

# 基于深度视觉的SLAM算法研究与实现

李策, 魏豪左, 卢冰, 陈晓雷

(兰州理工大学 电气工程与信息工程学院, 甘肃 兰州 730050)

**摘要:** 针对传统激光扫描仪在地图创建中的缺点, 提出一种基于 Kinect 深度数据模拟激光扫描的 SLAM 算法, 对室内未知环境进行二维地图创建。通过装载 Kinect 传感器的移动机器人采集室内环境的彩色图和深度图, 将深度图转换为三维点云, 通过模拟激光扫描方法获得对应的激光扫描数据; 利用 SIFT 算法与 GTM 算法相结合的方法进行特征点提取与匹配, 实现室内未知环境的二维地图创建。实验对比结果表明, 该方法对于室内二维地图创建有较好的效果, 具有一定的理论和实际应用价值。

**关键词:** 二维地图创建; Kinect 传感器; 激光扫描; SIFT 算法; 图变换匹配算法

中图分类号: TP242.6+2 文献标识码: A 文章编号: 1000-7024 (2017) 04-1062-05

doi: 10.16208/j.issn1000-7024.2017.04.040

## Research and implementation of SLAM algorithm based on depth vision

LI Ce, WEI Hao-zuo, LU Bing, CHEN Xiao-lei

(College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** Aiming at the disadvantages of traditional laser scanner in map building, a SLAM algorithm for 2D map creation was proposed in unknown environment, which based on the Kinect depth data simulation of laser scanning. The indoor environmental color image and depth image were sampled by loading Kinect sensor of the mobile robot, depth image was converted to 3D point cloud and the corresponding laser scanned data were obtained using simulating laser scanning method. Through the combination of SIFT algorithm and GTM algorithm, the feature points were extracted and matched. Experimental comparison results show that this method has good effects on indoor 2D map building, and has certain theoretical and practical application value.

**Key words:** 2D map creation; Kinect sensor; laser scanning; SIFT algorithm; GTM algorithm

### 0 引言

目前, 同步定位与地图创建 (simultaneous localization and mapping, SLAM) 系统大多采用了激光扫描仪作为环境信息采集装置<sup>[1-4]</sup>。虽然激光测距仪视角较宽, 并且测距精度较高, 环境感知结果准确, 但是其价格昂贵, 不能大量投入使用, 且使用固定的激光扫描仪只能检测与其等高的一个特定平面中的障碍物, 忽略了环境中的高度信息<sup>[5]</sup>。这可能导致移动机器人所创建的环境地图不准确, 容易缺失环境中较为重要的障碍物信息, 同时也可能会使移动机器人在自主行驶期间与障碍物发生碰撞, 从而导致了导航和避障的局限性。

近年来, 随着计算机视觉的发展和视觉传感器的广泛应用, 视觉 SLAM 逐渐兴起, 成为 SLAM 研究领域中的重要研究方向。Kinect 是微软公司开发的一种具有深度视觉感知的体感摄像机, 目前在计算机视觉图像处理, 三维空间探索等方面有着广泛的应用<sup>[6]</sup>。Kinect 传感器的最大优势在于采集速度快, 采集图像精度高, 但成本低廉等, 因此具有广泛的应用前景, 在本文中用来代替激光扫描仪对于移动机器人周边的环境信息进行采集。通过 Kinect 设备进行深度图像采集和 SLAM 算法进行二维地图创建是当前的研究热点和难点之一。

针对上述关于激光扫描仪中 SLAM 算法中的问题, 本文提出了一种基于 Kinect 深度数据模拟激光扫描数据的

收稿日期: 2016-03-15; 修订日期: 2016-04-21

基金项目: 国家自然科学基金项目 (61365003、61302116); 甘肃省基础研究创新群体基金项目 (1506RJJA031)

作者简介: 李策 (1974-), 男, 辽宁营口人, 教授, CCF 会员, 研究方向为计算机视觉、图像处理、模式识别与人工智能; 魏豪左 (1989-), 男, 山东济南人, 硕士, 研究方向为计算机视觉、图像处理; 卢冰 (1989-), 男, 河南郑州人, 硕士, 研究方向为计算机视觉、图像处理; 陈晓雷 (1979-), 男, 甘肃兰州人, 博士, 研究方向为图像处理。E-mail: xjtulice@gmail.com

