

文章编号 1004-924X(2014)05-1379-09

基于运动恢复的双目视觉三维重建系统设计

王 欣^{1*}, 袁 坤¹, 于 晓², 章明朝³

(1. 吉林大学 计算机科学与技术学院, 吉林 长春 130012;

2. 空军航空大学, 吉林 长春 130022;

3. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033)

摘要:对基于双目立体视觉的一种三维重建系统进行了改进和扩展。将视差细化处理环节引入现有系统,使原视差及相邻视差的匹配代价拟合为一条二次曲线,并为该曲线重新寻找更加精确的视差。进一步将运动恢复计算环节引入系统,通过估计当前视角的摄像机运动矩阵和以跟踪点和摄像机运动矩阵为参数构造能量函数,对能量函数进行优化来有效缩小误差,恢复出准确的运动矩阵。实验结果表明:新增的视差细化处理环节有效提升了重建点云的精度,使细化前后三维重建结果误差平均减少了 16.3%,避免了片状点云现象;新增的运动恢复优化环节,能够精确地恢复摄像机的运动矩阵,优化后三维重建结果平均重投影误差减少了 95.5%;重构后不同视角的点云之间不再孤立,重建模型整体拼接自然。

关 键 词:立体视觉;双目视觉;三维重建;视差细化;运动恢复

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/OPE.20142205.1379

Design of binocular vision 3D reconstruction system based on motion recovery

WANG Xin^{1*}, YUAN Kun¹, YU Xiao², ZHANG Ming-chao³

(1. College of Computer Science and Technology, Jilin University, Changchun 130012;

2. Aviation University of Air Force, Changchun 130022, China;

3. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)

* Corresponding author, E-mail: w_x@jlu.edu.cn

Abstract: An 3D reconstruction system based on binocular stereo vision was improved to obtain higher accuracy and larger scale of 3D reconstruction. A disparity refinement procedure was introduced to the system, so that the matching cost between the original parallax and the adjacent disparity was fitted into a quadratic curve and to re-find a more accurate parallax for the curve. Then, the motion recovery calculation was further applies to estimation of the camera motion matrices and the tracking points and camera motion matrices were taken as the parameters to construct the energy function. Furthermore, the energy function was optimized to reduce errors effectively and reconstruct the motion matrix accu-

收稿日期:2013-05-24;修订日期:2013-07-03.

基金项目:吉林省科技发展计划资助项目(No. 201105016)

rately. Experimental results indicate that the proposed refinement procedure effectively improves the accuracy of reconstruction of point clouds and reduces the error by 16.3 percent in average with avoiding the scale point cloud phenomenon. As the motion recovery calculation steps recover the camera motion matrix accurately, the optimized method reduces the mean re-projection error by 95.5 percent. The point clouds achieved from different angle images are no longer isolated and the reconstructed models are integrated naturally.

Key words: stereo vision; binocular vision; 3D reconstruction; disparity refinement; motion recovery

1 引言

三维重建是一种利用二维投影恢复物体三维信息(形状等)的计算机技术^[1]。随着计算机软硬件的飞速发展,大规模、高精度三维场景的绘制实时性不断提高,难度大幅降低。同时,具有高计算复杂度的三维重建技术也取得了长足进步,一些经典的三维重建算法已被成功地应用在逆向工程、影视娱乐、工业设计和文物保护等领域。

最早提出的三维重建方法是利用建模工具,采用人机交互的方式完成的。该方法的优点是重建模型的自由度高、描述完备。该方法的缺点是当进行大规模现实场景三维重建时,工作量较大。另外,由于该方法以间接的方式从现实场景中获取信息,因而重建场景的真实感不强。随着工业的发展,出现了利用激光扫描仪主动获取物体三维模型的重建方法。该方法获取的三维模型精度高、最为接近真实模型,可满足大多数应用需求。但是该方法只考虑场景的空间信息,导致重建的场景不具有纹理,丧失了一定的真实感;另外,该方法所依赖的光学仪器往往受环境和现实场景规模的限制;而且所采用的部分光学部件价格昂贵,不适合广泛应用。

目前,基于双目立体视觉的三维重建方法是三维重建技术中的热点和难点问题^[2]。不同于上述 2 类方法,双目立体视觉法以自然光下的现实场景图像为参考,所重建的场景具有非常好的真实感。双目立体视觉法所依据的数学理论已较为完善,并已形成了一套完整可行的实施方案,应用前景广阔。然而,该方法仍然面临着诸多挑战。首先,面向特征点的三维重建^[3]虽然可以保证点云数据具有较高的精度,但只能获得稀疏的点云数据;而面向像素点的三维重建^[4]虽然可以保证点云数据具有很高的密度,但却损失了一定

