

面向 AR 虚拟沙盘的虚拟化身定位定向技术研究

余田甜¹, 王晓萍², 李蔚清³, 苏智勇¹

(1. 南京理工大学 自动化学院, 江苏 南京 210094; 2. 北方信息控制研究所集团有限公司, 江苏 南京 211103;
3. 南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

摘要: 在利用增强现实虚拟沙盘进行多人异地协同指挥或研讨时, 本地操作者需要知道异地操作者的具体方向与位置来避免方向错误、位置冲突以更好地与对方交流。针对此问题, 设计了适合 AR 环境的三维标记物, 研究并实现了基于视觉的定位定向算法。采用 PnP 算法对标记物进行位姿估计, 使用面积比值法计算方位角。实验结果表明该算法可以实现快速精确地定位定向, 能够满足 AR 虚拟沙盘实际应用需求。

关键词: 增强现实; 定位定向; 三维标记物; PnP 算法

中图分类号: TP391.9

文献标志码: A

文章编号: 1004-731X (2021) 02-0324-07

DOI: 10.16182/j.issn1004731x.joss.19-0420

Research on Virtual Avatar Position and Orientation Technology for Augmented Reality Virtual Sandbox

Yu Tiantian¹, Wang Xiaoping², Li Weiqing³, Su Zhiyong¹

(1. School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China;

2. Northern Information Control Research Institute Group Co., Ltd, Nanjing 211103, China;

3. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: When using augmented reality (AR) virtual sandbox for multi-person collaborative command or discussion, the local operator needs to know the specific direction and location of the remote operator to avoid wrong direction, position conflict and to better communicate with the other. Aiming at this problem, a three-dimensional marker suitable for AR environment is designed, and a vision-based position and orientation algorithm is researched and implemented. The PnP algorithm is used to estimate the pose of the marker, and the area ratio method is used to calculate the azimuth. The experimental results show that the proposed algorithm can achieve fast and accurate position and orientation, meeting the practical application requirements of AR virtual sandbox.

Keywords: augmented reality; position and orientation; three-dimensional marker; PnP algorithm

引言

增强现实是一种实时计算相机位置及角度, 可以将真实的物体或场景与虚拟的 3D 模型叠加在一起, 实现对真实场景有效扩充和增强的技术^[1]。AR 技术可以广泛地应用于军事、教育、医疗、销

售、救援抢险、维修装配等领域。本文的研究背景是增强现实虚拟沙盘。指挥人员通过头戴式 AR 终端实现多人协同指挥与研讨。AR 远程协作可以跨越空间的限制, 使人们能“面对面”地交流或进行远程技术指导等。为了保障 AR 虚拟沙盘在多人异地作战指挥中发挥作用, 让本地操作

收稿日期: 2019-08-15

修回日期: 2019-11-03

基金项目: “十三五”装备预研基金(61409230104, 315100104, 31511040202), 中央高校基本科研业务费专项资金(30918012203), 上海航天科技创新基金(SAST2019009)

第一作者: 余田甜(1996-), 女, 硕士生, 研究方向为增强现实、计算机视觉。Email: 2531219218@qq.com

者实时观察异地操作者的位置及方向, 以免产生位置冲突, 需要在 AR 协同环境中对异地操作者的虚拟化身进行准确定位和定向, 从而实现指挥人员对虚实融合环境的准确感知, 提高指挥效率。

为了实现室内 AR 环境中的定位定向, 目前主要采用 WiFi 定位技术以及基于视觉的定位定向方法。Yadav^[2]提出一种基于 AR 的室内定位定向系统, 通过 WiFi 指纹实现定位、数字罗盘实现定向并在 AR 环境中显示出定位定向的结果信息。Jarkko Polvi 等^[3]提出一种应用于手持 AR 设备的三维定位方法 SlidAR, 依靠光线投射和触摸屏手势实现虚拟物体的定位, 解决了无标记或者自主建图困难情况下三维物体定位的问题。Jens Muller 等^[4]针对 AR 环境中协作者交互需要定位定向的问题, 通过在 AR 环境中添加虚拟地标的方法, 有效地实现了目标的定向。SLAM 是常用的无标记定位技术, Ruwan 等^[5]针对多个目标在 AR 环境中相对位置未知时的定位定向问题, 提出一种协调 AR 架构, 每个目标作为独立的节点进行地图构建并在此过程中确定各自的位置。相机姿态测量是 AR 中实现基于标记的跟踪注册技术的必要手段, 相机的方向位置信息为注册虚拟物服务^[6]。桂振文等^[7]提出一种基于二维码的移动 AR 系统, 对大规模场景中的二维码标记进行跟踪并实时测量其姿态信息。刘经纬^[8]针对三维标志物跟踪算法的要求, 在标志物设计阶段就对可能的标志物识别和跟踪算法进行研究, 采用 POSIT 算法确定三维标志物在空间中的位姿从而实现定向。

