



(12)发明专利

(10)授权公告号 CN 104376552 B

(45)授权公告日 2017.12.29

(21)申请号 201410480514.1

(22)申请日 2014.09.19

(65)同一申请的已公布的文献号

申请公布号 CN 104376552 A

(43)申请公布日 2015.02.25

(73)专利权人 四川大学

地址 610065 四川省成都市武侯区一环路
南一段24号

(72)发明人 李新胜 冯子亮

(51)Int.Cl.

G06T 7/33(2017.01)

(56)对比文件

CN 101706957 A,2010.05.12,

EP 2622572 A1,2013.08.07,

郭昌达.增强现实三维配准技术方法研究.

《中国优秀硕士学位论文全文数据库信息科技
辑》.2014,第I138-888页.

郭昌达.增强现实三维配准技术方法研究.
《中国优秀硕士学位论文全文数据库信息科技
辑》.2014,第I138-888页.

审查员 张文君

权利要求书1页 说明书11页 附图3页

(54)发明名称

一种3D模型与二维图像的虚实配准方法

(57)摘要

本发明为了解决二维图像与三维图形配准的虚实融合中计算效率低和需要预先标定相机内参数的问题,提供一种3D模型与二维可见光图像配准算法。它通过有多视几何关系的多幅图像,根据多视几何理论,用估计加优化的方法求解多个摄像机内参数(焦距)和外参数(外参数包括摄像机视点位置及旋转方向);再通过GPS位置变换三维恢复坐标下的摄像机视点到3D模型坐标的视点,这个3D模型是根据真实场景预先建好的;最后渲染显示三维模型的二维虚图,它是与二维图像是一致的。本方法可以将视点所观察到的三维场景(虚图)与图片配准,它只需预先测量得到摄像机位置,不需要进行摄像机的预标定,也不需要用户手动设置初始的模型对应于各幅二维图像的摄像机视角。相对传统配准方法,有较高的效率。

1. 一种3D模型与二维图像配准方法,其特征是具体步骤如下:

1) 对通过有多视几何关系的多幅图像求解摄像机内参数和外参数

1-1) 提取SIFT特征描述;

1-2) 再用KD-树进行特征点匹配;

1-3) 采用RANSAC框架通过8点法估计基础矩阵来去掉噪声匹配点对;

1-4) 选择共有特征点最多的两幅图像,再通过上步得到的基本矩阵F,按估计的摄像机参数算出本质矩阵E,估计一对摄像机之间的相对位置,得到摄像机位置旋转和平移参数;

1-5) 初始化摄像机内参数,将上述外参数以及内参数、所有特征点的投影坐标和世界坐标作为初始值,用稀疏Levenberg-Marquardt方法进行集束优化;

1-6) 再选择另一幅图片,根据已算出的空间特征点及特征点跟踪关系用RANSAC方法求解此摄像机投影矩阵,得到摄像机内外参数初始值,将其加入到优化框架下,转步骤1-5)继续优化;

1-7) 直到所有的图片处理完毕,得到所有摄像机的外参数,外参数为多幅图片视点的旋转平移的相对位置;

2) 变换三维重建坐标下的摄像机视点到3D模型坐标的视点

根据预先测量好的摄像机的经纬度坐标找到多个视点的相对位置到3D模型的相对变换关系,变换关系为旋转、平移、缩放,将视点位置变换到3D模型中;将测量的摄像机经纬度坐标转换成模型坐标系中的坐标时,采用高斯-克吕格投影(Gauss-Kruger projection);用最小二乘法求 $B_i H b_i = 0$ 中转换矩阵H的初始值,其中, b_i 为摄像机位置的归一化的相对三维坐标, B_i 为经纬度转换后的模型坐标系中的坐标;再用Levenberg-Marquardt非线性求得最优的转换矩阵H,在非线性的迭代中H保持转换后坐标轴的正交性和比例的一致性;

3) 在预先生成所监视场景的3D模型中设置变换后的视点,得到相应视点下三维图形渲染显示后的二维视图,即对应二维图像的已经配准的虚像。

一种3D模型与二维图像的虚实配准方法

技术领域

[0001] 本发明涉及计算机应用技术和计算机视觉中的增强现实领域,具体涉及一种3D模型与二维图像的虚实配准方法。

背景技术

[0002] 增强现实是指通过将计算机生成的虚拟场景、文字注释等信息实时、精确地叠加到使用者所观察到的真实世界景象中,对人的视觉系统进行延伸和扩充的一种技术。如何实时、精确地计算用户视点相对于真实世界的位置、姿态信息,并利用这些信息将虚拟场景正确地叠加到它所应处的位置上,即虚实配准,是增强现实系统中的关键问题之一。

[0003] 所以,虚实配准在增强现实中有极大的应用需求。同样,在三维空间进行二维图像与三维图形的虚实配准技术,也有着广泛的应用前景,但是它没有通用成熟配准框架。

[0004] 在三维空间进行虚实配准,常用的技术框架是通过多幅图片形成的多目视觉系统得到场景的稠密点云,通过稠密点云提取场景的三维几何特征,如平面、直线、特征点、曲线等,再与3D模型中相应的特征进行配准。这种方法的缺点一是生成稠密点云需要的图片多,一般要几十张图片,而在机场应用场景中难以布置如此多指向同一场景的摄像机;二是计算速度慢,根据情况不同,至少在10分钟以上。工程应用中,稠密点云常采用激光雷达高速测量的方法,一分钟可测得上百万的点云数据,但是激光雷达成本高,它主要用于场景3D模型建模,难与视频结合的虚实配准结合使用。并且,现有的二三维配准研究中,基本上都要求用户手动设置初始的配准参数,即模型的初始视角位置,然后再优化求解精确配准位置,如果不手动设置视角初始值,配准问题将变得无法求解。

[0005] 目前的解决方法,还有的需要预先标定相机。预标定方法可以提前测量摄像机所有的内参数,但在有的应用环境中(如:机场,山区等)较难测量实际标志物的三维世界坐标,即摄像机的外参数,尤其是摄像机旋转参数;针对可变焦距可旋转的PTZ (Pan-Tilt-Zoom) 相机就无法应用。

