结果。本申请解决了相关技术中存在的准确率较低、泛用性和鲁棒性差的问题。

基于双目视觉与粒子群优化的目标六自由度位姿估计方法

技术领域

本申请涉及计算机视觉领域，尤其涉及一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

背景技术

目标相对位姿估计在航天、工业装配中有大量应用，在航天领域，两艘航天器之间的相对姿态估计是空间交会对接的关键问题。在行星探测方面，星体探测器在着陆过程中需要高可靠性、高自主性和高精度的相对位置和姿态估计算法来实现着落任务。在工业领域，实现自动装配的工作任务的前提是需要估计物件相对摄像机的位姿。

根据选取图像特征的不同，可以将姿态估计的方法大致分为五类：基于点、基于模板、基于分割、基于边缘、基于基本几何特征，研究人员已经注意到，目标物体上明确的基本几何特征，例如 3-D 直线段和 3-D 圆，具有良好的鲁棒性。

然而，现有技术中的位姿估计方法都限制了基本几何特征的数量并需要基本几何特征之间的几何限制作为先验知识，导致方法的泛用性差，此外，在现实应用场景中，部分基本几何特征受到光照、遮挡的影响，在图像中无法成像，导致方法的鲁棒性差。

因此，急需一种准确率高、泛用性和鲁棒性高的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

发明内容

本申请各实施例提供一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法、装置、电子设备及存储介质，以解决相关技术中存在的准确率较低、泛用性和鲁棒性差的问题。

为解决上述技术问题，本申请所采用的技术方案为：

根据本申请的一个方面，一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法，所述方法包括获取目标的双目图像；所述双目图像中至少包括目标的旋转矩阵、位移向量和几何特征；通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度，对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集；根据所述几何特征的几何性质和所述数据点集，建立所述旋转矩阵与位移向量之间的数学关系；通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果；所述目标函数用于估计目标的姿态；所述优化算法用于优化位姿估计结果；所述位姿估计结果用于展示目标的六自由度姿态以及确定目标的位置和朝向。

在一示例性实施例，通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度通过以下步骤实现：通过罗德里格斯公式将旋转矩阵分解为旋转向量和旋转角度。

在一示例性实施例，对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集通过以下步骤实现：通过汤姆逊方法在单位球面对旋转向量进行六自由度的均匀采样；对 0 到 π 之间的旋转角度进行均匀采样，得到数据点集，确保所述数据点集中包括目标在六个自由度上的姿态变化；所述六个自由度包括三个平移自由度和三个旋转自由度。

在一示例性实施例，根据所述几何特征的几何性质和所述数据点集，建立所述旋转矩阵与位移向量之间的数学关系通过以下步骤实现：通过采样桑普森距离计算目标在所述双目图像上的投影点到数据点集的几何距离；所述几何距离为几何特征的几何性质；所述几何特征包括点、直线和面。根据所述投影点到数据点集的几何距离建立所述旋转矩阵与位移向量之间的数学关系。

在一示例性实施例，优化算法包括粒子群优化算法；通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果通过以下步骤实现：基于所述双目图像在立体视觉中的几何关系和旋转姿态设计目标函数；所述目标函数包含立体视觉中的几何关系和旋转姿态的约束条件；通过粒子群优化算法在所述旋转矩阵与位移向量之间的参数空间中寻找最优的旋转矩阵和位移向量，根据寻找到的最优的旋转矩阵和位移向

量对所述目标函数进行最小化；根据所述粒子群优化算法最小化后的目标函数，估计目标相对于目标坐标系的相对位姿，得到所述位姿估计结果；所述目标坐标系用于量化目标的位姿变化。

在一示例性实施例，设计目标函数通过以下步骤实现：将所述目标的旋转矩阵参数化，得到目标的直线和圆的几何特征在所述双目图像中的几何距离；根据所述几何距离得到对应的桑普森距离，根据所述桑普森距离构建目标函数。

在一示例性实施例，得到所述位姿估计结果之后，所述方法还包括以下步骤：通过列文伯格马夸特算法对所述位姿估计结果中所述目标的相对位姿进行优化。

根据本申请的一个方面，一种基于双目视觉与粒子群优化的目标六自由度位姿估计装置，所述装置包括数据获取模块，用于获取目标的双目图像；所述双目图像中至少包括目标的旋转矩阵、位移向量和几何特征；均匀采样模块，用于通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度，对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集；数学关系构建模块，用于根据所述几何特征的几何性质和所述数据点集，建立所述旋转矩阵与位移向量之间的数学关系；位姿估计模块，用于通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果。

根据本申请的一个方面，一种电子设备，包括至少一个处理器以及至少一个存储器，其中，所述存储器上存储有程序指令或代码；所述程序指令或代码被所述处理器加载并执行，使得电子设备实现如上所述的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

根据本申请的一个方面，一种存储介质，其上存储有程序指令或代码，所述程序指令或代码被处理器加载并执行，以实现如上所述的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

根据本申请的一个方面，一种计算机程序产品，计算机程序产品包括程序指令或代码，程序指令或代码存储在存储介质中，电子设备的处理器从存储介质读取程序指令或代码，加载并执行该程序指令或代码，使得电子设备实现如上所述的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

本申请提供的技术方案带来的有益效果是：

在上述技术方案中，本申请解决了相关技术中准确率较低、泛用性和鲁棒性差的问题。

具体而言，本申请首先获取目标的双目图像，双目图像中至少包括目标的旋转矩阵、位移向量和几何特征，通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角度，然后对旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集，确保全方位的姿态变化被覆盖，再根据几何特征的几何性质和数据点集，建立旋转矩阵与位移向量之间的数学关系，最后通过设计目标函数和应用优化算法对目标的双目图像进行目标位姿估计，得到位姿估计结果，通过粒子群优化算法及非线性优化算法，可以同时处理点、直线、圆三种基本几何特征，并且特征数量与相互间位置的关系没有限制，具有相当高的准确性、鲁棒性，进而解决了现有技术中存在的准确率较低、泛用性和鲁棒性差的问题。

应当理解的是，以的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本申请。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本申请的实施例，并于说明书一起用于解释本申请的原理。

图 1 是根据本申请所涉及的实施环境的示意图；

图 2 是根据一示例性实施例示出的一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法的流程图；

图 3 是图 2 对应实施例中步骤 270 在一个实施例的流程图；

图 4 是一示例性实施例中目标坐标系上的点在立体视觉中的投影示意图；

图 5 是一示例性实施例中设计目标函数的示意图；

图 6 是一示例性实施例中构建旋转矩阵与位移向量之间的数学关系的示意图；

图 7 是一示例性实施例中处理圆特征的示意图；

图 8 是一示例性实施例中目标位姿优化的示意图；

图 9 是一示例性实施例与一现有技术的对比试验结果折线图；

图 10 是一示例性实施例与另一现有技术的对比试验结果折线图；

图 11 是一示例性实施例的直线特征和圆特征数量实验结果折线图；

图 12 是一应用场景中一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法在模型 1、模型 2 和模型 3 中的实现结果示意图；

图 13 是根据一示例性实施例示出的一种基于双目视觉与粒子群优化的目标六自由度位姿估计装置的框图；

图 14 是根据一示例性实施例示出的一种服务器的结构示意图；

图 15 是根据一示例性实施例示出的一种电子设备的框图。

具体实施方式

下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，仅用于解释本申请，而不能解释为对本申请的限制。

本技术领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解，当我们称元件被“连接”或“耦接”到另一元件时，它可以直接连接或耦接到其他元件，或者也可以存在中间元件。此外，这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或多个相关联的列出项的全部或任一单元和全部组合。

下面是对本申请涉及的几个名词进行的介绍和解释：

双目视觉 (Binocular Vision)：是指通过两只眼睛同时观察同一场景，以获取深度和立体感的一种视觉方式，每只眼睛拍摄到的图像视角略有不同，这种差异被称为视差 (disparity)，由大脑解释为场景中物体的深度信息，双目视觉是人类和许多其他脊椎动物的一种常见的深度感知机制。

粒子群优化 (Particle Swarm Optimization, PSO)：是一种元启发式算法，

灵感来源于鸟群或鱼群等群体行为，该算法最初由 Russell C. Eberhart 和 James Kennedy 于 1995 年提出，PSO 的基本思想是通过模拟群体中个体之间的合作与竞争关系来搜索最优解，在 PSO 中，问题的解被称为粒子，这些粒子在问题的解空间中搜索最优解，每个粒子具有一个位置和速度，通过根据其自身历史最优位置和群体历史最优位置，以及当前位置和速度，来调整自身的移动方向和速度，这样，整个粒子群通过不断地调整位置和速度在解空间中找到最优解，PSO 的性能受到参数设置的影响，包括粒子数量、惯性权重、加速因子等，合适的参数设置可以使算法更好地收敛到全局最优解，PSO 广泛应用于函数优化、神经网络训练、图像处理等领域。

目标六自由度位姿估计：是指对物体或目标在空间中的位置和姿态（位姿）进行六自由度的估计，在三维空间中，物体的六自由度姿态包括三个平移自由度（沿 x 、 y 、 z 轴的平移）和三个旋转自由度（绕 x 、 y 、 z 轴的旋转），因此，目标六自由度位姿估计的目标是确定物体的位置和朝向，在计算机视觉、机器人学、虚拟现实等领域，六自由度位姿估计是一个关键问题，涉及到从传感器数据或图像中获取目标在空间中的准确位置和朝向。

几何特征：是指在几何学上描述物体或图像形状和结构的可测量属性，提供了有关物体或场景的几何信息，例如面积、周长、中心位置直径等。

基于单目摄像机和基本几何特征的姿态估计与跟踪算法已经被证明可行并且得到大量研究，双目摄像机与单目摄像机的差别在于，为了获得没有歧义的位姿估计，基于单目相机的位姿估计算法需要更多的几何特征，倘若目标上几何特征不足，则需要从不同角度拍摄目标的两幅以上图像，相比之下，基于双目相机的位姿估计算法由于从立体图像对中获取目标的基本几何特征的图像信息，因此需要的几何特征较少，而且鲁棒性较好。

Hu 提出一种基于一个圆和两条直线的立体视觉位姿估计算法，该方法通过立体相机重建空间圆和空间直线段，通过空间直线段的方向和空间圆的旋转轴方向估计位姿；Jian 提出一种基于一个圆和四个特征点的立体视觉位姿估计算法，该方法通过立体相机重建空间圆，并结合与圆共面的四个特征点，重建旋转矩阵，最后用正交投影迭代优化算法估计出目标的位姿；上述两种方法的缺点在于限制了基本几何特征的数量，并且基本几何特征之间需