本文研究虚拟化身在 AR 虚拟沙盘环境中的定位定向问题。首先目标是虚拟的, 无法安装实物传感器; 其次虚拟沙盘在使用时场景是变化的, 无法利用虚拟沙盘图像做特征定位, 此外实际应用时虚拟沙盘还应该能够根据需求调整摆放位置, 这会导致图像特征丢失, 之前的位置信息失效, 需要重新定位。目前尚无学者面向 AR 虚拟沙盘做过虚拟化身定位定向的研究, 前面提到的这些方法基本上可以实现 AR 环境中对特定目标的定位或者定向, 但

也存在着一些局限性:

- (1) 用于定向的传感器如数字罗盘等无法应用到本文的虚拟化身上;
- (2) 从固定场景中选择图像特征进行定位, 不适用于虚拟沙盘这种变化的场景;
- (3) 在某一固定位置才能打开 AR 场景, 无法满足用户在任意地点打开 AR 场景的需求。

根据需求分析, 本文设计了合适的虚拟三维标记物并对其检测识别与解码, 然后利用 PnP 算法进行位姿估计, 计算出相机(即人)相对于虚拟沙盘的距离及方向, 在实际 AR 虚拟沙盘环境下进行了实验, 结果表明本文方法的精度能满足实际应用的要求。

1 三维标记设计与识别

1.1 三维标记物的设计

设计标记物的目的在于借助虚拟标记物与虚拟沙盘的固定关系来解算出操作者相对于虚拟沙盘的方向和位置信息。本文将虚拟标记物置于虚拟沙盘正上方以避免其遮挡沙盘物体, 并设计成三维标记以实现 360° 的定位定向。

图 1 所示为三维标记物的 4 个面分别对应东南西北 4 个方向, 每个面分别是一个二维标记。它们都是由 7×7 块小方格组成, 二维标记的实际尺寸为 25 cm×25 cm。小方格均为黑色或者白色, 外围的 24 个小方格均设计为黑色, 便于定位标记物; 内部 5×5 的方形区域包含编码信息。虚拟沙盘和标记物如图 2 所示, 图中标记物和沙盘都是虚拟的。相对于传统的二维标记而言, 本文设计的三维标记识别速度快, 识别范围广且鲁棒性高, 标记的设计简单。