的精度。其次,不同视角下重建的点云数据只有自身的三维结构信息,不具备相对的三维位置信息,这使得无论多大规模的现实场景,其三维重建的点云数据均被局限在一个视角上。

针对双目立体视觉三维重建方法中存在的问题,对该方法进行了扩展和改进。首先,选取一种面向像素的双目视觉三维重建系统,为该系统加入一个视差细化环节,用于在保留原有点云数据高密度的基础上,提升重建精度;其次,为系统加入了一个运动恢复环节,并运用(Structure From Motion, SFM)算法使系统较准确地定位场景相对位置,从而能够让不同视角下场景的点云数据相互联系、拼接自然,突破了单一视角的限制。

2 双目立体视觉原理

双目立体视觉的三维重建方法所依据的原理来源于生物的双目视觉系统^[5-6]。具体过程为:选取同一场景、不同视角的 2 幅图像,用其中的二维信息恢复出场景的可见表面的三维信息。

图 1 为理想状态下的双目立体视觉原理模型。

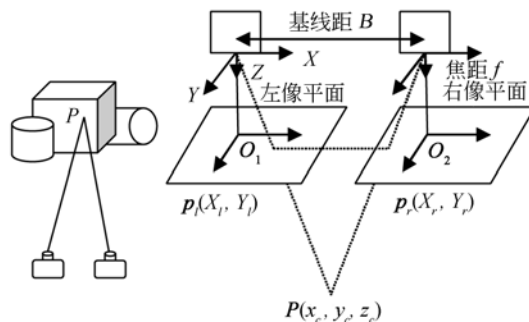


图 1 双目立体视觉原理

Fig. 1 Principle of binocular stereo vision

图中两部摄像机的内参数一致, f 为焦距, 光心

O_1, O_2 为 2 摄像机光轴与像平面的交点, 基线距 B 为 2 摄像机投影中心的直线距离, 2 部摄像机的光轴平行, 像平面重合。场景中的一物点 $P(x_c, y_c, z_c)$ 在 2 部摄像机像平面上分别形成了像点 $p_l(X_l, Y_l), p_r(X_r, Y_r)$ 。由投影几何知识可得, 物点 p 和像点 p_l, p_r 之间的数学关系如下:

$$\begin{cases} X_l = f \frac{x_c}{z_c} \\ X_r = f \frac{(x_c - B)}{z_c} \\ Y = f \frac{y_c}{z_c} \end{cases} \quad (1)$$

由式(1)可见, 2 个像点的 Y 坐标相等, 只在 X 坐标上有所差异。定义视差为: $D = X_l - X_r$, 根据三角测量原理导出物点 P 的三维坐标, 如式(2)所示:

$$\begin{cases} x_c = \frac{B \cdot X_l}{D} \\ y_c = \frac{B \cdot Y}{D} \\ z_c = \frac{B \cdot f}{D} \end{cases} \quad (2)$$

根据双目立体视觉原理, 在 2 台摄像机内外参数一致的前提下, 对于某台摄像机像平面上的任意一点, 只要在另一台摄像机像平面上找到其对应的匹配点, 就可以计算出该点的三维坐标。

3 双目立体视觉三维重建系统

如图 2 所示, 双目立体视觉三维重建系统一般由图像获取、摄像机标定、图像校正、立体匹配和深度值计算 5 部分组成^[7-8]。

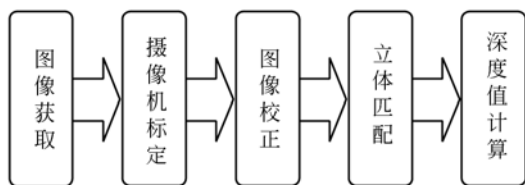


图 2 双目立体视觉三维重建系统组成

Fig. 2 3D reconstruction system of binocular stereo vision

(1) 图像获取。二维图像的获取是三维重建的前提, 系统使用 2 台摄像机获取原始图像。理论上对摄像机的位置没有严格要求, 但为了便于

后续计算, 2 台摄像机应尽可能向前平行对准, 并保证适当长度的基线。此外, 在实际拍摄过程中要考虑视点位置、光照条件、摄像机性能以及景物特点等因素的影响, 以利于后面的立体匹配运算。

(2) 摄像机标定^[9]。目的是为了确定摄像机的内外参数、建立摄像机成像模型, 进而确定物点和像点的数学关系。这种数学关系是二维像点信息反解得出物点三维信息的基础。

(3) 图像校正^[10]。目的是通过一定的变换, 使图像对中的共轭极线位于同一水平线上, 从而使二维空间的匹配问题简化为一维空间的求解。图像经过校正后, 摄像机达到了数学上的理想状态, 有利于提高后续立体匹配的效率和鲁棒性。