SLAM 算法, 对室内未知环境进行二维地图创建。

### 1 Kinect 传感器与本文算法介绍

#### 1.1 Kinect 工作原理

Kinect 传感器是一种基于 RGB-D 传感器的图像采集设备, 即可以同时获得图像的 RGB 彩色信息以及深度信息<sup>[7]</sup>。

如图 1 所示, Kinect 结构包括深度传感器, RGB 摄像头以及电机底座, 深度传感器包括了红外线发射器和红外深度感应器。Kinect 的水平视角为 57 度 (即以传感器为中心, 左右各 28.5 度), 垂直视角为 43 度 (同理, 以传感器为中心, 上下各 21.5 度), 再加上其底部电机底座的马达可以上下旋转 27 度, 其垂直视角可达 97 度。利用这一点, 可以获得不同水平面更加丰富的环境信息。Kinect 深度信息检测距离在 0.5 m~7 m 之间, 采集频率为 30 帧每秒。图 2 为 Kinect 传感器对实际实验室环境所采集的彩色图和对应的深度图示例。

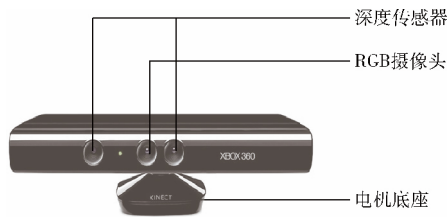


图 1 Kinect 结构



图 2 Kinect 所采集的彩色图和深度图 (左图为彩色图, 右图为深度图)

#### 1.2 本文算法介绍

##### 1.2.1 Kinect 深度数据模拟激光扫描数据

将 Kinect 采集得到的深度图像转换为点云数据, 然后根据深度数据的几何模型来模拟实现二维激光扫描的功能。深度数据的几何模型如图 3 所示。

图中所示平面为一个与地面平行的平面, 其中 Kinect 为红外深度感应器的中心。

Kinect 所获得的场景深度图像中每个像素点的像素值均对应了一个深度值, 根据图 3 所示, 此深度值并不是相机到物体的直线距离  $R$ , 而是该物体所在的平面到 Kinect 相机所在平面的垂直距离  $D$ 。需要经过相应的几何转换才能得到该物体的具体测距信息。

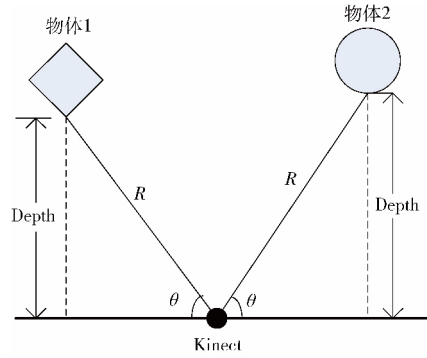


图 3 深度数据几何模型

几何转换的方法如式 (1) 所示

$$D = depth[i][j]$$
$$R = \frac{D}{\sin\theta} \tag{1}$$

式中:  $depth[i][j]$  表示深度图像中第  $i$  行和第  $j$  列像素点对应的深度值,  $\theta$  表示深度图像中该像素点的坐标在几何模型坐标中的角度。

通过式 (1), 可以快速将 Kinect 采集得到的深度图像转换为三维点云模拟得到激光扫描数据<sup>[8]</sup>。

##### 1.2.2 特征点提取和匹配

在本文中, 为减少 SIFT 算法<sup>[9]</sup> (scale invariant feature transform) 对前后两帧进行特征点提取和匹配时存在的一些错误匹配点, 本文给出了一种基于 SIFT 算法和 GTM<sup>[10]</sup> 算法融合的方法对 Kinect 所得到的图像进行特征点的提取和匹配。

GTM (graph transformation matching) 算法, 即图变换匹配算法, 其算法过程主要分为如下两步:

(1) 初始图形构建。将 SIFT 算法得到的特征点集合作为一个图的各个顶点, 设图形的顶点为  $U_i, i=1, 2, \dots, N$ , 其中  $N$  为顶点的数目。对于任意的两个顶点  $i, j$ , 如果满足  $\|U_i - U_j\| \leq \bar{d}$ , 则  $U_i$  和  $U_j$  相连接, 其中  $\bar{d}$  为图形中所有顶点两两距离的均值。最终将得到一个如图 4 的初始图形。

