

融合2D/3D摄像机的方法与获取高精度三维视觉信息的装置

余洪山^{1,2†}, 赵科¹, 王耀南¹, 唐延东², 张红良³

(1. 湖南大学 电气与信息工程学院, 湖南 长沙 410082; 2. 机器人学国家重点实验室, 辽宁 沈阳 110016;

3. 国防科学技术大学 航天科学与工程学院, 湖南 长沙 410000)

摘要: 三维飞行时间摄像机可实时同步获取场景三维信息和灰度图像信息. 虽然它存在图像分辨率和质量较差等问题, 但它可作为二维摄像机的互补. 本文借鉴立体视觉技术, 提出了一种2D/3D摄像机融合的三维视觉信息获取方案. 论文首先基于固定空间关系和相近视野原则, 设计2D/3D立体摄像机系统对空间场景同步成像. 结合三维ToF摄像机成像特性, 论文借鉴立体视觉技术完成二维摄像机的高质量二维彩色图像与插补后的三维摄像机深度图像的匹配关联. 因此, 本方法可实现场景的高精度彩色图像和对应三维空间信息的实时同步获取, 同时保留了二维摄像机的高质量彩色二维成像和三维摄像机的快速稠密三维信息获取的优势. 2D/3D摄像机图像融合匹配算法复杂度低, 匹配精度和准确度取决于二维摄像机和三维摄像机自身性能、摄像机标定参数精度和深度图像插补算法, 不会引入新的运算误差. 试验结果验证了本文算法的有效性和精确度.

关键词: 2D/3D图像匹配; 图像融合; 3D摄像机; 立体视觉; 三维重建

中图分类号: TP273 **文献标识码:** A

Fusion of 2D camera with 3D camera and equipment for acquiring high resolution 3D visual information

YU Hong-shan^{1,2†}, ZHAO Ke¹, WANG Yao-nan¹, TANG Yan-dong², ZHANG Hong-liang³

(1. Institute of Electrical and Information Engineering, Hunan University, Changsha Hunan 410082, China;

2. State Key Laboratory of Robotics, Shenyang Liaoning 110016, China;

3. Institute of Space Science and Engineering, National Defense Science and Technology University, Changsha Hunan 410000, China)

Abstract: Time of flight (ToF) camera can deliver grayscale image and 3D information simultaneously with high frame rates. Although its image resolution and quality are very poor, it provides a good supplement with traditional 2D camera image. Inspired by the stereo vision, we present a 3D visual information acquiring method by fusing 3D information of ToF camera with the high resolution color image. Firstly, we setup a 2D/3D stereo camera with fixed spatial relation and similar visual field to capture the environment information synchronously. Based on characteristics of ToF camera and the principles of stereo vision, the interpolated 3D information of ToF camera is registered with high resolution color image. Consequently, the proposed method can provide high quality 2D image as well as corresponding 3D space information, which preserves the advantages of 2D camera's high quality image and ToF camera's fast acquisition for dense 3D information. The proposed 2D/3D camera fusion method has very low computation cost while the matching accuracy is only determined by parameters of the 2D/3D stereo camera system and interpolation algorithm for range image without extra computation error. Experimental results have demonstrated the feasibility and accuracy of the proposed method.

Key words: 2D/3D image registration; image fusion; 3D camera; stereo vision; 3D reconstruction

1 引言(Introduction)

目前, 在精密生产线加工装配、自动焊接、机器人智能导航、极限或危险环境下自主探测等领域, 激光、声纳、摄像机等传感器已成为相关设备或机器人的基本配置, 以获取相应工作场景或工件的几何或图像信

息, 从而辅助进行决策. 随着机器人任务和工作环境复杂化的增加, 机器人或相关装置期望能像人类视觉系统一样同时获取三维空间范围内目标的图像和几何信息, 从而进一步提升对环境的三维感知和识别能力, 以进一步提升其智能化处理水平.

收稿日期: 2014-01-16; 录用日期: 2014-06-06.

†通信作者. E-mail: yuhongshancn@hotmail.com.

基金项目: 国家自然科学基金资助项目(61005068); 国家高技术研究发展计划("863"计划)资助项目(2012AA112312); 机器人学国家重点实验室开放课题基金资助项目(2013O09); 国家重大科学仪器设备开发专项项目(2013YQ140517); 广东省教育部产学研结合资助项目(2011B090400074); 中央高校基本科研业务费专项资金资助项目.

为实现空间三维信息和二维图像的同步获取,主要方案有立体视觉、激光扫描仪与二维摄像机融合方案、三维摄像机等获取方案,其主要特性如表1所示.

3D激光测距仪三维测量精度高,但是存在重量

大、功率高和价格昂贵等问题,且无法提供场景的图像信息^[1-3]. 为此,部分研究机构提出了激光测距仪和传统二维摄像机的融合方案,但是二维图像和三维信息的精确匹配仍存在很多困难.

表1 当前三维信息获取技术方案对比
Table 1 Current 3D information acquiring solutions

方案	优点	存在问题
立体视觉	可同时获取场景二维图像信息和三维信息	三维获取计算复杂,疏密程度严重依赖于场景中图像特征信息量,测量精度随距离增加迅速降低.
激光测距仪+二维摄像机	可获取场景的较高精度的二维图像信息和某一高度方向的距离信息,受外界环境干扰小、稳定,准确性高.	▲ 二维激光测距仪需借助额外装置和多次扫描,实现较大范围三维信息获取,计算复杂,实时性差; ▲ 三维激光测距仪非常昂贵;图像信息和三维信息的匹配和融合是个难题
TOF三维摄像机	可实时获取环境三维信息和匹配对应的二维图像,成像距离适中	二维灰度图像的质量非常差、分辨率低,受光线的影响很大,无颜色信息,很难用于后期的图像处理.
Kinect摄像机	可实时获取环境三维信息和二维彩色图像	二维图像和三维信息精度不高、成像距离近,仅可用于室内环境下近距离工作.

相对于激光测距传感器,立体视觉系统获取的深度信息受场景特征疏密程度和特征匹配程度影响;此外当超过一定深度范围后,其精度随距离增长迅速下降,但是该类系统可同时提供场景的高质量图像信息^[4-5]. 近年来,3D摄像机的出现引起了广泛关注,如微软Kinect摄像机、Swiss-ranger的SR系列三维摄像机等,其特点是可同时实时获取空间场景的图像信息以及图像像素点对应的空间三维信息,从而在场景目标分割和识别方面具有巨大的应用潜力,如目标识别、场景三维重建等.但是该类3D摄像机的局限在于其提供的场景图像分辨率比较低,如SR3000摄像机图像的分辨率仅为176*144,从而严重制约了其应用范围^[6-9].为提高三维摄像机精度,D.faile等针对深度信息误差进行了分析,针对结构化环境提出了一些解决方案^[10].为解决分辨率低的问题,Prasad等提出利用传统插补方法提高分辨率^[11],Sigurjon等提出SR3000摄像机与立体摄像机的融合方案^[12],Kuhnert等提出了PMD三维摄像机与立体摄像机的融合方案^[13],B. Huhle等提出利用超分辨率技术实现PMD摄像机与高精度彩色摄像机融合提高分辨率的方法^[14].

