



(12) 发明专利申请

(10) 申请公布号 CN 118784816 A

(43) 申请公布日 2024. 10. 15

(21) 申请号 202411202744.1

H04N 13/128 (2018.01)

(22) 申请日 2024.08.29

(71) 申请人 深圳市金沃德科技有限公司

地址 518100 广东省深圳市宝安区新安街
道兴东社区71区新政公司厂房C栋
601、301

(72) 发明人 魏伟坚

(74) 专利代理机构 广州品诚智产知识产权代理
事务所(普通合伙) 441215

专利代理师 赵增侠

(51) Int. Cl.

H04N 13/275 (2018.01)

H04N 13/271 (2018.01)

H04N 13/243 (2018.01)

H04N 13/167 (2018.01)

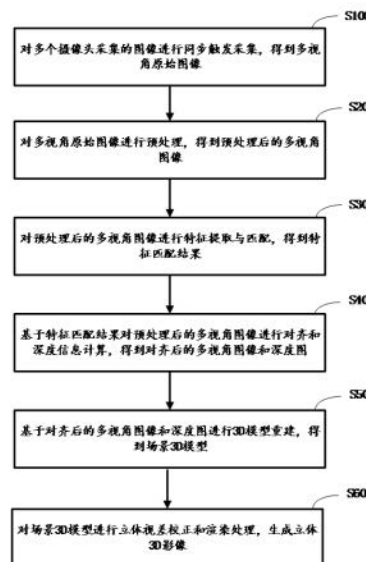
权利要求书4页 说明书18页 附图2页

(54) 发明名称

基于多摄像头拍摄的3D影像生成方法及系
统

(57) 摘要

本申请涉及3D影像生成技术领域,公开了一种基于多摄像头拍摄的3D影像生成方法及系统,该方法包括:对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;对所述多视角原始图像进行预处理,得到预处理后的多视角图像;对所述多视角原始图像进行特征提取与匹配,得到特征匹配结果;基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型;对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像,进而实现了高精度、高细节的3D模型重建,生成了高质量、高沉浸感的立体3D影像。



1. 一种基于多摄像头拍摄的3D影像生成方法,其特征在于,所述方法包括:
 - 对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;
 - 对所述多视角原始图像进行预处理,得到预处理后的多视角图像;
 - 对所述预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;
 - 基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;
 - 基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型;
 - 对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。
2. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像,包括:
 - 对所述多个摄像头进行阵列布置,得到摄像头阵列,并对所述摄像头阵列中的每个摄像头配置WiFi模块,得到带WiFi模块的摄像头阵列;
 - 设置中央控制单元,通过WiFi网络与所述带WiFi模块的摄像头阵列建立连接,构建带中央控制单元的摄像头系统;
 - 基于所述中央控制单元生成微秒级触发信号,得到微秒级同步触发信号,并通过WiFi网络将所述微秒级同步触发信号传输至各摄像头;
 - 基于所述传输后的微秒级同步触发信号对所述带中央控制单元的摄像头系统进行触发,得到初步同步的多视角图像;
 - 在拍摄场景中设置高速LED阵列,并基于所述微秒级同步触发信号控制所述高速LED阵列发射特定模式的光脉冲,得到带时间标记的场景;
 - 对所述带时间标记的场景进行拍摄,得到带光脉冲模式的多视角图像,并通过WiFi网络将所述带光脉冲模式的多视角图像传输至云端APP;
 - 在云端APP中基于所述带光脉冲模式的多视角图像中的光脉冲模式进行亚微秒级时间对齐校正,得到多视角原始图像。
3. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述对所述多视角原始图像进行预处理,得到预处理后的多视角图像,包括:
 - 对所述多视角原始图像进行去噪,得到初步降噪的多视角图像,并在拍摄场景中放置标准色卡,基于所述标准色卡对所述初步降噪的多视角图像进行分析,得到全局色彩映射模型;
 - 基于所述全局色彩映射模型对所述初步降噪的多视角图像进行联合色彩校正,得到色彩校正后的多视角图像,并对所述色彩校正后的多视角图像进行直方图分析,得到每张图像的像素值分布特征;
 - 基于所述像素值分布特征构建自适应曝光均衡函数,并基于所述自适应曝光均衡函数对所述色彩校正后的多视角图像进行像素值调整,得到曝光均衡后的多视角图像;
 - 对所述曝光均衡后的多视角图像进行多项式模型的畸变校正,得到畸变校正后的多视角图像,并对所述畸变校正后的多视角图像进行超分辨率处理,得到高分辨率多视角图像;
 - 对所述高分辨率多视角图像进行边缘锐化和细节增强处理,得到预处理后的多视角图像。
4. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述对所

述预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果,包括:

对所述预处理后的多视角图像进行特征点提取,得到初步特征点集,并基于所述初步特征点集构建仿射不变特征集;

对所述仿射不变特征集进行全局语义特征提取,得到语义增强特征集,并基于所述语义增强特征集构建特征快速索引树;

基于所述特征快速索引树对多视角图像间的特征点进行粗匹配,得到初步匹配对集合,并基于所述初步匹配对集合构建特征点局部关系图;

根据所述特征点局部关系图生成优化后的匹配对集合,并基于所述优化后的匹配对集合建立全局几何一致性模型;

基于所述全局几何一致性模型对匹配结果进行细化,得到精确匹配对集合,并基于所述精确匹配对集合进行多视图特征融合和传播,得到特征匹配结果。

5. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图,包括:

基于所述特征匹配结果构建多视图几何关系图,得到初始视图拓扑结构,并对所述初始视图拓扑结构进行优化,得到优化后的视图拓扑结构;

基于所述优化后的视图拓扑结构建立分层稀疏束调整模型并生成初始相机位姿参数,对所述初始相机位姿参数进行迭代优化,最小化重投影误差,得到目标相机位姿参数;

基于所述目标相机位姿参数对所述预处理后的多视角图像进行几何变换,得到粗对齐的多视角图像,并对所述粗对齐的多视角图像进行多视角立体匹配代价体分析,得到初始深度估计;

基于所述初始深度估计进行多视图一致性传播,得到优化后的深度图,并对所述优化后的深度图进行细化处理,得到高精度深度图;

基于所述高精度深度图对所述粗对齐的多视角图像进行精细对齐,得到精确对齐的多视角图像,并对所述精确对齐的多视角图像和所述高精度深度图进行时空一致性分析,识别并处理动态物体,得到对齐后的多视角图像和深度图。

6. 根据权利要求5所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述基于所述目标相机位姿参数对所述预处理后的多视角图像进行几何变换,得到粗对齐的多视角图像,并对所述粗对齐的多视角图像进行多视角立体匹配代价体分析,得到初始深度估计,包括:

基于所述目标相机位姿参数计算摄像头间的本质矩阵 E ,其中 $E=[t] \times R$, $[t] \times$ 为平移向量 t 的反对称矩阵, R 为旋转矩阵;对所述预处理后的多视角图像执行极线校正,得到校正后的图像对;

对所述校正后的图像对执行相位一致性变换,计算相位相关函数,得到亚像素级位移场,并基于所述亚像素级位移场,计算仿射变换矩阵,得到粗对齐的多视角图像;

对所述粗对齐的多视角图像进行各向异性扩散滤波,得到边缘保持的平滑图像集,并对所述平滑图像集执行多尺度Log-Gabor滤波,得到特征张量场;

基于所述特征张量场计算结构张量,对所述结构张量进行特征值分解,得到结构一致性映射,并基于所述结构一致性映射对所述初步对齐的多视角图像集进行非刚性配准,得

到精细对齐的多视角图像集；

对所述精细对齐的多视角图像集构建四元体分割,得到超像素集,基于所述超像素集计算多视图匹配代价体,得到初始深度图,并对所述初始深度图应用变分法求解,通过迭代优化得到初始深度估计。

7. 根据权利要求5所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述基于所述高精度深度图对所述粗对齐的多视角图像进行精细对齐,得到精确对齐的多视角图像,并对所述精确对齐的多视角图像和所述高精度深度图进行时空一致性分析,识别并处理动态物体,得到对齐后的多视角图像和深度图,包括:

对所述高精度深度图和所述粗对齐的多视角图像执行双边联合上采样,得到高分辨率深度图;

基于所述高分辨率深度图计算视差梯度场,并基于所述视差梯度场构建仿射变换矩阵,对所述粗对齐的多视角图像进行局部变形,得到精细对齐的多视角图像;

对所述精细对齐的多视角图像进行结构张量分析,得到边缘保持的平滑图像,并基于所述边缘保持的平滑图像计算多尺度局部相位一致性度量,得到局部结构相似性映射;

基于所述局部结构相似性映射构建光度一致性约束项,并结合几何平滑约束项,求解位移场,得到精确对齐的多视角图像;

对所述精确对齐的多视角图像和所述高分辨率深度图执行时空体素化,构建4D时空体素网格,并计算每个体素的颜色和深度特征向量,得到特征化时空体素场;

基于所述特征化时空体素场构建时空图,得到时空分割结果;

对所述时空分割结果进行连通域分析,计算每个连通域的时空一致性得分,根据所述时空一致性得分对连通域进行聚类,得到动态物体候选区域集合R;

基于所述动态物体候选区域集合,构建光流约束和深度一致性约束,联合优化,得到精细运动场;

根据所述精细运动场对所述精确对齐的多视角图像和所述高分辨率深度图进行时序插值和外推,得到对齐后的多视角图像和深度图。

8. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型,包括:

对所述对齐后的多视角图像和所述深度图进行体素化处理,得到初始低分辨率体素模型,并基于所述初始低分辨率体素模型进行八叉树数据结构分析,得到多分辨率体素表示;

