



(12) 发明专利

(10) 授权公告号 CN 103500013 B

(45) 授权公告日 2016. 05. 11

(21) 申请号 201310490129. 0

[0010], [0013], [0016] 段.

(22) 申请日 2013. 10. 18

CN 102938142 A, 2013. 02. 20, 全文.

(73) 专利权人 武汉大学

审查员 郑嘉青

地址 430072 湖北省武汉市武昌区珞珈山武汉大学

(72) 发明人 吕维 朱欣焰 刘异 陈呈辉 胡涛

(74) 专利代理机构 武汉科皓知识产权代理事务所(特殊普通合伙) 42222

代理人 张火春

(51) Int. Cl.

G06F 3/01(2006. 01)

(56) 对比文件

CN 103220543 A, 2013. 07. 24, 说明书第 [0035]-[0064] 段、图 1, 2.

CN 102824176 A, 2012. 12. 19, 说明书第

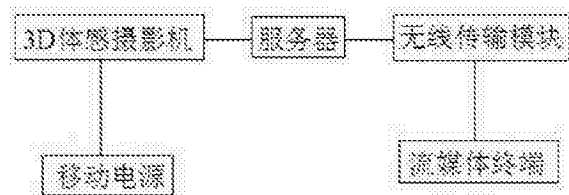
权利要求书1页 说明书6页 附图3页

(54) 发明名称

基于 Kinect 和流媒体技术的实时三维测图方法

(57) 摘要

本发明公开了一种基于 Kinect 和流媒体技术的实时三维测图系统及方法, 本发明系统包括: 3D 体感摄影机、流媒体终端、服务器和移动电源, 流媒体终端固定于 3D 体感摄影机上, 移动电源给 3D 体感摄影机提供电源, 3D 体感摄影机与服务器相连, 服务器通过无线传输模块与流媒体终端相连。本发明以 3D 体感摄影机作为室内三维场景的扫描设备, 使得三维测图更轻便、灵活、成本低廉, 尤其适用于某些较拥挤、难携带大型设备及细节丰富的室内区域的实时三维测图。



1. 基于Kinect和流媒体技术的实时三维测图方法,其特征是,包括:

通过3D体感摄影机采集室内场景的扫描数据,并将扫描数据传递至服务器,所述的扫描数据包括深度信息和颜色信息;

服务器基于扫描数据构建当前帧表面模型,基于当前帧表面模型实时更新全局体元模型并绘制反馈帧,将反馈帧发送至流媒体终端显示;本步骤具体为:

将扫描数据中的深度信息去噪后转换为三维点云,即当前帧表面模型;

将当前帧表面模型与全局体元模型的空间位置匹配,并反解出3D体感摄影机空间位置参数;

融合当前帧表面模型与全局体元模型获得更新后的全局体元模型,根据3D体感摄影机空间位置参数对更新后的全局体元模型进行光线渲染,光线渲染后的全局体元模型的截屏即为反馈帧;

使用者根据流媒体终端显示的反馈帧调整扫描过程;

所述的流媒体终端固定于3D体感摄影机上,移动电源给3D体感摄影机提供电源,3D体感摄影机与服务器相连,服务器通过无线传输模块与流媒体终端相连。

2. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征是:

通过流媒体终端开启、暂定或关闭3D体感摄影机扫描程序。

3. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征在于:

所述的使用者根据流媒体终端显示的反馈帧调整扫描过程,具体为:

若流媒体终端显示的反馈帧存在因遮挡造成的空洞,则移动3D体感摄影机并调整扫描角度以回避遮挡。

4. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征在于:

所述的使用者根据流媒体终端显示的反馈帧调整扫描过程,具体为:

以流媒体终端显示的反馈帧为目标进行重复扫描或采用“靠近”方式继续扫描。

5. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征在于:

所述的3D体感摄影机为微软公司的Kinect扫描设备。

6. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征在于:

所述的流媒体终端为智能手机。

7. 如权利要求1所述的基于Kinect和流媒体技术的实时三维测图方法,其特征在于:

所述的服务器为便携式电脑。

## 基于Kinect和流媒体技术的实时三维测图方法

### 技术领域

[0001] 本发明属于三维测图领域,特别涉及一种基于Kinect和流媒体技术的实时三维测图系统及方法。

### 背景技术

[0002] 随着我国地理信息公共服务平台的不断深入,三维地理信息系统进一步从宏观走向精细,从室外走向室内。在这种形势下,传统技术不再适用,发展快速、轻便、全自动的新型室内三维测图技术(3D Mapping)具有重要现实意义和应用前景。

[0003] 当前室内场景三维测图技术主要基于LiDAR,LiDAR(Light Detection And Ranging)是激光探测及测距系统的简称,其主要原理是用激光传感器通过计算激光回波时间,精确记录传感器与地物回波点之间的距离,由此可直接测量地面及地物各点的三维坐标。目前LiDAR在室内场景三维建模的使用中,主要是通过多点架设固定站进行360度扫描,同时在通视位置布置标靶,将各固定站的场景进行配准和合并。(参考文献:V Verma,3d building detection and modeling from aerial lidar data,Computer Vision and Pattern,2006)

[0004] 相对于LiDAR,Kinect在形态和原理上具有较大不同。Kinect是一种微软开发并商品化生产的体感设备,它主要由一个彩色摄像头(RGB Camera)、一个红外线摄像头(IR Camera)和一个红外线投影仪(IR Projector)构成。通过红外线投影仪和摄像头的配合,Kinect能分析红外线立体像对以获取扫描对象的深度(距离)信息,从而建立对象的三维模型。因此Kinect能同时采集对象的深度和颜色信息,从而全自动完成模型的纹理映射。和笨重昂贵的LiDAR不同,Kinect廉价且轻便,价格约150美元,重量仅450克。(参考文献:[http://www.ros.org/wiki/kinect\\_calibration/technical](http://www.ros.org/wiki/kinect_calibration/technical) ;Remondino, F.,Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. Remote Sensing, 2011.)

[0005] 作为Kinect的开发商,微软及其合作伙伴提供的基于Kinect的一种三维建模算法Kinect Fusion,可以实现利用一台围绕物体移动的kinect实时重建物体的三维模型。不同于简单的三维点云的拼接,该三维建模技术吸引人的特性在于:如果对物体进行持续扫描,三维重建精度可由粗到细逐渐提高,尤其适合应用在增强现实领域。该三维建模技术同时支持CPU和GPU级别的模型计算,核心思想是基于帧到帧表面追踪、实时建模和预测模式。(参考文献:Richard A.Newcombe,Shahram Izadi,KinectFusion:Real-Time Dense Surface Mapping and Tracking,2011)

[0006] 实时测图是一种新型测图概念,其理念与离线测图相对,在测图过程中允许人的参与,具有很高的灵活性。传统的离线测图在现场扫描完成后,再进行数据的离线建模。实时测图为移动测图,实时扫描、实时建模,并实时反馈当前建模结果给使用者。使用者根据当前建模结果,可对扫描过程进行动态调整,发挥主观能动性。传统离线测图对测图流程采用事先布置、事先计划的模式,当扫描数据出现问题,如遮挡、缺失、重影与粗误差等时,使

用者并不实时知晓。当对数据进行离线建模得到反馈以后,补测工作已经较难进行,而实时测图则回避这种弊端。(参考文献:DL Tulloch,Many, many maps: Empowerment and online participatory mapping,2007;Kurt Konolige and Motilal Agrawal,Frame-Frame Matching for Realtime Consistent Visual Mapping,IEEE ,2007)。

## 发明内容

[0007] 针对现有技术中离线测图存在的不足,本发明提供了一种轻便、灵活、成本低廉的基于Kinect和流媒体技术的实时三维测图系统及方法。

[0008] 为解决上述问题,本发明采用如下技术方案:

[0009] 一、一种基于Kinect和流媒体技术的实时三维测图系统,包括:3D体感摄影机、流媒体终端、服务器和移动电源,流媒体终端固定于3D体感摄影机上,移动电源给3D体感摄影机提供电源,3D体感摄影机与服务器相连,服务器通过无线传输模块与流媒体终端相连。

[0010] 上述3D体感摄影机为微软公司的Kinect扫描设备。

[0011] 上述流媒体终端为智能手机,优选安卓手机。

[0012] 上述服务器为便携式电脑。

[0013] 二、一种基于Kinect和流媒体技术的实时三维测图方法,包括:

[0014] 通过3D体感摄影机采集室内场景的扫描数据,并将扫描数据传递至服务器,所述的扫描数据包括深度信息和颜色信息;

[0015] 服务器基于扫描数据构建当前帧表面模型,基于当前帧表面模型实时更新全局体元模型并绘制反馈帧,将反馈帧发送至流媒体终端显示;

[0016] 使用者根据流媒体终端显示的反馈帧调整扫描过程。

[0017] 上述实时三维测图方法还包括特点:通过流媒体终端开启、暂定或关闭3D体感摄影机的扫描程序。

[0018] 上述服务器和流媒体终端基于无线局域网络进行信息交互。

[0019] 上述服务器基于扫描数据构建当前帧表面模型,基于当前帧表面模型实时更新全局体元模型并绘制反馈帧,将反馈帧发送至流媒体终端显示,具体为:

[0020] 将扫描数据中的深度信息去噪后转换为三维点云,即当前帧表面模型;

[0021] 将当前帧表面模型与全局体元模型的空间位置匹配,并反解出3D体感摄影机空间位置参数;

[0022] 融合当前帧表面模型与全局体元模型获得更新后的全局体元模型,根据3D体感摄影机空间位置参数对更新后的全局体元模型进行光线渲染,光线渲染后的全局体元模型的截屏即为反馈帧。

[0023] 上述使用者根据流媒体终端显示的反馈帧调整扫描过程,具体为:

[0024] 若流媒体终端显示的反馈帧存在因遮挡造成的空洞,则移动3D体感摄影机并调整扫描角度以回避遮挡;或,以流媒体终端显示的反馈帧为目标进行重复扫描或采用“靠近”方式继续扫描,以加强当前帧表面模型重建精度。

[0025] 目前,室内场景三维测图常采用昂贵且笨重的LiDAR激光扫描仪,需提前设置拍摄站点和标靶,扫描前期准备工作较长,且在拥挤环境中不易布站;所布扫描站不能移动,对于形态复杂、细节丰富的室内环境,易出现因遮挡导致的数据缺失等问题。由于扫描数据离

线处理,不能实时检验扫描效果,导致扫描数据易存在粗误差、重影等缺陷。若采用手持LiDAR补测,也常因设备拍摄范围小、有效距离短等因素,导致采集效率较低;并且还存在测图工序相对复杂、后期拼接困难等问题。

[0026] 与上述传统测图技术相比,本发明更灵活,且可实时发现扫描数据存在的问题,从而及时进行补测工作。

[0027] 本发明的有益效果如下:

[0028] 一、轻便、灵活、成本低廉。

[0029] 本发明以3D体感摄影机作为室内三维场景的扫描设备。微软公司的Kinect扫描设备即为一种3D体感摄影机,Kinect扫描设备定价约150美元,重量仅450克,加上其他部件,价格不超过300美元,且系统重量较轻,单人即可轻松携带。对于某些较拥挤、难携带大型设备及细节丰富的室内区域(例如犯罪现场等场景)均也可进行测图,受空间环境影响小;同时,由于价格低廉,普通公众也具有购买能力;对于室内环境多变区域,可由公众自发进行数据更新,提高数据现势性。