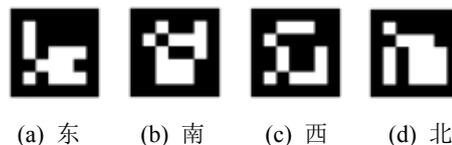


图 1 二维标记

Fig. 1 Two-dimensional marker

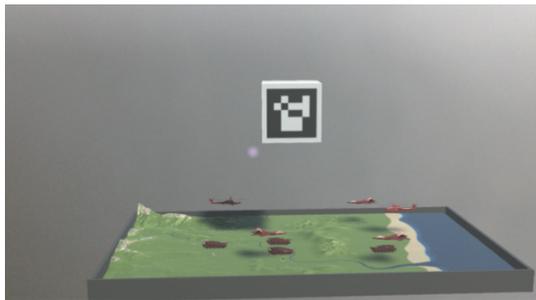


图 2 虚拟沙盘和虚拟标记物
Fig. 2 Virtual sandbox and virtual marker

1.2 三维标记物的识别

为满足定向要求, 标记识别算法需要能够同时检测多个标记。标记检测的流程图如图 3 所示。

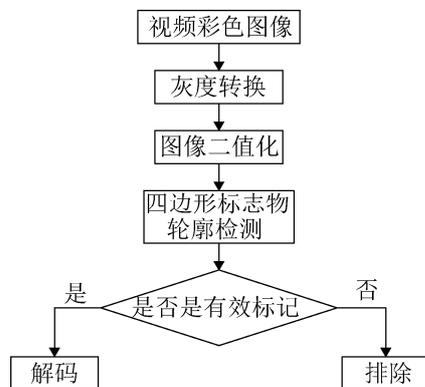


图 3 三维标记物检测流程
Fig.3 Flow chart of detection for 3D marker

标记检测的工作是找出可能的标记并进行筛选。检测输入的二值图像的轮廓, 返回一个多边形列表。每个多边形表示一个轮廓, 保留大轮廓, 舍去面积较小的小轮廓, 因为小轮廓可能是真实标记内部的一部分。对每一个轮廓进行四边形逼近, 然后判断是否是凸四边形, 排除不满足条件的情况。判断四边形各顶点之间的距离是否满足要求(四边形是否足够大)。以上条件筛选过后, 仍可能会有 2 个四边形距离较近、一个包含着另一个或者由于边界不清晰导致重复检测的情况, 比较这 2 个四边形的周长, 保留周长大的。对检测得到的标记进一步验证是否为本文设计的标记。首先校正检测出来的标记, 通过四边形顶点对得到透视变换矩阵, 将标记图像变换为正方形

图像; 然后利用 Otsu 算法进行二值化处理, 最后检查标记边界是否为黑色, 对满足条件的标记再进行解码。

每一种标记都有唯一的可以用来代表自己的编码, 内部 5×5 的方形区域每一行都是用 5 位(bit) 二进制表示, 其中两位(第 2 位和第 4 位)是信息位, 剩下 3 位是错误检测码。这种编码方式与海明码类似, 但区别在于这种编码方式会对第一位取反, 因为如果编码为 0, 则标记编码为 10 000, 这样就防止一个全黑的正方形被检测为一个有效标记。

由图 4 可见, 同一幅图由于相机视点不同会有 4 种不同的表示, 必须从这 4 种可能的情况中找到正确的标记方向。每 2 个信息位有 3 个奇偶校验位, 校验位能对每个可能的标记定位, 计算海明距离。海明距离为 0 说明是正确的标记方向, 而旋转成其他方向的标记其海明距离不为 0。将识别出来的内部 5×5 的网格区域矩阵最多旋转 3 次, 计算每一次的海明距离, 若海明距离为 0, 则可计算出该标记的编码即 ID; 若不为 0, 则这是一个错误标记。



图 4 4 种可能的标记方向
Fig. 4 Four possible directions of markers

2 定位定向算法

在三维空间中可以通过标记角点的精确位置来估算相机与标记物之间的相对变换关系, 这个过程称为二维到三维的姿态估计, 多用于跟踪注册技术^[9]。姿态估计旨在寻找一个物体与相机之间的变换, 由旋转变换和平移变换组成。接着由旋转和平移变换矩阵就可以求解相机相对于标志物的方位和距离。此外, 本文还运用面积比值法这种几何方法求解方位角。

2.1 定向算法原理

在可供多人异地同时操作的虚拟沙盘中, 沙盘

有东南西北 4 个方向, 操作者需要知道异地协作者的具体方位, 以防位置冲突。本文定向的需求就是求解相机(即操作者)相对于标记物的方向角。

在第 1 部分已经设计并识别了有 4 种不同 ID 的标记, 他们分别对应东南西北 4 个方向。考虑到在识别标记过程中会有两种情况出现, 一种是只识别出 1 个标记的情况, 即相机方向接近于正方向, 偏离正方向的角度小; 另一种是能同时识别出 2 个标记, 即相机偏离正方向的角度较大。本文中第一种情况用 PnP 姿态估计求解出来的变换关系计算平面小角度; 第二种情况用几何方法计算角度。

已知三维空间中若干个点的坐标以及它们在图像中所对应的二维点坐标, 将它们一一组合成点对, 通过几何关系可求解相机坐标系和目标坐标系之间的转换关系, 即 PnP(Perspective-n-Point)问题。3D-2D 方法可以在匹配点较少时获得较好的姿态估计而不需使用对极约束, 因此 PnP 是一种重要的运动姿态估计算法^[10]。如图 5 所示, 标记的 4 个角点作为特征点, 在获得 4 个角点在图像、世界坐标系的坐标后, 便可得到相机与目标的相对位姿关系。