要一些位置关系，例如直线必须互相平行等，均存在限制条件，难以推广。

由上可知，相关技术中仍存在准确率较低、泛用性和鲁棒性差的缺陷。

为此，本申请提供的基于双目视觉与粒子群优化的目标六自由度位姿估计方法，能够有效地提升基于双目视觉与粒子群优化的目标六自由度位姿估计的准确率、泛用性和鲁棒性，相应地，该基于双目视觉与粒子群优化的目标六自由度位姿估计方法适用于基于双目视觉与粒子群优化的目标六自由度位姿估计装置，该基于双目视觉与粒子群优化的目标六自由度位姿估计装置可部署于配置冯诺依曼体系结构的电子设备，例如，该电子设备可以是台式电脑、笔记本电脑、服务器等等。

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

图 1 为一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法所涉及的实施环境的示意图。该实施环境包括采集端 110 和服务端 130。

具体地，采集端 110 获取目标的双目图像，该采集端 110 可以是任意具有双目图像采集功能的电子设备，在此不进行限定。

采集端 110 与服务端 130 之间可以通过有线或者无线等方式建立的通信连接，以实现二者之间的数据传输。例如，传输的数据可以是双目图像等。

其中，服务端 130，也可以认为是云端、云平台、平台端、服务端等等，此服务端 130 可以是一台服务器，也可以是由多台服务器构成的一个服务器集群，或者是由多台服务器构成的云计算中心，以便于更好地向海量采集端 110 提供后台服务。例如，后台服务包括基于双目视觉与粒子群优化的目标六自由度位姿估计服务。

随着采集端 110 与服务端 130 的交互，在一个应用场景中，以服务端 130 提供基于双目视觉与粒子群优化的目标六自由度位姿估计服务为例，采集端 110 获取到目标的双目图像后，将双目图像中包括的目标的旋转矩阵、位移向量和几何特征发送至服务端 130，那么，服务端 130 便能够接收到该采集端 110 发送的目标的旋转矩阵、位移向量和几何特征，进而基于该目标的旋转矩阵、位移向量和几何特征提供基于双目视觉与粒子群优化的目

标六自由度位姿估计服务。具体而言，服务器端 130 获取到目标的旋转矩阵、位移向量和几何特征后，便通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角度，再对旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集，确保全方位的姿态变化被覆盖，再根据几何特征的几何性质和数据点集，建立旋转矩阵与位移向量之间的数学关系，最后通过设计目标函数和应用优化算法对目标的双目图像进行目标位姿估计，得到位姿估计结果。

请参阅图 2，本申请实施例提供了一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法，该方法适用于电子设备，该电子设备可以是图 1 所示出实施环境中的服务器端 130，还可以是台式电脑、笔记本电脑、服务器等等。

在下述方法实施例中，为了便于描述，以该方法各步骤的执行主体为电子设备为例进行说明，但是并非对此构成具体限定。

如图 2 所示，该方法可以包括以下步骤：

步骤 210，获取目标的双目图像。

其中，双目图像是通过两个摄像头（左摄像机和右摄像机）同时捕捉同一场景或对象的图像，这种摄像头布置模拟了人类的双眼视觉系统，能够提供深度信息，每个摄像头捕捉到的图像被称为双目立体图像对，其中一张图像对应于左眼视角，另一张图像对应于右眼视角，这两个图像之间的差异，也称为视差（disparity），是在深度感知中非常重要的信息，通过计算视差，可以得出场景中不同点的深度或距离。

在一种可能的实现方式，双目图像中至少包括目标的旋转矩阵、位移向量和几何特征。

在一种可能的实现方式，获取目标的双目图像的可以通过首先选择和安装双目摄像头，然后进行双目摄像头的校准，确保两个摄像头之间的参数和位置准确，再使用相应的软件或编程语言（如 Python 中的 OpenCV 库）来捕获双目摄像头的图像，确保同步捕获两个摄像头的图像，以便后续的深度计算，使用立体匹配算法来计算深度信息，根据立体匹配的结果，生成深度图。

步骤 230，通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角

度，对旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集。

在一种可能的实现方式，通过罗德里格斯公式将旋转矩阵分解为旋转向量和旋转角度。

具体的计算公式如下：

$$\begin{aligned}\mathbf{R} &= \cos \theta \mathbf{I} + (1 - \cos \theta) \mathbf{n} \mathbf{n}^T + \sin \theta [\mathbf{n}]_{\times} \\ &= \mathbf{I} + \sin \theta [\mathbf{n}]_{\times} + (1 - \cos \theta) [\mathbf{n}]_{\times}^2,\end{aligned}$$

其中，参数左上角符号表示坐标名称，例如 l 、 r 、 o 分别表示左摄像机坐标系、右摄像机坐标系、目标坐标系；普通字体表示标量；粗体字表示矩阵或向量； $\mathbf{R}$ 和 $\mathbf{t}$ 分别表示旋转矩阵和位移向量； $\mathbf{I}$ 表示三维单位矩阵。

具体地，旋转向量 $\mathbf{n}$ 的计算公式如下：

$$\begin{aligned}\mathbf{n} &= [n_1 \quad n_2 \quad n_3]^T \\ &= [\cos \alpha \sin \beta \quad \sin \alpha \sin \beta \quad \cos \beta]^T, 0 \leq \alpha < 2\pi, 0 \leq \beta < \pi \\ [\mathbf{n}]_{\times} &= \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}.\end{aligned}$$

在一种可能的实现方式，通过汤姆逊方法在单位球面对旋转向量进行六自由度的均匀采样，对 0 到 π 之间的旋转角度进行均匀采样，得到数据点集，确保数据点集中包括目标在六个自由度上的姿态变化，其中，六个自由度包括三个平移自由度和三个旋转自由度。

通过此种方式，本申请实施例通过汤姆逊方法在单位球面上对旋转向量进行均匀采样，汤姆逊方法旨在最小化系统的电势能，采样的旋转向量在球面上更加均匀，避免了其他非均匀采样方法可能引入的偏差，实现了更加稳定和有序的点分布，汤姆逊方法背后的物理原理基于库仑定律，模拟点之间的电荷相互作用，这种基于物理原理的方法更容易理解和解释，还可以通过调整库仑定律中的常数和初始配置优化过程，具有很强的可调性。

在一种可能的实现方式，通过上述过程实现目标的运动表示之后，根据左摄像机和右摄像机为目标构建立体相机模型。

具体地，如图 4 所示，以左摄像机和右摄像机的光心为原点，光轴为轴，建立 ${}^lO-{}^lX{}^lY{}^lZ$ 和 ${}^rO-{}^rX{}^rY{}^rZ$ 坐标系，图像平面在 ${}^lZ=f_l$ 和 ${}^rZ=f_r$ 处，其中 f_l 和 f_r 分别表示左摄像机和右摄像机的焦距，图像平面坐标系设成

$o_{imgl} - {}^l u {}^l v$ 和 $o_{imgl} - {}^r u {}^r v$, 其中 ${}^l u$ 、 ${}^l v$ 、 ${}^r u$ 、 ${}^r v$ 分别与 ${}^l X$ 、 ${}^l Y$ 、 ${}^r X$ 、 ${}^r Y$ 平行。

其中, 向量符号上面加一横表示这个向量代表的基本几何特征在图像中的投影, 例如: ${}^l \bar{\mathbf{p}}$ 表示 ${}^o \mathbf{p}$ 在左图像的投影, 在分别给定左摄像机和右摄像机的内参数矩阵和条件下, ${}^l \bar{\mathbf{p}}$ 和 ${}^r \bar{\mathbf{p}}$ 与 ${}^o \mathbf{p}$ 的具体关系如下:

$${}^l \bar{\mathbf{p}} = \frac{{}^l \mathbf{K}_{12} {}^l \mathbf{p}}{{}^l \mathbf{K}_3 {}^l \mathbf{p}} = \frac{{}^l \mathbf{K}_{12} ({}^l \mathbf{R}_o {}^o \mathbf{p} + {}^l \mathbf{t}_o)}{{}^l \mathbf{K}_3 ({}^l \mathbf{R}_o {}^o \mathbf{p} + {}^l \mathbf{t}_o)},$$

$${}^r \bar{\mathbf{p}} = \frac{{}^r \mathbf{K}_{12} {}^r \mathbf{p}}{{}^r \mathbf{K}_3 {}^r \mathbf{p}} = \frac{{}^r \mathbf{K}_{12} ({}^r \mathbf{R}_o {}^o \mathbf{p} + {}^r \mathbf{t}_o)}{{}^r \mathbf{K}_3 ({}^r \mathbf{R}_o {}^o \mathbf{p} + {}^r \mathbf{t}_o)},$$

其中, ${}^l \mathbf{K}_{12}$ 和 ${}^r \mathbf{K}_{12}$ 分别表示 ${}^l \mathbf{K}$ 和 ${}^r \mathbf{K}$ 的前两行所构成的矩阵, ${}^l \mathbf{K}_3$ 和 ${}^r \mathbf{K}_3$ 分别表示 ${}^l \mathbf{K}$ 和 ${}^r \mathbf{K}$ 的第三行所构成的矩阵; ${}^l \mathbf{R}_o$ 和 ${}^l \mathbf{t}_o$ 表示目标坐标系相对左摄像机坐标的旋转矩阵和位移向量, ${}^r \mathbf{R}_o$ 和 ${}^r \mathbf{t}_o$ 表示目标坐标系相对右摄像机坐标的旋转矩阵和位移向量。在给定两个摄像机的旋转矩阵 ${}^r \mathbf{R}_l$ 和位移向量 ${}^r \mathbf{t}_l$ 的条件下, ${}^r \mathbf{R}_o$ 和 ${}^r \mathbf{t}_o$ 的具体表示方式如下:

$${}^r \mathbf{R}_o = {}^r \mathbf{R}_l {}^l \mathbf{R}_o,$$

$${}^r \mathbf{t}_o = {}^r \mathbf{R}_l {}^l \mathbf{t}_o + {}^r \mathbf{t}_l.$$

通过上述过程, 本申请实施例构建了目标的立体相机模型, 通过模拟人类双眼视觉系统, 获取了场景中目标的深度信息, 可以推断目标距离相机的远近, 提供了更为精确的物体定位信息, 能够弥补单目相机在缺乏深度信息的情况下可能无法准确估计物体的距离这一缺陷, 提高了深度感知的准确性, 对于光照变化和纹理缺乏的环境更为鲁棒。

步骤 250, 根据几何特征的几何性质和数据点集, 建立旋转矩阵与位移向量之间的数学关系。

具体地, 步骤 250 可以包括以下步骤:

步骤 251 通过采样桑普森距离计算目标在所述双目图像上的投影点到数据点集的几何距离。

其中, 几何距离为几何特征的几何性质, 几何特征包括点、直线和面。

步骤 253, 根据投影点到数据点集的几何距离建立旋转矩阵与位移向量之间的数学关系。

步骤 270, 通过设计目标函数和应用优化算法, 对目标的双目图像进行目标位姿估计, 得到位姿估计结果。

其中，目标函数用于估计目标的姿态，优化算法用于优化位姿估计结果，位姿估计结果用于展示目标的六自由度姿态以及确定目标的位置和朝向。

在一种可能的实现方式，优化算法包括粒子群优化算法。

具体地，如图 3 所示，步骤 270 可以包括以下步骤：

步骤 310，基于双目图像在立体视觉中的几何关系和旋转姿态设计目标函数。

在一种可能的实现方式，设计目标函数通过首先将目标的旋转矩阵参数化，得到目标的直线和圆的几何特征在双目图像中的几何距离，然后根据几何距离得到对应的桑普森距离，最后根据桑普森距离构建目标函数。

具体地，将旋转矩阵 $\mathbf{R}$ 参数化后，记作 $\Omega = [\alpha \ \beta \ \theta]^T$ ，设计目标函数如下：

$$f(\Omega) = \sum_{i=1}^{N_L} d_i(\Omega) + \sum_{i=1}^{N_C} s_i(\Omega),$$

其中， d_i 表示数据点到线段的几何距离， s_i 表示数据点到椭圆的桑普森距离， N_L 和 N_C 分别表示图像上直线的数据点数量和椭圆的数据点数量； d_i 的计算方式如下：

$$d_i(\Omega) = \|q_i - \bar{p}_i\|_2,$$

其中， q_i 表示第 i 个直线数据点， $\bar{p}_i$ 表示该线段上距离 q_i 最近的点， $\bar{p}_i$ 的计算方法如下：

$$\begin{aligned} \bar{p}_i &= \bar{p}_s + \lambda(\bar{p}_e - \bar{p}_s), \\ \lambda &= \text{Median}(0 \ \lambda^* \ 1), \\ \lambda^* &= \frac{(p_e - p_s)^T (q_i - p_s)}{(p_e - p_s)^T (p_e - p_s)}. \end{aligned}$$

其中， $\bar{p}_s$ 和 $\bar{p}_e$ 分别表示模型直线在图像上的像， $\text{Median}()$ 表示取中位数，如图 5 和图 6 所示，引入 λ 和 λ^* 是为了避免数据点落在模型直线段的像所在的直线上，因为准确的位姿估计应该使数据点落在模型直线段的像上。

在一种可能的实现方式，目标函数中的 $s_i(\Omega)$ 表示数据点到椭圆的桑普森距离，具体计算公式如下：

$$s_i = \frac{\|\tilde{\mathbf{q}}_i^T \mathbf{C}' \tilde{\mathbf{q}}_i\|_2}{2\sqrt{(\mathbf{C}' \tilde{\mathbf{q}}_i)_1^2 + (\mathbf{C}' \tilde{\mathbf{q}}_i)_2^2}},$$

其中, $\tilde{\mathbf{q}}_i$ 表示图像中圆的数据点 q_i 的齐次坐标形式; $(\mathbf{C}' \tilde{\mathbf{q}}_i)_1$ 和 $(\mathbf{C}' \tilde{\mathbf{q}}_i)_2$ 分别表示 $(\mathbf{C}' \tilde{\mathbf{q}}_i)$ 的第一、第二个特征; $\mathbf{C}'$ 表示空间圆根据姿态计算得到的在图像上像的矩阵形式, 具体计算公式如下:

$$\mathbf{C}' = \kappa \mathbf{K}^{-T} \left((\delta \mathbf{I} - \boldsymbol{\eta} \mathbf{C}^T) (\delta \mathbf{I} - \mathbf{C} \boldsymbol{\eta}^T) - r^2 \boldsymbol{\eta} \boldsymbol{\eta}^T \right) \mathbf{K}^{-1},$$

其中, κ 表示一个非零常数; $\mathbf{K}$ 表示摄像机内参数矩阵 (在本申请实施例中表示 ${}^w\mathbf{K}$ 或 ${}^c\mathbf{K}$); δ 表示光心到圆的支撑平面的距离; $\boldsymbol{\eta}$ 表示空间圆的支撑平面的单位法向量; $\mathbf{C}$ 表示空间圆在摄像机坐标系中的坐标; r 表示空间圆的半径。

通过上述过程, 本申请实施例得到了有利于获取目标准确位姿的目标函数, 通过 $d_i(\Omega)$ 表示数据点到空间直线段在图像中的像的距离, 实现了尽可能地保证模型的像被数据点包围, 通过 $s_i(\Omega)$ 表示桑普森距离, 比直接求解几何距离更加高效准确。

步骤 330, 通过粒子群优化算法在旋转矩阵与位移向量之间的参数空间中寻找最优的旋转矩阵和位移向量, 根据寻找到的最优的旋转矩阵和位移向量对目标函数进行最小化。

步骤 350, 根据粒子群优化算法最小化后的目标函数估计目标相对于目标坐标系的相对位姿。

其中, 目标坐标系用于量化目标的位姿变化。

具体地, 粒子群优化算法 (Particle Swarm Optimization, PSO) 将每个优化问题的解抽象成搜索空间中一个没有质量和体积的微粒, 所有粒子都有一个由目标函数决定的适应值、一个决定粒子移动方向和距离的速度, 粒子们知道自己到目前为止发现的最好位置和现在的位置, 此外每个粒子还知道到目前为止整个群体中所有粒子发现的最好位置, 粒子就是通过自己发现的最好位置和群体发现的最好位置来决定自己的移动速度, 最终找到全局最优解。

本申请实施例通过 PSO 对旋转矩阵 $\mathbf{R}$ 与平移向量 $\mathbf{t}$ 进行估计, 旋转矩阵的自由度和平移向量的自由度决定了粒子群搜索空间的维度 $d=6$, 现假设 d 维搜索空间中第 i 个粒子的位置和速度分别为 $X_i = [x_{i1} \ x_{i2} \ \dots \ x_{id}]$ 和

$V_i = [v_{i1} \ v_{i2} \ \dots \ v_{id}]$ ，所有粒子的位置 X 和速度 V 的初始值随机生成。在每次迭代中，粒子通过跟踪两个最优解来更新自己，其一是粒子本身找到的最优解，记作 $Pbest_i = [pbest_{i1} \ pbest_{i2} \ \dots \ pbest_{id}]^T$ ；其二是整个群体目前找到的最优解，记作 $Gbest_g = [gbest_1 \ gbest_2 \ \dots \ gbest_d]^T$ ，粒子根据如下公式更新自己的速度和位置：

$$v_{ij}^{(k+1)} = wv_{ij}^{(k)} + c_1r_1(pbest_{ij} - x_{ij}^{(k)}) + c_2r_2(gbest_{gj} - x_{ij}^{(k)}),$$

$$x_{ij}^{(k+1)} = x_{ij}^{(k)} + v_{ij}^{(k+1)},$$

其中， $j=1,2,\dots,d$ ， k 表示当前迭代次数， w 表示惯性权重， c_1 和 c_2 表示正的学习因子， r_1 和 r_2 为 0 到 1 之间的均匀分布的随机数。

本申请实施例所构建的 t 关于 R 的函数关系的物理意义是：将 CAD 模型根据姿态（ R 和 t ）投影到图像上，并使几何特征尽可能靠近其图像上所对应的数据点；直线特征方面，直线特征根据姿态的投影应该尽可能落在对应数据点（像素点）集内，如图 6 所示，如果用投影点到数据点集的拟合直线的几何距离作为该函数关系，则可能出现投影点落在数据点集区域以外，但仍在该点集的拟合直线上的情况，因此，如图 5 所示，将投影点到数据点集的拟合直线段的距离作为该函数关系，能够保证投影点尽可能落在数据点集区域以内；圆特征方面，如图 7 所示，本申请实施例采样桑普森距离，原因在于直接计算投影点到数据点集的拟合曲线的几何距离需要大量时间，会导致效率大幅降低，而桑普森距离可以近似几何距离并且便于计算。

在一种可能的实现方式，根据粒子群优化算法最小化后的目标函数，得到位姿估计结果之后，通过列文伯格马夸特算法对位姿估计结果中目标的相对位姿进行优化。

具体地，在列文伯格马夸尔特算法优化过程中，第 k 次迭代中的更新公式如下：

$$\nabla_{x_2^T} {}^l \bar{p}_i = \frac{K_{12} - {}^l \bar{p}_i K_3}{K_3 (R_e {}^o p_i + t_e)}.$$

其中，如图 8 所示，在每次迭代都需要计算几何特征在图像中距离数据点最近的像，直线的最近点可以直接计算得到，而由于计算空间圆上在图像

中距离数据点最近的像需要大量的计算，所以本申请实施例引进中间变量 φ ，以近距离点取代最近点，从而避免每次迭代都计算空间圆上距离数据点最近的点，通过本申请实施例设计的更新公式，目标函数收敛时对应的位姿（ $\mathbf{R}$ 和 $\mathbf{t}$ ）即为最后的目标位姿估计结果。

通过上述过程，本申请实施例通过首先获取获取目标的双目图像，双目图像中至少包括目标的旋转矩阵、位移向量和几何特征，通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角度，然后对旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集，确保全方位的姿态变化被覆盖，再根据几何特征的几何性质和数据点集，建立旋转矩阵与位移向量之间的数学关系，最后通过设计目标函数和应用优化算法对目标的双目图像进行目标位姿估计，得到位姿估计结果，通过粒子群优化算法及非线性优化算法，可以同时处理点、直线、圆三种基本几何特征，并且特征数量与相互间位置的关系没有限制，具有相当高的准确性、鲁棒性，进而解决了现有技术中存在的准确率较低、泛用性和鲁棒性差的问题。

图 9 展示了一示例性实施例与一现有技术的对比试验结果折线图。

具体地，如图 9-1 所示，本申请实施例与一现有技术进行了距离测试的对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）在旋转矩阵和位移向量的误差上均明显比该现有技术更好、更稳定、更鲁棒。

如图 9-2 所示，本申请实施例与一现有技术进行了噪声测试的对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）在旋转矩阵估计方面，明显优于该现有技术，且在位移向量估计方面，噪声增大至 7 以后，比该现有技术误差更小、鲁棒性明显增强。