对所述多分辨率体素表示中的每个体素进行带符号距离场值计算,得到初始SDF模型,并基于所述初始SDF模型进行多视角深度融合,得到优化后的SDF模型;

对所述优化后的SDF模型进行自适应细分,得到高精度表面网格模型,并基于所述高精度表面网格模型进行多视图一致性约束,得到表面优化目标函数;

对所述表面优化目标函数进行迭代求解,得到精细化表面模型,并基于所述精细化表面模型提取多视角纹理集合;

对所述多视角纹理集合进行图割的全局优化,得到无缝拼接的纹理映射,并基于所述无缝拼接的纹理映射生成场景3D模型。

9. 根据权利要求1所述的基于多摄像头拍摄的3D影像生成方法,其特征在于,所述对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像,包括:

基于人眼立体视觉感知模型对所述场景3D模型进行深度分析,得到场景深度分布特征,并根据所述场景深度分布特征构建自适应立体视差调整模型;

基于所述自适应立体视差调整模型对所述场景3D模型进行视差调整,得到初步立体视差校正结果,并基于所述初步立体视差校正结果构建视差连续性优化模型;

通过所述视差连续性优化模型对大视差区域进行处理,得到优化后的立体视差校正结果,并基于所述优化后的立体视差校正结果生成稠密光场采样数据;

对所述稠密光场采样数据进行压缩,得到轻量化光场模型,并基于所述轻量化光场模型构建初始渲染结果;

对所述初始渲染结果进行图像增强,得到高质量渲染图像,并基于所述高质量渲染图像进行时空分析和抖动抑制处理,得到立体3D影像。

10. 一种基于多摄像头拍摄的3D影像生成装置,其特征在于,用于执行如权利要求1-9中任一项所述的基于多摄像头拍摄的3D影像生成方法,所述装置包括:

采集模块,用于对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;

预处理模块,用于对所述多视角原始图像进行预处理,得到预处理后的多视角图像;

匹配模块,用于对所述预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;

计算模块,用于基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;

重建模块,用于基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型;

生成模块,用于对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。

基于多摄像头拍摄的3D影像生成方法及系统

技术领域

[0001] 本申请涉及3D影像生成技术领域,尤其涉及一种基于多摄像头拍摄的3D影像生成方法及系统。

背景技术

[0002] 随着虚拟现实、增强现实和混合现实技术的快速发展,对高质量、高真实感的3D内容需求日益增长。传统的3D内容制作方法往往依赖于复杂的建模和渲染过程,耗时耗力且难以实现真实场景的精确重建。基于多摄像头拍摄的3D影像生成方法应运而生,旨在通过多个摄像头同时捕捉真实场景,并通过计算机视觉和图像处理技术重建高质量的3D模型。然而,这种方法在实际应用中仍面临诸多挑战。

[0003] 现有技术中,多摄像头系统的同步问题一直是一个技术难点。不同摄像头之间的时间差异会导致捕获的图像存在时间不一致性,影响后续的3D重建效果。其次,由于拍摄环境的复杂性,如光照变化、遮挡等因素,获取的多视角图像往往存在噪声、色差和畸变等问题,这些都会对3D重建的精度造成不利影响。此外,在特征提取与匹配、图像对齐、深度信息计算等关键环节中,如何在保证精度的同时提高计算效率也是一个重要问题。特别是在处理动态场景时,如何准确识别和处理运动物体,保证重建结果的时空一致性,仍然是该领域面临的重大挑战。如何生成高质量的立体3D影像,使其能够在各种显示设备上呈现出良好的立体感和沉浸感,也是需要解决的关键问题。

发明内容

[0004] 本申请提供了一种基于多摄像头拍摄的3D影像生成方法及系统,进而实现了高精度、高细节的3D模型重建,生成了高质量、高沉浸感的立体3D影像。

[0005] 本申请第一方面提供了一种基于多摄像头拍摄的3D影像生成方法,所述基于多摄像头拍摄的3D影像生成方法包括:

对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;

对所述多视角原始图像进行预处理,得到预处理后的多视角图像;

对所述预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;

基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;

基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型;

对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。

[0006] 本申请第二方面提供了一种基于多摄像头拍摄的3D影像生成装置,所述基于多摄像头拍摄的3D影像生成装置包括:

采集模块,用于对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;

预处理模块,用于对所述多视角原始图像进行预处理,得到预处理后的多视角图

像;

匹配模块,用于对所述预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;

计算模块,用于基于所述特征匹配结果对所述预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;

重建模块,用于基于所述对齐后的多视角图像和所述深度图进行3D模型重建,得到场景3D模型;

生成模块,用于对所述场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。

[0007] 本申请第三方面提供了一种电子设备,包括:存储器和至少一个处理器,所述存储器中存储有指令;所述至少一个处理器调用所述存储器中的所述指令,以使得所述电子设备执行上述的基于多摄像头拍摄的3D影像生成方法。

[0008] 本申请的第四方面提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行上述的基于多摄像头拍摄的3D影像生成方法。

[0009] 与现有技术相比,本申请具有以下有益效果:通过WiFi模块进行时钟同步和中央控制单元生成微秒级触发信号,结合高速LED阵列发射特定模式的光脉冲,实现了多摄像头的亚微秒级时间对齐校正,大大提高了多视角图像的时间一致性。采用全局色彩映射模型、自适应曝光均衡函数、多项式模型的畸变校正等技术,有效解决了多视角图像的噪声、色差、曝光不均和畸变等问题,通过构建仿射不变特征集、全局语义特征提取和特征快速索引树,实现了多视角图像间的高效精确匹配,利用分层稀疏束调整模型、多视角立体匹配代价体分析等技术,实现了多视角图像的精确对齐和高精度深度图的生成。通过时空体素化、时空一致性分析和精细运动场优化,有效识别和处理了动态物体,保证了重建结果的时空一致性。采用多分辨率体素表示、带符号距离场优化和自适应表面细分等技术,实现了高精度、高细节的3D模型重建。基于人眼立体视觉感知模型和自适应立体视差调整,结合稠密光场采样和渲染增强技术,生成了高质量、高沉浸感的立体3D影像。

附图说明

[0010] 为了更清楚地说明本发明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简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,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它的附图。

[0011] 本说明书附图所绘示的结构、比例、大小等,均仅用以配合说明书所揭示的内容,以供熟悉此技术的人士了解与阅读,并非用以限定本发明可实施的限定条件,故不具技术上的实质意义,任何结构的修饰、比例关系的改变或大小的调整,在不影响本发明所能产生的功效及所能达成的目的下,均应仍落在本发明所揭示的技术内容得能涵盖的范围内。

[0012] 图1是本发明实施例提供的基于多摄像头拍摄的3D影像生成方法的流程示意图;
图2是本发明实施例提供的基于多摄像头拍摄的3D影像生成装置的结构示意性框图。

具体实施方式

[0013] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0014] 附图中所示的流程图仅是示例说明,不是必须包括所有的内容和操作/步骤,也不是必须按所描述的顺序执行。例如,有的操作/步骤还可以分解、组合或部分合并,因此实际执行的顺序有可能根据实际情况改变。

[0015] 还应当理解,在此本申请说明书中所使用的术语仅仅是出于描述特定实施例的目的而并不意在限制本申请。如在本申请说明书和所附权利要求书中所使用的那样,除非上下文清楚地指明其它情况,否则单数形式的“一”、“一个”及“该”意在包括复数形式。

[0016] 还应当进一步理解,在本申请说明书和所附权利要求书中使用的术语“和/或”是指相关联列出的项中的一个或多个的任何组合以及所有可能组合,并且包括这些组合。请参阅图1,本申请实施例中基于多摄像头拍摄的3D影像生成方法的一个实施例包括:

步骤100、对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;

可以理解的是,本申请的执行主体可以为基于多摄像头拍摄的3D影像生成装置,还可以是终端或者服务器,具体此处不做限定。本申请实施例以服务器为执行主体为例进行说明。

[0017] 具体的,根据场景的需求和视角覆盖范围,将摄像头按照阵列进行布置,最大限度地覆盖场景的不同视角,确保从多个角度捕捉到足够的信息。在摄像头阵列中,每个摄像头都配置WiFi模块,从而通过WiFi网络与中央控制单元进行通信,形成一个带WiFi模块的摄像头阵列。设置中央控制单元,通过WiFi网络与带WiFi模块的摄像头阵列建立连接。中央控制单元的作用是生成和分发高精度的同步触发信号,以确保所有摄像头在同一时间点同时启动拍摄。为了实现微秒级别的时间同步,中央控制单元需要生成微秒级触发信号。微秒级的同步触发信号能够保证摄像头之间的拍摄时间差小于1微秒,避免图像中由于时间差导致的运动模糊或失真。这种触发信号通过WiFi网络传递给各个摄像头。传输后的微秒级同步触发信号能够确保每个摄像头在接收到信号的瞬间开始曝光和拍摄,由此得到初步同步的多视角图像。为了提高时间同步的精度,并为后续的图像处理提供可靠的时间基准,在拍摄场景中设置高速LED阵列。高速LED阵列能够通过发射光脉冲,为场景提供明确的时间标记。这些光脉冲由中央控制单元基于微秒级同步触发信号来控制,确保光脉冲与摄像头的拍摄同步。在摄像头接收到同步触发信号并开始拍摄后,带有时间标记的场景被记录下来,形成带有光脉冲模式的多视角图像。这些图像通过WiFi网络传输至云端APP进行进一步处理。在云端APP中,利用图像中的光脉冲标记对图像进行亚微秒级的时间对齐校正。时间对齐校正通过分析图像中光脉冲的出现时间,确保各视角图像在时间轴上的精确同步,消除由于摄像头硬件或信号传输导致的微小时间偏差,最终得到高精度同步的多视角原始图像。