发明内容

[0006] 本发明要解决的技术问题是:为了解决二维图像与三维图形配准的虚实融合中计算效率低和需要预先标定相机内参数的问题,提供一种3D模型与二维可见光图像配准方法。此方法不需要预标定相机,而且计算时间在两三分钟以内,可以达到实用的要求。

[0007] 本发明解决其技术问题所采用的技术方案是:一种3D模型与二维图像配准方法,它通过有多视几何关系的多幅图像,用估计加优化的方法求解多个摄像机内参数(焦距)和外参数,外参数包括摄像机视点位置及旋转方向;再变换三维恢复坐标下的摄像机视点到3D模型坐标的视点,这个3D模型是根据真实场景预先建好的;最后渲染显示三维模型的二维虚图,它是与二维图像是一致的。

[0008] 本发明的创新点在于:1)不需要像以往方法一样预先标定相机的内参数;2)不需要其它激光雷达等复杂设备进行场景的三维测量,也不需要限制相机的个数,只需要简单

测量摄像机的位置,而摄像机的焦距和旋转方向靠多视几何关系自动计算出来;3)提出了一个自动配准的方法解决3D模型与二维图像的配准。

[0009] 方法具体步骤是:

[0010] 1)通过有多视几何关系的多幅图像求解摄像机内参数和外参数。

[0011] 1-1)提取SIFT特征描述(它具有尺度和旋转不变性)。

[0012] 1-2)再用KD-树进行特征点匹配。

[0013] 1-3)采用RANSAC框架通过8点法估计基础矩阵F,同时去掉了噪声匹配点对。

[0014] 1-4)选择共有特征点最多的两幅图像,再通过上步得到的基本矩阵F,按估计的摄像机参数算出本质矩阵E,计算出第一对摄像机之间的相对位置,得到摄像机位置旋转和平移参数。

[0015] 1-5)初始化摄像机内参数,将上述外参数以及内参数、所有特征点的投影坐标和世界坐标作为初始值,用稀疏Levenberg-Marquardt方法进行集束优化。

[0016] 1-6)再选择另一幅图片,根据已算出的空间特征点及特征点跟踪关系用RANSAC方法求解此摄像机投影矩阵,得到摄像机内外参数初始值,将其加入到优化框架下,转步骤1-5)继续优化。

[0017] 1-7)反复执行1-6)直到每一个摄像机(所有的图片)处理完毕,得到所有摄像机的外参数(多幅图片视点的旋转平移的相对位置)

[0018] 2)变换三维重建坐标下的摄像机视点到3D模型坐标的视点。

[0019] 根据预先好的测量摄像机的经纬度坐标找到多个视点的相对位置到3D模型的相对变换关系(旋转、平移、缩放),可以将视点位置变换到3D模型中。我们采用的方法是经纬度转换法,它分为2个步骤:

[0020] 2-1)将测量的摄像机经纬度坐标转换成3D模型坐标系中的XYZ坐标。

[0021] 2-2)设摄像机位置的归一化的相对三维坐标为 b_i ,经纬度转换后的欧氏坐标为 B_i ,则矩阵 H 是它们的转换的关系。对公式 $B_i H b_i = 0$ 用最小二乘法求转换矩阵 H 的初始值,再用Levenberg-Marquardt非线性求得更好的转换矩阵 H , H 需保持转换后坐标轴的正交性和比例的一致性。若给定多对匹配点可以解出线性的最小二乘解。进一步,需加入旋转、平移、等比例缩放约束再用Levenberg-Marquardt算法进行非线性优化,即得到最终 H 。

[0022] 3)在3D模型中设置变换后的视点,得到相应视点下三维图形渲染显示后的二维视图(如图2~7的b)所示),即对应二维图像的已经配准的虚像。

[0023] 本发明的有益效果是,本发明的一种3D模型与二维图像配准方法,它利用多视几何理论恢复3D模型在相对坐标中的视点,只需预先测量得到摄像机位置,将二维真实可见光图像与3D模型投影虚拟图进行自动配准。它不需要进行摄像机的预标定,也不需要用户手动设置初始的模型对应于各幅二维图像的摄像机视角,而是通过有多视几何关系的多幅图像,用估计加优化的方法求解摄像机内参数和外参数,再变换三维恢复坐标下的摄像机视点到3D模型坐标的视点位置及视角。配准后可以进行虚实图像融合,得到更佳的效果。一旦焦距或是摄像机内参、旋转方向发生变化,重新将本方法执行一次即可。

附图说明

[0024] 图1是一个示例中摄像机位置及世界坐标系下的特征点位置。图中:重建的摄像机

位置(*)及世界坐标系下的特征点(+)位置。

- [0025] 图2是一个示例中第1幅虚实图像对的视点配准结果。
 [0026] 图3是一个示例中第2幅虚实图像对的视点配准结果。
 [0027] 图4是一个示例中第3幅虚实图像对的视点配准结果。
 [0028] 图5是一个示例中第4幅虚实图像对的视点配准结果。
 [0029] 图6是一个示例中第5幅虚实图像对的视点配准结果。
 [0030] 图7是一个示例中第6幅虚实图像对的视点配准结果。

具体实施方式

[0031] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。

[0032] 本发明的方法的使用方式是:在用多台PTZ摄像机进行场景监视的环境中,预先生成监视环境的3D模型,再测量出各个摄像机的经纬度坐标。通过本发明的方法,找到摄像机在3D模型中对应的视点及相关摄像机参数,然后按上述参数进行3D模型渲染,得到配准视图。图片2至7给出了一个监视场景的配准的示例。

[0033] 本发明的方法首先要假定采用哪种摄像机模型,本发明中认为摄像机符合小孔成像的标准摄像机模型(Pinhole model),采用以下透视投影公式描述:

$$[0034] \quad X \cong K[R \mid -Rt_c]b$$

$$[0035] \quad K = \begin{bmatrix} a_r f & s & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

[0036] 式中, x, b 均为齐次坐标,分别为图像坐标和世界坐标。世界坐标系与摄像机坐标系之间的关系用一个旋转矩阵 R 和一个平移向量 t_c 来描述。 a_r 和 s 分别是摄像机的像素纵横比(Aspect ratio)和倾斜度(Skew); f 为等效焦距(像素为单位); (u_0, v_0) 为图像平面坐标中心点,即光轴与图像成像平面的交点。为了进一步简化模型,假设:

- [0037] (1) 像素纵横是同一个尺度, $a_r=1$ 。
 [0038] (2) 图像坐标X轴与Y轴相互垂直, $s=0$ 。
 [0039] (3) 图像没有畸变。
 [0040] (4) 当焦距变化时,假设光轴和图像平面坐标中心点不变,即 (u_0, v_0) 保持不变。

[0041] 由于现在摄像机的制作工艺已经很成熟,这些假设一般情况下均成立。

[0042] 在上述假设条件下,本发明的具体实施步骤如下:

[0043] 步骤1)通过有多视几何关系的多幅图像求解摄像机内参数和外参数

[0044] 步骤1-1)提取SIFT特征点

[0045] SIFT特征是图像的局部描述量,对尺度缩放、平移、旋转、噪声和亮度变化等具有良好的不变性,对仿射变换、视觉变化也保持一定的稳定性。特征维数较高(128维)。

[0046] SIFT算法提取特征点主要分4个步骤:

[0047] (1) 首先利用高斯差分(DoG)以及影像金字塔找出尺度空间中的极值点。

[0048] (2) 求亚像素精度的极值点。因为DoG算子会产生边缘响应,还要去除低对比度的关键点和不稳定的边缘响应点。

[0049] (3)将剩余的极值点定义为特征点后,开启一窗口,利用窗口内的影像梯度和幅值定义出特征点的方向。

[0050] (4)按方向进行旋转,生成SIFT特征向量。

[0051] 这个向量就是所求的SIFT特征描述符。SIFT特征向量已经去除了尺度变化、旋转等几何变形因素的影响,如归一化特征向量,则进一步减弱光照变化的影响。

[0052] 步骤1-2)特征匹配

[0053] 这里采用最近邻与次近邻的距离比率作为判断准则。最近邻特征点指与样本特征点欧氏距离最短的特征点,次近邻特征点是具有只比最近邻距离大的欧氏距离(即距离第二小)的特征点。

[0054] 设两幅图像中检测出的两个特征点向量分别为 $R_i=(r_{i,1}, r_{i,2}, \dots, r_{i,128})$ 和 $S_i=(s_{i,1}, s_{i,2}, \dots, s_{i,128})$,是它们的欧氏距离为:

$$[0055] \quad d(R_i, S_i) = \sum_{j=1}^n (r_{i,j} - s_{i,j})^2 \quad (1)$$

[0056] 设 S_i' 是次近邻特征点,给定阈值T,当

$$[0057] \quad \frac{d(R_i, S_i)}{d(R_i, S_i')} < T \quad (2)$$

[0058] 则认为 R_i, S_i 匹配,是同一特征,其中, $d(R_i, S_i)$ 为最近邻的特征距离, $d(R_i, S_i')$ 为次近邻的特征距离。可以调整T改变检测出的特征点的个数,一般取 $T=0.8$ 。本步骤采用的均为高清摄像机,图像分辨率达300万,所以特征点较多,可能排除的部分正确匹配对结果影响不大。

[0059] 一般考虑匹配采用穷举法进行遍历搜索,但SIFT描述符的维数较大,而且一幅300万像素的图像可能有10000个特征,特征点很多,如果采取一一对应查找最近邻域点的时间复杂度为 $O(n^2)$,很费时,效率较低。为了提高匹配速度,本发明采用构造特征点K-D树的方法,可以将查找时间复杂度降低到 $O(\log n)$ 。

[0060] 其步骤是:为第一幅图像特征描述符分别建立K-D树。虽然建立K-D树比较费时,但在特征点比较匹配时会大大提高效率。依次查找图像每个特征点 R_i ,在待第一幅图像的特征描述的K-D树上的最近邻 S_i 和次近邻特征点 S_i' ,若满足公式(2)则 R_i, S_i 就是一对匹配点对。

[0061] 步骤1-3)RANSAC框架下8点法估计基本矩阵F

[0062] 双目视觉系统的外极几何系统中,对于第一个摄像机像面上任意一个图像点 x_{i1} ,在第二个摄像机像面上对应的像点 x_{i2} 一定位于该摄像机上的极线 l_2 上。这就是所谓的“外极”约束。其中, b_i 为一个三维空间中的点, O_1 和 O_2 是两个摄像机的光心。 e, e' 是两个光心的直线与图像平面的交点,称为极点。

[0063] 外极约束可以用基本矩阵(Fundamental matrix)F公式表示:

$$[0064] \quad x_{i2}^T F x_{i1} = 0 \quad (3)$$

[0065] 其中, x_{i1} 和 x_{i2} 是图像点的齐次坐标。F因为包含了一个尺度因子,所以是一个秩为2的 3×3 矩阵,有8个参数。

[0066] 步骤1-3-1)8点法估计基本矩阵F

[0067] 根据前面的特征点匹配,得到不同图像帧之间的匹配点对后就可以计算两帧之间的F。令 $F=(f_{ij})$,设匹配点对坐标是 $x_{i1}=(u_{i1},v_{i1},1)^T$, $x_{i2}=(u_{i2},v_{i2},1)^T$,根据公式(3),可以写成如下形式:

$$[0068] \quad u_{i2}u_{i1}f_{11} + u_{i2}v_{i1}f_{12} + u_{i2}f_{13} + v_{i2}u_{i1}f_{21} + v_{i2}v_{i1}f_{22} + v_{i2}f_{23} + u_{i1}f_{31} + v_{i1}f_{32} + f_{33} = 0 \quad (4)$$

[0069] 只要给定大于8对的点对 x_{i1} , x_{i2} ,基本矩阵F即可由方程(4)解出。通常,可以将 $f_{11} \sim f_{33}$ 做为参数,而前面的系统组成矩阵A。