如图 9-3 所示，本申请实施例与一现有技术进行了遮挡率测试的对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）在旋转矩阵方面，明显比该现有技术更准确、更稳定、更鲁棒，且在位移向量方面，当遮挡率小于 0.2 时，估计误差略大于该现有技术，当遮挡率大于 0.2 时，估计误差明显优于该现有技术。

图 10 展示了一示例性实施例与另一现有技术的对比试验结果折线图。

具体地，如图 10-1 所示，本申请实施例与一现有技术进行了距离测试的

对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）在位移向量估计误差相近，而在旋转矩阵估计方面明显比该现有技术更准确、更稳定、更鲁棒。

如图 10-2 所示，本申请实施例与一现有技术进行了噪声测试的对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）在位移向量方面当噪声小于 8 时，误差略大于该现有技术，当噪声大于 8 时，误差明显优于该现有技术；而在旋转矩阵估计方面明显比该现有技术更准确、更稳定、更鲁棒。

如图 10-3 所示，本申请实施例与一现有技术进行了遮挡率测试的对比实验，实验结果表明本申请实施例（pso 线和 pso+lm 线）的位移向量估计方面，在遮挡率小时（遮挡率小于 0.3），误差略大于该现有技术，但遮挡率大时（遮挡率大于 0.3），误差明显优于该现有技术；而在旋转矩阵估计方面明显比该现有技术误差更小、更稳定、更鲁棒。

图 11 展示了一示例性实施例的直线特征和圆特征数量实验结果折线图。

具体地，如图 11-1 所示，本申请实施例进行了直线特征数量实验，可以看出随着直线特征数量增加，旋转矩阵和位移向量的误差均减小。

如图 11-2 所示，本申请实施例进行了圆特征数量实验，可以看出随着圆特征数量增加，旋转矩阵和位移向量的误差均减小。

通过上述过程，本申请实施例提高了旋转矩阵和位移向量的准确度，提高了位移向量估计的准确度，可以同时处理点、直线、圆三种基本几何特征，并且特征数量与相互间位置的关系没有限制，具有相当高的准确性、鲁棒性，进而解决了现有技术中存在的准确率较低、泛用性和鲁棒性差的问题。

图 13 是一应用场景中一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法在模型 1、模型 2 和模型 3 中的实现结果示意图。

具体地，如图 13-1 所示，本申请实施例在模型 1 上进行了实物实验，图（a）为左摄像机图像，图（b）为右摄像机图像，圆形线条是模型 1 根据本申请实施例估计的姿态投影到各个图像中的像；如图 13-2 所示，本申请实施例在模型 2 上进行了实物实验，图（a）为左摄像机图像，图（b）为右摄像机图像，圆形线条是模型 2 根据本申请实施例估计的姿态投影到各个图像中的像；如图 13-3 所示，本申请实施例在模型 3 上进行了实物实验，图（a）

为左摄像机图像，图（b）为右摄像机图像，圆形线条是模型 3 根据本申请实施例估计的姿态投影到各个图像中的像，从各图中均可看出模型的像跟实物已经基本重合，说明本申请提供的方法估计出来的姿态非常接近真实姿态。

下述为本申请装置实施例，可以用于执行本申请所涉及的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。对于本申请装置实施例中未披露的细节，请参照本申请所涉及的基于双目视觉与粒子群优化的目标六自由度位姿估计方法的方法实施例。

请参阅图 13，本申请实施例中提供了一种基于双目视觉与粒子群优化的目标六自由度位姿估计装置 800，包括但不限于：数据获取模块 810、均匀采样模块 830、数学关系构建模块 850 以及位姿估计模块 870。

其中，数据获取模块 810，用于获取目标的双目图像；双目图像中至少包括目标的旋转矩阵、位移向量和几何特征。

均匀采样模块 830，用于通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角度，对旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集。

数学关系构建模块 850，用于根据几何特征的几何性质和数据点集，建立旋转矩阵与位移向量之间的数学关系。

位姿估计模块 870，用于通过设计目标函数和应用优化算法，对目标的双目图像进行目标位姿估计，得到位姿估计结果。

需要说明的是，上述实施例所提供的基于双目视觉与粒子群优化的目标六自由度位姿估计装置在进行基于双目视觉与粒子群优化的目标六自由度位姿估计时，仅以上述各功能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模块完成，即基于双目视觉与粒子群优化的目标六自由度位姿估计装置的内部结构将划分为不同的功能模块，以完成以上描述的全部或者部分功能。

另外，上述实施例所提供的基于双目视觉与粒子群优化的目标六自由度位姿估计装置与基于双目视觉与粒子群优化的目标六自由度位姿估计方法的实施例属于同一构思，其中各模块执行操作的具体方式已经在方法实施例中进行了详细描述，此处不再赘述。

图 14 根据一示例性实施例示出的一种服务器的结构示意图。该服务器适用于图 1 所示出实施环境中的服务器端 130。

需要说明的是，该服务器只是一个适配于本申请的示例，不能认为是提供了对本申请的使用范围的任何限制。该服务器也不能解释为需要依赖于或者必须具有图 14 示出的示例性的服务器 2000 中的一个或者多个组件。

服务器 2000 的硬件结构可因配置或者性能的不同而产生较大的差异，如图 14 所示，服务器 2000 包括：电源 210、接口 230、至少一存储器 250、以及至少一中央处理器（CPU, Central Processing Units）270。

具体地，电源 210 用于为服务器 2000 上的各硬件设备提供工作电压。

接口 230 包括至少一有线或无线网络接口，用于与外部设备交互。例如，进行图 1 所示出实施环境中终端 100 与服务器 200 之间的交互。

当然，在其余本申请适配的示例中，接口 230 还可以进一步包括至少一串并转换接口 233、至少一输入输出接口 235 以及至少一 USB 接口 237 等，如图 14 所示，在此并非对此构成具体限定。

存储器 250 作为资源存储的载体，可以是只读存储器、随机存储器、磁盘或者光盘等，其上所存储的资源包括操作系统 251、应用程序 253 及数据 255 等，存储方式可以是短暂存储或者永久存储。

其中，操作系统 251 用于管理与控制服务器 200 上的各硬件设备以及应用程序 253，以实现中央处理器 270 对存储器 250 中海量数据 255 的运算与处理，其可以是 Windows Server™、Mac OS X™、Unix™、Linux™、FreeBSD™ 等。

应用程序 253 是基于操作系统 251 之上完成至少一项特定工作的计算机程序，其可以包括至少一模块（图 14 未示出），每个模块都可以分别包含有对服务器 2000 的计算机程序。例如，基于双目视觉与粒子群优化的目标六自由度位姿估计装置可视为部署于服务器 2000 的应用程序 253。

数据 255 可以是存储于双目相机中的双目图像等，存储于存储器 250 中。

中央处理器 270 可以包括一个或多个以上的处理器，并设置为通过至少一通信总线与存储器 250 通信，以读取存储器 250 中存储的计算机程序，进而实现对存储器 250 中海量数据 255 的运算与处理。例如，通过中央处理器 270 读取存储器 250 中存储的一系列计算机程序的形式来完成基于双目视觉

与粒子群优化的目标六自由度位姿估计方法。

此外，通过硬件电路或者硬件电路结合软件也能同样实现本申请，因此，实现本申请并不限于任何特定硬件电路、软件以及两者的组合。

请参阅图 15，本申请实施例中提供了一种电子设备 4000，该电子设备 400 可以包括：（需要根据本申请具体情况适应性修改）台式电脑、笔记本电脑、服务器等。

在图 15 中，该电子设备 4000 包括至少一个处理器 4001、至少一条通信总线 4002 以及至少一个存储器 4003。

其中，处理器 4001 和存储器 4003 相连，如通过通信总线 4002 相连。可选地，电子设备 4000 还可以包括收发器 4004，收发器 4004 可以用于该电子设备与其他电子设备之间的数据交互，如数据的发送和/或数据的接收等。需要说明的是，实际应用中收发器 4004 不限于一个，该电子设备 4000 的结构并不构成对本申请实施例的限定。

处理器 4001 可以是 CPU（Central Processing Unit，中央处理器），通用处理器，DSP（Digital Signal Processor，数据信号处理器），ASIC（Application Specific Integrated Circuit，专用集成电路），FPGA（Field Programmable Gate Array，现场可编程门阵列）或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请公开内容所描述的各种示例性的逻辑方框，模块和电路。处理器 4001 也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，DSP 和微处理器的组合等。

通信总线 4002 可包括一通路，在上述组件之间传送信息。通信总线 4002 可以是 PCI（Peripheral Component Interconnect，外设部件互连标准）总线或 EISA（Extended Industry Standard Architecture，扩展工业标准结构）总线等。通信总线 4002 可以分为地址总线、数据总线、控制总线等。为便于表示，图 10 中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

存储器 4003 可以是 ROM（Read Only Memory，只读存储器）或可存储静态信息和指令的其他类型的静态存储设备，RAM（Random Access Memory，随机存取存储器）或者可存储信息和指令的其他类型的动态存储设备，也可以是 EEPROM（Electrically Erasable Programmable Read Only Memory，电可擦可编程只读存储器）、CD-ROM（Compact Disc Read Only Memory，只读光

盘)或其他光盘存储、光碟存储(包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等)、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质,但不限于此。

存储器 4003 上存储有计算机程序,处理器 4001 通过通信总线 4002 读取存储器 4003 中存储的计算机程序。

该计算机程序被处理器 4001 执行时实现上述各实施例中的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

此外,本申请实施例中提供了一种存储介质,该存储介质上存储有计算机程序,该计算机程序被处理器执行时实现上述各实施例中的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

本申请实施例中提供了一种计算机程序产品,该计算机程序产品包括计算机程序,该计算机程序存储在存储介质中。计算机设备的处理器从存储介质读取该计算机程序,处理器执行该计算机程序,使得该计算机设备执行上述各实施例中的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

与相关技术相比,本申请的有益效果是:

1.本申请提出了一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法。本申请通过首先获取获取目标的双目图像,双目图像中至少包括目标的旋转矩阵、位移向量和几何特征,通过旋转矩阵分解方法将旋转矩阵分解为旋转向量和旋转角度,然后对旋转向量和旋转角度进行均匀采样,确保全方位的姿态变化被覆盖,再根据几何特征的几何性质,建立旋转矩阵与位移向量之间的数学关系,最后通过设计目标函数和应用优化算法进行目标位姿估计,通过粒子群优化算法及非线性优化算法,可以同时处理点、直线、圆三种基本几何特征,并且特征数量与相互间位置的关系没有限制,具有相当高的准确性、鲁棒性,进而解决了现有技术中存在的准确率较低、泛用性和鲁棒性差的问题。