(4) 立体匹配^[4]。该步骤是双目立体视觉三维重建系统的核心, 决定了三维重建的最终效果。立体匹配算法在另一幅图像上为每一个像素点搜索到对应像素, 并将结果保存为整数视差值形式, 算法最终获得一个与原图像分辨率一致的视差图。立体匹配算法的计算量通常很高, 所以通常会在立体匹配之前进行图像校正, 以提高算法效率。

(5) 深度值计算。深度值计算的依据是三角测量原理, 利用摄像机模型和视差图重建出场景的三维点云信息。重建精度主要受匹配视差精度和基线宽度的影响, 其一般与匹配视差精度成正比, 与摄像机基线长度成反比, 而当视差精度不足时会出现片层状点云瑕疵。

4 三维重建系统的改进

为了提升系统的重建精度, 突破单一视角的限制, 对双目立体视觉三维重建系统进行了改进, 为系统扩展了视差细化和运动恢复两个处理环节, 扩展后的系统如图 3 所示。

4.1 视差细化

目前, 双目视觉三维重建系统一般采用在离散的像素空间中进行立体匹配, 匹配出的视差值用整数表示。根据图像采样原理, 每幅图像帧中的像素点几乎不可能准确的落在另一幅图像帧的对应像素点上, 因此以整数表示的视差值即意味着精度的损失。在精度需求不高的应用场合, 这种精度损失还可以接受, 但在三维重建中, 精度

的损失有时会严重影响重建模型的视觉效果,甚至会出现片层状点云缺陷^[9]。

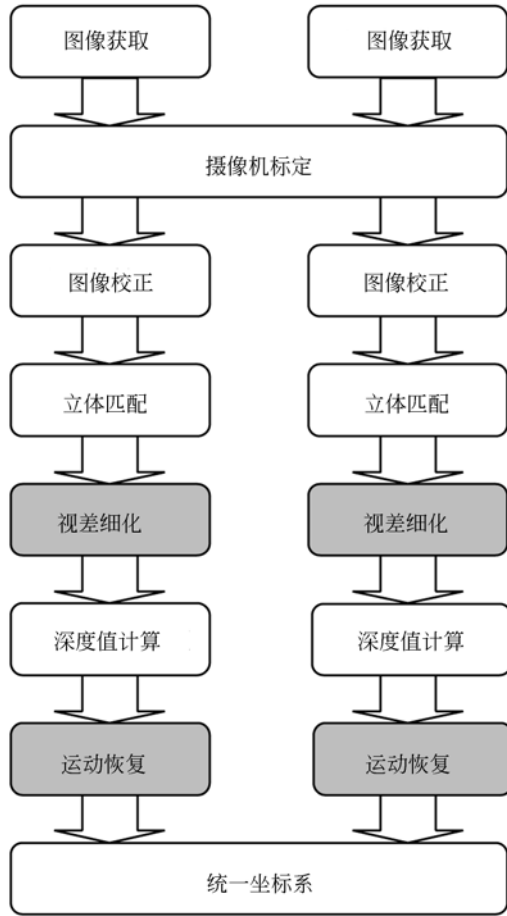


图 3 改进的双目立体视觉三维重建系统

Fig. 3 3D reconstruction system of improved binocular stereo vision

为降低双目视觉三维重建系统中的精度损失,在立体匹配处理环节后插入基于插值函数的视差细化环节^[11]。

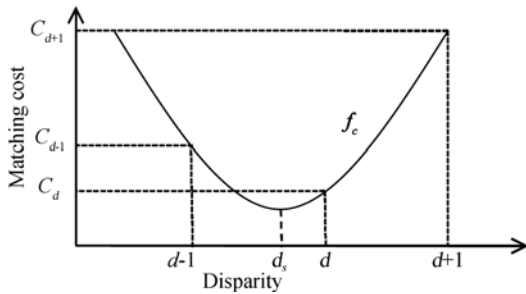


图 4 视差细化中的插值函数

Fig. 4 Interpolation function of disparity refinement

如图 4 所示,视差 d 为某像素点的原整数视差;令 C_d 为视差 d 的匹配代价, C_{d-1} 和 C_{d+1} 分别为 d 左右相邻 2 整数视差值 $d-1$ 和 $d+1$ 的匹配代价。

在匹配代价的亚像素极值附近,匹配代价与视差的函数关系可以近似为一条抛物线,因此,采用二次插值函数 $f(x) = Ax^2 + Bx + C$ 进行视差细化,对 C_d 、 C_{d-1} 和 C_{d+1} 进行插值,得到 3 个等式:

$$\begin{cases} C_d = Ad^2 + Bd + C \\ C_{d-1} = A(d-1)^2 + B(d-1) + C \\ C_{d+1} = A(d+1)^2 + B(d+1) + C \end{cases} \quad (3)$$