[0070] 线性方程组(4),给定8个以上的点对, A一般会是一个秩9的矩阵。这样的方程组只有零解。这时需要求解约束条件 $\|f\|=1$ 下的最小二乘解,即求解下列问题:

$$[0071] \quad \begin{cases} \min \|Af\| \\ \|f\|=1 \end{cases} \quad (5)$$

[0072] 为求该最小二乘问题的解A,将A奇异值分解为 $A=UDV^T$,取V的最后一个列向量为F的值,即 $f=V_9$,重新排列为 3×3 矩阵即为F。

[0073] 我们需要找到一个矩阵 \bar{F} 去逼近F作为基本矩阵的估计以保证F秩为2,即求解如下最优化问题:

$$[0074] \quad \begin{cases} \min \|F - \bar{F}\| \\ \text{rank}(\bar{F}) = 2 \end{cases} \quad (6)$$

[0075] 由上述方程得到的解是基本矩阵F的未用归一化坐标的最终估计。为了达到更高的精度,本步骤需对输入匹配点的图像坐标进行了归一化,得到 \bar{x} 后又解除了归一化。

[0076] 本步骤用RANSAC估计F时,每一轮迭代中 $N=8$ 。

[0077] 步骤1-3-2)用RANSAC估计F

[0078] RANSAC方法可以从有错误的匹配点对集合中估计基本矩阵F的鲁棒结果。其算法的主要步骤如下:

[0079] (1)用1-2)SIFT特征匹配的方法得到初始的匹配点对(其中一般有一些误匹配结果)。

[0080] (2)从第(1)步的匹配点对中,随机抽取8个点对,用1-3-1)基本矩阵F的8点算法计算得到基本矩阵F。

[0081] (3)用第(2)步结果F计算内点集合S(F)。

[0082] (4)如果内点集合S(F)中集合残差值 err 小于原先的最优内点集合 $S(F)_{best}$,则采用当前内点集合S(F)及对应的F,即将 $S(F)_{best}$ 设为S(F),对应的最优 F_{best} 设为F。

[0083] (5)当迭代次数小于2048次,回到第(2)步迭代。

[0084] (6)当迭代次数达到2048次,终止迭代估计。 F_{best} 是RANSAC迭代得到的最优结果。

[0085] 一对匹配点 (x_{i1}, x_{i2}) 是否为内点的判断标准是 x_{i2} 到由 x_{i1} 和F计算出的对应极线 l_2' 的距离,记为 $d(x_{i2}, l_2')$:

$$[0086] \quad d(x_{i2}, l_2') = \frac{x_{i2}^T F x_{i1}}{\|(l_2^1(1), l_2^1(2))\|}, l_2' = F x_{i1} \quad (7)$$

[0087] 同理,记 x_{i1} 到由 x_{i2} 和F计算出的对应极线 l_1' 的距离是 $d(x_{i1}, l_1')$ 。

[0088] 如果 $d(x_{i2}, l_2') + d(x_{i1}, l_1') < 4$, 则认为 (x_{i1}, x_{i2}) 是匹配的内点。

[0089] 内点集合 $S(F)$ 中集合残差值的定义如下:

$$[0090] \quad err = \sum_x (d(x_{i2}, l_2') + d(x_{i1}, l_1')) \quad (8)$$

[0091] 其中, $x \in S(F)$ 即 x 是所有的内点匹配对, err 是计算所有内点匹配到计算出的极线的距离之和。

[0092] 步骤1-3-3) 将所有内点作为输入再求 F 更优解

[0093] 得到上步RANSAC框架下最优的基本矩阵后, 为了得到更优的结果, 再采用1-3-1) (基本矩阵 F 的8点算法), 将所有匹配的内点(内点个数 ≥ 8)做为输入, 得到所有内点匹配意义下外极几何约束下的 F 优化解。

[0094] 步骤1-3-4) 用LM算法优化得到 F 最终结果

[0095] 得到上述内点匹配意义下外极几何约束下的 F 优化解后, 再以公式(8)为目标函数, 用Levenberg-Marquardt进行非线性优化, 得到更优的基本矩阵 F 。这就是RANSAC框架下8点法估计基本矩阵 F 的最终结果。

[0096] 步骤1-4) 估计本质矩阵 E

[0097] 如果已知内参数矩阵置, 像平面使用归一化坐标 x'_{i1}, x'_{i2} , 则称归一化坐标下的基本矩阵为本质矩阵。记为本质矩阵 E , 设双目视觉对应的摄像机内分别是 K_1, K_2 ,

$$[0098] \quad x_{i2}^T K_2^{-T} E K_1^{-1} x_{i1} = 0 \quad (9)$$

$$[0099] \quad F = K_2^{-T} E K_1^{-1} = K_2^{-T} [t] \times R K_1^{-1} \quad (10)$$

$$[0100] \quad E = [t] \times R \quad (11)$$

[0101] 基本矩阵 F 包含了两个摄像机的内参数矩阵 K 和摄像机本质矩阵 E 之间的相对关系。本质矩阵描述了两幅规范化图像间的极几何, 它与基本矩阵的不同之处是它仅与摄像机的外参数有关, 即 t, R 。

[0102] 步骤1-4-1) 估计本质矩阵 E

[0103] 同理, 本质矩阵也是秩2的, 但它仅有5个自由度, 所以求解此 E 矩阵有5点法、6点法、7点法至8点法。一般由8个图像点对应, 从特征点对应用公式(9)可建立关于本质矩阵的线性方程线性求解。由于公式的齐次性, 只能获得相差一个非零尺度因子意义下的 E 。

[0104] 对于实际图像, 常用的方法是将 E 的最小奇异值置为零, 用另外两个奇异值的均值作为两个相等的奇异值, 这样的矩阵作 E 最小二乘意义下的最佳近似。即分解 $E = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V^T$, E 的近似矩阵是:

$$[0105] \quad E = U \text{diag}((\sigma_1 + \sigma_2) / 2, (\sigma_1 + \sigma_2) / 2, 0) V^T$$

[0106] 步骤1-4-2) 由 F 估计本质矩阵 E

[0107] 由公式(10)可以推出基本矩阵 F 与本质矩阵的关系:

$$[0108] \quad E = K_2^T F K_1 \quad (12)$$

[0109] 假设图片的大小为宽度 w , 高度 h , 按照本发明设定的针孔摄像机投影模型, 则摄像机的内参数初始值为:

$$[0110] \quad K = \begin{bmatrix} w & 0 & w/2 \\ 0 & w & h/2 \\ 0 & 0 & 1 \end{bmatrix} \quad (13)$$