针对上述问题,论文深入分析三维摄像机和二维摄像机成像原理和特性,综合已有研究成果,研制了基于二维/三维摄像机信息融合的高精度三维视觉信息获取方法及装置.该装置将三维摄像机与二维摄像机构成类立体视觉系统的复合摄像机,并基于立体视觉技术、三维摄像机特性和图像插补技术实现同一场景高精度高质量彩色图像和对应三维

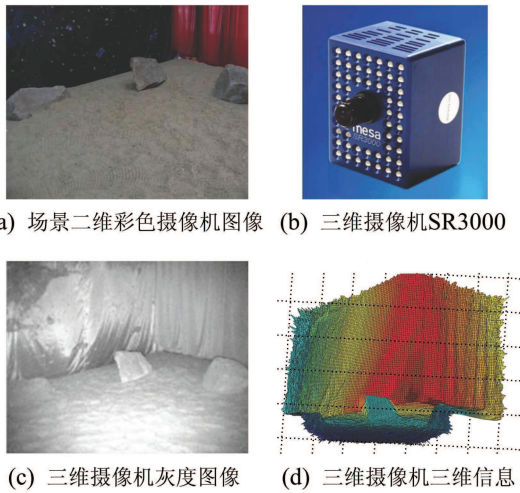
空间信息的实时同步获取,同时保留了二维摄像机的高质量彩色二维成像和三维摄像机的快速三维信息获取的优势.以下论文将详细介绍其原理和实现,具体安排如下:第2节介绍复合摄像机视觉信息融合基本原理;第3节介绍复合摄像机系统硬件设计及实现;第4节介绍二维摄像机和三维摄像机图像匹配融合算法实现;第5节进行实验验证分析,最后对论文工作进行分析总结.

2 复合摄像机图像融合原理(The fusion principle of 2D/3D stereo image)

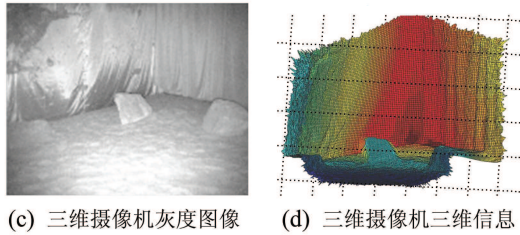
2.1 二维和三维摄像机信息互补性分析(Complementary analysis on 2D camera and 3D camera information)

三维摄像机与传统二维摄像机信息具有良好的互补性.如图1所示,三维摄像机可同时获取同一场景的相同分辨率的场景灰度图像和深度图像,从而获得场景以摄像机为坐标系的三维信息.三维摄像机获取的三维信息质量高,对环境和光线的依赖性不大,如图1(b)所示;而三维摄像机灰度图像质量比较差,造成很多图像细节损失,且容易受噪声污染,对环境光线非常敏感,成像不稳定.因此很难直接利用三维摄像机提供的灰度图像进行后续的场景分析和目标识别处理.而目前二维摄像机具有成像清晰、分辨率高、畸变小优势,易于获取目标纹理、色彩特征.因此将三维摄像机与二维彩色摄像机对同一场景同时获取的成像信息进行融合,以高质量的二维彩色图像替代三维摄像机的灰度图像,并完成二维彩色图像与三维摄像机三维信息的逐一匹配关

联, 即可实现高精度彩色图像和三维空间信息的同步获取, 从而提高其三维空间场景理解与识别的准确度和可靠性.



(a) 场景二维彩色摄像机图像 (b) 三维摄像机SR3000



(c) 三维摄像机灰度图像 (d) 三维摄像机三维信息

图 1 二维彩色摄像机和三维摄像机获取信息对比

Fig. 1 2D camera/3D camera information comparison

2.2 复合摄像机三维视觉信息获取(2D/3D stereo camera visual information acquiring)

对于传统立体视觉系统, 通过左右摄像机图像匹配, 借助立体视觉空间几何关系模型, 求取像素点对应的三维信息. 对于二维摄像机和三维摄像机组成的复合摄像机, 虽类似于立体视觉系统, 但本文的复合摄像机系统希望基于三维摄像机提供的三维信息获取二维和三维摄像机图像的关联映射模型.

如图2所示, 设定左摄像机对应的摄像机为三维摄像机, 右摄像机对应于二维摄像机. 根据三维摄像机特性, 对于左摄像机的二维图像上任意一点像素 p_1 , 可直接确定其对应空间点 P 的三维位置. 同时, 由空间点 P 在二维摄像机图像中的成像点 p_2 , 则可由 P 点的空间位置和两摄像机的相对空间位置关系(立体摄像机标定参数 R 和 T)唯一确定, 从而实现 p_1 和 p_2 的关联匹配. 因此, 复合摄像机中三维摄像机的三维信息和二维摄像机的二维图像间存在固定的关联映射模型.

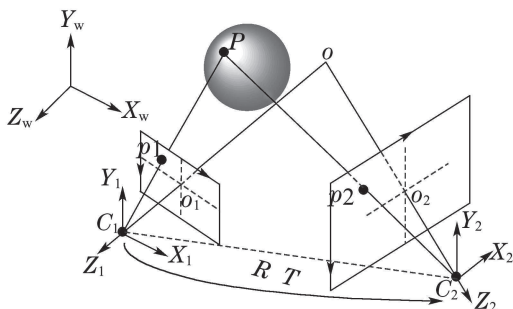


图 2 2D/3D复合视觉系统成像模型

Fig. 2 The projection model of 2D/3D stereo camera

基于该原理, 本论文提出了一种全新的基于二维/三维摄像机信息融合的高精度三维彩色图像获取方案, 并研制了相应的复合摄像机硬件系统.

3 复合摄像机设计及实现(Design and realization of 2D/3D stereo camera)

3.1 复合摄像机系统设计(2D/3D stereo camera system design)

本文设计的复合摄像机成像系统是由二维摄像机和三维摄像机组成的特殊立体视觉系统. 为实现三维摄像机的三维信息与二维彩色摄像机的图像信息的完全匹配关联, 即三维摄像机中的成像区域中的任意一点均能在二维摄像机图像中找到唯一匹配关联点, 因此本文设计的装置要求保证三维摄像机的成像范围完全包含于二维摄像机成像范围中, 并且两者尽可能大的比例重合, 以充分利用两者信息. 为此, 该装置基本设计要求如下:

1) 二维摄像机焦距与三维摄像机焦距参数相同或接近, 二维摄像机视野成像范围稍大于三维摄像机, 保证三维摄像机成像范围在二维摄像机成像范围内.

2) 二维摄像机和三维摄像机空间水平方向保持平行关系, 同时保证成像不相互遮挡和电子干扰, 在此条件下尽量减小两摄像机的空间相对平移量, 以保证尽可能大比例的重合成像区域.

3.2 复合摄像机系统实现(The mechanical realization of 2D/3D stereo camera system)

基于上述分析, 在参考已有立体双目视觉系统设计基础之上, 结合三维摄像机(SR-3000摄像机)与二维彩色摄像机的硬件特点, 本文设计的复眼摄像机机械装置如图3所示.