利用式(3)计算出二次插值函数 $f_c(x)$:

$$f_c(x) = \frac{C_{d+1} + C_{d-1} - 2C_d}{2}x^2 + \frac{(1-2d)C_{d+1} - (1+2d)C_{d-1} + 4dC_d}{2}x + \frac{(d^2-d)C_{d+1} + (d^2+d)C_{d-1} - (2-2d^2)C_d}{2} \quad (4)$$

最后,在函数 $f_c(x)$ 中重新查找最小值 $f_c(d_s)$, d_s 所对应的值即为最终获取的亚像素级的视差值。

$$d_s = \frac{C_{d-1} - C_{d+1}}{2(C_d - C_{d+1} - C_{d-1})} + d \quad (5)$$

4.2 运动恢复

双目视觉三维重建系统所恢复出的三维点云数据处在一个局部坐标系中,该坐标系在不同视角下不具备一致性,因此点云数据无法实现不同视角下的自然拼接,这使得同时用于描述某个场景的点云数据规模受到单一视角的限制。

为解决单一视角限制问题,在三维重建系统中的深度值计算环节前加入基于 SFM 算法的运动恢复环节。SFM 算法最早被用在稀疏条件下的三维重建^[12],而在本文系统中,将 SFM 算法用于不同视角下摄像机的运动旋转矩阵和运动平移矩阵恢复。SFM 算法对图像序列进行特征点的提取与匹配,建立起图像特征点与三维重建点对应的数学关系,进而恢复出摄像机的运动矩阵。摄像机在局部坐标系中有固定的位置,因此,其运动矩阵对应着点云数据的位置。本文三维重建系统原理如图 5 所示。首先,利用若干(此处为 4)个不同视角的双目摄像机分别获取场景不同侧面的 2 幅图像;第二,每对图像经过三维重建后,

得到场景局部坐标系下的一个侧面的点云数据; 第三, 利用 SFM 算法恢复出不同视角下摄像机的运动矩阵; 最后, 在运动矩阵的变换下, 将局部坐标系下的点云数据统一于一个世界坐标系之中。

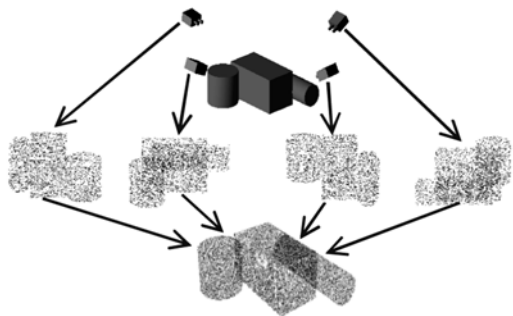


图5 基于运动恢复的双目三维重建系统示意图

Fig. 5 Diagram of binocular 3D reconstruction system based on motion recovery

为了增强实时性和鲁棒性, 本文在 SFM 算法^[12]的基础上进行了部分改进。算法描述如下: 首先, 利用第一个视角的 2 幅图像进行算法的初始化; 然后, 对每一个新拍摄视角, 考虑其左眼图像: a. 提取图像特征点并与跟踪点集进行匹配, 依据匹配关系估计视角的运动矩阵; b. 然后利用稀疏光束平差法对已恢复出的运动矩阵集和跟踪点集进行整体优化; c. 更新和维护跟踪点集。

4.2.1 特征点提取和匹配

SFM 算法在不同阶段都要对图像的特征点进行提取与匹配。尺度不变特征变换 (Scale Invariant Feature Transform, SIFT) 特征点对光照和仿射变换等具有较强的不变性, 适于处理摄像机拍摄的图像^[13]。SIFT 特征点中包含有描述自身属性的特征向量, 当匹配特征点时, 直接计算其特征向量的欧式距离即可。匹配得出特征点对集合, 使用基于 RANSAC 的 8 点算法^[14]计算出相应的基础矩阵, 利用匹配点对与基础矩阵的几何关系滤除误匹配, 进而增强匹配结果的鲁棒性。

获取了具有几何一致性的匹配特征点集合后, SFM 算法将对这些特征点进行三维重建并组织成跟踪点集合。跟踪点实质上即是重建的三维场景点, 同时也关联着能够观测到该场景点的多幅图像上该特征点的像点。

4.2.2 算法初始化

SFM 算法对每一个新加入的视角都进行了一次运动矩阵集合的优化, 优化的实质是一个能量最小化的迭代算法。这种迭代算法通常需要对初始参数有良好估计, 否则极有可能收敛到一个不理想的局部最优解上。文中讨论的双目视觉三维重建系统具有一对严格标定后的摄像机, 因而对系统中 SFM 算法的初始化策略可进行适当精简。具体初始化过程如下:

(1) 将左右眼第一帧图像视为由一部摄像机在不同视角下获取的连续 2 帧图像, 并将第一视角下的运动矩阵固定为单位矩阵;

(2) 利用已知的外参数初始化第二视角的运动矩阵;

(3) 对 2 幅图像帧进行特征点提取与匹配;

(4) 对匹配的特征点进行三维重建, 并作为初始的跟踪点集合。

4.2.3 运动矩阵估计

初始化完毕, 提取每一个新视角左眼图像的特征点, 并将其与跟踪点集进行匹配; 然后, 利用匹配点对直接线性变换法^[12]计算摄像机运动参数。

空间上一点 (X, Y, Z) 在像平面 (u, v) 上的成像方程为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (6)$$

对上式整理变形得到:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{L} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & l_3 & l_4 \\ l_5 & l_6 & l_7 & l_8 \\ l_9 & l_{10} & l_{11} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (7)$$

由上式得到直接线性变换法的基本关系式:

$$\begin{cases} u - \frac{l_1 X + l_2 Y + l_3 Z + l_4}{l_9 X + l_{10} Y + l_{11} Z + 1} = 0 \\ v - \frac{l_5 X + l_6 Y + l_7 Z + l_8}{l_9 X + l_{10} Y + l_{11} Z + 1} = 0 \end{cases}, \quad (8)$$

其中: 坐标 (u, v) 和 (X, Y, Z) 分别为像点和物点坐标, 系数 $l_i (i=1, 2, 3, \dots, 11)$ 是摄像机外参数

(运动参数)、内参数、坐标轴不正交系数和坐标比例不一致系数的函数。由于本文系统中所采用的摄像机已经过严格标定和图像校正,因此除了摄像机外参数以外的所有参数均已确定,也即将 $l_i (i=1,2,3,\dots,11)$ 视为仅由外参数决定的函数。由式(6)可以列出 l_i 系数做未知数的方程:

$$\begin{cases} Xl_1 + Yl_2 + Zl_3 + l_4 + xXl_9 + xYl_{10} + xZl_{11} + x = 0 \\ Xl_5 + Yl_6 + Zl_7 + l_8 + yXl_9 + yYl_{10} + yZl_{11} + y = 0 \end{cases} \quad (9)$$

由 n 个像点与物点的匹配点对可以列出 $2n$ 个关于 l_i 系数的线性方程,共有 11 个未知数,所以至少需要 6 个匹配点对。求解出 l_i 之后,进而由 l_i 与外参数的函数关系式求解出外参数,确定摄像机的运动矩阵。

4.2.4 运动矩阵优化

接下来,SFM 算法要对包含了新估计的运动矩阵的运动矩阵集合和跟踪点集实施稀疏光束平差法(Sparse Bundle Adjustment, SBA)^[15]进行优化。通过对运动矩阵和跟踪点集进行细微的数据调整,使总体的再投影误差最小化,从而恢复出更为准确的运动矩阵。总体再投影误差 e 的定义如下:

$$e = \sum_{i \in \mathcal{I}} \sum_{j \in x(i)} (r_{ij})^2, \quad (10)$$

其中: $\mathcal{I}$ 为图像集合; $x(i)$ 为能够被图像 i 观测到的跟踪点集合; r_{ij} 为跟踪点 j 再投影到图像 i 上的误差:

$$r_{ij} = m_{ij} - u_{ij}, \quad (11)$$

其中: m_{ij} 为跟踪点 j 在图像 i 上对应像点(特征点)的坐标,而 u_{ij} 为跟踪点 j 在图像 i 上的再投影坐标:

$$\tilde{u}_{ij} = K_i(R_i X_i + t_i), \quad (12)$$

其中: K_i 和 R_i, t_i 分别为图像 i 对应摄像机的内参数矩阵和外参数矩阵。

再投影误差最小化问题是一个非线性最小二乘问题,在这个问题的解决上稀疏光束平差法采取了 LM 迭代算法。其中的每一次迭代都进行了如下形式计算:

$$(J^T J + \mu I) \Delta \Phi = J^T r, \quad (13)$$

其中: $\Phi = [\theta, X]$ 为摄像机运动矩阵集合(θ)和跟踪点集合(X)的参数化表示, I 为单位矩阵; μ 为阻尼量,用于减少误差, J 为一个雅克比(Jacobi-

an)矩阵, $J = \partial M / \partial \Phi$,其中: M 为像点(跟踪点)矩阵。每次 LM 迭代都将获得一个 $\Delta \Phi$,进而得到一个 $\Phi_{\text{new}}, \Phi_{\text{new}} = \Phi_{\text{old}} + \Delta \Phi$, Φ_{new} 会使投影误差缩减,每一步迭代获得的 Φ_{new} 都将参与到下一次迭代中去,最终算法收敛,此时所获得的 Φ 可以使再投影误差最小化,运动矩阵集合(θ)包含了新恢复出的运动矩阵。