[0018] 步骤200、对多视角原始图像进行预处理,得到预处理后的多视角图像;

具体的,对多视角原始图像进行去噪处理,以减少图像中的噪声干扰,得到初步降噪的多视角图像。为了提高图像的一致性和色彩准确性,在拍摄场景中放置标准色卡,通过

标准色卡的已知色彩值,对初步降噪的多视角图像进行分析,以建立全局色彩映射模型。该模型用于校正多视角图像之间的色差,确保不同摄像头所捕捉到的图像在色彩上具有一致性。基于全局色彩映射模型,对初步降噪的多视角图像进行联合色彩校正,得到色彩校正后的多视角图像。为了改善图像质量,对色彩校正后的多视角图像进行直方图分析,分析每张图像的像素值分布特征,获得关于亮度和对比度的详细信息。基于像素值分布特征,构建一个自适应曝光均衡函数,该函数能够针对每张图像的具体情况,调整其曝光和亮度,使得最终的图像更具均衡性。基于自适应曝光均衡函数对色彩校正后的多视角图像进行像素值调整,得到曝光均衡后的多视角图像,使图像的整体亮度更加均匀,避免因曝光不均导致的图像质量问题。为了消除由于摄像头镜头特性造成的图像畸变,对曝光均衡后的多视角图像进行多项式模型的畸变校正。通过多项式模型可以精确地校正镜头引入的几何变形,得到畸变校正后的多视角图像,使得所有视角图像的几何结构一致。对畸变校正后的多视角图像进行超分辨率处理,得到高分辨率的多视角图像。超分辨率处理技术能够通过插值和细节补偿,提高图像的分辨率,使得最终生成的3D影像更加清晰和细致。为了确保图像边缘的清晰度和整体视觉效果,对高分辨率多视角图像进行边缘锐化和细节增强处理,增强图像中的关键细节,最终得到预处理后的多视角图像。

[0019] 步骤300、对预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;

需要说明的是,对预处理后的多视角图像进行特征点提取,通过算法从每一张图像中识别出具备独特性质的初步特征点集,这些特征点通常代表图像中具有高度辨识度的关键部分。基于初步特征点集,构建仿射不变特征集,利用图像处理技术确保这些特征集在仿射变换下保持不变性,提高不同视角下特征匹配的鲁棒性。对仿射不变特征集进行全局语义特征提取,提取过程中考虑图像中物体的语义信息,使得这些特征点不仅仅依赖于几何形状,同时具备一定的语义理解能力,形成语义增强特征集。通过融入语义信息,能够有效提升匹配的准确性。基于语义增强特征集构建特征快速索引树,该索引树结构能够加速特征点之间的匹配过程,减少计算时间。在特征快速索引树的基础上,进行多视角图像间的特征点粗匹配,得到初步匹配对集合。粗匹配过程通过快速索引树对特征点进行有效筛选,确定可能对应的特征对。基于初步匹配对集合构建特征点局部关系图,利用图论方法对特征点之间的关系进行建模,分析它们之间的局部几何关系,提高匹配的可靠性。根据特征点局部关系图生成优化后的匹配对集合,通过优化算法剔除错误匹配和增加新的正确匹配对,形成更准确的特征点匹配集。基于该优化后的匹配对集合建立全局几何一致性模型,通过分析整个图像集合的几何关系,确保特征点匹配在全局尺度上保持一致性。基于全局几何一致性模型对匹配结果进行细化处理,得到精确匹配对集合。通过对初步匹配对的微调和调整,确保匹配的每一个特征点对在全局和局部层面都是一致的。基于精确匹配对集合,进行多视图特征的融合和传播,将不同视角下的特征信息综合考虑,得到最终的特征匹配结果。

[0020] 步骤400、基于特征匹配结果对预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;

具体的,基于特征匹配结果构建多视图几何关系图,得到初始视图的拓扑结构。拓扑结构表示不同视角图像之间的几何关系,通过分析这些关系,初步了解摄像头在拍摄时的相对位置和姿态。对初始视图拓扑结构进行优化,通过消除冗余的视角关系或增强关键

视角之间的连接,得到优化后的视图拓扑结构。基于优化后的视图拓扑结构建立分层稀疏束调整模型,该模型可以有效地处理多个视角之间的复杂几何关系,并生成初始相机的位姿参数,即每个摄像头在拍摄时的空间位置和朝向。对初始相机位姿参数进行迭代优化,通过最小化重投影误差,逐步调整相机位姿,使得所有视角图像在同一三维空间中达到最佳对齐,得到目标相机位姿参数。基于目标相机位姿参数,对预处理后的多视角图像进行几何变换,将图像中的物体在空间中进行调整,以便它们在所有视角下尽可能一致,得到粗对齐的多视角图像。对粗对齐的多视角图像进行多视角立体匹配代价体分析,通过计算视差信息,初步估计出各个视角图像中对应像素点的深度,得到初始深度估计。基于初始深度估计进行多视图一致性传播,通过综合考虑多个视角之间的深度信息,传播和调整这些深度值,获得优化后的深度图。对优化后的深度图进行细化处理,消除噪声和误差,得到高精度深度图。基于高精度深度图对粗对齐的多视角图像进行精细对齐,调整图像,使得它们在所有视角下的细节都完全对齐,得到精确对齐的多视角图像。对精确对齐的多视角图像和高精度深度图进行时空一致性分析,识别并处理场景中的动态物体,这些物体可能在不同视角下有不同的位置和形态,通过专门的处理,确保它们在最终对齐结果中保持一致。得到对齐后的多视角图像和深度图。

[0021] 基于目标相机位姿参数计算摄像头间的本质矩阵 E ,其中 E 由平移向量 t 的反对称矩阵 $[t]_{\times}$ 和旋转矩阵 R 组成。本质矩阵 E 是多摄像头系统中描述摄像头相对运动的关键元素,提供了多视角图像之间的几何关系。基于此,对预处理后的多视角图像进行极线校正,确保每对图像的对应点位于已校正的极线上,得到校正后的图像对。对校正后的图像对进行相位一致性变换,通过计算相位相关函数,提取图像之间亚像素级别的位移场,这一步骤精确捕捉到微小的位移误差,并通过这些信息计算仿射变换矩阵,从而将图像进行几何调整,得到粗对齐的多视角图像。对粗对齐的多视角图像进行各向异性扩散滤波,在保留图像边缘的同时,对图像进行平滑处理,生成边缘保持的平滑图像集。对平滑图像集执行多尺度Log-Gabor滤波,该滤波方法通过在多个尺度下分析图像特征,能够生成特征张量场,描述图像的局部结构信息。基于特征张量场计算结构张量,通过对结构张量进行特征值分解,得到结构一致性映射,揭示图像在不同视角下的一致性程度。利用结构一致性映射对初步对齐的多视角图像集进行非刚性配准,对图像进行更加精细的对齐,确保它们在所有视角下高度一致,生成精细对齐的多视角图像集。基于精细对齐的多视角图像集构建四元体分割,通过对图像进行分割,将其分解为多个超像素集。每个超像素集包含一组具有相似特征的像素。基于超像素集计算多视图匹配代价体,通过综合多个视角的图像信息,得到初始深度图。对初始深度图应用变分法进行求解,通过迭代优化,不断调整和细化深度信息,得到高精度的初始深度估计。

[0022] 对高精度深度图和粗对齐的多视角图像执行双边联合上采样,在保持边缘信息的同时提高深度图的分辨率,得到高分辨率深度图。基于高分辨率深度图计算视差梯度场,视差梯度场能够反映不同图像之间视差的变化率。在此基础上,构建仿射变换矩阵,对粗对齐的多视角图像进行局部变形,使得各视角图像在空间上更加精确地对齐,得到精细对齐的多视角图像。为了优化对齐后的图像,对精细对齐的多视角图像进行结构张量分析,提取出图像的边缘信息,同时保持图像的平滑性,生成边缘保持的平滑图像。基于边缘保持的平滑图像计算多尺度局部相位一致性度量,能够反映图像在不同尺度下的结构相似性,得到局

部结构相似性映射,帮助识别出图像中的微小差异。基于局部结构相似性映射构建光度一致性约束项,并结合几何平滑约束项,通过求解位移场精确对齐图像,得到精确对齐的多视角图像。对精确对齐的多视角图像和高分辨率深度图执行时空体素化,将图像和深度信息映射到四维空间中,构建4D时空体素网格,并计算每个体素的颜色和深度特征向量,形成特征化的时空体素场。基于此时空体素场构建时空图,并得到时空分割结果,通过这些分割结果,识别出不同时间和空间上的物体区域。对时空分割结果进行连通域分析,计算每个连通域的时空一致性得分。通过这些得分,识别和聚类出不同的动态物体区域,得到动态物体候选区域集合R。动态物体候选区域集合包含可能存在动态物体的区域。基于动态物体候选区域集合,构建光流约束和深度一致性约束,通过联合优化得到精细的运动场,捕捉动态物体的移动轨迹和形态变化。根据精细运动场,对精确对齐的多视角图像和高分辨率深度图进行时序插值和外推,确保在时间轴上处理动态物体时保持一致性,最终得到对齐后的多视角图像和深度图。

[0023] 步骤500、基于对齐后的多视角图像和深度图进行3D模型重建,得到场景3D模型;