[0030] 二、实现实时测图。

[0031] 使用时,使用者手持固定有流媒体终端的3D体感摄影机,将服务器和移动电源置于背包内背于身上。使用者可根据实际需要主动选择扫描路线和角度,可避免因遮挡而造成的场景数据的缺失。扫描数据实时发送至服务器,服务器处理扫描数据并将处理后获得的更新当前帧表面模型反馈至流媒体终端并显示。使用者根据流媒体终端上反馈的处理结果对扫描流程进行调整,以避免扫描数据的粗误差、重影等缺陷。

[0032] 三、自动化程度高,测图精度高,测图效率高。

[0033] 服务器根据接收到的实施扫描数据,将当前帧表面模型与总体场景模型之间自动实时匹配,相比于布置标靶等传统匹配方式,自动化程度更高,测图精度和测图效率也得到极大的改善。

[0034] 四、前期工作耗时少,扫描效果好。

[0035] 本发明无需在场景进行站点和标靶布置,单人携带本发明系统即可进场扫描,对场景环境无任何要求,且无前期准备工作;扫描过程可根据实际需求灵活取舍,对于感兴趣区域可通过重复扫描、靠近扫描等方式进行加强测图,对于墙壁、地板可快速扫描。

## 附图说明

[0036] 图1为本发明系统结构框图;

[0037] 图2为本发明系统的具体实施方式示意图;

[0038] 图3为当前帧表面模型构建流程图;

[0039] 图4为服务器将反馈帧以流媒体形式发送至客户端的流程图;

[0040] 图5为流媒体终端处理交换指令的流程图。

[0041] 图中,1-背包,2-电源线,3-数据线,4-移动电源,5-服务器,6-流媒体终端,7-3D体感摄影机,8-电缆。

## 具体实施方式

[0042] 见图1~2,本发明系统包括:3D体感摄影机(7)、流媒体终端(6)、服务器(5)和移动

电源(4)。流媒体终端(6)固定于3D体感摄影机(7)上,移动电源(4)给3D体感摄影机(7)提供电源,3D体感摄影机(7)与服务器(5)相连,服务器(5)通过无线传输模块与流媒体终端(6)相连。

[0043] 固定有流媒体终端(6)的3D体感摄影机(7)为手持设备,服务器(5)和移动电源(4)置于背包(1)中。使用者背上背包(1),手持3D体感摄影机(7),对室内场景进行扫描。流媒体终端(6)与3D体感摄影机(7)通过无线网络方式进行信息交换。3D体感摄影机(7)通过电缆(8)与背包内置的服务器(5)和移动电源(4)相连,电缆(8)由数据线(3)和电源线(2)共线合成,电缆(8)一端与3D体感摄影机(7)相连,另一端分为电源接口和数据接口,电源接口与移动电源(4)相连,数据接口与服务器(5)的USB接口相连。

[0044] 为达到便携的目的,本具体实施中采用的移动电源(4)外形尺寸为20cm\*8cm\*3cm,仅重约1.2kg;服务器(5)为便携式电脑;3D体感摄影机(7)为微软公司的Kinect扫描设备;流媒体终端(6)为智能手机,优选安卓手机。

[0045] 3D体感摄影机(7)用来采集室内场景信息以获得扫描数据,并将扫描数据通过电缆(8)的数据接口传输至服务器(5),服务器(5)对接收的扫描数据实时建模获得当前帧表面模型,并将建模结果通过无线网络反馈于流媒体终端(6),使用者通过流媒体终端(6)显示的当前帧表面模型对扫描过程进行调整和控制。本发明中,流媒体终端除了能实时显示当前帧表面模型外,工作者还可以通过流媒体终端控制3D体感摄影机的扫描过程。

[0046] 基于上述实时三维测图系统,本发明的实时三维测图方法包括步骤:

[0047] 步骤一,通过流媒体终端开启3D体感摄影机,使3D体感摄影机进入扫描状态。

[0048] 步骤二,移动3D体感摄影机采集室内场景的扫描数据,包括深度信息和颜色信息,并将扫描数据传递至服务器;本具体实施中,扫描数据以每秒30帧的速率传递至服务器。

[0049] 步骤三,服务器实时处理扫描数据,基于扫描数据构建当前帧表面模型,基于当前帧表面模型实时更新全局体元模型并绘制反馈帧。

[0050] 步骤四,将反馈帧在无线网络中发布,流媒体终端接收并实时显示。

[0051] 步骤五,使用者根据流媒体终端显示的反馈帧调整扫描过程,例如:若流媒体终端显示的反馈帧存在因遮挡造成的空洞,则移动Kinect扫描设备并调整扫描角度以回避遮挡;若对细节对象等感兴趣,则以该细节对象为目标进行重复扫描,逐渐加强模型重建精度,也可采用“靠近”方式提高目标的扫描精度。

[0052] 步骤六,移动Kinect扫描设备并重复步骤二~五,直到完成室内全部场景的扫描,通过流媒体终端关闭Kinect扫描设备的扫描状态。

[0053] 下面将详细说明本发明的具体实现方式。

[0054] 下述,流媒体终端以安卓手机为例,服务器以便携式电脑为例,3D体感摄影机以Kinect扫描设备为例。

[0055] (1)建立服务器和客户端环境。

[0056] 本具体实施中,客户端指流媒体终端。首先,建立无线局域网络环境,具体建立方法可根据现场实际情况决定。

[0057] 检索并记录便携式电脑的IP地址。在便携式电脑上的Kinect扫描程序上新建两个服务器进程:流媒体服务器进程和交互服务器进程。流媒体服务器用来将便携式电脑更新后的当前帧实时发送至客户端,设置流媒体服务器独立端口号(如5556),模式为“PUB”,PUB

模式表示单向数据分发(Publish),即,流媒体服务器推送一组更新当前帧到客户端,不需要客户端回应。交互服务器用来回应客户端并发送各种指令,设置交互服务器独立端口号(如5555),模式为“REP”(Reply)。“REP”模式与下述“REQ”模式配合,为一种双向消息模式,必须得到对方回应后才能继续发送,从而可保证数据安全性。

[0058] 同理,在安卓手机上的Kinect扫描程序建立两个客户端线程:流媒体客户端线程和交互客户端线程。流媒体客户端绑定便携式电脑IP的流媒体服务器端口(如5556),流媒体客户端模式为“SUB”(Subscribe),流媒体客户端用来接收便携式电脑发送的更新后的当前帧。交互客户端绑定便携式电脑的交互服务器端口(如5556),交互客户端模式为“REQ”(Request),交互客户端用来与便携式电脑发送并回应各种指令。

[0059] (2)Kinect扫描设备的扫描和建模。

[0060] Kinect扫描设备扫描的每幅深度信息可理解成场景的帧表面模型,本具体实施中,便携式电脑采用体重构技术基于扫描数据构建帧表面模型,具体为:

[0061] (a)预定义三维空间,将该三维空间细分成连续规整的体元,获得初始全局体元模型。初始时,各体元中均不包含任何数据。对当前深度信息去噪后转换为三维点云,即当前帧表面模型;通过多规模ICP(Iterative Closest Point,迭代最近点)配准算法,将当前帧表面模型与全局体元模型的空间位置匹配,从而将各当前帧表面模型归化于同一坐标系下,并由此反解出Kinect扫描设备位置参数。

[0062] (b)通过TSDF(Truncated Signed Distance Function,截断有向距离函数)将各当前帧表面模型与全局体元模型进行融合得到更新后的全局体元模型,具体为:利用TSDF计算三维空间任意点的值,从而生成有向距离场;场的正负相交处则为重建物体表面。该方法能很好地延续表面的弯曲趋势,并且能够很好地修复模型表面的空洞。同一个场景表面从不同方向有多幅深度数据,因此数据融合的时候应对多幅深度数据加权求平均,这样还可逐渐提高该场景表面的重建精度。