4.2.5 管理跟踪点集合

SFM 算法开始更新跟踪点集合,以保证与后面的视角衔接。与算法初始化策略类似,针对双目立体视觉三维重建系统的特点,对 SFM 算法的跟踪点集更新策略进行了适当的裁剪。因为系统获取的是有序图像帧,所以相邻 2 帧图像具有足够多的匹配特征点。当前帧与之前所有帧的匹配特征点多数集中在当前帧及上一帧中,因而仅对当前帧与上一帧的左眼图像进行特征点匹配,对匹配特征点进行三维重建,并从中挑选出新的跟踪点组织到跟踪点集合中。

接下来,为了提高算法的鲁棒性,对跟踪点集合进行一次鲁棒性检查,将每一个跟踪点再投影在图像上,求出再投影误差。对于一个跟踪点,若其与某个对应特征点的再投影误差大于指定值,则采取以下策略处理这个跟踪点:

(1) 若该跟踪点已被标记为可疑或者与跟踪点关联的特征点数目等于 2(最少为 2 个),则直接剔除这个跟踪点;

(2) 若该跟踪点未被标记为可疑,且跟踪点关联的特征点数目大于 2,则将该点标记为可疑,并解除跟踪点与这个再投影误差过大的特征点的关联;

(3) 若该跟踪点关联的再投影误差过大或特征点数目大于 2 个,则直接剔除这个跟踪点。

5 实验结果及分析

为便于量化重建结果,采用仿真实验对文中提出的三维重建系统进行检验。实验采用的摄像机为虚拟的双目摄像机,拍摄的物体为一个三维立方体模型。实验中对三维立方体进行了 4 个不同视角下的拍摄和三维重建,图 6 为三维模型 4 个视角的左右眼图像。

图 7(a)为 4 个视角下所计算出的整数值视

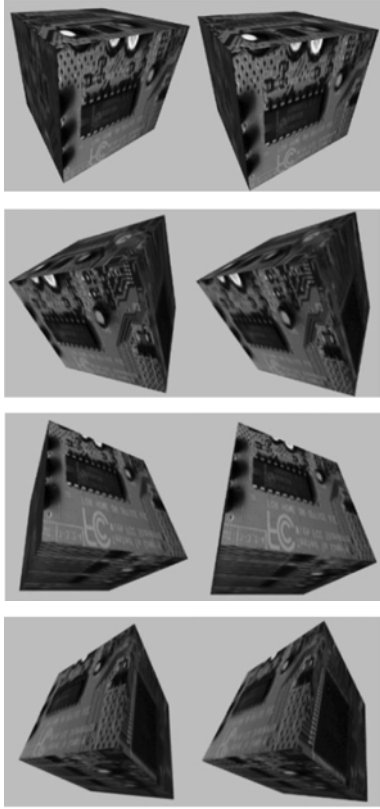


图6 4个视角的左右眼图像

Fig. 6 Binocular images from four views

差图,图7(b)为对4个视差图进行视差细化处理所获得的亚像素级视差图。

由图7可见,视差细化后的视差图更加平滑,去除了原视差图由于精度不足所产生的纹路。图8为对视差图进行深度值计算的重建结果对比。由图8可见,视差细化后的重建模型更加细腻逼真,更接近实际的三维模型。

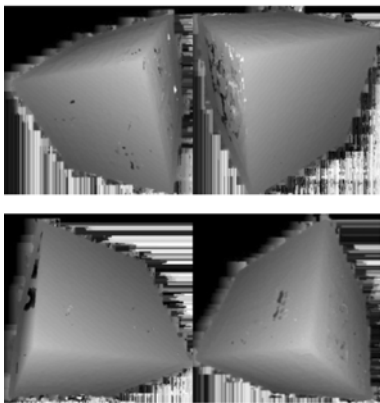
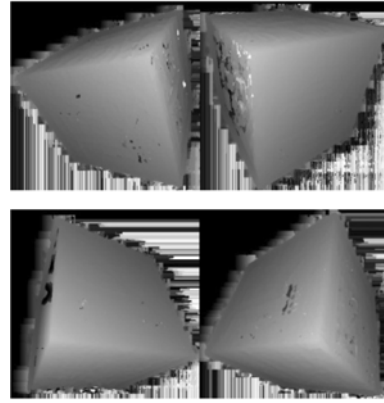
(a) 整数值视差图
(a) Integer value disparity map(b) 亚像素视差图
(b) Sub-pixel disparity map