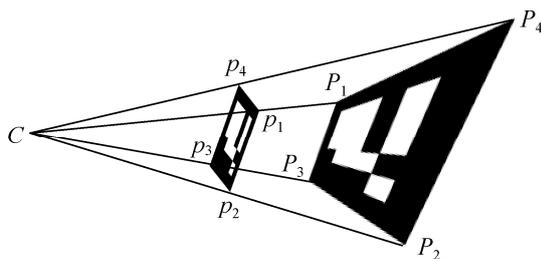


图 5 PnP 原理图
Fig. 5 PnP principle

图 5 中 C 表示相机中心, 世界坐标系中的三维点 $P_1 \sim P_4$ 在成像平面的投影点分别是二维点 $p_1 \sim p_4$ 。容易看出图中有 4 对三角形存在相似关系, 利用余弦定理可以得到式(1):

$$\begin{cases} CP_1^2 + CP_2^2 - 2CP_1 \cdot CP_2 \cdot \cos\langle CP_1, CP_2 \rangle = P_1P_2^2 \\ CP_3^2 + CP_2^2 - 2CP_3 \cdot CP_2 \cdot \cos\langle CP_3, CP_2 \rangle = P_3P_2^2 \\ CP_3^2 + CP_4^2 - 2CP_3 \cdot CP_4 \cdot \cos\langle CP_3, CP_4 \rangle = P_3P_4^2 \\ CP_1^2 + CP_4^2 - 2CP_1 \cdot CP_4 \cdot \cos\langle CP_1, CP_4 \rangle = P_1P_4^2 \end{cases} \quad (1)$$

对式(1)中四式同时除以 CP_4 , 并且令:

$$\begin{aligned} x &= \frac{CP_1}{CP_4}, y = \frac{CP_2}{CP_4}, z = \frac{CP_3}{CP_4}, m = \frac{P_1P_2^2}{CP_4^2}, \\ mn &= \frac{P_3P_2^2}{CP_4^2}, mk = \frac{P_3P_4^2}{CP_4^2}, mp = \frac{P_1P_4^2}{CP_4^2} \end{aligned} \quad (2)$$

整理成式(3):

$$\begin{aligned} x &= \frac{CP_1}{CP_4}, y = \frac{CP_2}{CP_4}, z = \frac{CP_3}{CP_4}, m = \frac{P_1P_2^2}{CP_4^2}, \\ mn &= \frac{P_3P_2^2}{CP_4^2}, mk = \frac{P_3P_4^2}{CP_4^2}, mp = \frac{P_1P_4^2}{CP_4^2} \end{aligned} \quad (3)$$

已知图像中 2D 点的位置以及余弦角, 同时 $n = P_2P_3^2 / P_1P_2^2$, $k = P_4P_3^2 / P_1P_2^2$, $p = P_1P_4^2 / P_1P_2^2$, 可利用点 $P_1 \sim P_4$ 的坐标求出, 变换到相机坐标系后的比值不变。因此, 该方程组是关于 x, y, z 的三元二次方程, 利用吴消元法求解该方程组, 得到点 $P_1 \sim P_4$ 在相机坐标系下的 3D 坐标。最后根据 3D-3D 点对用迭代最近点(ICP)求解计算出相机运动的旋转矩阵 R 和平移矩阵 T 。

以标记的中心为坐标系原点(0,0,0), 坐标平面为 XOY 平面, Z 轴正方向垂直于坐标平面向里建立右手坐标系。为统一单位方便后续计算距离, 设置标记边长为 25 cm, 则标记的 4 个顶点在世界坐标系中的坐标分别为(-12.5,-12.5,0.0), (12.5,-12.5,0.0), (12.5,12.5,0.0), (-12.5,12.5,0.0), 如图 6 所示。

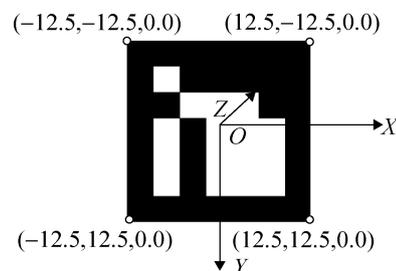


图 6 顶点在世界坐标系中的坐标
Fig. 6 Coordinates of the vertices in the world coordinate system