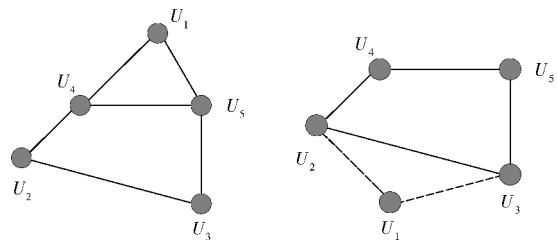


图 4 SIFT 特征点匹配初始图形

(2) 错误匹配点搜索和剔除。针对初始构建的两个图进行特征点连接比对, 可以得到两个图形的顶点  $U_1$  之间存在差异, 因此将两个图像  $U_1$  对应的连接进行剔除, 可以得到图 5 所示图形。

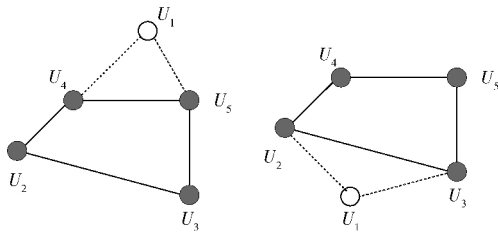


图 5 SIFT 特征点匹配图形错误匹配点剔除

具体差异性搜索方法如下：

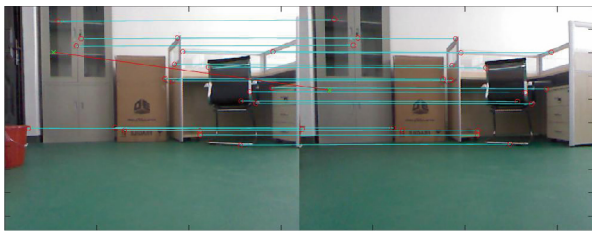
(1) 构建两个图对应的矩阵，初始值为 0，如果存在  $i$  和  $j$  两个顶点满足  $\|U_i - U_j\| \leq \bar{d}$ ，则在矩阵的第  $i$  行，第  $j$  列赋值为 1，最后可以将图 4 的图转换为矩阵。这里以一个包含 4 个顶点的图为例，对应的可以获得如下矩阵表达式

$$A_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

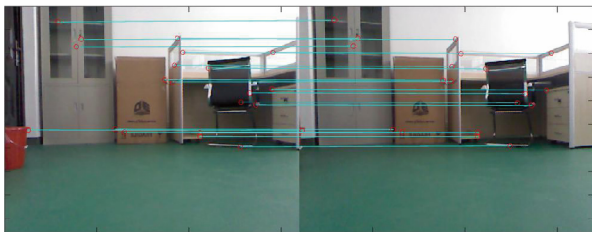
(2) 将两个矩阵做差，则可以得到如下结果

$$A = |A_1 - A_2| = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3)$$

(3) 根据式 (3)，其非零元素就是对应的错误匹配点。SIFT 和 GTM 算法融合后的结果如图 6 所示。



(a) SIFT 特征点匹配效果



(b) SIFT+GTM 特征点匹配效果

图 6 特征提取与匹配结果

图 6 (a) 表示前后两帧相似点的匹配结果，从图 6 (a) 结果我们可以看出，其中交叉部分连线的匹配点出现了错误匹配。因此，本文利用 SIFT 算法和 GTM 算法相结合的方法对错误匹配点进行剔除。对比图 6 (a) 和图 6 (b) 两个结果，我们可以看到通过本文方法可有效的将其中的错

误匹配点进行剔除。

将 Kinect 得到的相邻帧进行连续匹配，将获得一个完整的空间环境匹配。设连续的三帧图像分别为  $I_{t-1}, I_t, I_{t+1}$ ，图像匹配过程函数为  $SIFTGTM$ ，匹配后的图像为  $Match$ ，则连续匹配可以通过如下过程进行计算：