图7 视差细化实验结果

Fig. 7 Experimental results of disparity refinement

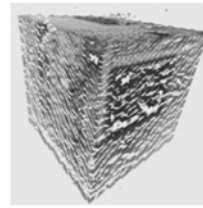
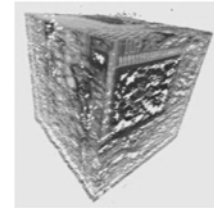
(a) 实际 3D 模型
(a) Actual 3D model(b) 无视差细化的重建模型
(b) Reconstruction model without refinement(c) 视差细化的重建模型
(c) Reconstruction model with refinement

图8 深度值计算的重建结果对比

Fig. 8 Comparison of reconstruction results for depth value calculation

为定量分析算法的重建精度,这里采用平均重建误差对不同视角的三维重建结果误差进行计算。即通过计算重建点与理想立方体模型的8个面中最近面的距离和求平均得到三维重建结果误差:

$$Err = \frac{1}{N} \sum_{j=0}^N \min_{1 \leq i \leq 8} \left\{ \frac{|A_i x_j + B_i y_j + C_i z_j + D_i|}{\sqrt{A_i^2 + B_i^2 + C_i^2}} \right\}, \quad (14)$$

其中: N 为重建点个数,立方体8个平面的方程为 $A_i x + B_i y + C_i z + D_i = 0 (i=1, 2, \dots, 8)$,重建点坐标值为 $(x_j, y_j, z_j), j=0, 1, \dots, N$ 。

4个视角的三维重建误差如表1所示。

表 1 4 个视角三维重建结果误差

Tab. 1 3D reconstruction errors of four views

	视角 1	视角 2	视角 3	视角 4
无细化	0.024 011	0.018 749	0.020 270	0.017 300
细化后	0.020 587	0.015 377	0.016 276	0.015 030

由表 1 可以得出,加入视差细化后,三维重建结果的误差平均减少了 16.3%。

表 2 运动恢复实验结果

Tab. 2 Experimental results of motion recovery

实际的运动矩阵				
1	0.854 845	0.232 296	0.463 981	0
	0.027 412	0.872 725	-0.487 441	0
	-0.518 159	0.429 406	0.739 677	0
	1.56 8991	-1.311 446	0.734 570	1
2	0.861 152	0.089 720	-0.506 792	0
	-0.814 093	0.825 450	-0.469 019	0
	0.399 702	0.563 076	0.723 314	0
	-1.185 222	-1.693 201	0.880 737	1
3	0.913 997	-0.157 845	0.878 756	0
	-0.076 820	0.887 226	0.541 435	0
	-0.398 881	-0.528 582	0.758 096	0
	1.203 744	1.586 580	0.703 334	1
4	0.798 114	0.170 246	-0.577 954	0
	0.186 250	0.842 558	0.505 386	0
	0.572 997	-0.510 999	0.640 745	0
	-1.698 802	1.515 973	1.135 560	1
恢复的运动矩阵				
1	0.854 607	0.226 494	0.467 277	0
	0.031 196	0.875 843	-0.481 586	0
	-0.518 338	0.426 144	0.741 436	0
	1.567 067	-1.308 140	0.724 500	1
2	0.862 354	0.039 711	-0.504 745	0
	-0.313 088	0.825 285	-0.469 980	0
	0.397 895	0.563 319	0.724 121	0
	-1.185 555	-1.698 668	0.878 147	1
3	0.914 108	-0.161 329	0.371 995	0
	-0.074 605	0.834 854	0.545 392	0
	-0.398 549	-0.526 300	0.751 110	0
	1.202 759	1.588 896	0.701 355	1
4	0.810 525	0.173 458	-0.559 430	0
	0.182 808	0.882 505	0.522 989	0
	0.556 445	-0.526 168	0.643 056	0
	-1.674 620	1.535 322	1.109 525	1

对 4 个视角进行运动恢复,4 个视角的实际运动矩阵及恢复后的运动矩阵如表 2 所示。

由表 2 可见,恢复出的运动矩阵基本上与实际的运动矩阵相符合,这说明该算法可用于局部点云的坐标变换。

为定量比较实施与未实施运动矩阵时,SBA 优化算法所恢复运动矩阵之间的差异,在立方体上选取 1 094 个点,采用公式(10)计算重构点的平均再投影误差。其中,SBA 优化前的平均再投影误差为 1.413;SBA 优化后的平均再投影误差为 0.063,运动恢复优化环节平均再投影误差减少了 95.5%。

在运动矩阵的作用下,将 4 个视角的重建点云统一于一个坐标系下的三维点云模型。(如图 9 所示)