[0063] (c)同时,根据Kinect扫描设备位置参数,对融合后的更新后的全局体元模型进行光线跟踪(Ray Casting),由此对当前影像范围的更新后的全局体元模型进行光影渲染;并将渲染后的全局体元模型作为反馈帧,输出至流媒体终端并显示,以使使用者对扫描效果进行检验。扫描结束后,最终全局体元模型即为扫描获得的室内三维场景模型。

[0064] (3)服务器将更新帧以流媒体形式发送至客户端并在客户端显示。

[0065] 见图4。由“流媒体服务器”监听Kinect扫描程序主进程,每当更新帧生成便将更新帧内存位置传递进来。更新帧数据流为30Hz三通道640 x480图像,其高达27M每秒,网络难以负担,所以,本具体实施采用Jpeg压缩模式对更新帧数据进行压缩,压缩后每秒数据流降至150K。接下来“流媒体服务器”将压缩数据进行编码(encode)并发送,具体编码方式为:根据压缩数据长度对压缩数据分组,各组根据已有协议加上表头信息,并在流媒体服务器端口进行发布。

[0066] 与之对应,流媒体终端的“流媒体客户端”监听流媒体服务器端口,当流媒体服务器端口有新数据发布时则接收并解码数据,接收模式为异步通信模式(Asynchronous Message Passing)。具体接收方式为分组接收数据,即,根据表头信息将各小组重新组合成完整的数据,即压缩图像;然后,将压缩图像的内存位置发送给主线程,主线程反解码压缩文件,获得位图(Bitmap)形式的图像;最后,在流媒体终端设定区域显示图像。

[0067] (4)流媒体终端接收并处理各种交互指令。

[0068] 见图5。流媒体终端接受使用者的各种指令,如开始、结束、暂停等。交互客户端将指令相应的编号在交互服务器端口以“请求”(Request)形式发送,交互服务器接收到交互客户端发送的编号后与操作指令库进行匹配,若无对应操作指令则“回应”(Reply)交互客户端其指令无效;如有相应操作指令则执行(如,开启或关闭相应的线程),并回应交互客户端“处理完毕”。