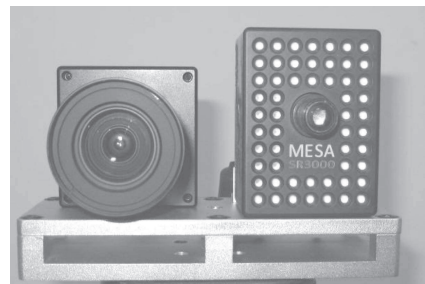


图 3 本文设计的复合摄像机系统

Fig. 3 The 2D/3D stereo mechanical system

如图4所示, 该装置机械部分包括: 主固定板1、扩展应用接口板2、三维摄像机固定连接板3. 二维摄像机和三维摄像机平行固定在主固定板1上, 并通过扩展应用接口板2与外部应用装置对接.

本装置选用的二维摄像机为底部固定方式, 直

接水平固定于主固定板1的上表面;而SR3000三维摄像机为背部固定(固定面的法线方向与镜头方向相同)方式. 为保证复合摄像机中二维摄像机和三维摄像机光轴平行,三维摄像机通过垂直L型固定连接板3水平固定于装置上固定板表面. 此外,在三维摄像机与主固定板1表面固接区域设计三维摄像机散热孔11,满足SR3000摄像机底部散热的特殊需求,同时满足主固定板1对三维摄像机的支撑作用,如图4所示.

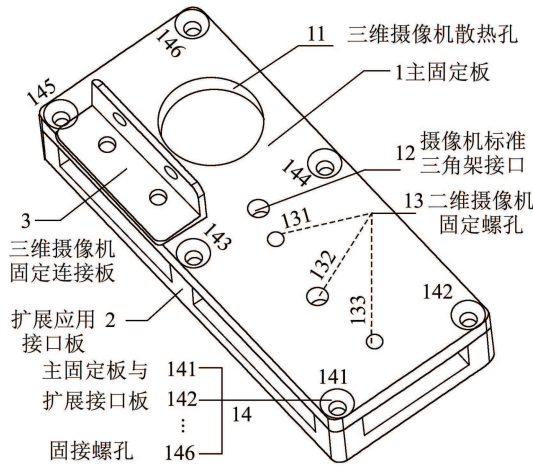


图4 复合摄像机系统机械装配平台

Fig. 4 Assemble platform of 2D/3D stereo system

为满足不同应用中复合摄像机装置的特定装配和固定要求,本文设计预留了扩展应用接口板2. 如图5所示,扩展应用接口板2通过主固定板接口端22与主固定板固定;扩展应用接口板2可根据具体应用需求,设计相应的外部连接端口. 作为应用实例,图5中设计了本装置与Denso机械手的固定接口24和通用摄像机三脚架接口23等.

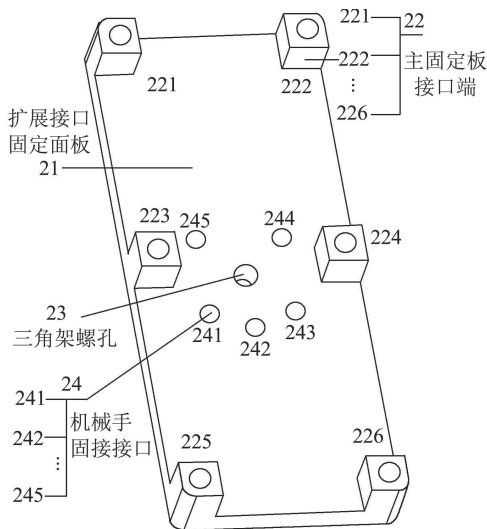


图5 复合摄像机平台扩展装配接口板

Fig. 5 Extended assemble interface of 2D/3D stereo system

4 复合摄像机匹配融合方法(Image matching and fusion of 2D/3D stereo system)

4.1 融合匹配算法总体流程(The scheme of proposed infusion method)

基于上述复合摄像机系统,本文进一步提出三维摄像机和二维摄像机的融合匹配方法,总体流程如图6所示.

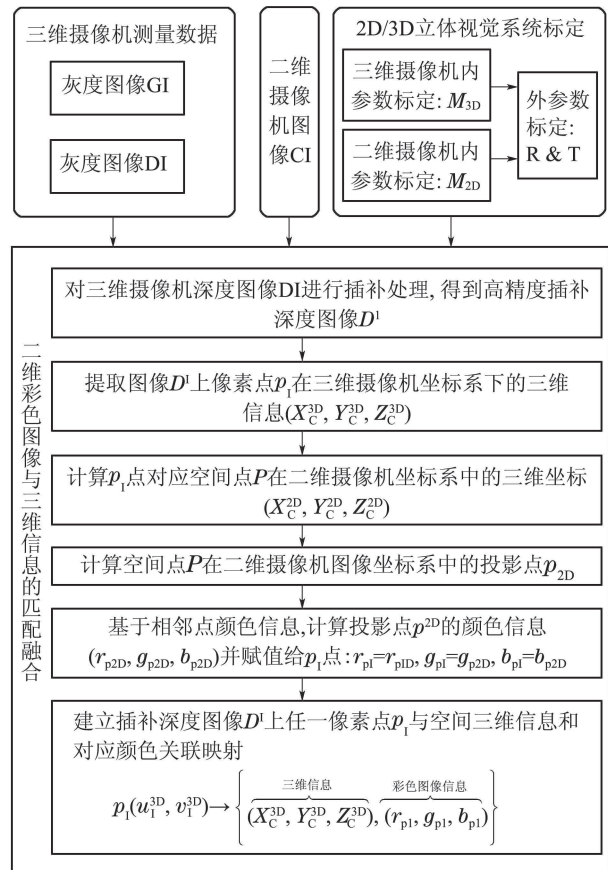


图6 二维摄像机和三维摄像机图像融合匹配流程

Fig. 6 Schematic of the 2D/3D stereo camera fusion

首先对三维摄像机和二维摄像机组成的立体视觉系统进行离线标定,分别获取三维摄像机内参数矩阵 M_{3D} 、二维摄像机内参数矩阵 M_{2D} ,以及二维摄像机坐标系和三维摄像机坐标系之间对应的相对三维空间变换参数(空间旋转关系变换矩阵 R 和空间平移关系变换矩阵 T). 采用二维图像插补算法对三维摄像机的深度图像进行插补成高分辨率深度图像. 对插补后的深度图像,基于三维摄像机成像模型,计算插补后深度图像像素点对应的三维信息;采用高质量高精度二维彩色图像与三维摄像机深度图像像素点之间的匹配映射模型,计算深度图像点对应的二维彩色图像信息,从而完成场景的高质量彩色图像及其对应的三维信息的同步获取. 以下章节详细阐述算法的具体实现如下.

4.2 复合摄像机匹配融合算法实现(Registration and fusion realization of 2D/3D stereo camera)

1) 复合摄像机标定.