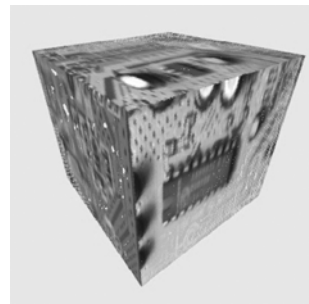


图 9 统一坐标系的三维点云模型

Fig. 9 3D point cloud model in unified coordinate system

5 结 论

本文对双目立体视觉三维重建系统进行了改进,增加了视差细化和运动恢复 2 个处理环节,并给出了重建流程和具体实现方法。首先,对获取的图像进行图像校正和立体匹配,以获得视差图;其次,采用二次函数对得到的视差图进行细化,得到亚像素级的视差值;最后,根据视差值恢复三维点云数据,并经过改进稀疏光束平差法优化后,得出运动矩阵,将点云数据统一到一个世界坐标系中,完成三维重建。实验结果表明,与无视差细化相比,本文改进系统的重建误差平均减小了 16.3%;与无运动矩阵 SBA 优化算法相比,本文改进系统的平均重投影误差减少了 95.5%,从而保证了三维重建的精确性和可靠性,克服了三维重建系统的精度瓶颈,

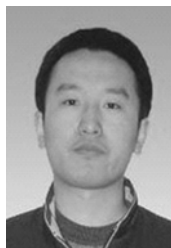
参考文献:

- [1] 全燕鸣,黎淑梅,麦青群.基于双目视觉的工件尺寸在机三维测量[J].光学精密工程,2013,21(4):1054-1061.
QUAN Y M, LI SH M, MAI Q Q. On-machine 3D measurement of workpiece dimensions based on binocular vision [J]. *Opt. Precision Eng.*, 2013, 21(4):1054-1061. (in Chinese)
- [2] 姜宏志,赵慧洁,梁宵月,等.基于极线校正的快速相位立体匹配[J].光学精密工程,2011,19(10):2520-2525.
JIANG H ZH, ZHAO H J, LIANG X Y, *et al.* Phase-based stereo matching using epipolar line rectification[J]. *Opt. Precision Eng.*, 2011, 19(10):2520-2525. (in Chinese)
- [3] BROWN M, LOWE D G. Unsupervised 3D object recognition and reconstruction in unordered datasets [C]. *Fifth International Conference on 3D Digital Imaging and Modeling*, 2005:56-63.
- [4] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International Journal of Computer Vision*, 2002, 47(1-3):7-42.
- [5] SHRIVASTHAVA P, VUNDAVILLI P R, PRATHAR D K. An approach for 3D reconstruction of environment using stereo-vision system [C]. *IEEE Region 10 and The Third International Conference on Industrial and Information Systems*, 2008, 1-7.
- [6] JIA W D, YI W J, JAFAR S. *et al.* 3D image reconstruction and human body tracking using stereo vision and kinect technology [C]. *2012 IEEE International Conference on Electro/Information Technology*, 2012:1-4.
- [7] YOON K J, SHIN M G. Recognizing 3D objects with 3D information from stereo vision [C]. *International Conference on Pattern Recognition*, 2010, 4020-4023.
- [8] SUN J H, JEON B S, LIM J W, *et al.* Stereo vision based 3D modeling system for mobile robot [C]. *International Conference on Control, Automation and Systems*, 2010:71-75.
- [9] ZHANG ZH Y. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11):1330-1334.
- [10] FUSIELLO A, TRUCCO E, VERRI A. A compact algorithm for rectification of stereo pairs [J]. *Machine Vision and Applications*, 2000, 12(1):16-22.
- [11] SHIMIZU M, OKUTOMI M. Precise sub-pixel estimation on area-based matching [J]. *systems and computers in Japan*, 2002, 33(7):1-10.
- [12] HARTLEY R, ZISSERMAN A. *Multiple View Geometry in Computer Vision* [M]. Cambridge: Cambridge University Press, 2004.
- [13] DAVID G, LOWE. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2):91-110.
- [14] HARTLEY R I. In Defense of the 8-point algorithm [C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(6):580-593.
- [15] LOURAKI M I A, ARGYROS A. SBA: A software package for generic sparse bundle adjustment [J]. *ACM Transactions on Mathematical Software*, 2009, 36(1):1-30.

作者简介:



王 欣(1975—),女,吉林白山人,博士,副教授,2006年于吉林大学获得博士学位,主要从事计算机图形学、数字图像处理及信息融合方面的研究。
E-mail:w_x@jlu.edu.cn



袁 坤(1989—),男,内蒙古呼伦贝尔人,硕士研究生,2011年于吉林大学获得学士学位,主要从事计算机图形学、数字图像处理等方面的研究。
E-mail:yuankund@qq.com