在只能识别一个有效标记的情况下, 设由 PnP 求解出来的旋转矩阵 R 如式(4)所示。

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (4)$$

$$R = \begin{bmatrix} \cos \theta \cos \phi & \sin \psi \sin \theta \cos \phi - \cos \psi \sin \phi & \cos \psi \sin \theta \cos \phi + \sin \psi \sin \phi \\ \cos \theta \sin \phi & \sin \psi \sin \theta \sin \phi + \cos \psi \cos \phi & \cos \psi \sin \theta \sin \phi - \sin \psi \cos \phi \\ -\sin \theta & \sin \psi \cos \theta & \cos \psi \cos \theta \end{bmatrix} \quad (5)$$

已知旋转矩阵 R , 则 3 个姿态角的求解公式如式(6)所示。

$$\begin{cases} \psi = \arctan(r_{32}/r_{33}) \\ \theta = \arctan(-r_{31}/\sqrt{r_{32}^2 + r_{33}^2}) \\ \phi = \arctan(r_{21}/r_{11}) \end{cases} \quad (6)$$

在能有效识别出 2 个标记时, 计算这 2 个标记的面积。图 7 是标记物的俯视图, 相机(即人)在南偏东 θ 角处。以识别出南和东 2 个标记的情况为例, 可以看出标记(南)和标记(东)在成像平面上的投影面积之比就是相机方位角 θ 的正切角即 $\tan \theta = S_{\text{south}} / S_{\text{east}}$, 则相机的所在方位角为南偏东 $\theta = \arctan(S_{\text{south}} / S_{\text{east}})$ 。

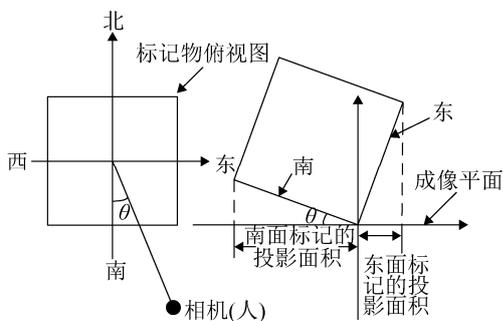


图 7 面积比值法求解方位角示意图

Fig. 7 Schematic diagram of solving azimuth with areas

2.2 定位算法原理

求得姿态角之后, 利用 2.1 节得到的平移向量 T 求解相机和目标的距离。 T 代表从世界坐标系原点到相机坐标系原点的平移关系, 则 2 个坐标系原点间的距离可用该平移向量的模表示; 设平移矩阵 $T = [a \ b \ c]^T$, 则目标到相机的距离 $dist$ 可表示

为: $dist = \sqrt{a^2 + b^2 + c^2}$ 。

旋转矩阵 R 与对应的姿态角: 绕 x 轴旋转的角度 ψ , 绕 y 轴旋转的角度翻滚角和绕 z 轴旋转的角度角 ϕ 之间的数学关系如式(5)所示。

为: $dist = \sqrt{a^2 + b^2 + c^2}$ 。

当识别出一个标记时只计算出一个距离为最终距离。当识别出 2 个有效标记时选择面积较大的标记与相机之间的距离为最终距离。

3 实验

本文在 AR 虚拟沙盘环境下对提出的定位定向算法进行了实验。两名实验者分别佩戴 HoloLens 头盔在 2 个不同的房间部署 AR 虚拟沙盘, 头盔之间通过 WiFi 互联。如图 8 所示, 虚拟沙盘尺寸为 1.25 m×0.6 m, 虚拟三维标记物放置在 AR 虚拟沙盘正上方, 其底部距离 AR 沙盘表面 0.5 m。实验过程中, 在 AR 终端利用本文算法计算操作者相对于虚拟沙盘的位置和方向信息。为了降低对 AR 终端计算资源的占用率, 考虑到沙盘作业的实际情况, 本文选择每隔 0.5 s 采样当前 AR 终端图像信息, 并进行定位定向计算。