具体的,对对齐后的多视角图像和深度图进行体素化处理,通过将图像和深度信息映射到三维体素网格中,得到初始低分辨率体素模型,使得场景数据在空间中得以表示。基于该初始低分辨率体素模型进行八叉树数据结构分析,通过将三维空间递归划分为更小的子空间,得到多分辨率体素表示。多分辨率体素表示不仅能够有效压缩数据,还能在不同分辨率下提供细节信息。对多分辨率体素表示中的每个体素进行带符号距离场值计算,这些值表示场景中每个体素与物体表面的距离,并用正负号区分体素处于物体内外,得到初始SDF模型。基于该初始SDF模型进行多视角深度融合,将来自不同视角的深度信息整合起来,消除各视角间的偏差和不一致,得到优化后的SDF模型。对优化后的SDF模型进行自适应细分,通过在需要更高分辨率的区域生成更多细节,得到高精度的表面网格模型。基于该高精度表面网格模型,进行多视图一致性约束,确保模型表面在不同视角下的一致性,建立表面优化的目标函数。对表面优化目标函数进行迭代求解,通过逐步调整网格的顶点位置,使得模型表面更加精细,得到精细化的表面模型。基于精细化表面模型提取多视角纹理集合,通过从不同视角的图像中获取纹理信息,为模型表面添加丰富的细节。对多视角纹理集合进行图割的全局优化,通过优化算法选择最佳的纹理拼接路径,消除图像之间的边界,得到无缝拼接的纹理映射。基于无缝拼接的纹理映射,将其应用到精细化的表面模型上,生成最终的场景3D模型。

[0024] 步骤600、对场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。

[0025] 具体的,基于人眼立体视觉感知模型对场景3D模型进行深度分析,通过模拟人眼对深度的感知,获取场景的深度分布特征,反映不同物体在空间中的相对距离和层次感。根据场景的深度分布特征,构建自适应立体视差调整模型,该模型能够动态调整视差,以适应不同深度范围内的物体,确保最终影像的立体效果与人眼视觉体验相一致。基于自适应立体视差调整模型对场景3D模型进行视差调整,得到初步的立体视差校正结果。基于初步立体视差校正结果构建视差连续性优化模型,通过该模型对大视差区域进行细致处理,消除因视差变化过大而引起的不自然效果,得到优化后的立体视差校正结果。基于优化后的立体视差校正结果生成稠密光场采样数据,包含场景中丰富的光线信息和视角细节。为了提高数据处理效率,对稠密光场采样数据进行压缩,生成轻量化光场模型,该模型在保证光场

信息完整性的同时,大大减少了存储和计算开销,便于在实际应用中进行快速处理。基于轻量化光场模型,构建初始渲染结果,将光场数据转化为可视化的影像。对初始渲染结果进行图像增强处理,提升影像的细节表现力和对比度,得到高质量的渲染图像。对高质量渲染图像进行时空分析和抖动抑制处理,有效减少因视角或显示设备变化带来的抖动现象,得到流畅且稳定的立体3D影像。

[0026] 本申请实施例中,通过WiFi模块进行时钟同步和中央控制单元生成微秒级触发信号,结合高速LED阵列发射特定模式的光脉冲,实现了多摄像头的亚微秒级时间对齐校正,大大提高了多视角图像的时间一致性。采用全局色彩映射模型、自适应曝光均衡函数、多项式模型的畸变校正等技术,有效解决了多视角图像的噪声、色差、曝光不均和畸变等问题,通过构建仿射不变特征集、全局语义特征提取和特征快速索引树,实现了多视角图像间的高效精确匹配,利用分层稀疏束调整模型、多视角立体匹配代价体分析等技术,实现了多视角图像的精确对齐和高精度深度图的生成。通过时空体素化、时空一致性分析和精细运动场优化,有效识别和处理了动态物体,保证了重建结果的时空一致性。采用多分辨率体素表示、带符号距离场优化和自适应表面细分等技术,实现了高精度、高细节的3D模型重建。基于人眼立体视觉感知模型和自适应立体视差调整,结合稠密光场采样和渲染增强技术,生成了高质量、高沉浸感的立体3D影像。

[0027] 在一具体实施例中,执行步骤100的过程可以具体包括如下步骤:

对所述多个摄像头进行阵列布置,得到摄像头阵列,并对所述摄像头阵列中的每个摄像头配置WiFi模块,得到带WiFi模块的摄像头阵列;

设置中央控制单元,通过WiFi网络与所述带WiFi模块的摄像头阵列建立连接,构建带中央控制单元的摄像头系统;

基于所述中央控制单元生成微秒级触发信号,得到微秒级同步触发信号,并通过WiFi网络将所述微秒级同步触发信号传输至各摄像头;

基于所述传输后的微秒级同步触发信号对所述带中央控制单元的摄像头系统进行触发,得到初步同步的多视角图像;

在拍摄场景中设置高速LED阵列,并基于所述微秒级同步触发信号控制所述高速LED阵列发射特定模式的光脉冲,得到带时间标记的场景;

对所述带时间标记的场景进行拍摄,得到带光脉冲模式的多视角图像,并通过WiFi网络将所述带光脉冲模式的多视角图像传输至云端APP;

在云端APP中基于所述带光脉冲模式的多视角图像中的光脉冲模式进行亚微秒级时间对齐校正,得到多视角原始图像。

[0028] 具体的,根据场景的需求和视角覆盖范围,将摄像头按照阵列进行布置,最大限度地覆盖场景的不同视角,确保从多个角度捕捉到足够的信息。在摄像头阵列中,每个摄像头都配置WiFi模块,从而通过WiFi网络与中央控制单元进行通信,形成一个带WiFi模块的摄像头阵列。通过WiFi网络建立中央控制单元与摄像头阵列之间的连接。中央控制单元的作用是生成和分发高精度的同步触发信号,以确保所有摄像头在同一时间点同时启动拍摄。为了实现微秒级别的时间同步,中央控制单元需要生成微秒级触发信号。微秒级的同步触发信号能够保证摄像头之间的拍摄时间差小于1微秒,避免图像中由于时间差导致的运动模糊或失真。这种触发信号通过WiFi网络传递给各个摄像头。虽然WiFi传输可能引入一些

延迟,但通过在信号中嵌入时间戳信息,可以在接收端对这些延迟进行补偿。传输后的微秒级同步触发信号能够确保每个摄像头在接收到信号的瞬间开始曝光和拍摄,由此得到初步同步的多视角图像。为了提高时间同步的精度,并为后续的图像处理提供可靠的时间基准,在拍摄场景中设置高速LED阵列。高速LED阵列能够通过发射光脉冲,为场景提供明确的时间标记。这些光脉冲由中央控制单元基于微秒级同步触发信号来控制,确保光脉冲与摄像头的拍摄同步。时间标记通过在场景中引入可识别的光学信号,能够在后续的图像处理中用作时间对齐的参照。在摄像头接收到同步触发信号并开始拍摄后,带有时间标记的场景被记录下来,形成带有光脉冲模式的多视角图像。由于每个摄像头所拍摄的图像中都包含这些光脉冲标记,可以在后续的处理过程中利用这些标记对图像进行亚微秒级的时间对齐校正。带有光脉冲模式的多视角图像通过WiFi网络传输至云端APP。在云端APP中,利用图像中的光脉冲标记对图像进行亚微秒级的时间对齐校正。时间对齐校正通过分析图像中光脉冲的出现时间,确保各视角图像在时间轴上的精确同步,消除由于摄像头硬件、WiFi传输或信号处理导致的微小时间偏差,最终得到高精度同步的多视角原始图像。为了更好地理解这种时间同步过程,我们可以通过一个简单的数学模型来描述。设定每个摄像头*i*接收到的触发信号时间为 t_i ,其中*i*=1,2,...,n代表不同的摄像头。通过对所有摄像头的时间信号进行平均,得到一个系统平均时间 T_{avg} 作为基准时间:

$$T_{avg} = (1/n) * \sum (t_i) \quad (i \text{ 从 } 1 \text{ 到 } n \text{ 求和});$$

其中, T_{avg} 是各摄像头时间的平均值,用作全系统的时间基准。如果某个摄像头*i*的时钟有偏差,通过调整其内部时钟使得 t_i 等于 T_{avg} ,实现全系统的时间同步。

[0029] 在处理光脉冲标记时,定义光脉冲到达摄像头的时间差 Δt ,这种差异通过相应的校正算法进行修正,确保所有图像在相同的时间点被记录下来。这种时间校正可以表示为:

$$t_i' = t_i + \Delta t_i;$$

其中 t_i' 是校正后的时间戳,通过调整每个摄像头的时间戳,确保最终的图像在时间上完全一致。

[0030] 通过以上步骤,能够实现高精度的时间同步和视角对齐,最终在云端APP中生成准确的多视角原始图像,为后续的3D影像或动态TIF图像生成奠定基础。这种基于WiFi和云端处理的方法,不仅简化了硬件要求,还提供了更大的灵活性和可扩展性,使得系统可以更容易地适应不同的应用场景和需求。

[0031] 在一具体实施例中,执行步骤200的过程可以具体包括如下步骤:

对多视角原始图像进行去噪,得到初步降噪的多视角图像,并在拍摄场景中放置标准色卡,基于标准色卡对初步降噪的多视角图像进行分析,得到全局色彩映射模型;

基于全局色彩映射模型对初步降噪的多视角图像进行联合色彩校正,得到色彩校正后的多视角图像,并对色彩校正后的多视角图像进行直方图分析,得到每张图像的像素值分布特征;

基于像素值分布特征构建自适应曝光均衡函数,并基于自适应曝光均衡函数对色彩校正后的多视角图像进行像素值调整,得到曝光均衡后的多视角图像;

对曝光均衡后的多视角图像进行多项式模型的畸变校正,得到畸变校正后的多视角图像,并对畸变校正后的多视角图像进行超分辨率处理,得到高分辨率多视角图像;