[0111] 本发明采用这种算法在只有2张图片时估计相机的本质矩阵和内参数,虽然这种模型过于理想,与实际情况有所差距,但是经过基于稀疏Levenberg-Marquardt的集束优化,优化结果可以弥补理想模型的误差。

[0112] 步骤1-4-3)由本质矩阵E估计出R和t

[0113] 如果第一个摄像机矩阵规定为 $M_1=(I,0)$,令 $E=Udiag(\sigma,\sigma,0_3)V^T$ 是它的一个奇异值分解,则第二个摄像机投影矩阵 $M_2=(R,t)$ 的解可能为四组:

$$[0114] \quad \begin{aligned} (1) [t]_x &= USU^T, R = UZV^T \\ (2) [t]_x &= US^T U^T, R = UZ^T V^T \\ (3) [t]_x &= USU^T, R = UZ^T V^T \\ (4) [t]_x &= US^T U^T, R = UZV^T \end{aligned}$$

$$[0115] \quad \text{其中, } S = \text{diag}(\sigma, \sigma, 0) \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

[0116] 四个解的几何解释中只有一个是合理的,即这个重构点同时位于两个摄像机前方的一对摄像机是合理的。

[0117] 步骤1-4-4)用三角法解出特征点的世界坐标

[0118] 由上述方法求解出每一个摄像机的本质矩阵 E_i 后,结合摄像机内参矩阵 K_i ,得到摄像机投影矩阵 M_i ,再解出每一个特征点的世界坐标。

[0119] 步骤1-5)基于稀疏Levenberg-Marquardt的集束优化摄像机外参数以及内参数、所有特征点的世界坐标

[0120] 设 x_{ij} 是第 i 个三维点投影到第 j 个图像上(由第 j 个视点看到)的坐标, a_j 是摄像机的第 j 个内外参数, b_i 是第 i 个世界坐标系下三维点的坐标。集束优化的目标函数是:

$$[0121] \quad (\hat{a}, \hat{b}) = \arg \min_{a, b} \sum_{i=1}^n \sum_{j=1}^m v_{ij} d(Q(a_j, b_i), x_{ij})^2 \quad (14)$$

[0122] $Q(a_j, b_i)$ 是第 i 个点投影到第 j 幅图像上的投影函数,可以由这个投影函数估计 x_{ij} ; $d(Q(a_j, b_i), x_{ij})$ 是由投影函数估计的二维图像中坐标和真实的二维图像坐标 x_{ij} 的欧氏距离; v_{ij} 表示第 i 个坐标点是否可以由第 j 个摄像机看到,即是否在第 j 幅图像中可见。

[0123] 如果摄像机内外参数有 k 个,而三维坐标点需3个坐标,则此集束优化的需求解的参数为 $mk+3n$ 个。本发明中摄像机的内外参数为 $k=9$ 个,分别是:摄像机在世界坐标系中的位置3个,表示摄像机平移;旋转分量3个,表示摄像机的朝向;缩放尺度1个;径向畸变参数2个。

[0124] 设求解参数的矢量为 $P, P=(a_1^T, a_2^T, \dots, a_m^T, b_1^T, b_2^T, \dots, b_m^T)^T$, 矢量 X 是图像上特征点的坐标,记为 $P=(x_{11}^T, x_{12}^T, \dots, x_{1m}^T, x_{21}^T, x_{22}^T, \dots, x_{2m}^T, \dots, x_{n1}^T, x_{n2}^T, \dots, x_{nm}^T)^T$,对图像上特征点坐标量测值 X ,有相应的通过投影函数计算得到的估计值 \hat{X} ,

$$\hat{X} = (\hat{x}_{11}^T, \dots, \hat{x}_{1m}^T, \hat{x}_{21}^T, \dots, \hat{x}_{2m}^T, \dots, \hat{x}_{n1}^T, \dots, \hat{x}_{nm}^T)^T, \hat{X} = Q(a_j, b_i)$$

[0125] 并且相应的误差为:

[0126] $\varepsilon = (\varepsilon_{11}^T, \dots, \varepsilon_{1m}^T, \varepsilon_{21}^T, \dots, \varepsilon_{2m}^T, \dots, \varepsilon_{n1}^T, \dots, \varepsilon_{nm}^T)^T, \varepsilon_{ij} = x_{ij} - \hat{x}_{ij}, \forall i, j$

[0127] 那么最优化问题变成了求解 $\sum_{i=1}^n \sum_{j=1}^m \|\varepsilon_{ij}\|^2$ 的最小值, $P = P^* = \arg \min_P \sum_{i=1}^n \sum_{j=1}^m \|\varepsilon_{ij}\|^2$, P^* 为所求的参数结果。

[0128] 设雅可比矩阵 $J = \frac{\partial \hat{X}}{\partial P}$, 将J分块为 $[A \mid B], A = \begin{bmatrix} \frac{\partial \hat{X}}{\partial a} \end{bmatrix}, B = \begin{bmatrix} \frac{\partial \hat{X}}{\partial b} \end{bmatrix}$ 。由此, LM算法的

更新向量 δ_p 可以写成 $(\delta_a^T, \delta_b^T)^T$, 则标准的迭代求解步骤 $J^T J \delta_p = J^T \varepsilon$ 变成:

[0129]
$$\begin{bmatrix} A^T A & A^T B \\ B^T A & B^T B \end{bmatrix} \begin{pmatrix} \delta a \\ \delta b \end{pmatrix} = \begin{pmatrix} A^T \varepsilon \\ B^T \varepsilon \end{pmatrix} \quad (15)$$

[0130] 实际的特征点量测位置可以组织成一维向量 X ,

[0131] $X = (x_{11}^T, x_{12}^T, \dots, x_{1m}^T, x_{21}^T, x_{22}^T, \dots, x_{2m}^T, \dots, x_{n1}^T, x_{n2}^T, \dots, x_{nm}^T)^T$