(a) A 房间的 AR 场景 (b) B 房间的 AR 场景

图 8 HoloLens 中的增强现实场景效果

Fig. 8 AR scenes in HoloLens

实验中进行标记物的检测与信息解读, 完整的定位定向算法流程图如图 9 所示。

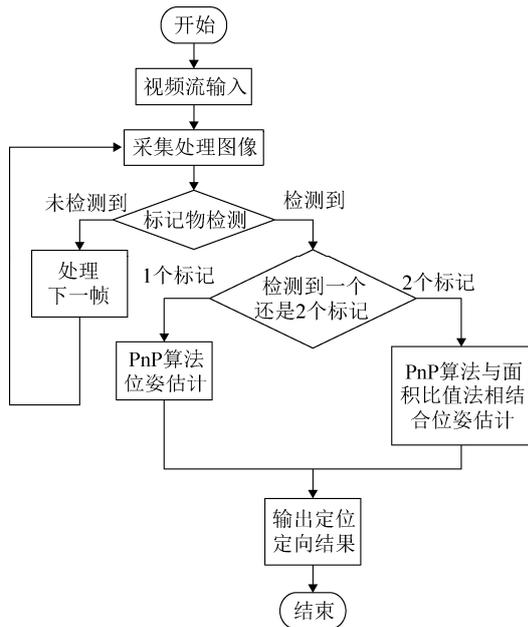


图 9 定位定向算法流程图

Fig. 9 Flow chart of position and orientation

3.1 定向实验

为定量研究, 操作者头戴 HoloLens 围绕标记物沿圆轨迹移动, 每隔 0.5 s 用本文的定位定向算法计算方向和距离, 结果如图 10 所示。图 10 中 * 点表示实际计算结果, 它们离标记中心的距离和角度都是由算法测得。图中的虚线为理想情况下的轨迹。可以看出, 除去个别误差大的点, 实际测得的点都在理想值附近, 可见算法的精确度能达到要求。计算这组实际角度数据的平均误差为 1.5°, 角度的最大误差为 4°, 这表明定向算法的精度能达到协同研讨设计的要求。同时这组数据方差为 3.25(度的平方), 标准差为 1.8°, 这表明测量的数据集中且稳定。

3.2 定位实验

定位结果即相机与标记物的距离。测试时, 操作者头戴 HoloLens 相机沿着直线行走, 每隔 10 cm 记录下用本文的定位定向算法测得的结果。图 11 中圆圈代表理想值, * 点代表实际测量值。经实验发现, 定位算法在 320 cm 以内的距离检测结果精度较高, 大于该值的距离检测结果误差较大, 检测范围满足实际应用时的距离需求。这组数据误差的均值为

3.13 cm, 方差为 14 cm²。可以看出距离越远角度和距离的精度偏差就越大, 其中也包括人为原因引起的误差, 但总体上还是能满足课题要求。定位定向效果图如图 12 所示, 一名实验者在其房间与另一名异地实验者的虚拟化身进行协同指挥、研讨。