对高分辨率多视角图像进行边缘锐化和细节增强处理,得到预处理后的多视角图

像。

[0032] 具体的,对多视角原始图像进行去噪,消除拍摄过程中由于各种因素引入的噪声,例如传感器噪声或低光照条件下的随机噪声。常见的去噪方法包括高斯滤波、中值滤波以及更高级的非局部均值滤波。通过分析像素之间的相似性或空间关系,去除噪声的同时尽可能保留图像的细节和纹理,得到初步降噪的多视角图像。在拍摄场景中放置标准色卡。标准色卡是一种具有已知颜色参考值的工具,通常用于图像处理中的色彩校正。通过将标准色卡放置在拍摄场景中,利用已知的颜色参考值对初步降噪的多视角图像进行分析,建立全局色彩映射模型。该模型的构建是通过比较标准色卡在图像中的实际色彩值与预期的标准色彩值来实现的,能够识别出图像中的色偏并进行相应的校正。基于所建立的全局色彩映射模型,对初步降噪的多视角图像进行联合色彩校正,确保所有视角图像的色彩表现一致,避免由于不同摄像头或拍摄条件导致的色彩不一致问题。通过将每张图像的色彩值映射到统一的色彩空间中,得到色彩校正后的多视角图像。对色彩校正后的多视角图像进行直方图分析。通过统计每张图像中不同亮度级别的像素数量,能够反映图像的整体亮度分布和对比度。直方图的形状能够帮助识别图像是否存在曝光不足或过度的问题。通过对每张图像的直方图进行分析,提取出像素值的分布特征。基于像素值分布特征,构建自适应曝光均衡函数。该函数通过调整图像中不同亮度区域的曝光量,使得图像在整体上具有更加均匀的亮度分布。自适应曝光均衡通过拉伸或压缩图像的直方图,使得亮度值更广泛地分布在整个可用的亮度范围内,提高图像的对比度和清晰度。通过该函数对色彩校正后的多视角图像进行像素值调整,得到曝光均衡后的多视角图像。对曝光均衡后的多视角图像进行多项式模型的畸变校正,消除由于摄像头镜头引起的几何畸变,特别是在广角镜头或鱼镜头中,畸变尤为明显。多项式模型可以用来近似描述这种畸变,通过调整图像中每个像素的位置,将其恢复到一个无畸变的视角中。假设原始图像中的坐标为 (x, y) ,校正后的坐标为 (x', y') ,则多项式模型可以表示为:

$$x' = a_0 + a_1x + a_2x^2 + a_3xy + a_4y + \dots;$$

$$y' = b_0 + b_1y + b_2y^2 + b_3xy + b_4x + \dots;$$

其中, $a_0, a_1, \dots, b_0, b_1, \dots$ 是模型的系数,通过对这些系数进行拟合,实现畸变的校正,得到畸变校正后的多视角图像。对畸变校正后的多视角图像进行超分辨率处理。通过算法将低分辨率图像转换为高分辨率图像,常见的方法包括基于插值的传统方法和基于深度学习的现代方法。超分辨率处理能够恢复图像中的细节和纹理,使得最终的图像更加清晰,得到高分辨率的多视角图像。对高分辨率的多视角图像进行边缘锐化和细节增强处理。边缘锐化通过增强图像中边缘部分的对比度,使得物体轮廓更加清晰分明。细节增强则是通过提高图像中纹理和微小结构的清晰度,使得图像看起来更加细腻,最终得到预处理后的多视角图像。

[0033] 在一具体实施例中,执行步骤300的过程可以具体包括如下步骤:

对预处理后的多视角图像进行特征点提取,得到初步特征点集,并基于初步特征点集构建仿射不变特征集;

对仿射不变特征集进行全局语义特征提取,得到语义增强特征集,并基于语义增

强特征集构建特征快速索引树；

基于特征快速索引树对多视角图像间的特征点进行粗匹配,得到初步匹配对集合,并基于初步匹配对集合构建特征点局部关系图；

根据特征点局部关系图生成优化后的匹配对集合,并基于优化后的匹配对集合建立全局几何一致性模型；

基于全局几何一致性模型对匹配结果进行细化,得到精确匹配对集合,并基于精确匹配对集合进行多视图特征融合和传播,得到特征匹配结果。

[0034] 具体的,对预处理后的多视角图像进行特征点提取。这些特征点代表图像中具有独特纹理或形状的关键位置,常用于后续的图像匹配和3D重建。常用的特征点提取算法包括SIFT(尺度不变特征变换)、SURF(加速稳健特征)、ORB(Oriented FAST and Rotated BRIEF)等,这些算法能够在多尺度下检测特征点,并为每个特征点计算其描述子,得到初步特征点集。为了应对不同视角之间的仿射变换,基于初步特征点集构建仿射不变特征集。仿射不变特征能够在图像发生缩放、旋转、平移等仿射变换时保持稳定。通过仿射变换矩阵对初步特征点集进行变换,确保在不同视角下的相同物体特征能够得到一致的描述。假设图像中的一个特征点在变换前的坐标为 (x, y) ,变换后的坐标为 (x', y') ,则仿射变换可以表示为:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix};$$

其中, a, b, c, d 是仿射变换矩阵的参数, e, f 是平移向量。通过对特征点应用这种变换,得到仿射不变特征集。为了提取更高层次的特征信息,对仿射不变特征集进行全局语义特征提取,捕捉图像中物体的语义信息,例如识别出图像中的物体类别、形状和空间关系。通过深度学习技术,将图像的局部特征提升为具有语义含义的特征,形成语义增强特征集。语义增强特征不仅考虑了像素级的相似性,还融入了物体识别和分类的信息,有助于提高匹配的准确性。基于语义增强特征集构建特征快速索引树。特征快速索引树是一种数据结构,能够高效地组织和检索特征点,在大量特征点中进行快速匹配。常见的索引结构包括KD树和LSH,这些结构能够通过分层或哈希方法,将特征点映射到特定的空间中,从而加快匹配过程。根据索引树,在多视角图像间对特征点进行粗匹配,得到初步匹配对集合。通过快速检索,找出不同图像中可能对应的特征点对。对初步匹配对集合进行筛选。基于初步匹配对集合构建特征点局部关系图。局部关系图通过分析特征点之间的空间关系,揭示它们在图像中的相对位置和几何布局。通过这种图论方法,识别出那些不符合几何关系的匹配对,并将其剔除,生成优化后的匹配对集合。基于优化后的匹配对集合建立全局几何一致性模型。全局几何一致性模型通过综合考虑所有视角图像的匹配结果,确保特征点匹配在全局尺度上保持一致。该模型通常通过最小化重投影误差或几何误差来优化匹配结果。重投影误差可以表示为:

$$E = \sum_{i=1}^n \|p_i - \hat{p}_i\|^2;$$

其中, p_i 是实际匹配点的位置, \hat{p}_i 是基于匹配模型计算得到的投影点位置, E 表示所有匹配点的重投影误差和。通过优化该误差函数,得到最优的全局几何一致性模型。基于

全局几何一致性模型对匹配结果进行细化。通过更精细的几何校正和优化,消除错误匹配,并增强正确匹配的置信度,得到精确匹配对集合。基于精确匹配对集合进行多视图特征融合和传播,得到特征匹配结果。多视图特征融合通过将来自不同视角的特征点信息进行整合,生成更加稳健和准确的匹配结果。特征传播则利用已知的匹配信息,在尚未匹配的视角中推断出新的匹配对,从而扩大匹配覆盖范围。

[0035] 在一具体实施例中,执行步骤400的过程可以具体包括如下步骤:

基于特征匹配结果构建多视图几何关系图,得到初始视图拓扑结构,并对初始视图拓扑结构进行优化,得到优化后的视图拓扑结构;

基于优化后的视图拓扑结构建立分层稀疏束调整模型并生成初始相机位姿参数,对初始相机位姿参数进行迭代优化,最小化重投影误差,得到目标相机位姿参数;

基于目标相机位姿参数对预处理后的多视角图像进行几何变换,得到粗对齐的多视角图像,并对粗对齐的多视角图像进行多视角立体匹配代价体分析,得到初始深度估计;

基于初始深度估计进行多视图一致性传播,得到优化后的深度图,并对优化后的深度图进行细化处理,得到高精度深度图;

基于高精度深度图对粗对齐的多视角图像进行精细对齐,得到精确对齐的多视角图像,并对精确对齐的多视角图像和高精度深度图进行时空一致性分析,识别并处理动态物体,得到对齐后的多视角图像和深度图。

[0036] 具体的,基于特征匹配结果构建多视图几何关系图,描述不同视角之间的相对几何关系。通过分析匹配特征点在各个视角之间的位置关系,构建多视图的拓扑结构,该拓扑结构代表摄像头相对位置和姿态的初始估计。对初始视图拓扑结构进行优化,剔除不可靠的视角连接、增强关键视角之间的联系,并通过优化算法调整视角间的相对位置,得到优化后的视图拓扑结构。基于优化后的视图拓扑结构建立分层稀疏束调整模型。分层稀疏束调整是一种常用的多视角重建技术,通过对摄像头的位姿和三维点的位置进行全局优化,能够同时校正图像之间的投影误差和摄像头位姿的不准确性。这个过程的核心是最小化重投影误差,即使得三维点在所有图像上的投影点与实际观测到的特征点尽可能一致。重投影误差可以用以下公式表示:

$$E = \sum_{i=1}^n \sum_{j=1}^m \|p_{ij} - \Pi(R_j, T_j, P_i)\|^2;$$

其中, E 是总的重投影误差, p_{ij} 表示第 i 个三维点在第 j 个图像上的实际观测位置, $\Pi(R_j, T_j, P_i)$ 表示通过相机的旋转矩阵 R_j 和平移向量 T_j 对三维点 P_i 进行投影后的计算位置。通过迭代优化误差函数,得到最优的相机位姿参数,即目标相机位姿参数。基于目标相机位姿参数对预处理后的多视角图像进行几何变换,将所有图像对齐到一个公共的三维空间中,得到粗对齐的多视角图像。对粗对齐的多视角图像进行多视角立体匹配代价体分析。立体匹配代价体分析通过计算视差图,初步估计出每个像素的深度信息。视差图表示的是同一场景中的物体在不同视角下的像素偏移量,这一偏移量与物体的深度直接相关。基于初步的视差估计,得到初始深度图。基于初始深度估计进行多视图一致性传播。通过将多个视角的深度信息进行整合和传播,消除不一致的深度值,得到优化后的深度图。对优化后的深度图进行细化处理,去除噪声和误差,提高深度图的精度,得到高精度深度图。基于高精度深度图,对粗对齐的多视角图像进行精细对齐,通过调整图像的几何变换参数,使得图