首先将  $I_{t-1}$  和  $I_t$  做匹配

$$Match = SIFTGTM(I_{t-1}, I_t) \quad (4)$$

然后将式 (4) 的匹配结果和下一帧进行匹配

$$Match = SIFTGTM(Match, I_{t+1}) \quad (5)$$

根据式 (4) 和式 (5)，可以获得连续帧之间的匹配结果。连续帧之间匹配结果如图 7 所示。

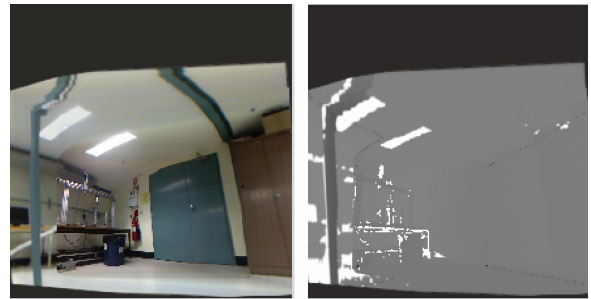


图 7 彩色图和深度图连续帧匹配

(左图为彩色图连续匹配结果，右图为深度图连续匹配结果)

### 1.2.3 二维地图创建

通过 SLAM 算法进行地图创建<sup>[1]</sup>，其步骤为：

(1) 机器人运动初始化。机器人运动初始化主要包括机器人初始位置协方差  $P_r$ ，机器人特征位置的初始协方差值  $P_{mm}$ ，机器人与特征之间的交叉协方差值  $P_m$ ，这三个初始变量构成初始状态协方差矩阵  $P = \begin{bmatrix} P_r, P_m \\ P_m^T, P_{mm} \end{bmatrix}$ 。通常情况下，取初始状态协方差矩阵为零矩阵。

(2) 数据匹配。数据匹配也称数据关联，主要是用来确定传感器的观测信息和预测观测值之间的对应关系。

(3) 状态更新。通常情况，SLAM 算法采用扩展卡尔曼滤波对状态进行更新。在实际信息监测过程中，机器人的轨迹可以简化为由有限个线段组成的线段。设单位时间机器人移动距离为  $\Delta d$ ，那么下一时刻机器人的坐标满足

$$x(t+1) = x(t) + \Delta d \times \sin(\theta) \quad (6)$$

$$y(t+1) = y(t) + \Delta d \times \cos(\theta) \quad (7)$$

其中， $\theta$  为机器人当前时刻的运动方向。通过 1.2.2 小节所述，利用 SIFT 算法和 GTM 算法相结合对前后两帧进行特征点提取与匹配求出其旋转矩阵和平移矩阵，得到  $\Delta d$  和  $\theta$ 。

(4) 新信息处理。设在  $t+1$  时刻，机器人的特征集合为  $L_1$ ， $t$  时刻为  $L_2$ ，不同时刻机器人的特征变换为  $L_0 = L_2 - L_1$ ，如果  $L_0$  为非空集，那么将  $L_0$  中的元素加入到地

图中, 即

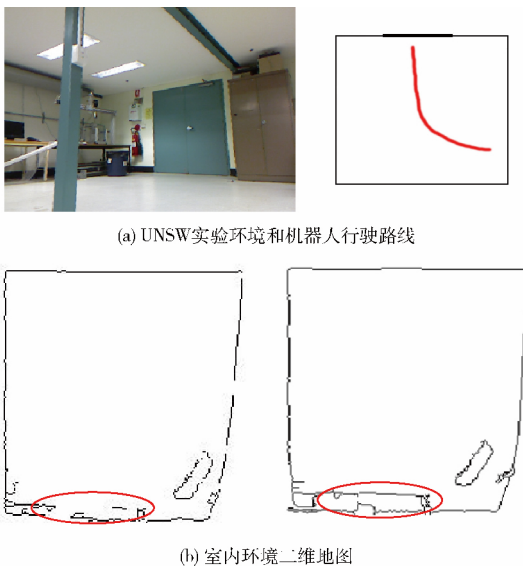
$$X(k+1) = \begin{bmatrix} X(k) \\ L_{new} \end{bmatrix} \quad (8)$$

## 2 实验结果及分析

为了验证所提算法, 本文首先在澳大利亚新南威尔士大学 (UNSW), Jose Guivant 和 Alicia Robledo 所建数据库<sup>[11]</sup>上进行了测试实验; 然后为了进一步验证实际环境下的有效性, 在我们自己搭建的实验平台上, 对所提算法进行了验证实验。

### 2.1 UNSW 数据集上的测试实验

该数据库所采集数据一共包括 1924 帧彩色图以及深度图, 在本文算法的测试实验中选取前 200 帧进行。如图 8 所示, 为本文算法在 UNSW 大学数据库中所做对比实验的相关实验结果。