2.本申请提出一种基于立体视觉和粒子群优化的目标六自由度位置与姿态(位姿)估计算法,该算法利用目标上的基本几何特征(圆、直线),根据基本几何特征的投影方程,精确计算目标位姿,并且不受几何特征数量与相互间位置关系的限制,实验表明本申请估计的位姿非常精确,为下游任务例

如装配、对接提供了可靠的位姿信息。

3.本申请通过粒子群优化算法及非线性优化算法，可以同时处理点、直线、圆三种基本几何特征，并且特征数量与相互间位置的关系没有限制，具有相当高的准确性、鲁棒性。

4.本申请提出一种基于立体视觉的粒子群优化算法，实现物体相对位姿的估计，首先通过罗德里格斯公式将旋转矩阵分解旋转向量和旋转角度，并利用汤姆逊方法，在单位球面对旋转向量进行均匀采样，旋转角度在到 π 之间均匀采样，再通过立体相机中基本几何特征本身的几何特性，建立旋转矩阵与位移向量之间的数学关系，然后结合粒子群优化算法和我们自己设计目标函数，可以估计出物体的姿态，最后，用莱文伯格-马夸特方法对粒子群优化算法估计的位姿进行优化，得到高精度、高鲁棒性的位姿估计。

应该理解的是，虽然附图的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头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，其可以以其他的顺序执行。而且，附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子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，其执行顺序也不必然是依次进行，而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

上述内容，仅为本申请的较佳示例性实施例，并非用于限制本申请的实施方案，本领域普通技术人员根据本申请的主要构思和精神，可以十分方便地进行相应的变通或修改，故本申请的保护范围应以权利要求书所要求的保护范围为准。

权 利 要 求 书

1、一种基于双目视觉与粒子群优化的目标六自由度位姿估计方法，其特征在于，所述方法包括：

获取目标的双目图像；所述双目图像中至少包括目标的旋转矩阵、位移向量和几何特征；

通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度，对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集；

根据所述几何特征的几何性质和所述数据点集，建立所述旋转矩阵与位移向量之间的数学关系；

通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果；所述目标函数用于估计目标的姿态；所述优化算法用于优化位姿估计结果；所述位姿估计结果用于展示目标的六自由度姿态以及确定目标的位置和朝向。

2、如权利要求1所述的方法，其特征在于，所述通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度，包括：

通过罗德里格斯公式将旋转矩阵分解为旋转向量和旋转角度。

3、如权利要求1所述的方法，其特征在于，所述对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集，包括：

通过汤姆逊方法在单位球面对旋转向量进行六自由度的均匀采样；

对0到 π 之间的旋转角度进行均匀采样，得到数据点集，确保所述数据点集中包括目标在六个自由度上的姿态变化；所述六个自由度包括三个平移自由度和三个旋转自由度。

4、如权利要求1所述的方法，其特征在于，所述根据所述几何特征的几何性质和所述数据点集，建立所述旋转矩阵与位移向量之间的数学关系，包括：

通过采样桑普森距离计算目标在所述双目图像上的投影点到数据点集的几何距离；所述几何距离为几何特征的几何性质；所述几何特征包括点、直线和面；

根据所述投影点到数据点集的几何距离建立所述旋转矩阵与位移向量之

间的数学关系。

5、如权利要求1所述的方法，其特征在于，所述优化算法包括粒子群优化算法；

所述通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果，包括：

基于所述双目图像在立体视觉中的几何关系和旋转姿态设计目标函数；所述目标函数包含立体视觉中的几何关系和旋转姿态的约束条件；

通过粒子群优化算法在所述旋转矩阵与位移向量之间的参数空间中寻找最优的旋转矩阵和位移向量，根据寻找到的最优的旋转矩阵和位移向量对所述目标函数进行最小化；

根据所述粒子群优化算法最小化后的目标函数，估计目标相对于目标坐标系的相对位姿，得到所述位姿估计结果；所述目标坐标系用于量化目标的位姿变化。

6、如权利要求5所述的方法，其特征在于，所述设计目标函数，包括：将所述目标的旋转矩阵参数化，得到目标的直线和圆的几何特征在所述双目图像中的几何距离；

根据所述几何距离得到对应的桑普森距离，根据所述桑普森距离构建目标函数。

7、如权利要求5所述的方法，其特征在于，所述得到所述位姿估计结果之后，所述方法还包括：

通过列文伯格马夸特算法对所述位姿估计结果中所述目标的相对位姿进行优化。

8、一种基于双目视觉与粒子群优化的目标六自由度位姿估计装置，其特征在于，所述装置包括：

数据获取模块，用于获取目标的双目图像；所述双目图像中至少包括目标的旋转矩阵、位移向量和几何特征；

均匀采样模块，用于通过旋转矩阵分解方法将所述旋转矩阵分解为旋转向量和旋转角度，对所述旋转向量和旋转角度进行六自由度的均匀采样，得到数据点集；

数学关系构建模块，用于根据所述几何特征的几何性质和所述数据点集，

建立所述旋转矩阵与位移向量之间的数学关系；

位姿估计模块，用于通过设计目标函数和应用优化算法，对所述目标的双目图像进行目标位姿估计，得到位姿估计结果。

9、一种电子设备，其特征在于，包括：至少一个处理器以及至少一个存储器，其中，

所述存储器上存储有程序指令或代码；

所述程序指令或代码被所述处理器加载并执行，使得电子设备实现如权利要求 1 至 7 中任一项所述的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

10、一种存储介质，其上存储有程序指令或代码，其特征在于，所述程序指令或代码被处理器加载并执行，以实现如权利要求 1 至 7 中任一项所述的基于双目视觉与粒子群优化的目标六自由度位姿估计方法。