像在像素级别上与深度信息完全一致,得到精确对齐的多视角图像。对精确对齐的多视角图像和高精度深度图进行时空一致性分析,识别和处理场景中存在的动态物体,这些物体在不同时间点上可能处于不同的位置,导致深度图和图像的不一致。通过分析这些不一致性,准确识别出动态物体,并通过相应的处理消除其对3D重建的影响,最终得到对齐后的多视角图像和深度图。

[0037] 在一具体实施例中,执行步骤基于目标相机位姿参数对预处理后的多视角图像进行几何变换,得到粗对齐的多视角图像,并对粗对齐的多视角图像进行多视角立体匹配代价体分析,得到初始深度估计的过程可以具体包括如下步骤:

基于目标相机位姿参数计算摄像头间的本质矩阵 E ,其中 $E=[t] \times R$, $[t] \times$ 为平移向量 t 的反对称矩阵, R 为旋转矩阵;对预处理后的多视角图像执行极线校正,得到校正后的图像对;

对校正后的图像对执行相位一致性变换,计算相位相关函数,得到亚像素级位移场,并基于亚像素级位移场,计算仿射变换矩阵,得到粗对齐的多视角图像;

对粗对齐的多视角图像进行各向异性扩散滤波,得到边缘保持的平滑图像集,并对平滑图像集执行多尺度Log-Gabor滤波,得到特征张量场;

基于特征张量场计算结构张量,对结构张量进行特征值分解,得到结构一致性映射,并基于结构一致性映射对初步对齐的多视角图像集进行非刚性配准,得到精细对齐的多视角图像集;

对精细对齐的多视角图像集构建四元体分割,得到超像素集,基于超像素集计算多视图匹配代价体,得到初始深度图,并对初始深度图应用变分法求解,通过迭代优化得到初始深度估计。

[0038] 具体的,基于目标相机位姿参数计算出摄像头之间的本质矩阵 E 。本质矩阵 E 描述了两个摄像头之间的相对运动关系并且与三维空间中的点及其在图像平面上的投影之间的关系密切相关。本质矩阵可以表示为 $E = [t] \times R$,其中 $[t] \times$ 是平移向量 t 的反对称矩阵, R 是旋转矩阵。反对称矩阵 $[t] \times$ 可以表示为:

$$[t] \times = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix};$$

其中, t_x, t_y, t_z 分别是平移向量 t 在 x, y, z 轴上的分量。通过将相机的旋转矩阵 R 和平移向量 t 代入上述公式,计算出本质矩阵 E ,该矩阵用于描述两个摄像头之间的几何关系。对预处理后的多视角图像执行极线校正,将图像中的对应点约束到同一条极线上,简化立体匹配的过程。通过极线校正,消除摄像头之间由于拍摄角度不同而导致的图像畸变,使得每对图像的对应点对齐在相应的极线上,得到校正后的图像对。对校正后的图像对执行相位一致性变换。相位一致性变换通过分析图像中不同频率分量的相位信息,能够有效地提取图像的边缘和纹理特征。通过计算相位相关函数,得到图像之间的相位一致性信息,计算出亚像素级别的位移场。相位相关函数通常用于在频域中比较图像间的相似性,能够提供比传统的像素级对齐更高的精度。通过这些信息,计算仿射变换矩阵,将图像进行几何校正,得到粗对齐的多视角图像。对粗对齐的多视角图像进行各向异性扩散滤波。各向异性扩散

滤波是一种能够在保留图像边缘信息的同时去除噪声的滤波方法。通过扩散过程中的方向性约束,使得图像在边缘附近保持锐利,而在平滑区域内则去除噪声,得到边缘保持的平滑图像集。对边缘保持的平滑图像集执行多尺度Log-Gabor滤波。Log-Gabor滤波器是一种在频域内表现优越的滤波器,能够在不同尺度下提取图像的局部特征。通过在多个尺度上应用Log-Gabor滤波,生成特征张量场,包含图像中丰富的特征信息。基于特征张量场计算结构张量。结构张量是一种描述图像局部结构信息的数学工具,通过对结构张量进行特征值分解,得到结构一致性映射,反映图像在不同区域内的结构相似性。基于结构一致性映射对初步对齐的多视角图像集进行非刚性配准。非刚性配准通过允许图像之间进行局部变形,使得它们在像素级别上达到更高的对齐精度。通过这种配准方式,得到精细对齐的多视角图像集。对精细对齐的多视角图像集构建四元体分割,生成超像素集。超像素集是一种将图像划分为多个具有相似特征的小区域的表示方式,能够有效减少图像处理的计算复杂度,并保持重要的边缘信息。基于超像素集计算多视图匹配代价体,生成初始深度图。多视图匹配代价体分析通过比较不同视角图像中的超像素,估计出每个像素的深度信息,得到初始深度图。对初始深度图应用变分法求解。变分法通过定义一个能量函数,并通过迭代优化能量函数,使得深度图的整体误差最小化,得到更加准确的初始深度估计。这个过程的核心是寻找一个最优的深度分布,使得所有视角图像的重投影误差达到最小。

[0039] 在一具体实施例中,执行步骤基于高精度深度图对粗对齐的多视角图像进行精细对齐,得到精确对齐的多视角图像,并对精确对齐的多视角图像和高精度深度图进行时空一致性分析,识别并处理动态物体,得到对齐后的多视角图像和深度图的过程可以具体包括如下步骤:

对高精度深度图和粗对齐的多视角图像执行双边联合上采样,得到高分辨率深度图;

基于高分辨率深度图计算视差梯度场,并基于视差梯度场构建仿射变换矩阵,对粗对齐的多视角图像进行局部变形,得到精细对齐的多视角图像;

对精细对齐的多视角图像进行结构张量分析,得到边缘保持的平滑图像,并基于边缘保持的平滑图像计算多尺度局部相位一致性度量,得到局部结构相似性映射;

基于局部结构相似性映射构建光度一致性约束项,并结合几何平滑约束项,求解位移场,得到精确对齐的多视角图像;

对精确对齐的多视角图像和高分辨率深度图执行时空体素化,构建4D时空体素网格,并计算每个体素的颜色和深度特征向量,得到特征化时空体素场;

基于特征化时空体素场构建时空图,得到时空分割结果;

对时空分割结果进行连通域分析,计算每个连通域的时空一致性得分,根据时空一致性得分对连通域进行聚类,得到动态物体候选区域集合R;

基于动态物体候选区域集合,构建光流约束和深度一致性约束,联合优化,得到精细运动场;

根据精细运动场对精确对齐的多视角图像和高分辨率深度图进行时序插值和外推,得到对齐后的多视角图像和深度图。

[0040] 具体的,对高精度深度图和粗对齐的多视角图像执行双边联合上采样。双边联合上采样是一种结合空间域和颜色域信息的滤波方法,能够在保持图像边缘细节的同时,有

效提升深度图的分辨率。双边联合上采样的公式为：

$$D'(x) = \frac{1}{W(x)} \sum_{y \in N(x)} D(y) \cdot \exp\left(-\frac{\|x-y\|^2}{2\sigma_s^2}\right) \cdot \exp\left(-\frac{\|I(x)-I(y)\|^2}{2\sigma_r^2}\right);$$

其中， $D'(x)$ 是上采样后的深度值， $D(y)$ 是原始深度值， $N(x)$ 是像素 x 的邻域， $I(x)$ 和 $I(y)$ 分别为像素 x 和 y 的颜色值， σ_s 和 σ_r 分别控制空间域和颜色域的平滑参数， $W(x)$ 是归一化因子。通过这个过程，生成与粗对齐的多视角图像分辨率一致的高分辨率深度图。基于高分辨率深度图计算视差梯度场。视差梯度场表示的是深度信息的变化率，能够揭示图像中物体的三维结构特征。基于视差梯度场，构建仿射变换矩阵，对粗对齐的多视角图像进行局部变形。仿射变换矩阵的公式可以表示为：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix};$$

其中， (x, y) 是原始坐标， (x', y') 是变换后的坐标， a, b, c, d 是控制变换的参数， e 和 f 是平移量。局部变形过程确保了不同视角的图像在几何结构上达到更高的对齐精度，从而得到精细对齐的多视角图像。对精细对齐的多视角图像进行结构张量分析。结构张量是一种能够捕捉图像局部结构特征的工具，通过分析图像中的梯度信息，得到边缘保持的平滑图像。结构张量 J 可以表示为：

$$J = \sum_{i=1}^n \begin{pmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \\ \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} & \left(\frac{\partial I}{\partial y}\right)^2 \end{pmatrix};$$