本文中所述的二维摄像机和三维摄像机均为定焦摄像机, 两摄像机相对空间关系保持固定, 因此本文的装置中由二维摄像机和三维摄像机组成特殊的立体视觉系统的内参数和外参数均保持不变, 只需进行一次离线标定处理获得该装置的内参数和外参数. 该装置采用线性摄像机成像模型, 采用文献[15]中立体摄像机标定算法, 离线标定获取三维摄像机内参数矩阵 M_{3D} 、二维摄像机的内参数矩阵 M_{2D} , 以及二维摄像机坐标系和三维摄像机坐标系之间对应的相对空间变换参数 R (三维空间旋转关系变换矩阵)和 T (空间平移关系变换矩阵):

$$M_{3D} = \begin{bmatrix} \alpha_x^{3D} & 0 & u_0^{3D} & 0 \\ 0 & \alpha_y^{3D} & v_0^{3D} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$M_{2D} = \begin{bmatrix} \alpha_x^{2D} & 0 & u_0^{2D} & 0 \\ 0 & \alpha_y^{2D} & v_0^{2D} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (1)$$

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}. \quad (2)$$

2) 同步成像步骤.

由三维摄像机和二维摄像机对空间场景同步成像, 其中三维摄像机获取的信息包括二维灰度图像 GI 和深度图像 DI , 其对应图像分辨率为 $U \times V$, 其中深度图像信息与灰度图像信息按像素坐标一一对应; 二维摄像机同步获取同一空间场景的二维彩色图像 CI .

3) 插补获取三维摄像机高分辨率深度图像.

由于三维摄像机提供的深度图像分辨率较低, 本文对获取的深度图像进行二维插补计算, 以获取高分辨率的深度信息. 由于深度图像与传统二维灰度图像数据结构类似, 可根据实际需要选择已有的二维灰度图像插补算法. 设定横向 X 方向插补系数为 K_x , 纵向 Y 方向插补系数为 K_y , 插补后深度图像表示为 D^I , 其分辨率为 $M \times N$, 其中: $M = K_x \times U$, $N = K_y \times V$. 插补深度图像 D^I 中任意点 $p_I(u_I^{3D}, v_I^{3D})$ 在原深度图像 DI 中对应亚像素坐标 $p_{3D}(u^{3D}, v^{3D})$ 计算关系为

$$u^{3D} = u_I^{3D}/K_x, v^{3D} = v_I^{3D}/K_y. \quad (3)$$

本文选用双线性插补算法^[16]对深度图像 DI 进行插补计算, 则插补深度图像 D^I 中任意点 $p_I(u_I^{3D}, v_I^{3D})$

的深度值 $D^I(u_I^{3D}, v_I^{3D})$ 计算如下:

$$D^I(u_I^{3D}, v_I^{3D}) = \frac{(u_2 - u_1^{3D}/K_x)(v_2 - v_I^{3D}/K_y)}{(u_2 - u_1)(v_2 - v_1)} * D(u_1, v_1) + \frac{(u_1^{3D}/K_x - u_1)(v_2 - v_I^{3D}/K_y)}{(u_2 - u_1)(v_2 - v_1)} * D(u_2, v_1) + \frac{(u_2 - u_1^{3D}/K_x)(v_I^{3D}/K_y - v_1)}{(u_2 - u_1)(v_2 - v_1)} * D(u_1, v_2) + \frac{(u_1^{3D}/K_x - u_1)(v_I^{3D}/K_y - v_1)}{(u_2 - u_1)(v_2 - v_1)} * D(u_2, v_2). \quad (4)$$

4) 计算插补后深度图像 D^I 中任意一个像素点 $p_I(u_I^{3D}, v_I^{3D})$ 对应的三维信息 $(X_C^{3D}, Y_C^{3D}, Z_C^{3D})$.

对插补深度图像 D^I 中任意一点 $p_I(u_I^{3D}, v_I^{3D})$, 假定其对应空间场景中某一三维点 P , P 点在三维摄像机坐标系下的三维坐标记为 $(X_C^{3D}, Y_C^{3D}, Z_C^{3D})$. 由于三维摄像机获取的灰度图像和深度图像在图像坐标上按像素坐标值一一对应, 因此 $p_I(u_I^{3D}, v_I^{3D})$ 点在原深度图像 DI 中对应的亚像素坐标 $p_{3D}(u^{3D}, v^{3D})$ 即等同为在灰度图像 GI 中对应的亚像素坐标 $p_{3D}(u^{3D}, v^{3D})$. 根据线性摄像机成像模型, 采用齐次坐标表示, 三维摄像机灰度图像像素点 p_{3D} 的图像坐标 $(u^{3D}, v^{3D}, 1)$ 与对应的以三维摄像机坐标系为参考的空间三维信息 $(X_C^{3D}, Y_C^{3D}, Z_C^{3D}, 1)$ 关系为

$$Z_C^{3D} \begin{bmatrix} u^{3D} \\ v^{3D} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x^{3D} & 0 & u_0^{3D} & 0 \\ 0 & \alpha_y^{3D} & v_0^{3D} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C^{3D} \\ Y_C^{3D} \\ Z_C^{3D} \\ 1 \end{bmatrix} = M_{3D}(X_C^{3D}, Y_C^{3D}, Z_C^{3D}, 1)^T. \quad (5)$$

对式(5)进一步展开计算可得式(6):

$$Z_C^{3D} \begin{bmatrix} u^{3D} \\ v^{3D} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x^{3D} * X_C^{3D} + u_0^{3D} * Z_C^{3D} \\ \alpha_y^{3D} * Y_C^{3D} + v_0^{3D} * Z_C^{3D} \\ Z_C^{3D} \end{bmatrix}. \quad (6)$$

对图像 D^I 中的任一像素点 $p_I(u_I^{3D}, v_I^{3D})$, 其深度信息 Z_C^{3D} 为该像素点的插补深度图像值:

$$Z_C^{3D} = D^I(u_I^{3D}, v_I^{3D}). \quad (7)$$

将深度信息 Z_C^{3D} 、坐标点 (u_I^{3D}, v_I^{3D}) 和标定参数 M_{3D} 代入式(3)–(6), 可求得三维空间 X 方向坐标 X_C^{3D} 和三维空间 Y 方向坐标 Y_C^{3D} :

$$\begin{cases} X_C^{3D} = \frac{Z_C^{3D}(u_I^{3D}/K_x - u_0^{3D})}{\alpha_x^{3D}}, \\ Y_C^{3D} = \frac{Z_C^{3D}(v_I^{3D}/K_y - v_0^{3D})}{\alpha_y^{3D}}. \end{cases} \quad (8)$$

因此,根据插补深度图像中的任意像素的图像坐标 (u_1^{3D}, v_1^{3D}) ,即可计算出该像素点在三维摄像机坐标系的三维信息 $X_C^{3D}, Y_C^{3D}, Z_C^{3D}$.以齐次坐标表示为

$$(u_1^{3D}, v_1^{3D}, 1)^T \rightarrow (X_C^{3D}, Y_C^{3D}, Z_C^{3D}, 1)^T. \quad (9)$$

5) 深度图像 D^I 与二维彩色图像 CI 的映射.