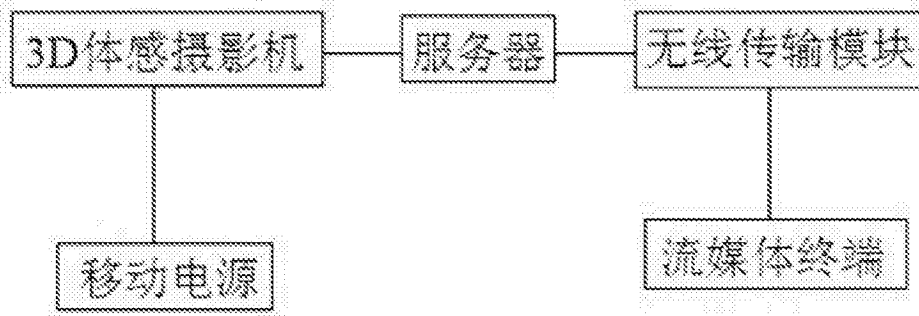


图1

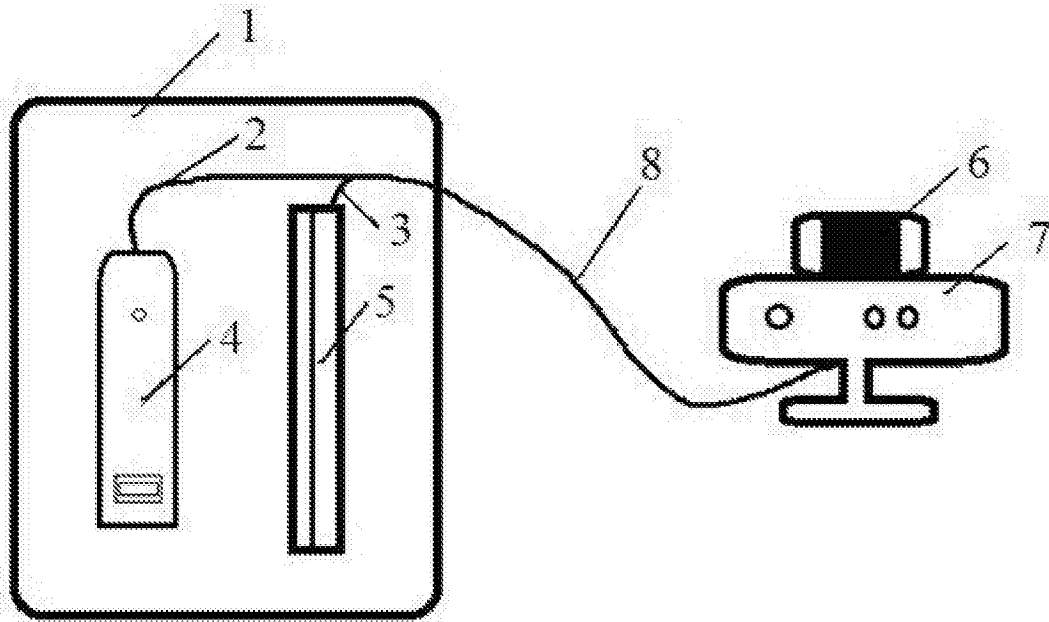


图2

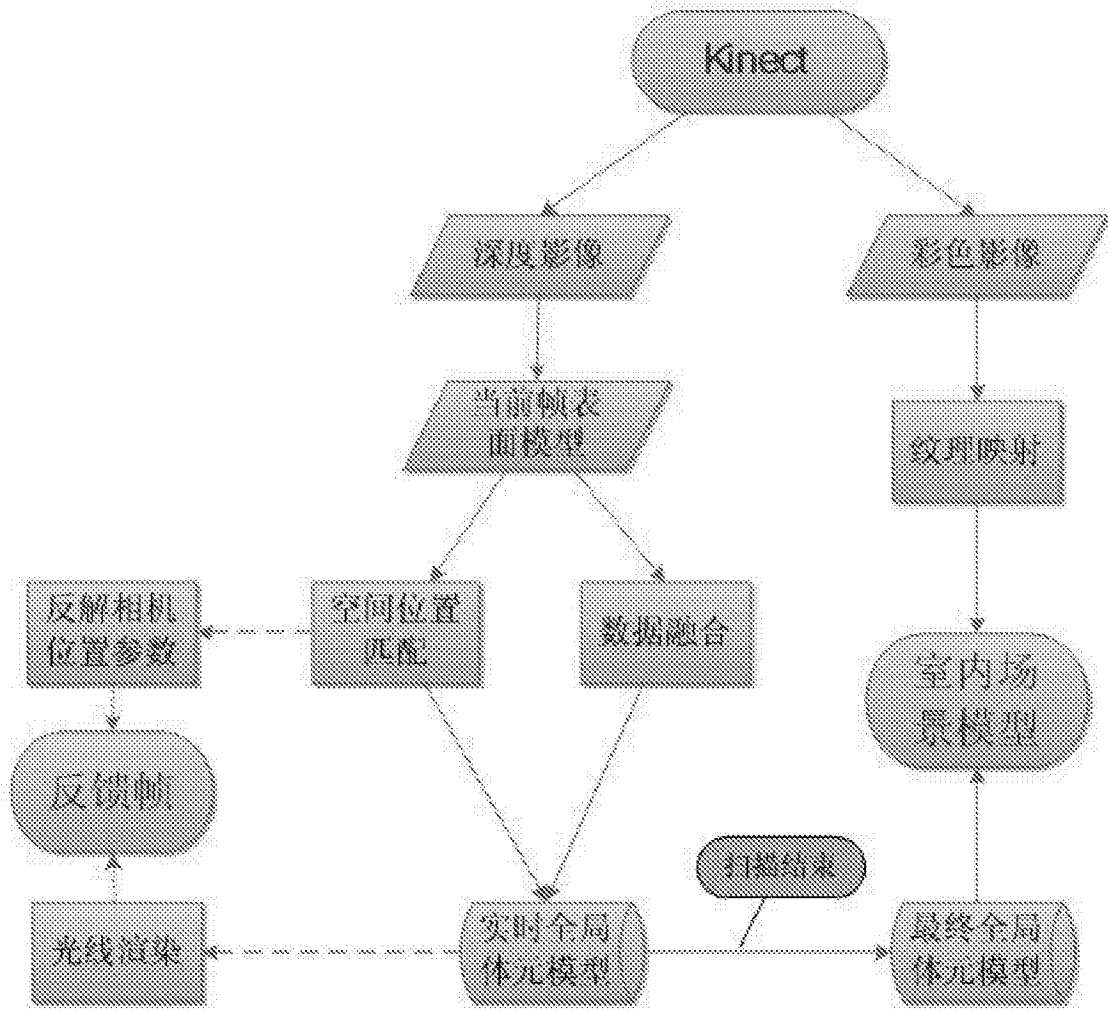