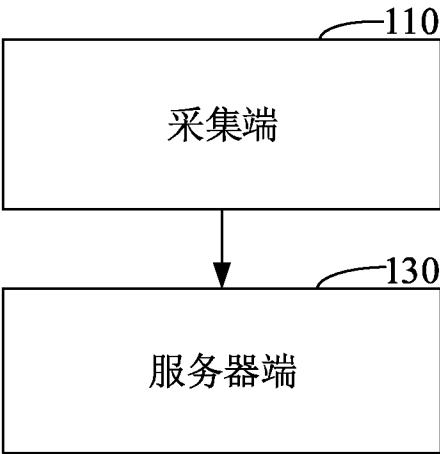


图 1

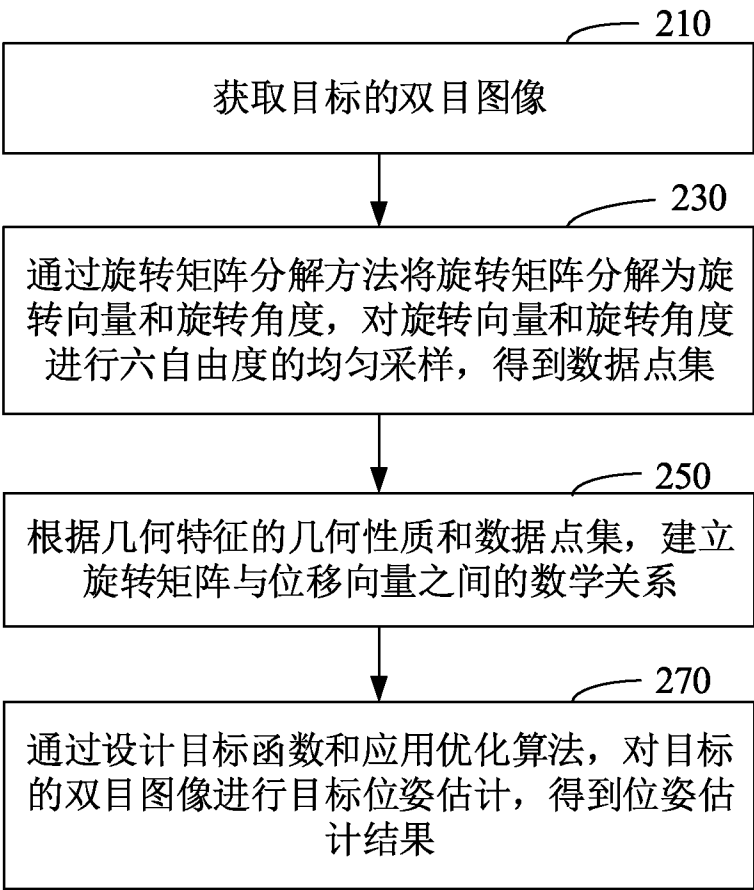


图 2

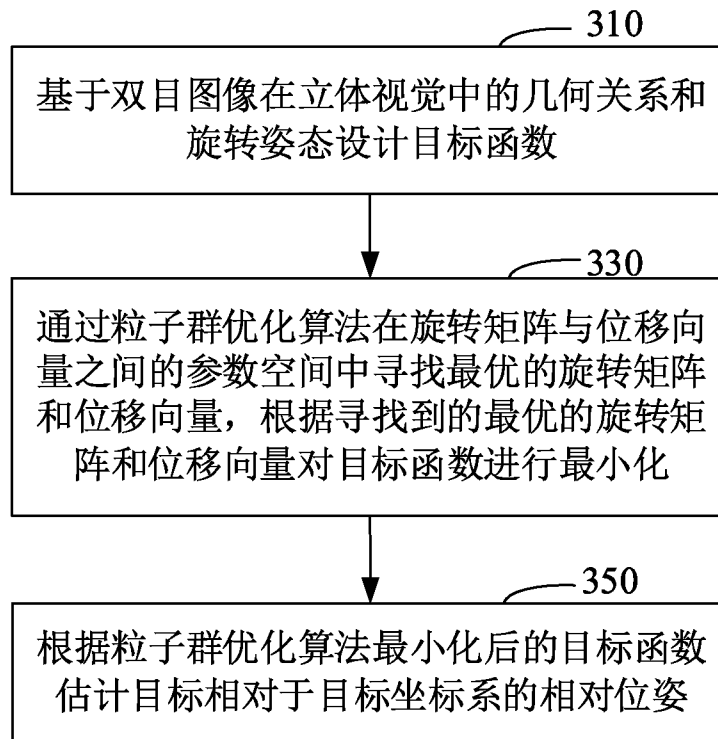


图 3

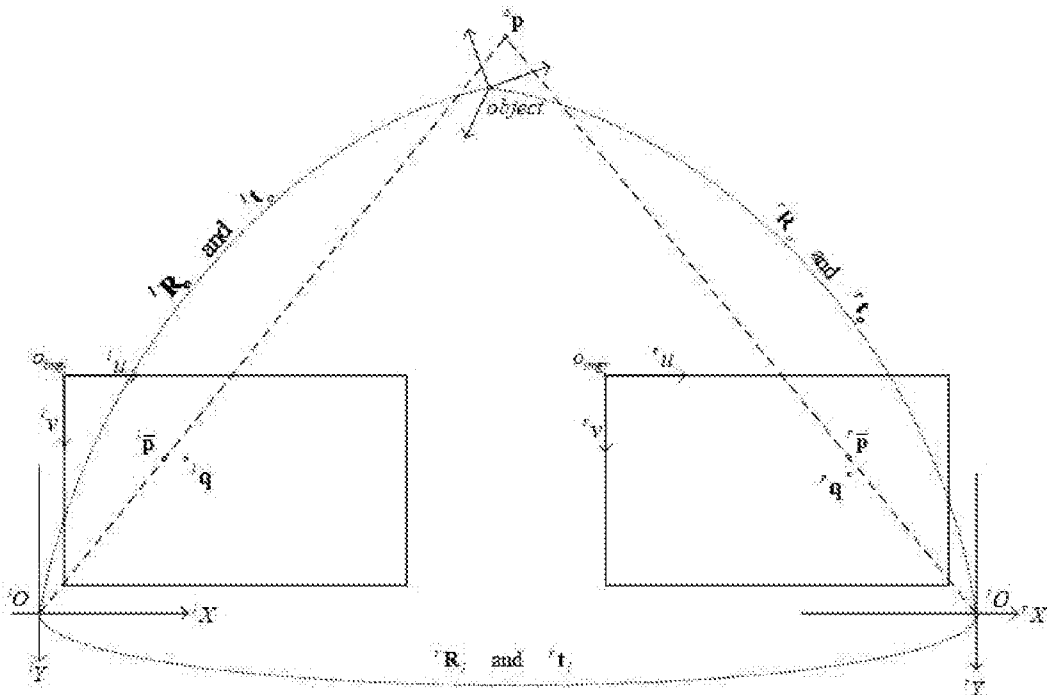


图 4

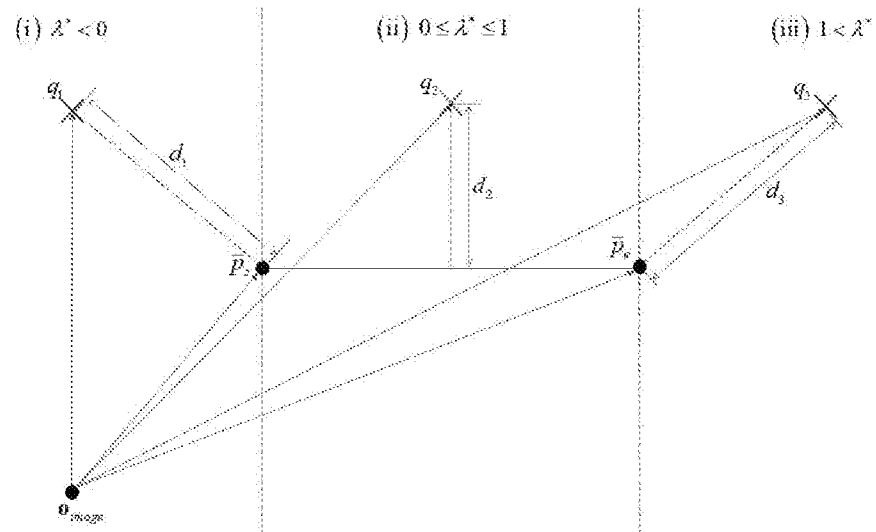


图 5

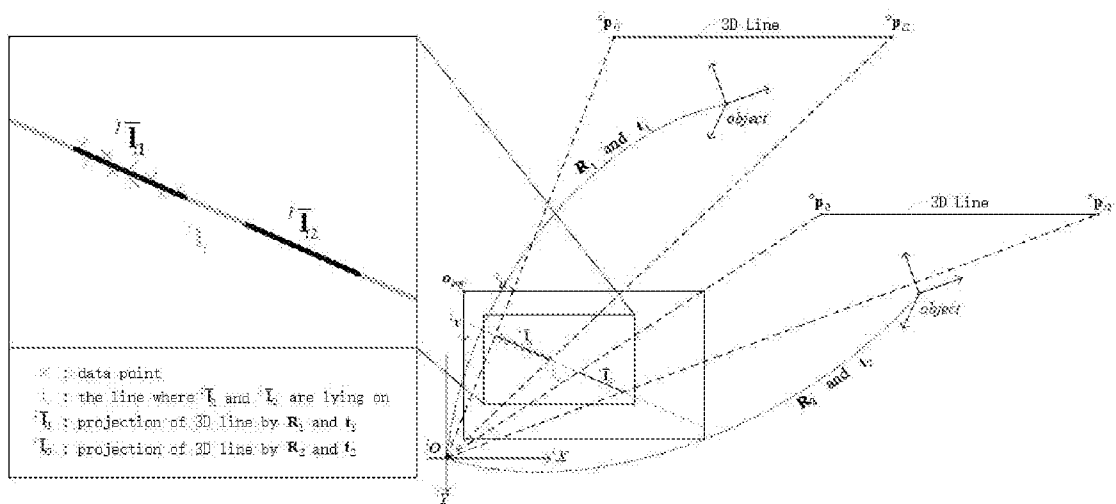


图 6

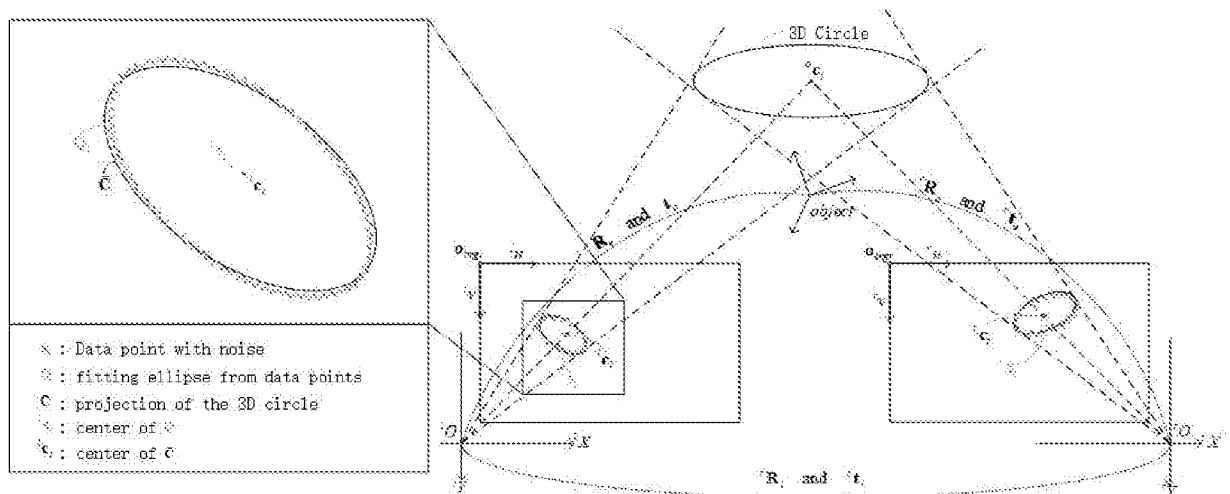


图 7

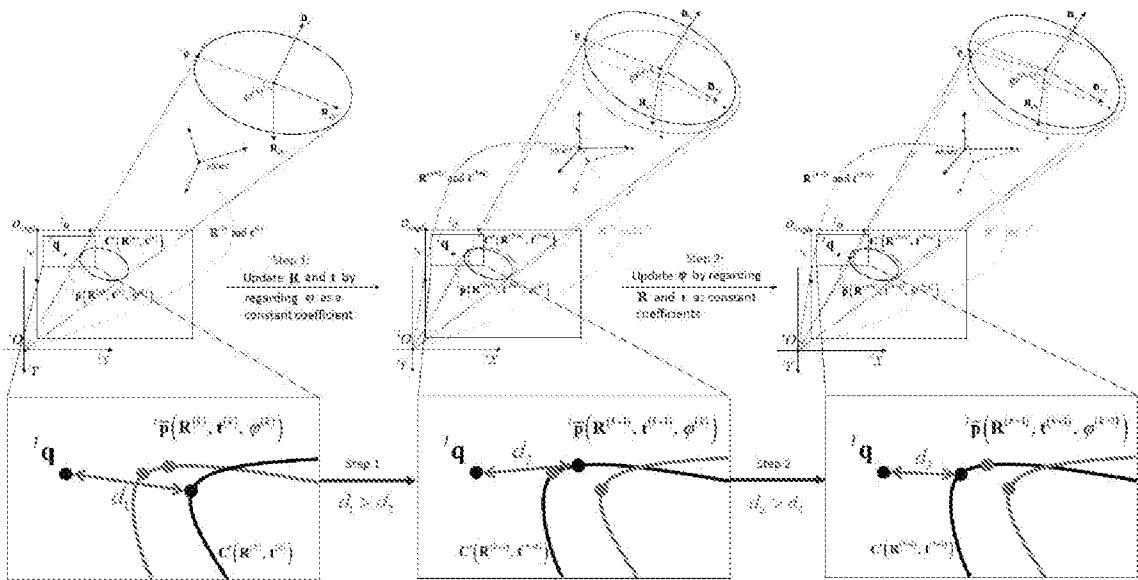


图 8

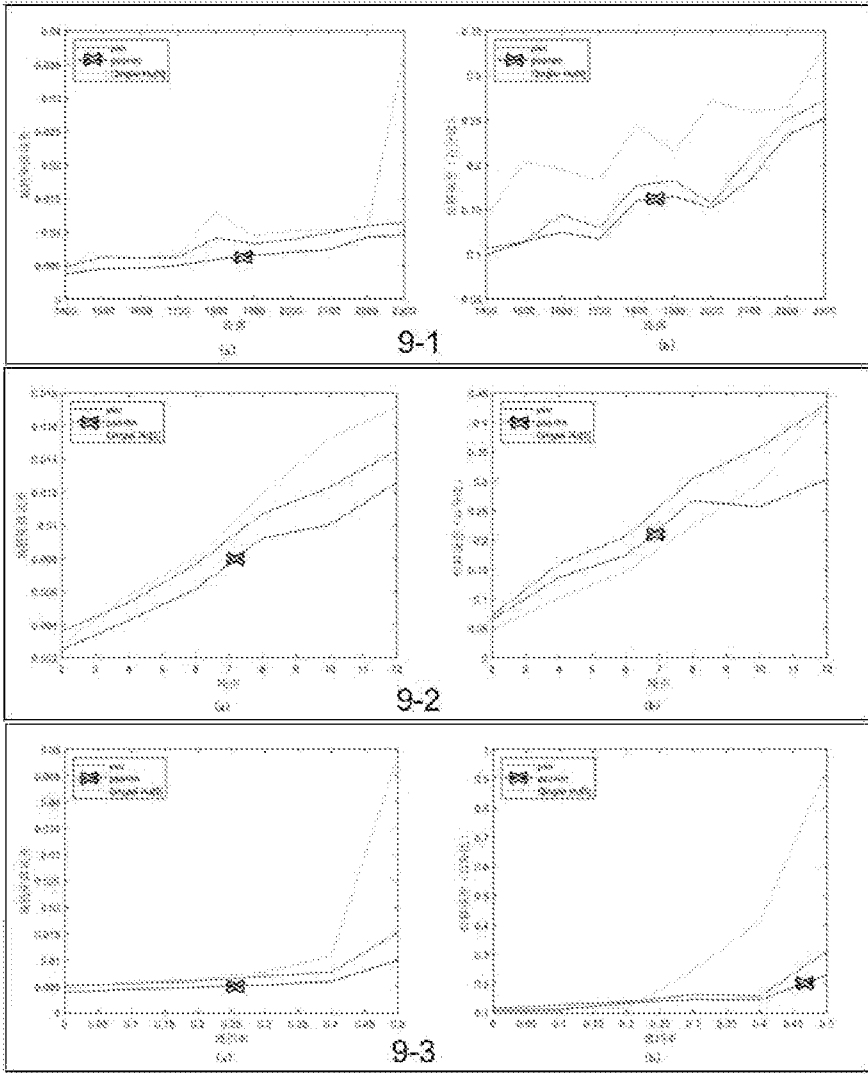


图 9

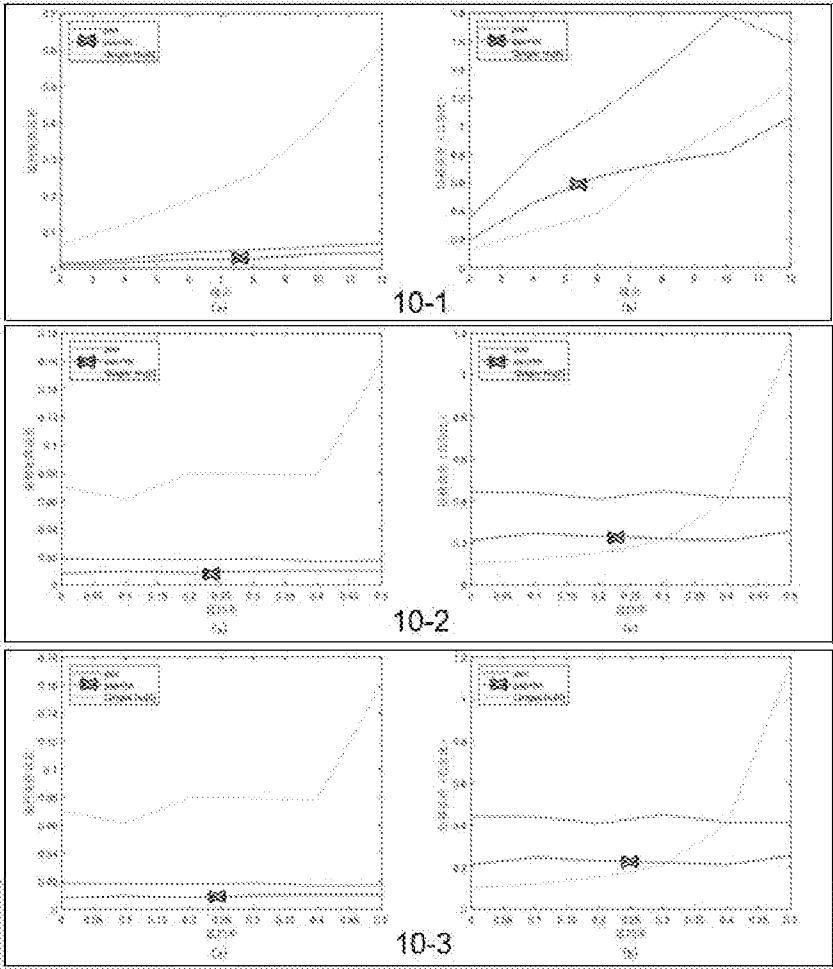


图 10

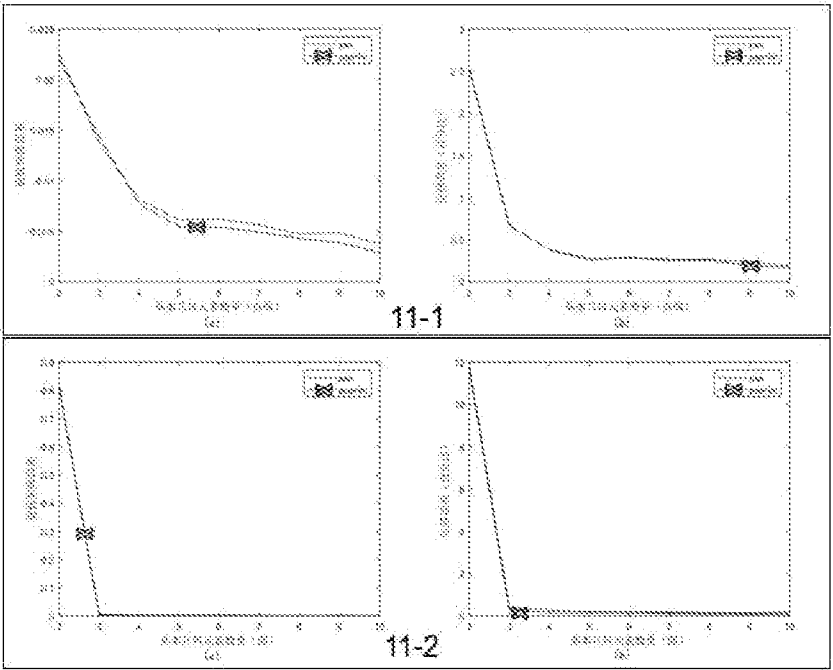


图 11

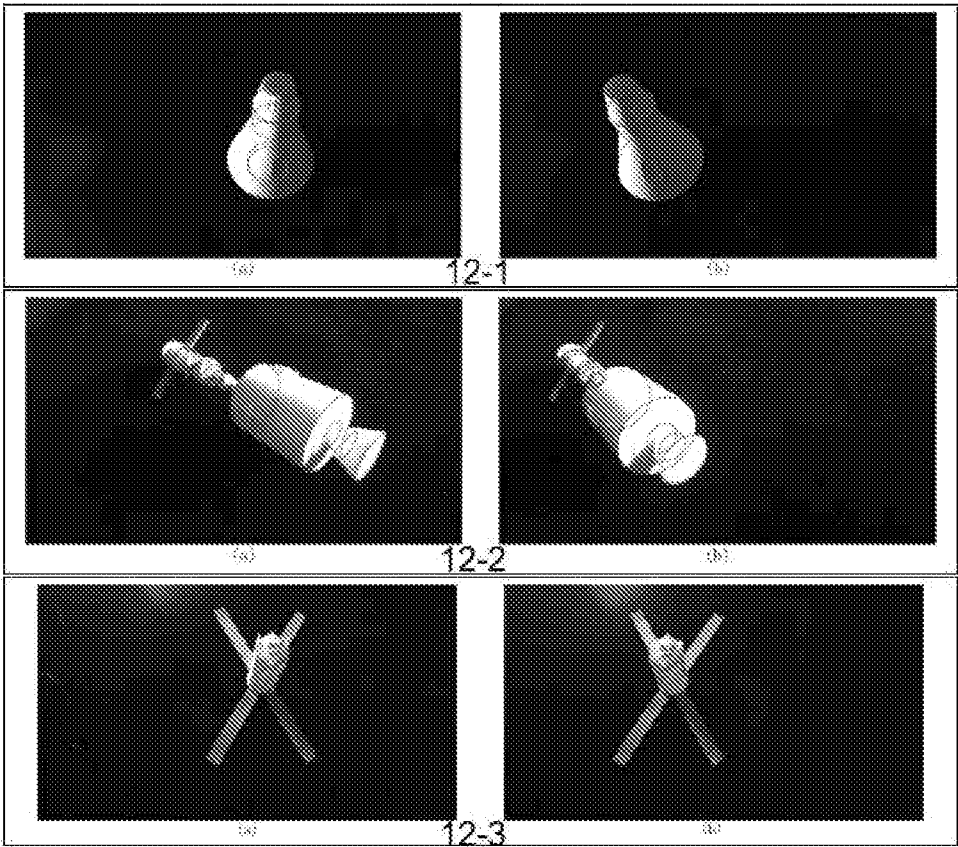


图 12

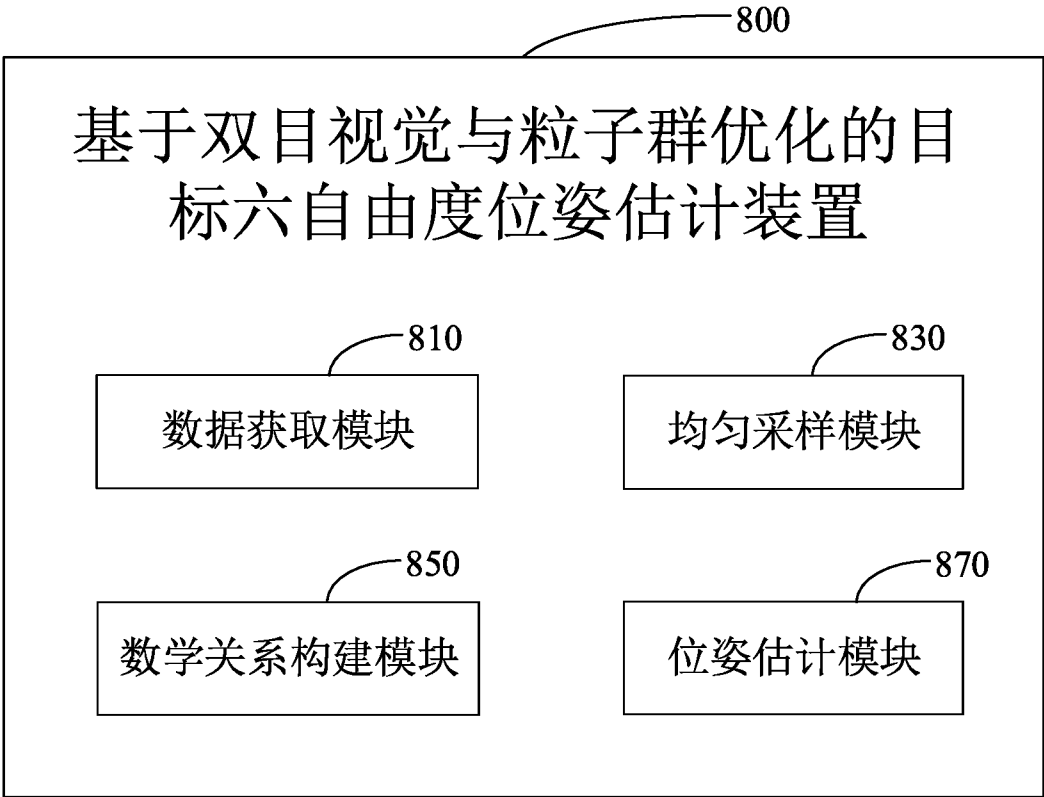


图 13

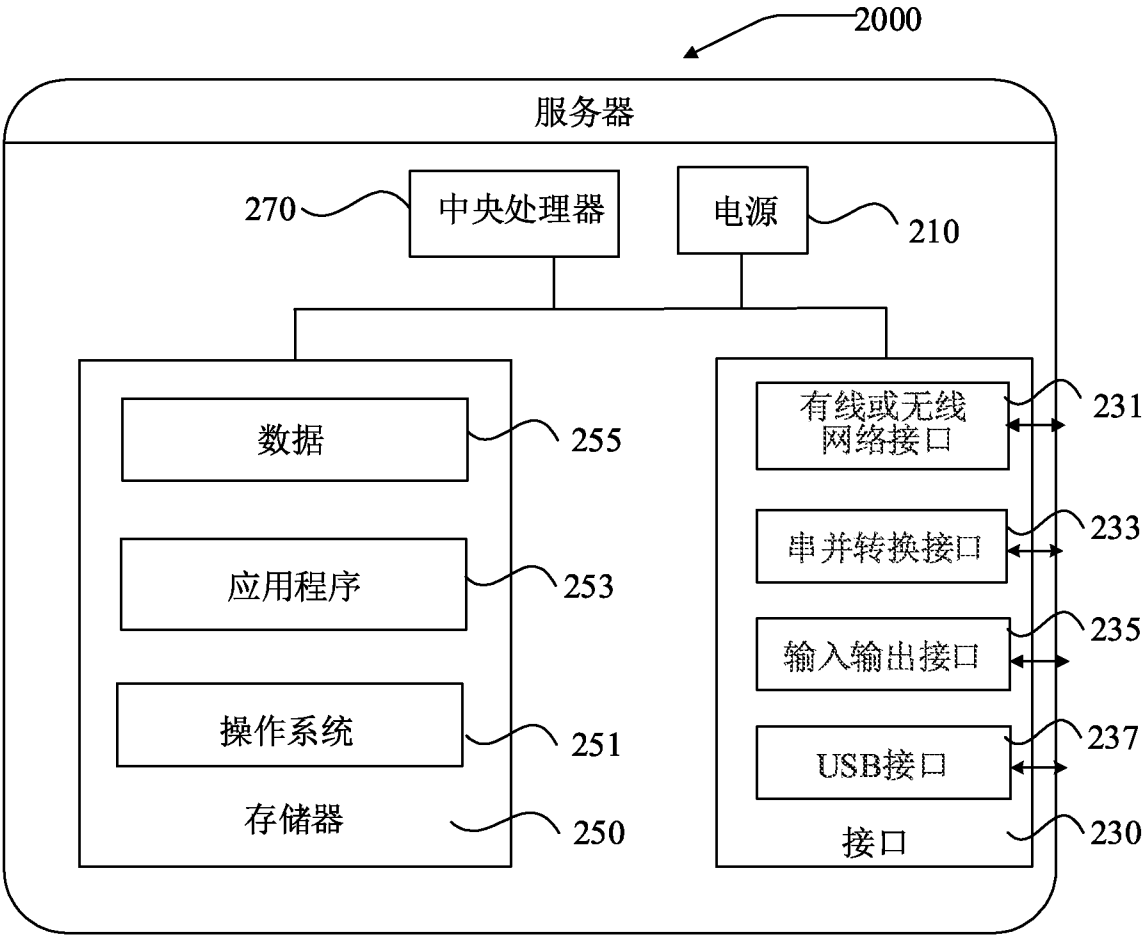


图 14

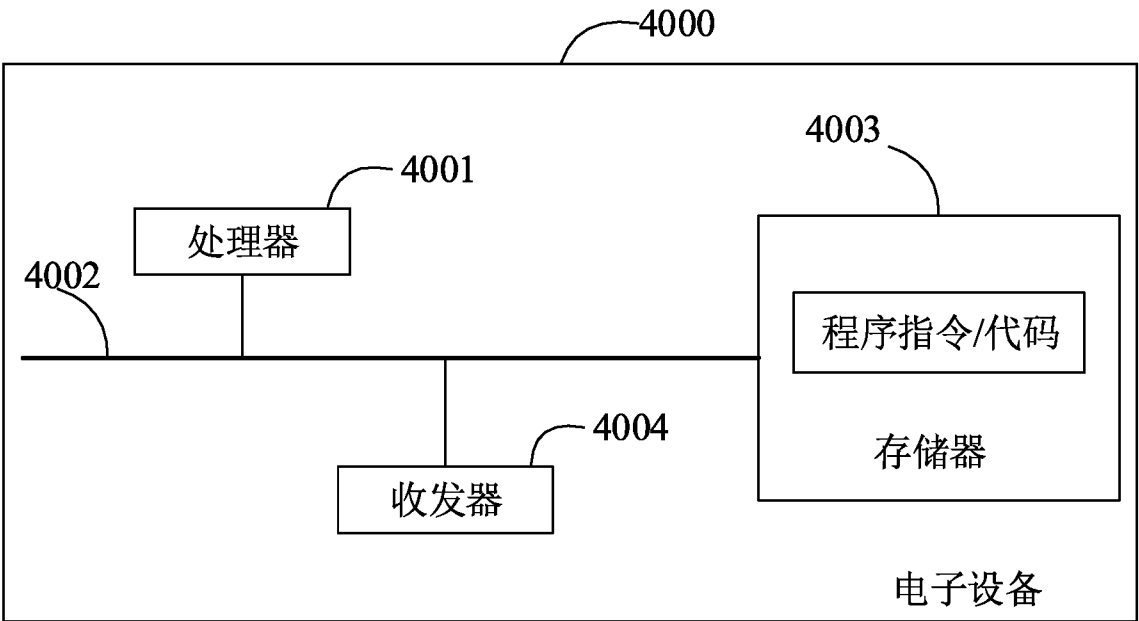


图 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/139298

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/73(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)

IPC: 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)

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 双目, 粒子群, 优化, 六自由度, 位姿, 姿态, 估计, 预测, 旋转, 矩阵, 位移, 几何, 向量, 角度, 点集, 点云, 目标函数, 位置, 朝向, 方向, 投影, binocular, particle swarm, optimization, six degrees of freedom, pose, estimate, prediction, rotation, matrix, displacement, geometry, vector, angle, point set, point cloud, objective function, position, orientation, direction, projection