即求取 D^I 中的任一像素点 (u_1^{3D}, v_1^{3D}) 对应的在二维彩色图像 CI 中的映射点坐标 (\hat{u}, \hat{v}) ;具体方法为:根据已标定的二维摄像机和三维摄像机组成的立体视觉模型,对于空间点 P ,在三维摄像机坐标系下的坐标 $(X_C^{3D}, Y_C^{3D}, Z_C^{3D}, 1)^T$ 与二维摄像机坐标系下的坐标 $(X_C^{2D}, Y_C^{2D}, Z_C^{2D}, 1)^T$ 之间的关系可表示为

$$\begin{bmatrix} X_C^{2D} \\ Y_C^{2D} \\ Z_C^{2D} \\ 1 \end{bmatrix} = \begin{bmatrix} R^T & T^T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_C^{3D} \\ Y_C^{3D} \\ Z_C^{3D} \\ 1 \end{bmatrix}, \quad (10)$$

其中: R 为二维摄像机坐标系和三维摄像机坐标系之间的相对三维空间旋转矩阵, T 为空间平移矩阵.

对于高精度彩色二维摄像机,根据摄像机成像模型,空间点 P 在二维彩色图像中投影点 p_{2D} 的图像坐标 $(u^{2D}, v^{2D}, 1)^T$ 与 P 点在二维摄像机坐标系下的三维坐标 $(X_C^{2D}, Y_C^{2D}, Z_C^{2D}, 1)$ 关系可表示为

$$Z_C^{2D} \begin{bmatrix} u^{2D} \\ v^{2D} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x^{2D} & 0 & u_0^{2D} & 0 \\ 0 & \alpha_y^{2D} & v_0^{2D} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C^{2D} \\ Y_C^{2D} \\ Z_C^{2D} \\ 1 \end{bmatrix} = M_{2D}(X_C^{2D}, Y_C^{2D}, Z_C^{2D}, 1)^T. \quad (11)$$

因此,三维摄像机的高分辨率插补深度图像 D^I 中任一像素点 $p_I(u_1^{3D}, v_1^{3D})$,将式(9)–(11)联立,即可求得 p_I 点在二维摄像机彩色图像中对应匹配点 p_{2D} 的图像坐标值 $(u^{2D}, v^{2D}, 1)^T$.由于两者为匹配点,因此 p_{2D} 像素点对应的颜色、纹理和灰度信息即可等同为 p_I 点的信息.然而由于两摄像机坐标系映射变换的计算处理、三维摄像机插补深度图像和二维彩色图像的分辨率的差异等原因,像素点 p_I 点的映射点 p_{2D} 的坐标值不一定是整数,即对应于二维彩色图像中的亚像素点,如图8所示.为此,本文对映射点 p_{2D} 的坐标值计算值 (u^{2D}, v^{2D}) 进行四舍五入处理,进行取整运算,得到 p_I 点在二维彩色图像上的近似匹配关联点 $\hat{p}_{2D}(\hat{u}, \hat{v})$,并将近似坐标值 (\hat{u}, \hat{v}) 作为 p_{2D} 点的最终图像坐标值:

$$\hat{u} = \lfloor u^{2D} + 0.5 \rfloor, \quad \hat{v} = \lfloor v^{2D} + 0.5 \rfloor, \quad (12)$$

$$u_{2D} = \hat{u}, \quad v_{2D} = \hat{v}, \quad (13)$$

其中 $\lfloor \cdot \rfloor$ 为下取整运算符.

对映射点 $p_{2D}(u^{2D}, v^{2D})$ 对应的二维图像颜色信息,为提高可靠性、降低随机干扰的影响,本文选取近似匹配关联点 $\hat{p}_{2D}(\hat{u}, \hat{v})$ 8邻域像素点的彩色信息平均值作为映射点的图像颜色信息.以RGB颜色模型为例, p_{2D} 点对应的颜色信息 $(r_{p_{2D}}, g_{p_{2D}}, b_{p_{2D}})$ 计算如下:

$$\begin{cases} r_{p_{2D}} = \sum_{\hat{u}-1}^{\hat{u}+1} \sum_{\hat{v}-1}^{\hat{v}+1} r_{u,v}/9, \\ g_{p_{2D}} = \sum_{\hat{u}-1}^{\hat{u}+1} \sum_{\hat{v}-1}^{\hat{v}+1} g_{u,v}/9, \\ b_{p_{2D}} = \sum_{\hat{u}-1}^{\hat{u}+1} \sum_{\hat{v}-1}^{\hat{v}+1} b_{u,v}/9, \end{cases} \quad (14)$$

其中: $(r_{u,v}, g_{u,v}, b_{u,v})$ 分别表示二维彩色图像的像素点 $CI(u, v)$ 对应的红色-R、蓝色-G和绿色-B的颜色信息.由于 p_I 和 p_{2D} 为空间同一点 P 在不同摄像机图像中的匹配关联点,且同时成像,因此 p_I 点颜色信息可等同于 p_{2D} 点的图像颜色信息,从而可以求出 $p_I(u_1^{3D}, v_1^{3D})$ 点对应的颜色信息 $(r_{p_I}, g_{p_I}, b_{p_I})$:

$$r_{p_I} = r_{p_{2D}}, \quad g_{p_I} = g_{p_{2D}}, \quad b_{p_I} = b_{p_{2D}}. \quad (15)$$

根据式(1)–(14),三维摄像机的插补深度图像 D^I 中任一像素点 $p_I(u_1^{3D}, v_1^{3D})$ 均可计算出其对应的三维信息 $(X_C^{3D}, Y_C^{3D}, Z_C^{3D})$ 和高质量彩色图像信息 $(r_{p_I}, g_{p_I}, b_{p_I})$,即如式(16)所示:

$$\text{If } p_I(u_1^{3D}, v_1^{3D}) \in D_{M \times N}^I, \text{ then} \\ p_I \rightarrow \left\{ \underbrace{(X_C^{3D}, Y_C^{3D}, Z_C^{3D})}_{3D}, \underbrace{(r_{p_I}, g_{p_I}, b_{p_I})}_{\text{color}} \right\}. \quad (16)$$

因此,基于本文可同步获取空间场景中高分辨率的三维信息及其对应的颜色信息,而分辨率取决于插补系数和二维彩色摄像机的图像分辨率,有效解决了实时场景三维信息和高质量二维彩色图像信息的融合难题.

5 试验结果与分析(Experimental results and analysis)

本节对本文提出的复合摄像机系统及其匹配融合算法的有效性和匹配精度进行试验验证和深入分析.如图3所示,复合摄像机中的三维摄像机为Swiss-ranger公司的SR3000三维摄像机,二维摄像机为CCD彩色摄像机.试验内容包括:

1) 复合摄像机中二维和三维图像映射匹配误差的量化分析.

2) 基于复合摄像机融合匹配图像的场景三维重建效果分析.