[0132] 则参数向量 P 为 $P = (a_1^T, a_2^T, \dots, a_m^T, b_1^T, b_2^T, \dots, b_n^T)^T$

[0133] 更新向量 δ_p 为 $\delta_p = (\delta_{a1}^T, \delta_{a2}^T, \dots, \delta_{am}^T, \delta_{b1}^T, \delta_{b2}^T, \dots, \delta_{bn}^T)^T$ 。

[0134] 定义 $A_{ij} = \frac{\partial \hat{X}_{ij}}{\partial a_j}, B_{ij} = \frac{\partial \hat{X}_{ij}}{\partial b_i}$, 则雅可比矩阵 J 可以写成:

[0135]
$$J = \frac{\partial \hat{X}}{\partial P} = \begin{matrix} & a_1 & a_2 & \dots & a_m & b_1 & b_2 & \dots & b_n \\ \begin{matrix} x_{11} \\ x_{12} \\ \dots \\ x_{1m} \\ \dots \\ x_{n1} \\ x_{n2} \\ \dots \\ x_{nm} \end{matrix} & \begin{pmatrix} A_{11} & 0 & 0 & 0 & B_{11} & 0 & 0 & 0 \\ 0 & A_{12} & 0 & 0 & B_{12} & 0 & 0 & 0 \\ \dots & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{1m} & B_{1m} & 0 & 0 & 0 \\ \dots & & & & & & & & \\ A_{n1} & 0 & 0 & 0 & 0 & 0 & 0 & B_{n1} \\ 0 & A_{n2} & 0 & 0 & 0 & 0 & 0 & B_{n2} \\ \dots & 0 & 0 & \dots & 0 & 0 & 0 & B_{n3} \\ 0 & 0 & 0 & A_{nm} & 0 & 0 & 0 & B_{nm} \end{pmatrix} \end{matrix}$$

[0136] 近似计算Hessian矩阵 $J^T J$ 分块为:

[0137]
$$J^T J = \begin{matrix} & a_1 & a_2 & \dots & a_m & b_1 & b_2 & \dots & b_n \\ \begin{matrix} a_1 \\ a_2 \\ \dots \\ a_m \\ b_1 \\ b_2 \\ \dots \\ b_n \end{matrix} & \begin{pmatrix} U_1 & 0 & 0 & 0 & W_{11} & W_{21} & \dots & W_{n1} \\ 0 & U_2 & 0 & 0 & W_{12} & W_{22} & \dots & W_{n2} \\ \dots & 0 & 0 & \dots & 0 & \dots & \dots & \dots \\ 0 & 0 & 0 & U_m & W_{1m} & W_{2m} & \dots & W_{nm} \\ W_{11}^T & W_{12}^T & \dots & W_{1m}^T & V_1 & 0 & 0 & 0 \\ W_{21}^T & W_{22}^T & \dots & W_{2m}^T & 0 & V_2 & 0 & 0 \\ \dots & \dots & \dots & \dots & 0 & 0 & \dots & 0 \\ W_{n1}^T & W_{n2}^T & \dots & W_{nm}^T & 0 & 0 & 0 & V_n \end{pmatrix} = \begin{pmatrix} U & W \\ W^T & V \end{pmatrix} \quad (16)$$

[0138] 其中, $U_j = \sum_{i=1}^n A_{ij}^T A_{ij}$, $V_j = \sum_{i=1}^n B_{ij}^T B_{ij}$, $W_{ij} = A_{ij}^T B_{ij}$, U 和 V 是对角块矩阵, W 是稀疏矩阵。

[0139] 因此公式(15) (16)引入 λ 后变成:

$$[0140] \begin{pmatrix} U^* & W \\ W^T & V^* \end{pmatrix} \begin{pmatrix} \delta a \\ \delta b \end{pmatrix} = \begin{pmatrix} \varepsilon_a \\ \varepsilon_b \end{pmatrix} \quad (17)$$

[0141] 实际计算时先 U , V , W ,再用 $1+\lambda$ 乘 U 和 V 的对角元素。

[0142] 计算逆矩阵 V^{-1} 并定义 $Y=W V^{-1}$ 。求解 $(U-WY^T) \delta a = \varepsilon_a - Y \varepsilon_b$,求出 δa ,因为 a 中的参数个数远少于 b 中的参数个数。求解 $\delta b = V^{-1} (\varepsilon_b - W^T \delta a)$,求出 δb 。通过增加更新矢量 $(\delta a^T, \delta b^T)^T$ 更新参数矢量 P ,并计算新的误差矢量 ε 。如果新的误差矢量 ε 小于老的误差,则接受这个新的参数值,把 λ 减少至 $1/10$,并且返回 U , V , W 的计算继续,否则终止计算。如果新的矢量 ε 大于老的误差,把 λ 增大10倍,并且继续计算实际计算时先 U , V , W ,再反复进行迭代计算。

[0143] 步骤1-6)RANSAC框架下单摄像机投影矩阵估计

[0144] 上述的步骤已经可以从双目视觉系统中求得两个摄像机的相对位置 R , t_c ,但是对于多目视觉中,如果有更多的摄像机指向了相同的场景,应该充分利用更多的立体视觉信息,得到更精确的视点相对位置。

[0145] 第一步,通过提取SIFT特征点,得到多视图的特征点对应关系,即

[0146] $(x_{i1}, x_{i2}, x_{i3}, \dots, x_{im})$ 。第二步,对于第3个摄像机,再根据 (x_{i1}, x_{i2}) 双目视觉得到的三维世界坐标 b_i 与 x_{i3} ,利用RANSAC框架求解摄像机投影矩阵 M_3 ,再解得第3个摄像机的求相对位置 R_3 , t_{c3} 。这样就得到了第3个摄像机相对位置的初始值。第三步,将代入基于稀疏Levenberg-Marquardt的集束优化方法即可得到第3个摄像机相对位置的优化结果。第四步,对于加入更多的指向同一场景的摄像机,可以采用继续采用上述几步计算相应摄像机的相对位置。

[0147] 下面分5个步骤描述本步骤1-6)所应用的方法:

[0148] 步骤1-6-1)6点法求得摄像机矩阵 M

[0149] 在经典立体视觉中标定方法是根据一些空间点在世界坐标系下的坐标 b 与其图像坐标 x 之间的对应,即 $x \sim Mb$,建立约束方程(14),再摄像机矩阵 M 。这里世界坐标的空间点 b_i 由 (x_{i1}, x_{i2}) 已经由上述的双目视觉计算初始值并迭代优化得到。

[0150] 在基于多幅图像的场景重建问题中,投影矩阵的估计是非常重要的环节。 x 与 b 关系表示为如下公式:

$$[0151] x \times Mb = 0 \quad (18)$$

[0152] 不妨设已知 n 组对应点 (b_i, x_i) , $x_i = (u_i, v_i, 1)^T$, $M = \begin{pmatrix} M_1^T \\ M_2^T \\ M_3^T \end{pmatrix}$, M_i^T 表示矩阵 M 的第 i

行向量。代入公式(18),消去尺度因子常数后,可得到下述公式:

$$[0153] A_i p = 0 \quad (19)$$

$$[0154] \quad \text{其中, } p = (M_1^T \quad M_2^T \quad M_3^T)^T, \quad A_i = \begin{pmatrix} 0 & b_i^T & -v_i b_i^T \\ b_i^T & 0 & u_i b_i^T \\ v_i b_i^T & -u_i b_i^T & 0 \end{pmatrix},$$

[0155] 给定n个点(6个或更多)对应,得到n个形如这样的矩阵 A_i ,再将这n个矩阵 A_i 组合起来得到一个 $3n \times 12$ 的矩阵 $A = (A_1^T, A_2^T, \dots, A_n^T)$,对A作奇异值分解 $A = UDV^T$,则V的最后一个列向量 $p = V_{12}$ 是方程 $Ap = 0$ 的最小二乘解,再将p写成 3×4 矩阵的形式就是摄像机矩阵M。

[0156] 步骤1-6-2)用RANSAC估计M

[0157] 本发明采用RANSAC估计世界坐标系下的坐标b与其图像坐标x之间的摄像机矩阵M,可以得到更好的结果。

[0158] (1)用1-1)步SIFT特征匹配的方法得到第j幅图片与前j-1幅图片的初始的匹配点对(同样,其中一般有一些误匹配结果)。

[0159] (2)从第(1)步的匹配点对中,随机抽取6个点对 (x_{ij}, b_i) ,用求摄像机矩阵M的6点法计算得到第i幅图片对应的摄像机矩阵 M_j 。

[0160] (3)用第(2)步结果 M_j 计算内点集合 $S(M_j)$ 。

[0161] (4)如果内点集合 $S(M_j)$ 中集合中内点的个数大于原先的最优内点集合 $S(M_j)_{best}$,则采用当前内点集合 $S(M_j)$ 及对应的 M_j ,即将 $S(M_j)_{best}$ 设为 $S(M_j)$,对应的最优 M_{jbest} 设为 M_j 。

[0162] (5)当迭代次数小于4096次,回到第(2)步迭代。

[0163] (6)当迭代次数达到4096次,终止迭代估计。 M_{jbest} 是第i个摄像机RANSAC迭代得到的最优结果。

[0164] 判断 (x_{ij}, b_i) 是否是内点的标准是 b_i 经过投影到图像上与像素坐标 x_{ij} 的欧氏距离,表示为 $d(x_{ij}, M_j b_i) = \|x_{ij} - M_j b_i\|$,如果二者之间的距离小于4个像素,则 (x_{ij}, b_i) 是内点。

[0165] 步骤1-6-3)将所有内点作为输入再求 M_j 更优解

[0166] 得到上步RANSAC框架下最优的摄像机矩阵 M_j 后,为了得到更优的结果,再采用6点法,将所有匹配的内点(内点个数 >6)做为输入,得到所有内点匹配下的 M_j 最小二乘优化解。

[0167] 步骤1-6-4)用LM算法优化得到M最终结果

[0168] 得到上述所有内点匹配下的 M_j 最小二乘优化解后,再以最大化内点集合中内点个数为目标函数,用Levenberg-Marquardt进行非线性优化,得到更优的摄像机矩阵 M_j 。这就是RANSAC框架下单摄像机投影矩阵估计 M_j 的最终结果。

[0169] 步骤1-6-5)根据投影矩阵 M_j 求相对位置

[0170] 本发明的方法通过投影矩阵M采用RQ分解求解摄像机内参数、外参数。

[0171] 将M表示成 $M = [m_1, m_2, m_3, m_4]$, m_i 表示M的第i列。令摄像机内参数矩阵为K,外参数矩阵为 (R, t_c) ,由于此时的摄像机矩阵是欧氏的,所以可写成:

$$[0172] \quad M = \kappa K [R \mid -Rt_c]$$

[0173] 首先,按下列公式计算出 t_c 。

$$[0174] \quad t_c = \frac{[\det(M_{234}) - \det(M_{134}) \det(M_{124}) - \det(M_{123})]}{-\det(M_{123})}$$

[0175] 上式中的 M_{123} 表示矩阵 $[m_1, m_2, m_3]$, m_i 表示M的第i列,其它 M_{234} 等为相同的意义。

[0176] 再利用RQ分解, $[K R]=RQ(M_{123})$, K 是对角元素均为正数的上三角矩阵, R 为正交旋转矩阵, $M_{123}=KR$ 。

[0177] 这样,就得到了摄像机外参数 (R, t_c) 。

[0178] 步骤1-7) 反复执行直到所有的图片处理完毕

[0179] 反复执行步骤1-5)~1-6) 直到每一个摄像机所对应的图片处理完毕,得到摄像机的内外参数后,优化求出特征点的三维世界坐标后,其一个示例结果如图2~7所示。

[0180] 步骤2) 图像摄像机视点到3D模型的视点配准

[0181] 根据预先好的测量摄像机的经纬度找到多个视点的相对位置到3D模型的相对变换关系(旋转、平移、缩放),可以将视点位置变换到3D模型中。我们采用的方法是经纬度转换法,它分为2个步骤:

[0182] 步骤2-1) 将测量的摄像机经纬度坐标转换成模型坐标系中的坐标,采用高斯-克吕格投影(Gauss-Kruger projection)。

[0183] 步骤2-2) 对公式(20)用最小二乘法求转换矩阵 H 的初始值,再用Levenberg-Marquardt非线性求得更好的转换矩阵 H , H 保持转换后坐标轴的正交性和比例的一致性,它仅由平移,旋转,缩放的7个参数确定。设摄像机位置的归一化的相对三维坐标为 b_i ,经纬度转换后的欧氏坐标为 B_i ,则它们与转换矩阵 H 的关系是:

$$[0184] \quad B_i H b_i = 0 \quad (20)$$

[0185] 若给定多对匹配点可以解出线性的最小二乘解。进一步,需加入旋转、平移、等比例缩放约束再用Levenberg-Marquardt算法进行非线性优化,即得到最终 H 。

[0186] 步骤3) 在3D模型中设置变换后的视点并进行渲染

[0187] 在3D模型中设置变换后的视点,得到相应视点下三维图形渲染显示后的二维视图(如图2~7中的b)图所示),即对应二维图像的已经配准的虚像。这时三维图形渲染显示后的二维视图与图像已经配准,方法执行完毕。

[0188] 经过本步骤的集束优化后,得到的所有摄像机内外参数和世界坐标系下的特征点位置,就是最后的结果。

[0189] 以上述依据本发明的理想实施例为启示,通过上述的说明内容,相关工作人员完全可以在不偏离本项发明技术思想的范围内,进行多样的变更以及修改。本项发明的技术性范围并不局限于说明书上的内容,必须要根据权利要求范围来确定其技术性范围。

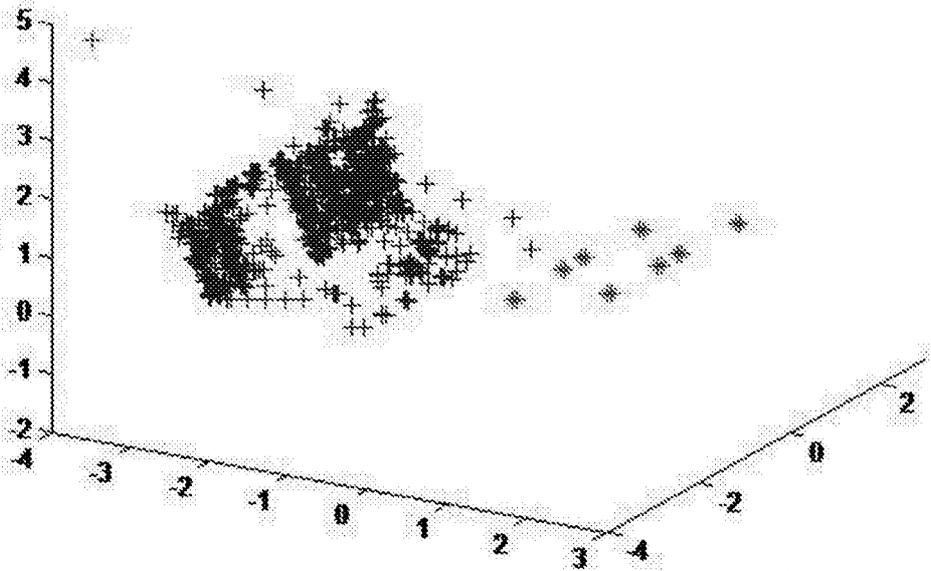


图1

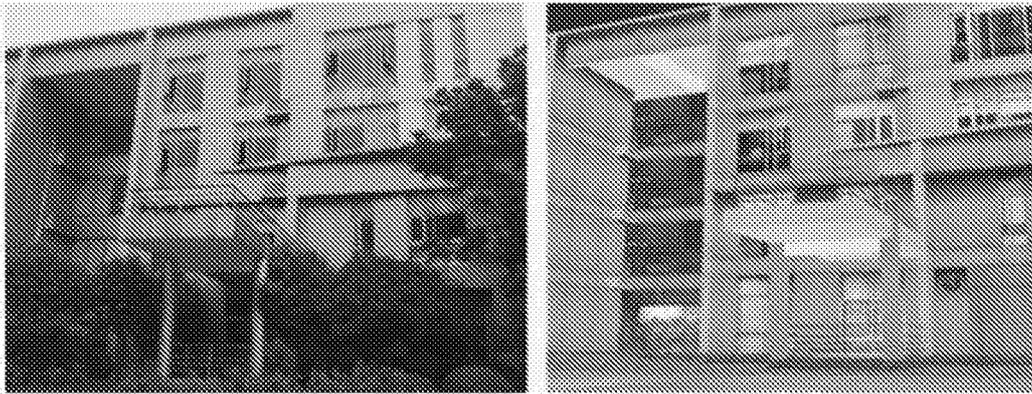


图2



图3



图4

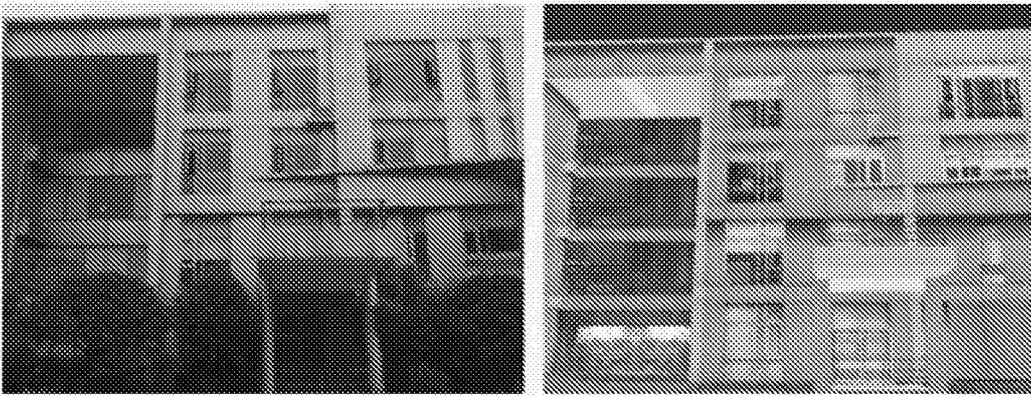


图5



图6



图7