(左图为激光扫描所得地图, 右图为本文方法所得地图)

图 8 本文算法在 UNSW 数据库上的测试实验

对比图 8 (b) 左、右两图圈出的部分, 我们可以发现, 左图激光扫描地图对于真实环境中一些较高的桌子没有扫描出来, 而右图中本文算法能够准确的将 UNSW 实验环境中的一些障碍物及其相对位置在地图上表示出来。

### 2.2 实验平台搭建与实验验证

#### 2.2.1 实验平台搭建

如图 9 所示, 为本文搭建的机器人实验平台。该平台处理核心为一个 4 GHz 的 CPU 的计算机, 平台所使用机器人是由利曼公司所开发的 PowerBot™ 型机器人, 在 PowerBot™ 型机器人平台前端安装有激光扫描传感器, 激光扫描传感器距离地面高度为 25 cm, 在本文实验中将 Kinect 传感器安装在机器人的前端激光扫描传感器一侧, 距离地面



图 9 在 PowerBot 机器人上自行搭建的测试平台

同样为 25 cm 以便于与激光扫描传感器采集创建的二维地图进行对比。

图 10 (a) 为本文实验中自行搭建的实际实验环境。图中所示物体为实验之前, 根据本文引言部分的描述, 人工设置的障碍物。本文所用激光扫描仪高度为 25 cm; 而该障碍物, 底部高度为 15 cm, 低于激光扫描仪所在平面, 而上层椅子底面高度为 50 cm, 高于激光扫描仪所在平面, 这样设置可以更容易在本文实验的结果中做出对比。

#### 2.2.2 实际验证实验结果

在实验过程中, 首先, 仅开启机器人平台所携带的激光扫描仪按图 10 (b) 中的轨迹进行移动, 同时创建基于激光扫描数据的二维地图; 其次, 关闭激光扫描仪打开 Kinect 传感器, 机器人按照同样路径采集数据, 然后通过本文方法建立二维地图。机器人在移动过程中, 当遇到转弯的情况, 则按每一时刻小于 1 度的旋转角度进行逐渐的调整方向, 且每一时刻机器人移动距离均小于 4 cm, 通过这种移动模式, 可以提高采集图像的空间连续性。

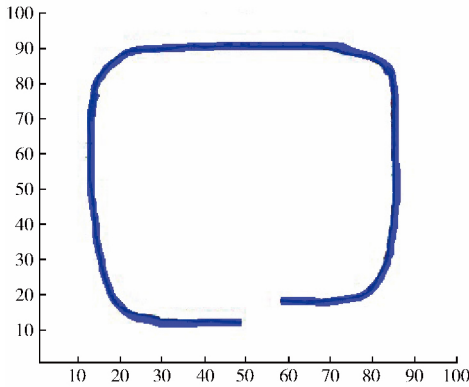
在本次实验中, 利用 Kinect 传感器分别采集了 500 帧实验室环境的彩色图像和深度图像。本文实验结果是从其中选取前 280 帧进行实验测试。最后通过两种方法所得到实验室的二维地图, 如图 10 (c) 和图 10 (d) 所示。

#### 2.2.3 实验分析

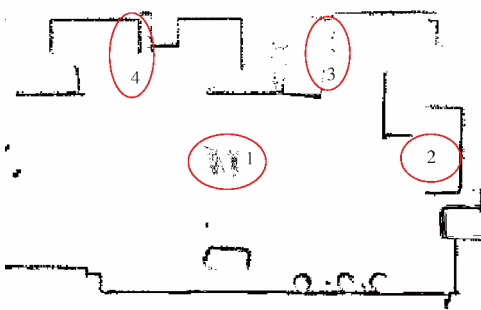
如图 10 (c) 和图 10 (d) 两图中用圆圈标出的 1、2、3 和 4 是在实验之前, 我们人工设置的一些障碍物, 其中, 第 1 处障碍物, 在 2.2.1 小节进行了简单描述, 第 2 处障碍物为一个椅子紧挨地面, 高度可以忽略, 第 3、4 处障碍物高度均为 15 cm。通过对比图 10 (c) 和图 10 (d) 两个结果, 我们可以发现, 通过激光扫描仪进行地图创建时, 一些不在其扫描平面上的障碍物, 尤其是高度较低的一些障碍物会被忽略, 从而导致所创建地图不准确; 而通过本文方法可以获得更加详细并且接近实际环境的二维地图。