5.1 复合摄像机标定 (The 2D/3D stereo camera calibration)

基于文献[16]立体摄像机标定算法, 对本文提出的复合摄像机系统进行标定. 其二维摄像机内参数为

$$M_R = \begin{bmatrix} \alpha_x^R & 0 & u_0^R & 0 \\ 0 & \alpha_y^R & v_0^R & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 799 & 0 & 557 & 0 \\ 0 & 801 & 313 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (17)$$

复合摄像机的外参数模型, 即二维摄像机和三维摄像机的空间关系模型为

$$R = \begin{bmatrix} 0.9999 & -0.0040 & 0.0117 \\ 0.0042 & 0.9998 & -0.0167 \\ -0.0117 & 0.0167 & 0.9997 \end{bmatrix}, \quad (18)$$

$$t = \begin{bmatrix} -67.7712 \\ -10.3068 \\ 7.6634 \end{bmatrix}. \quad (19)$$

本节主要进行三维摄像机二维图像和二维摄像机图像的映射匹配精度试验. 其中, 实验中三维摄像机二维图像为未插补信息, 其分辨率为 176×144 ; 二维摄像机彩色图像的分辨率为 1028×764 . 如图7所示, 论文用平面棋盘格标定板为参考, 以不同距离在复合摄像机系统的二维/三维摄像机中同时成像, 然后分别提取标定板图像中的角点特征, 并将三维摄像机中的图像角点特征根据本文算法映射到二维摄像机图像坐标系, 通过比较二维摄像机图像角点图像坐标和映射点图像坐标进行匹配误差分析.

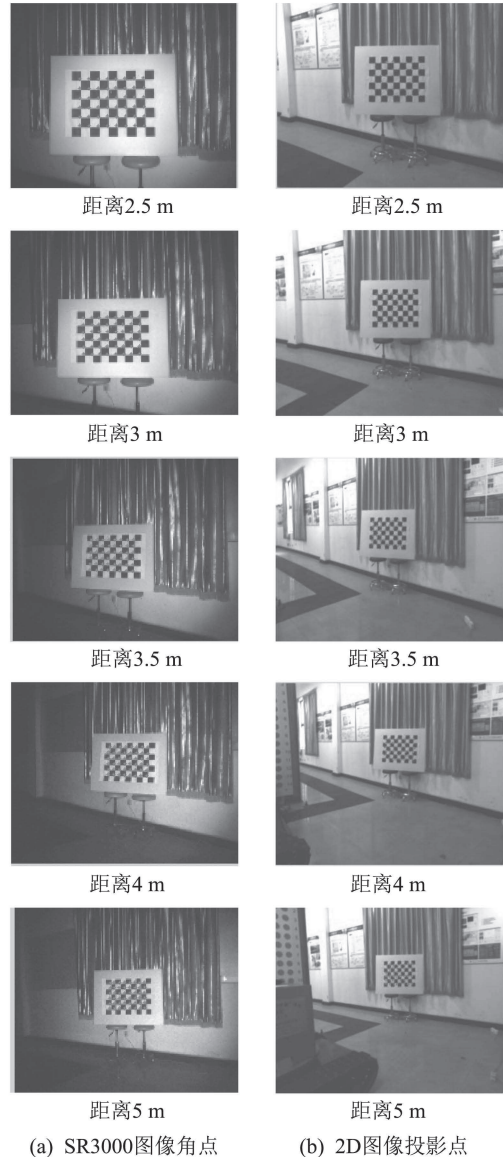
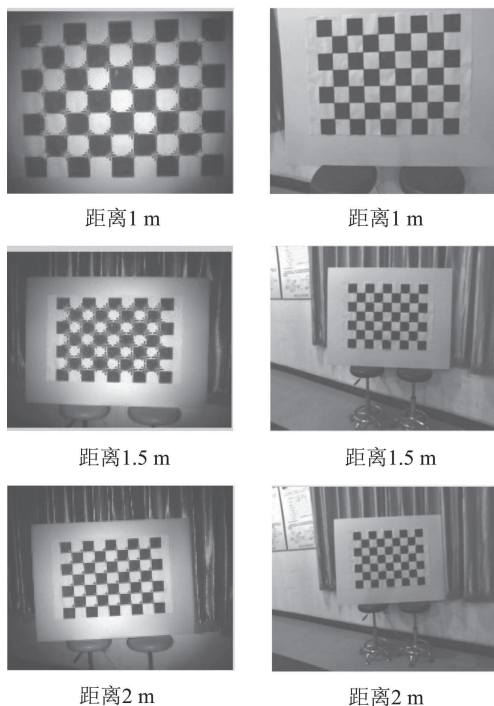


图 7 复合摄像机序列图像匹配测试
Fig. 7 Matching test of 2D/3D stereo image sequence

具体步骤如下:

1) 复合摄像机分别以不同距离 d 对棋盘格标定板进行成像, 图像序列分别标识为 I_d^{2D} (高精度彩色2D摄像机图像)和 I_d^{3D} (三维摄像机SR3000图像), 其中 $d = 1, 2, 3, 4, 5, 6, 7, 8$ 分别表示标定板至复合摄像机的距离为1 m, 1.5 m, 2 m, 2.5 m, 3 m, 3.5 m, 4 m, 5 m.

2) 利用Harris角点提取算法^[17], 分别提取对应棋盘格图像序列 I_d^{2D} 和 I_d^{3D} 中对应的角点序列, 并按从左到右、从上到下的顺序分别将提取的角点存储, 其中三维摄像机中角点表示为 $CSet_d^{3D}[1 : M]$, 二维摄像机中角点表示为 $CSet_d^{2D}[1 : M]$, M 代表提取的角点总数目. $CSet[m] = (i, j)$ 表示该角点在相应图像坐标系中的坐标.

3) 基于本文第4.2节匹配映射算法, 求取三维摄

像机角点序列 $CSet_d^{3D}[1:M]$ 在二维摄像机中的映射点序列 $PSet_d^{3D}[1:M]$, 如图7所示。

4) 假定三维摄像机图像的角点序列中的某一点 $CSet_d^{3D}[m] = (i_m^{3D}, j_m^{3D})$, 其映射对应点 $PSet_d^P[m] = (i_m^P, j_m^P)$, 二维摄像机中的检测角点 $CSet_d^{2D}[m] = (i_m^{2D}, j_m^{2D})$. 定义该点的匹配误差为

$$\delta_d[m] = \sqrt{|i_m^P - i_m^{2D}|^2 + |j_m^P - j_m^{2D}|^2}. \quad (20)$$

本文对所有的三维摄像机图像的角点序列 $CSet_d^{3D}[1:M]$ 执行上述处理, 从而得到不同成像距离下角点序列的匹配误差 $\delta_d[1:M]$, 如图8所示。

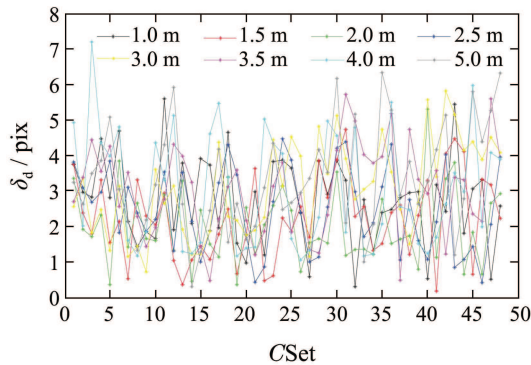


图8 复合摄像机图像特征点匹配误差

Fig. 8 Matching error of 2D/3D stereo image features

为进一步量化分析映射匹配误差, 对上述不同距离对应角点特征序列的匹配误差做均值化处理, 以描述不同距离 d 对应的平均映射匹配误差 $\bar{\delta}_d$, 如图9所示。