其中， $\frac{\partial I}{\partial x}$ 和 $\frac{\partial I}{\partial y}$ 是图像在 x 和 y 方向上的梯度。通过对结构张量的分析，得到边缘保持的平滑图像，进一步计算多尺度局部相位一致性度量，生成局部结构相似性映射。基于局部结构相似性映射，构建光度一致性约束项和几何平滑约束项。光度一致性约束项 E_{phot} 可表示为：

$$E_{\text{phot}} = \sum_p \left(I_1(p) - I_2(W(p)) \right)^2;$$

其中， $I_1(p)$ 和 $I_2(p)$ 分别是两幅图像中像素 p 的强度值， $W(p)$ 是位移场 W 将 p 映射到另一图像中的位置。几何平滑约束项 E_{smooth} 表示为：

$$E_{\text{smooth}} = \sum_p \sum_{q \in N(p)} \|W(p) - W(q)\|^2;$$

其中， $q \in N(p)$ 表示像素 p 的邻域像素 q 。通过联合求解这两个约束，得到精确

对齐的多视角图像。对精确对齐的多视角图像和高分辨率深度图执行时空体素化,构建4D时空体素网格,计算每个体素的颜色和深度特征向量,得到特征化时空体素场。基于该体素场构建时空图,并进行时空分割,得到时空分割结果。对时空分割结果进行连通域分析,计算每个连通域的时空一致性得分,识别动态物体。基于动态物体,构建光流约束和深度一致性约束,通过联合优化,得到精细运动场。利用精细运动场对精确对齐的多视角图像和高分辨率深度图进行时序插值和外推,得到对齐后的多视角图像和深度图。

[0041] 在一具体实施例中,执行步骤500的过程可以具体包括如下步骤:

对对齐后的多视角图像和深度图进行体素化处理,得到初始低分辨率体素模型,并基于初始低分辨率体素模型进行八叉树数据结构分析,得到多分辨率体素表示;

对多分辨率体素表示中的每个体素进行带符号距离场值计算,得到初始SDF模型,并基于初始SDF模型进行多视角深度融合,得到优化后的SDF模型;

对优化后的SDF模型进行自适应细分,得到高精度表面网格模型,并基于高精度表面网格模型进行多视图一致性约束,得到表面优化目标函数;

对表面优化目标函数进行迭代求解,得到精细化表面模型,并基于精细化表面模型提取多视角纹理集合;

对多视角纹理集合进行图割的全局优化,得到无缝拼接的纹理映射,并基于无缝拼接的纹理映射生成场景3D模型。

[0042] 具体的,对对齐后的多视角图像和深度图进行体素化处理,体素化是一种将三维空间划分为小立方体单元(即体素)的过程,通过将每个体素与图像中的像素或深度信息关联,生成初始低分辨率体素模型。低分辨率体素模型能够以较小的计算代价,初步捕捉场景的三维结构。基于初始低分辨率体素模型进行八叉树数据结构分析。八叉树是一种常用于三维空间数据划分的树形结构,通过递归地将空间划分为八个子空间,形成层次化的体素表示。每个体素被细分为更小的子体素,这些子体素的分辨率可以根据需要进行调整。多分辨率体素表示不仅能够在关键区域提供更高的细节分辨率,还能够通过压缩不重要的区域来降低计算复杂度。对多分辨率体素表示中的每个体素进行带符号距离场值的计算。带符号距离场值表示的是体素到物体表面的距离,并且用正负号来区分体素是位于物体内部还是外部。通过计算每个体素的带符号距离场值,生成初始的SDF模型,该模型描述了整个空间中的物体形状。基于初始SDF模型,进行多视角深度融合。通过整合来自不同视角的深度信息,消除单一视角带来的误差和不确定性,得到更加准确的三维形状表示。通过融合这些深度信息,生成一个优化后的SDF模型。对优化后的SDF模型进行自适应细分。通过在需要高精度表示的区域增加体素分辨率,以捕捉更细致的表面特征,生成高精度表面网格模型,该模型能够精确描述物体的三维形状。为了优化高精度表面网格模型,基于多视图一致性约束来构建表面优化目标函数。多视图一致性约束通过确保来自不同视角的图像在网格表面上的投影一致,能够提高表面模型的精度。表面优化目标函数可以表示为:

$$E_{\text{surface}} = \sum_i \sum_j \|I_i(p_j) - I'_i(p_j)\|^2 + \lambda \sum_p \|\nabla \text{SDF}(p)\|^2;$$

其中, $I_i(p_j)$ 表示第 i 个视角在点 p_j 处的像素值, $I'_i(p_j)$ 表示投影后的像素值,

第一项度量了投影误差,第二项 $\lambda \sum_p \|\nabla \text{SDF}(p)\|^2$ 是SDF梯度的平滑项,控制网格的光滑度。通过对该目标函数进行迭代求解,优化网格表面,得到精细化的表面模型。在精细化的表面模型基础上,提取多视角纹理集合。多视角纹理集合包含了来自不同摄像头拍摄的图像中与表面网格对应的纹理信息。对纹理进行图割进行全局优化。图割算法能够在多视角纹理中选择最佳的拼接路径,消除不同视角之间的接缝,生成无缝拼接的纹理映射。基于无缝拼接的纹理映射,将其应用到精细化的表面模型上,生成最终的场景3D模型。

[0043] 在一具体实施例中,执行步骤600的过程可以具体包括如下步骤:

基于人眼立体视觉感知模型对场景3D模型进行深度分析,得到场景深度分布特征,并根据场景深度分布特征构建自适应立体视差调整模型;

基于自适应立体视差调整模型对场景3D模型进行视差调整,得到初步立体视差校正结果,并基于初步立体视差校正结果构建视差连续性优化模型;

通过视差连续性优化模型对大视差区域进行处理,得到优化后的立体视差校正结果,并基于优化后的立体视差校正结果生成稠密光场采样数据;

对稠密光场采样数据进行压缩,得到轻量化光场模型,并基于轻量化光场模型构建初始渲染结果;

对初始渲染结果进行图像增强,得到高质量渲染图像,并基于高质量渲染图像进行时空分析和抖动抑制处理,得到立体3D影像。

[0044] 具体的,通过人眼立体视觉感知模型对场景3D模型进行深度分析。人眼立体视觉感知模型模拟了人类双眼如何通过视差感知物体的深度。当人眼观察一个场景时,左右眼之间的微小位置差异导致它们看到的图像略有不同,这种差异被称为视差。通过计算这些视差,人脑能够推断出物体的深度。基于这一模型,对场景3D模型进行深度分析,提取出场景的深度分布特征。场景的深度分布特征反映了物体在空间中的距离层次和深度变化。基于这些特征,构建自适应立体视差调整模型。自适应立体视差调整模型根据不同的深度区域,动态调整视差值,以确保最终的立体影像在视觉上符合人眼的感知习惯。例如,对于远处的物体,视差应较小,而对于近处的物体,视差应较大。这个调整过程的数学表示可以用以下公式来描述:

$$\Delta d = \alpha \cdot \frac{1}{Z};$$

其中, Δd 表示调整后的视差值, α 是调节参数, Z 表示物体到摄像头的深度距离。通过这个公式动态调整视差,得到适应不同深度场景的初步立体视差校正结果。基于初步立体视差校正结果构建视差连续性优化模型。视差连续性优化模型旨在处理由于视差变化过大而引起的不连续性问题。视差不连续会导致立体影像中出现视觉错觉或不自然的跳跃感。为了避免这些问题,视差连续性优化模型通过平滑视差值,使得视差变化更加自然和连贯。这一过程通常通过最小化以下目标函数来实现:

$$E_{\text{cont}} = \sum_i \sum_{j \in N(i)} (\Delta d_i - \Delta d_j)^2;$$

其中, E_{cont} 是视差连续性误差, Δd_i 和 Δd_j 分别是相邻像素*i*和*j*的视差值,

$N(i)$ 表示像素 i 的邻域。通过最小化这个误差,平滑视差变化,得到优化后的立体视差校正结果。基于优化后的立体视差校正结果生成稠密光场采样数据。稠密光场表示的是场景中每个点在多个方向上的光线信息,通过稠密光场采样,捕捉场景中的细微光照变化和视角差异。稠密光场采样数据量通常非常大,因此需要进行压缩处理,以便在有限的存储和计算资源下实现高效处理。对稠密光场采样数据进行压缩,通过减少冗余数据,得到轻量化光场模型。压缩方法可以包括主成分分析、变换编码、或其他高效的压缩算法。通过这些方法,在保留重要光场信息的同时,显著减少数据量。基于轻量化光场模型,构建初始渲染结果。初始渲染结果是基于压缩后的光场数据生成的场景图像,通过重构光场中的光线信息,再现场景的基本视觉效果。对初始渲染结果进行图像增强处理。图像增强可以包括对比度调整、色彩校正、锐化处理等,以提升图像的视觉效果。例如,可以通过以下公式调整图像对比度:

$$I' = I \times \beta + \gamma;$$

其中, I' 是增强后的图像像素值, I 是原始像素值, β 是对比度调整因子, γ 是亮度偏移值。通过增强操作,得到高质量的渲染图像。基于高质量渲染图像进行时空分析和抖动抑制处理。时空分析用于确保图像在时间序列上的一致性,避免由于动态变化引起的闪烁或不稳定问题。抖动抑制通过检测和补偿图像中的微小抖动,确保在播放立体3D影像时,图像保持平稳和自然。通过以上步骤,最终得到一个高质量且稳定的立体3D影像。

[0045] 上面对本申请实施例中基于多摄像头拍摄的3D影像生成方法进行了描述,下面对本申请实施例中基于多摄像头拍摄的3D影像生成装置10进行描述,请参阅图2,本申请实施例中基于多摄像头拍摄的3D影像生成装置10一个实施例包括:

采集模块11,用于对多个摄像头采集的图像进行同步触发采集,得到多视角原始图像;

预处理模块12,用于对多视角原始图像进行预处理,得到预处理后的多视角图像;

匹配模块13,用于对预处理后的多视角图像进行特征提取与匹配,得到特征匹配结果;

计算模块14,用于基于特征匹配结果对预处理后的多视角图像进行对齐和深度信息计算,得到对齐后的多视角图像和深度图;

重建模块15,用于基于对齐后的多视角图像和深度图进行3D模型重建,得到场景3D模型;

生成模块16,用于对场景3D模型进行立体视差校正和渲染处理,生成立体3D影像。

[0046] 通过上述各个组成部分的协同合作,通过WiFi模块进行时钟同步和中央控制单元生成微秒级触发信号,结合高速LED阵列发射特定模式的光脉冲,实现了多摄像头的亚微秒级时间对齐校正,大大提高了多视角图像的时间一致性。采用全局色彩映射模型、自适应曝光均衡函数、多项式模型的畸变校正等技术,有效解决了多视角图像的噪声、色差、曝光不均和畸变等问题,通过构建仿射不变特征集、全局语义特征提取和特征快速索引树,实现了多视角图像间的高效精确匹配,利用分层稀疏束调整模型、多视角立体匹配代价体分析等技术,实现了多视角图像的精确对齐和高精度深度图的生成。通过时空体素化、时空一致性分析和精细运动场优化,有效识别和处理了动态物体,保证了重建结果的时空一致性。采用

多分辨率体素表示、带符号距离场优化和自适应表面细分等技术,实现了高精度、高细节的3D模型重建。基于人眼立体视觉感知模型和自适应立体视差调整,结合稠密光场采样和渲染增强技术,生成了高质量、高沉浸感的立体3D影像。

[0047] 本申请还提供一种电子设备,所述电子设备包括存储器和处理器,存储器中存储有计算机可读指令,计算机可读指令被处理器执行时,使得处理器执行上述各实施例中的所述基于多摄像头拍摄的3D影像生成方法的步骤。

[0048] 本申请还提供一种计算机可读存储介质,该计算机可读存储介质可以为非易失性计算机可读存储介质,该计算机可读存储介质也可以为易失性计算机可读存储介质,所述计算机可读存储介质中存储有指令,当所述指令在计算机上运行时,使得计算机执行所述基于多摄像头拍摄的3D影像生成方法的步骤。

[0049] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0050] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台电子设备(可以是个人计算机,服务器,或者网络设备等)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(read-only memory, ROM)、随机存取存储器(random access memory, RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0051] 以上所述,以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

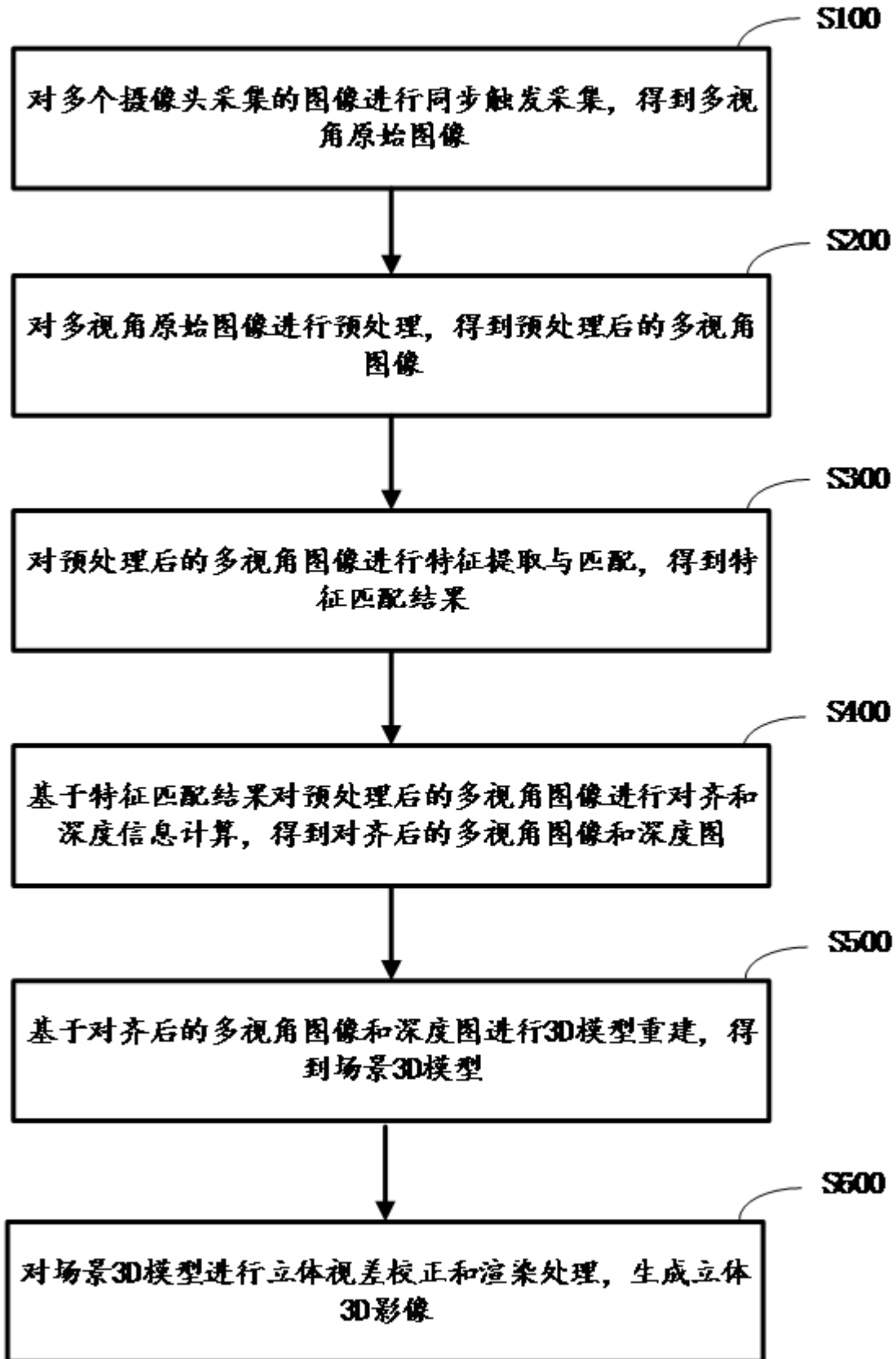


图 1

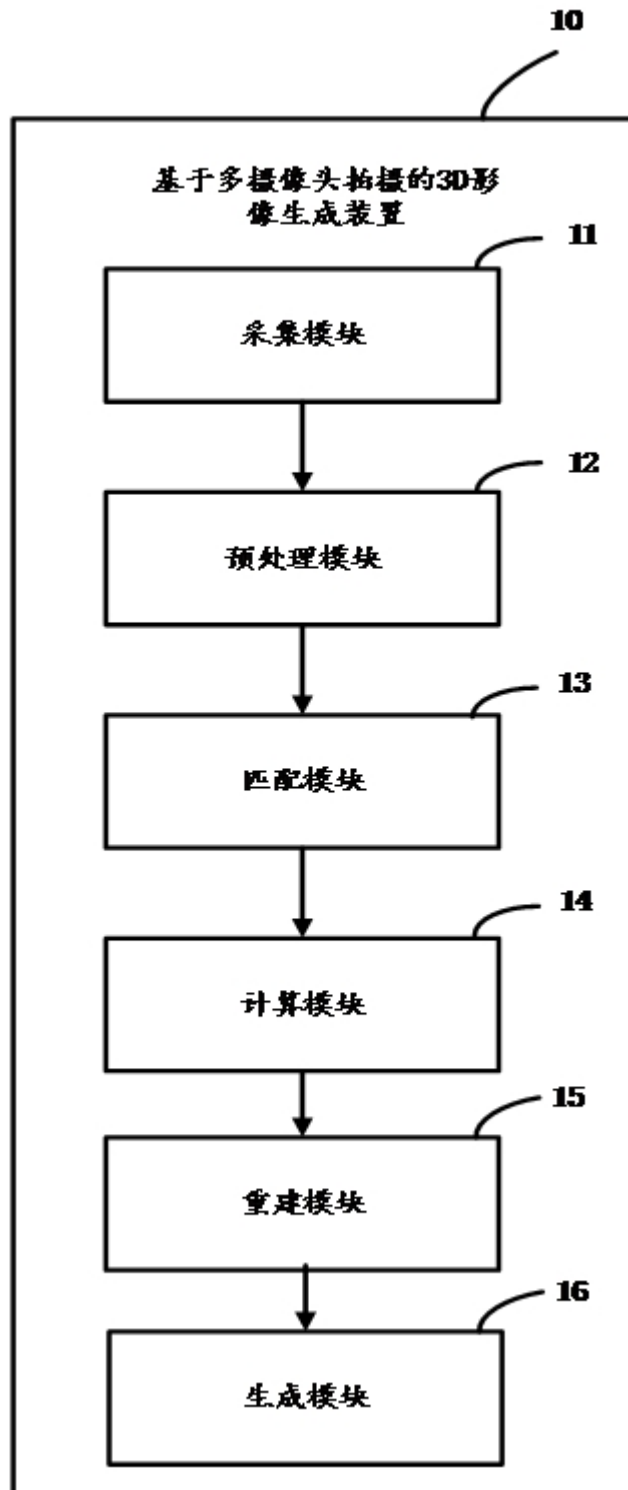


图 2