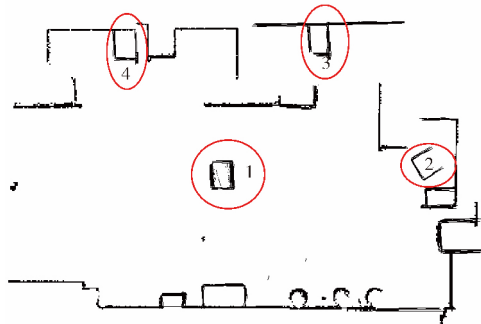
(a) 实际实验环境



(b) 机器人行驶路径



(c) 激光扫描仪所得地图



(d) 本文方法所得地图

图 10 室内环境二维地图

### 3 结束语

本文算法针对激光测距仪只能检测一个特定的平面中的障碍物，或者只检测到一个障碍物的其中一部分从而容易导致机器人与障碍物发生碰撞的问题，提出了一种基于 Kinect 深度数据模拟激光扫描数据的 SLAM 算法进行室内

环境的二维地图创建，所提方法通过 Kinect 传感器采集室内环境的彩色图像和深度图像，然后通过模拟激光扫描的方法获得对应的点云数据，其次通过 SIFT 算法以及 GTM 算法相结合，对 Kinect 所获得的每帧图像进行特征点的提取与匹配，最后建立室内未知环境的二维地图，通过实验分析可知，本文算法能够获得比普通利用激光扫描得到的地图更详细、更接近真实环境的二维地图，在今后的研究工作中，具有一定的理论和实际应用价值。

### 参考文献：

- [1] Smith R, Self M, Cheeseman P. Estimating uncertain spatial relationships in robotics [J]. Machine Intelligence & Pattern Recognition, 2014, 5 (5): 435-461.
- [2] Weng CC, Wang CC, Healey J. A segmentation and data association annotation system for laser-based multi-target tracking evaluation [C] //IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2012: 80-86.
- [3] GUO Lei. Navigation and capturing goal of robot [J]. Computer and Digital Engineering, 2015, 43 (11): 1959-1961 (in Chinese). [郭雷. 机器人导航与目标捕获 [J]. 计算机与数字工程, 2015, 43 (11): 1959-1961.]
- [4] Emter T, Stein A. Simultaneous localization and mapping with the Kinect sensor [C] //Proceedings of Robotics, 2012: 1-6.
- [5] Kamarudin K, Mamduh SM, Shakaff AYM, et al. Method to convert Kinect's 3D depth data to a 2D map for indoor SLAM [C] //IEEE International Conference on Signal Processing and its Applications, 2013: 247-251.
- [6] Herrera DC, Kannala J, Heikkila J. Joint depth and color camera calibration with distortion correction [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34 (10): 2058-2064.
- [7] Henry P, Krainin M, Herbst E, et al. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments [J]. International Journal of Robotics Research, 2012, 31 (5): 647-663.
- [8] CAO Hongjun, WU Panlong, YAO Xiang. On depth smoothing and fake laser rangefinder based on Kinect [C] //Proceedings of the 33rd Chinese Control Conference, 2014 (in Chinese). [曹洪君, 吴盘龙, 姚翔. 基于 Kinect 的深度数据处理及二维激光雷达功能实现研究 [C] //第 33 届中国控制会议, 2014.]
- [9] GONG Lun, LIU Can, HU Junpu. A new image stitching technology based on Harris corner and SIFT improved [J]. Computer and Digital Engineering, 2015, 43 (11): 2055-2060 (in Chinese). [龚伦, 刘璨, 胡均谱. 一种改进 Harris 角点与 SIFT 的图像拼接技术 [J]. 计算机与数字工程, 2015, 43 (11): 2055-2060.]
- [10] Wang S, Guo X, Mu X, et al. Advanced weight graph transformation matching algorithm [J]. Computer Vision, 2015, 9 (6): 960-966.
- [11] Guivant JE, Marden S, Pereida K. Distributed multi sensor data fusion for autonomous 3D mapping [C] //IEEE International Conference on Indoor Positioning and Indoor Navigation, 2012: 1-11.