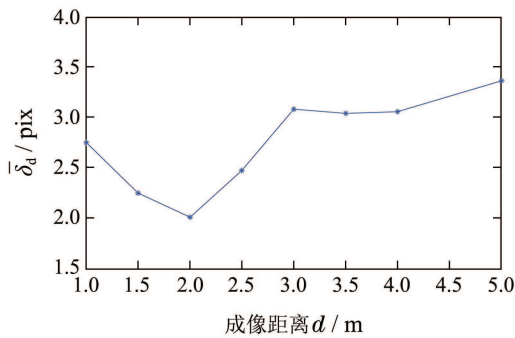


图9 复合摄像机图像特征点平均匹配误差

Fig. 9 Mean matching error of 2D/3D stereo image features

5.2 复合摄像机图像匹配误差分析(Quantitative analysis of 2D/3D stereo image matching)

如图10所示, 对于同一区域用不同分辨率图像表示时, 对于左图像中分辨率为 1×1 , 右图像分辨率为 4×4 , 对于左图像中像素点 $(1, 1)$, 在右图像中可能的像素点为 $(1, 1)$ 和 $(4, 4)$ 围成区域的任一像素点. 根据上节关于匹配误差的定义, 对于本文中复合摄像机系统, 左摄像机分辨率为 176×144 , 右图

像分辨率为 1024×768 时, 允许的匹配误差为

$$\delta = \sqrt{(1024/176)^2 + (768/144)^2} \approx 8.5.$$

由于误差成随机分布, 其最大误差均值约为 $\delta_{\text{mean}} = 4.2$.

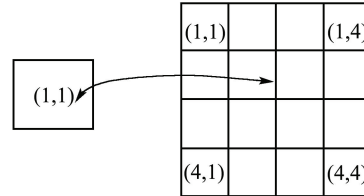


图10 图像间匹配误差理论分析

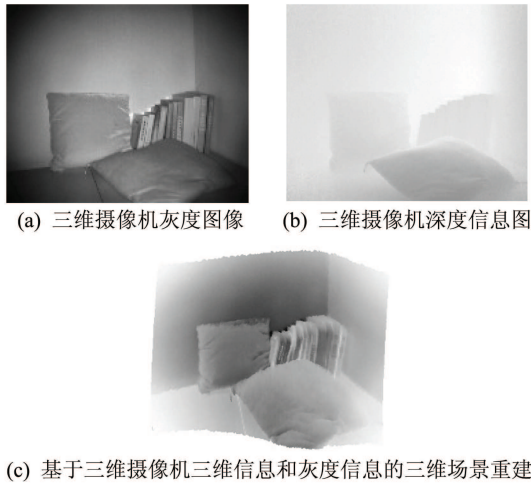
Fig. 10 Matching error analysis between different images

由图9结果可知, 所有测量序列的映射匹配误差均稳定在7个像素以内; 而其相应的误差均值在 $(2, 3.5)$ 像素范围内. 该结果与上述理论分析一致, 在允许误差范围. 此外, 从图9可知, 随感知距离增加, 匹配误差基本呈增长趋势. 其原因可分为两个方面: 1) 由式(4)可知, 由于三维摄像机距离信息的误差也会影响到匹配误差. 根据SR3000摄像机成像特性, 摄像机的距离信息误差随距离增大而增大, 且该误差不可避免; 2) 另外随着成像距离的增加, 图像成像的清晰度下降. 从而在图像角点检测环节可能会引入新的误差. 因此, 在实际图像映射处理时, 由于无需引入图像检测环节, 因此实际算法的匹配精度会比上述结果要高。

通过上述分析可知, 该算法可以实现2D/3D摄像机的图像映射匹配, 匹配精度仅取决于三维摄像机的物理成像性能指标, 而不会引入新的误差。

5.3 基于复合摄像机的三维重建(2D/3D stereo camera based 3D reconstruction)

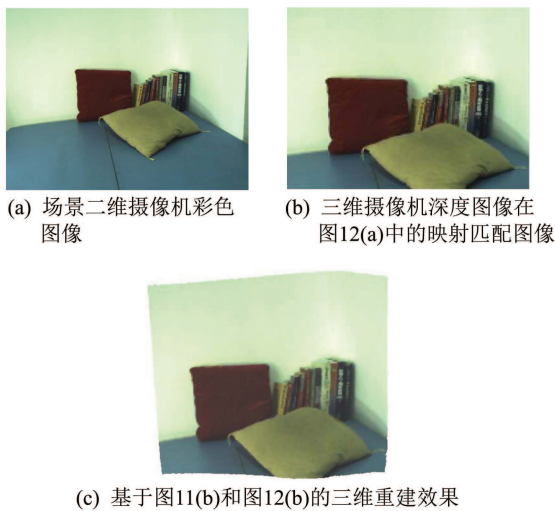
为验证本算法在实际场景中的信息获取优势, 本节分别以同一固定实验室场景为背景, 分别给出基于单三维摄像机成像方案的三维场景重建和基于本文方案的三维场景重建实验, 分别如图11-13所示. 图11给出基于三维摄像机获取的深度图像和灰度图像的三维重建效果, 其中: (a)为三维摄像机灰度图像; (b)为三维摄像机深度图像; (c)为基于三维摄像机三维信息和灰度信息的三维场景重建效果图. 为进一步比较分析本文融合方案有效性, 图12给出了针对三维摄像机原始灰度信息(即插补系数为1)和二维摄像机彩色图像的三维重建效果, 图13给出基于插补后的三维摄像机深度图像和二维摄像机彩色图像融合的三维重建效果, 其中插补算法为双线性插补算法, 横向和纵向插补倍率相同, 插补系数为2, 即 $K_x = K_y = 2$.



(c) 基于三维摄像机三维信息和灰度信息的三维场景重建

图 11 基于三维摄像机的场景重建

Fig. 11 3D environment reconstruction based on SR 3000

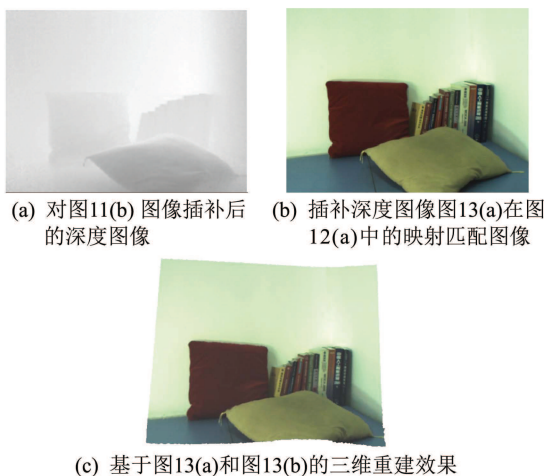


(a) 场景二维摄像机彩色图像 (b) 三维摄像机深度图像在图12(a)中的映射匹配图像

(c) 基于图11(b)和图12(b)的三维重建效果

图 12 复合摄像机三维场景效果(未插补)

Fig. 12 3D environment reconstruction based on the proposed method (No interpolation process)



(a) 对图11(b)图像插补后的深度图像 (b) 插补深度图像图13(a)在图12(a)中的映射匹配图像

(c) 基于图13(a)和图13(b)的三维重建效果

图 13 复合摄像机三维场景效果(插补系数为2)

Fig. 13 3D environment reconstruction based on the proposed method (interpolation coefficients is 2)

从上述试验结果看出, 相对于三维摄像机, 本文复合摄像机方案有效实现了二维摄像机图像和三维摄像机深度图像的匹配和融合, 为三维场景引入了二维彩色图像信息. 此外, 引入插补环节后, 二维彩色图像信息得以更大程度保留, 可以得到更高分辨率和更高成像质量的三维彩色图像信息. 试验结果进一步验证了本文方案的有效性和可行性.