图3

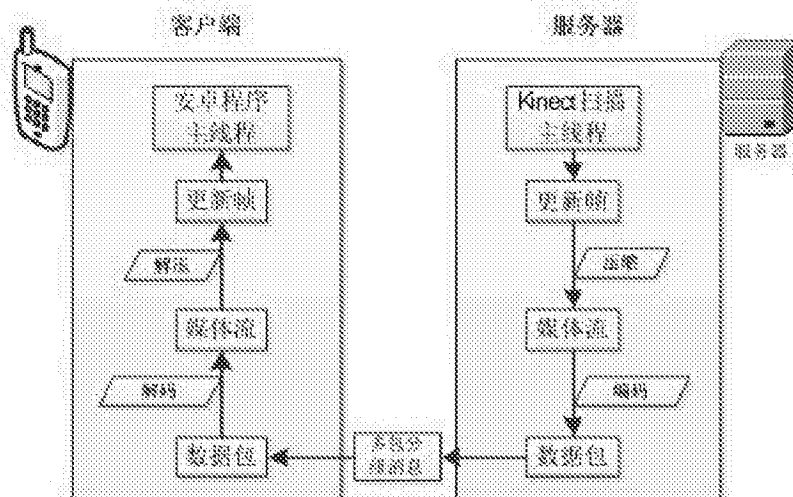


图4

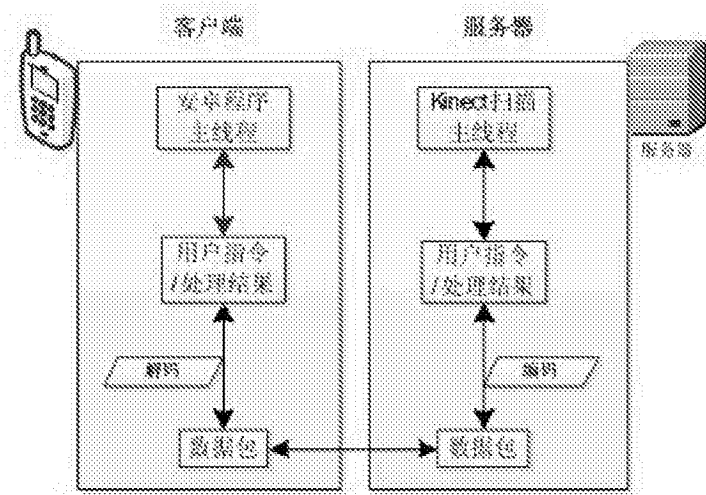


图5