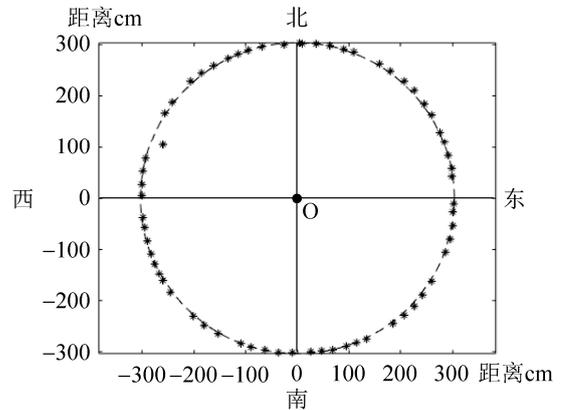


图 10 定向结果

Fig. 10 Result of orientation

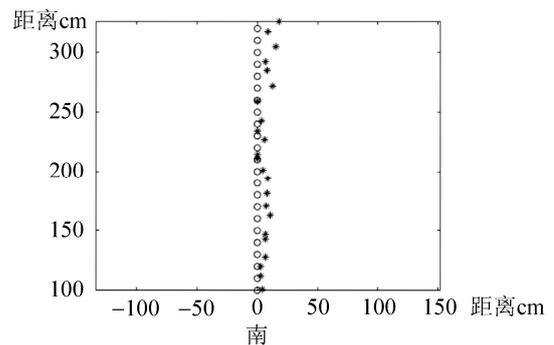


图 11 定位结果

Fig. 11 Result of position



图 12 定位定向效果图

Fig. 12 Position and orientation effect

3.3 分析比较

实验过程中存在一些环境及人为因素会限制定位定向结果的精确度、稳定性。如环境光线较强时,标记的几何特征不够明显导致标记的漏检;背景环境过于复杂时会增加干扰导致标记的误检等问题。此外,在检测虚拟标记物时,本文使用的方法是检测图像中的四边形,同时获取顶点坐标。但由于用户的观察视角持续变化以及光线透视等原因,检测到的四边形难以与标记外边框完全重合,进而导致计算的顶点坐标值存在一定的误差。

与现有的技术相比,本文提出的定位定向算法实现时不需要依托数字罗盘、陀螺仪等硬件;不需要在实际环境中设置地标、指示牌等标志物,在室内可即时使用,易于推广应用。定位定向的速度可以满足实时异地交流的要求。

4 结论

本文提出了基于视觉的虚拟化定位定向技术,设计出具有编码信息的三维虚拟标记物,运用 PnP 姿态估计算法估计方向和距离,提出了面积比值法计算在识别出 2 个标记情况下的角度,最后进行实验验证。实验结果表明本文提出的方法在增强现实协同操作中对虚拟化定位定向有较为准确的结果,且鲁棒性能够满足当前协同操作的需求。

参考文献:

[1] 杨俊勇. 基于增强现实技术的报刊富媒体化系统[D]. 北京: 北京邮电大学, 2013: 12.
Yang Junyong. Newspaper Rich Media System Based on Augmented Reality Technology[D]. Beijing: Beijing University of Posts and Telecommunications, 2013: 12.

[2] Yadav R, Chugh H, Jain V, et al. Indoor Navigation System Using Visual Positioning System with Augmented Reality[C]// 2018 International Conference on

Automation and Computational Engineering (ICACE). Greater Noida, India: IEEE, 2018: 52-56.

[3] Polvi J, Taketomi T, Yamamoto G, et al. SlidAR: A 3D Positioning Method for SLAM-based Handheld Augmented Reality[J]. Computers & Graphics (S0097-8493), 2016, 55(1): 33-43.

[4] Müller J, Butscher S, Feyer, S P, et al.. Studying Collaborative Object Positioning in Distributed Augmented Realities[C]// 16th International Conference on Mobile and Ubiquitous Multimedia, Stuttgart, Germany: ACM, 2017: 123-132.

[5] Egodagamage R, Tuceryan M. Distributed Monocular Visual SLAM as ABasis for A Collaborative Augmented Reality Framework[J]. Computers & Graphics (S0097-8493), 2018, 71(1): 113-123.

[6] 吴鸿飞. 增强现实中标记设计与识别方法研究[D]. 青岛: 青岛大学, 2010: 6.
Wu Hongfei. Research on Marker Design and Recognition Method in Augmented Reality[D]. Qingdao: Qingdao University, 2010: 6.

[7] 桂振文, 王涌天, 刘越, 等. 二维码在移动增强现实中的应用研究[J]. 计算机辅助设计与图形学学报, 2014, 26(1): 34-39.
Gui Zhenwen, Wang Yongtian, Liu Yue, et al. Study on the Application of Two-Dimensional Code in Mobile Augmented Reality[J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(1): 34-39.

[8] 刘经纬. 增强现实中基于三维标志物的跟踪技术研究[D]. 成都: 电子科技大学, 2011: 4.
Liu Jingwei. Research on Tracking Technology Based on 3D Markers in Augmented Reality[D]. Chengdu: University of Electronic Science and Technology, 2011: 4.

[9] 尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述[J]. 自动化学报, 2016, 42(10): 1466-1489.
Yin Hongpeng, Chen Bo, Chai Yi, et al. Vision-based Object Detection and Tracking: A Review[J]. Acta Automatica Sinica, 2016, 42(10): 1466-1489.

[10] 高翔, 张涛, 刘毅, 等. 视觉 SLAM 十四讲: 从理论到实践[M]. 北京: 电子工业出版社, 2017: 3.
Gao Xiang, Zhang Tao, Liu Yi, et al. Visual SLAM XIV: From Theory to Practice[M]. Beijing: Publishing House of Electronics Industry, 2017: 3.