6 分析与总结(Conclusions and analysis)

论文通过深入分析三维摄像机特性和二维/三维摄像机获取信息的互补特性, 借鉴立体视觉技术, 提出了一种复合摄像机的三维视觉信息获取方案. 该方案通过专用的二维摄像机和三维摄像机组成的复合摄像机对空间场景同步成像; 采用二维图像插补算法将三维摄像机的深度图像插补成期望的高分辨率图像, 结合三维摄像机成像特性, 完成二维摄像机的高质量二维彩色图像与插补后的三维摄像机深度图像的匹配关联, 实现同一场景高精度的二维彩色图像和对应三维空间信息的实时获取, 同时保留了二维摄像机图像质量好和三维摄像机三维信息易于获取的优势.

与现有立体视觉系统相比, 本文方案在三维视觉成像范围、实时性和稳定性方面具有显著优势. 在成像范围方面, 本方案可获取三维摄像机视野和成像距离范围的较高质量的匹配图像信息和三维信息, 而立体视觉系统深度信息获取受环境特征稀疏性影响很大, 且深度信息超过一定成像范围后精度迅速下降; 在实时性方面, 本方案不涉及传统立体摄像机匹配算法, 仅需离线复合摄像机系统标定处理、深度插补处理和简单代数运算, 具有很好的实时性. 但是本方案获取的三维信息精度受三维摄像机自身性能制约, 在特定测量应用中, 三维测量精度可能会低于专业立体视觉系统.

由于本文可实时获取场景高质量的二维彩色图像信息和对应三维信息, 从而可利用完整的三维信息、颜色和纹理信息解决场景目标的分割、识别和三维测量等难题, 满足移动机器人导航、工业加工和装配机器人、自动焊接等领域中对高速高质量三维视觉信息获取的迫切要求.

参考文献(References):

- [1] HEBERT M, DEANS M, HUBER D, et al. Progress in 3-d mapping and localization [C] // *Proceedings of the 9th International Symposium on Intelligent Robotic Systems*. Toulouse, France: IEEE, 2001: 145 - 154.
- [2] VASKEVICIUS N, BIRK A, PATHAK K, et al. Efficient representation in three-Dimensional environment modeling for planetary robotic exploration [J]. *Advanced Robotics*, 2010, 24(8/9): 1169 -

- 1197.
- [3] 闫飞, 庄严, 王伟. 移动机器人基于多传感器信息融合的室外场景理解 [J]. 控制理论与应用, 2011, 28(8): 1093 – 1098.
(YAN Fei, ZHUANG Yan, WANG Wei. Outdoor scene comprehension of mobile robot based on multi-sensor information fusion [J]. *Control Theory & Applications*, 2011, 28(8): 1093 – 1098.)
- [4] LAZAROS N, SIRAKOULIS G C, GASTERATOS A. Review of stereo vision algorithms: from software to hardware [J]. *International Journal of Optomechatronics*, 2008, 2(4): 435 – 462.
- [5] ROVIRA-MAS F, WANG Q, ZHANG Q. Design parameters for adjusting the visual field of binocular stereo cameras [J]. *Biosystems Engineering*, 2010, 105(1): 59 – 70.
- [6] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments [C] // *Proceedings of International Symposium on Experimental Robotics (ISER)*. New Delhi and Agra, India: Springer, 2010: 477 – 491.
- [7] LINDNER M, SCHILLER I, KOLB A, et al. Time-of-Flight sensor calibration for accurate range sensing [J]. *Computer Vision and Image Understanding*, 2010, 114(12): 1318 – 1328
- [8] MUFTI F, MAHONY R. Statistical analysis of signal measurement in time-of-flight cameras [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2011, 66(5): 720 – 731
- [9] MEHROTRA S, ZHANG Z, CAI Q, et al. Low-complexity, near-lossless coding of depth maps from kinect-like depth cameras [C] // *2011 IEEE 13th International Workshop on Multimedia Signal Processing (MMSP)*. Hangzhou, China: IEEE, 2011: 1 – 6.
- [10] FALIE D, BUZULOIU V. Noise characteristics of 3D time-of-flight cameras [C] // *2007 IEEE 9th International Symposium on Signals, Circuits and Systems*. San Francisco, USA: IEEE, 2007, 1: 1 – 4.
- [11] PRASAD T D A, HARTMANN K, WEIHS W, et al. First steps in enhancing 3d vision technique using 2d/3d sensors [C] // *2006 Proceedings of the Computer Vision Winter Workshop*. Telc, Czech Republic: Czech Society for Cybernetics and Informatics, 2006: 82 – 86.
- [12] GUDMUNDSSON S A, AANAES H, LARSEN R. Fusion of stereo vision and time-of-flight imaging for improved 3d estimation [J]. *International Journal of Intelligent Systems Technologies and Applications*, 2008, 5(3): 425 – 433.
- [13] KUHNERT K D, STOMMEL M. Fusion of stereo-camera and pmd-camera data for real-time suited precise 3d environment reconstruction [C] // *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Beijing, China: IEEE, 2006: 4780 – 4785.
- [14] HUHLE B, FLECK S, SCHILLING A. Integrating 3D time-of-flight camera data and high resolution images for 3DTV applications [C] // *2007 IEEE 3DTV Conference*. Tubingen: IEEE, 2007: 1 – 4.
- [15] BOUGUET J Y. *Camera Calibration Toolbox for Matlab*[EB/OL]. 2010, http://www.vision.caltech.edu/bouguetj/calib_doc.
- [16] PARKER J A, KENYON R V, TROXEL D. Comparison of interpolating methods for image resampling [J]. *IEEE Transactions on Medical Imaging*, 1983, 2(1): 31 – 39.
- [17] MIROSLAV T, MARK H. Fast corner detection [J]. *Image and Vision Computing*, 1998, 16(1): 75 – 87.

作者简介:

余洪山 (1980–), 男, 博士, 副教授, 从事智能移动机器人系统、机器视觉的研究, E-mail: yuhongshancn@hotmail.com;

赵科 (1985–), 男, 博士研究生, 主要从事机器视觉的研究, E-mail: zhaoke_hnu@hotmail.com;

王耀南 (1957–), 男, 教授, 博士生导师, 从事智能机器人、智能信息处理和智能控制的研究, E-Mail: yaonan@hnu.cn;

唐延东 (1962–), 男, 研究员, 博士生导师, 从事图像处理与模式识别、机器人视觉研究, E-mail: ytang@sia.cn;

张红良 (1981–), 男, 博士, 讲师, 从事摄像测量、视觉导航技术的研究, E-mail: mrzhanghongliang@163.com.