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 108053443 A (TECHNOLOGY AND ENGINEERING CENTER FOR SPACE UTILIZATION, CHINESE ACADEMY OF SCIENCES) 18 May 2018 (2018-05-18) description, paragraphs [0056]-[0126]	1-10
Y	CN 115187781 A (BEIJING INFORMATION SCIENCE & TECHNOLOGY UNIVERSITY et al.) 14 October 2022 (2022-10-14) description, paragraphs [0051]-[0140]	1-10
Y	CN 116776719 A (HARBIN ENGINEERING UNIVERSITY) 19 September 2023 (2023-09-19) description, paragraphs [0100]-[0306]	1-10
A	CN 101419055 A (BEIHANG UNIVERSITY) 29 April 2009 (2009-04-29) entire document	1-10
A	CN 114952832 A (TSINGHUA UNIVERSITY) 30 August 2022 (2022-08-30) 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

09 August 2024

Date of mailing of the international search report

14 August 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/CN2023/139298

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2019204084 A1 (GOERTEK INC.) 04 July 2019 (2019-07-04) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/139298

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108053443	A	18 May 2018	None			
CN	115187781	A	14 October 2022	None			
CN	116776719	A	19 September 2023	None			
CN	101419055	A	29 April 2009	None			
CN	114952832	A	30 August 2022	None			
US	2019204084	A1	04 July 2019	WO	2019062291	A1	04 April 2019
				US	10989540	B2	27 April 2021

A. 主题的分类

G06T 7/73(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)

IPC: 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))

VEN、CNABS、CNTXT、WOTXT、EPTXT、USTXT、CNKI、IEEE: 双目, 粒子群, 优化, 六自由度, 位姿, 姿态, 估计, 预测, 旋转, 矩阵, 位移, 几何, 向量, 角度, 点集, 点云, 目标函数, 位置, 朝向, 方向, 投影, binocular, particle swarm, optimization, six degrees of freedom, pose, estimate, prediction, rotation, matrix, displacement, geometry, vector, angle, point set, point cloud, objective function, position, orientation, direction, projection

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	CN 108053443 A (中国科学院空间应用工程与技术中心) 2018年5月18日 (2018 - 05 - 18) 说明书第[0056]-[0126]段	1-10
Y	CN 115187781 A (北京信息科技大学 等) 2022年10月14日 (2022 - 10 - 14) 说明书第[0051]-[0140]段	1-10
Y	CN 116776719 A (哈尔滨工程大学) 2023年9月19日 (2023 - 09 - 19) 说明书第[0100]-[0306]段	1-10
A	CN 101419055 A (北京航空航天大学) 2009年4月29日 (2009 - 04 - 29) 全文	1-10
A	CN 114952832 A (清华大学) 2022年8月30日 (2022 - 08 - 30) 全文	1-10
A	US 2019204084 A1 (GOERTEK INC.) 2019年7月4日 (2019 - 07 - 04) 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8月9日

国际检索报告邮寄日期

2024年8月14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

徐淑娴

电话号码 (+86) 010-53961535

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/CN2023/139298

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	108053443	A	2018年5月18日	无			
CN	115187781	A	2022年10月14日	无			
CN	116776719	A	2023年9月19日	无			
CN	101419055	A	2009年4月29日	无			
CN	114952832	A	2022年8月30日	无			
US	2019204084	A1	2019年7月4日	WO	2019062291	A1	2019年4月4日
				US	10989540	B2	2021年4月27日