

文献引用格式: 王连兵, 叶兵, 马伟东. 基于稠密匹配的单目立体视图生成[J]. 电视技术, 2018, 42(8): 5-9.

WANG L B, YE B, MA W D. Monocular stereo view generation based on dense matching[J]. Video engineering, 2018, 42(8): 5-9.

中图分类号: TP391.41

文献标志码: A

DOI: 10.16280/j.videoe.2018.08.002

基于稠密匹配的单目立体视图生成

王连兵, 叶兵, 马伟东

(合肥工业大学 电子科学与应用物理学院 安徽 合肥 230009)

摘要: 双目立体视觉在小型化系统中的运用存在尺寸限制、分辨率下降等问题, 文章尝试基于稠密匹配技术重建虚拟视图的方法解决单目情形下立体视图的获取问题。文章中用特征匹配信息求解出小幅视角差别和位移的序列图像间的空间信息, 并以此对图像进行匹配前的矫正, 将矫正后图像继承的匹配点作为种子点, 通过图割法优化的区域扩展法得到稠密视差, 根据双目视觉原理完成虚拟视角的重建。实验表明算法对实验场景中的复杂环境具有一定适应性, 在初始位移已知的情况下, 依据图像匹配信息得出实验场景预设位置具有较小影像畸形和缺失的虚拟相机图像。

关键词: 单目立体视觉; 区域增长; 稠密匹配; 图割法; 视图重建

Monocular stereo view generation based on dense matching

WANG Lianbing, YE Bing, MA Weidong

(School of Information Science and Technology, Hefei University of Technology, Hefei 230026, China)

Abstract: There are many problems such as size limitation and resolution degradation in applying the binocular stereo vision to miniaturization system. Virtual view reconstruction based on the dense matching is tried to solve the problem of obtaining the stereoscopic view under the monocular situation. It assumes that the difference between two sequence images in field and angle is little. The strategy proceeds in the following way: first, the feature matching information is used to solve the spatial information between the sequence images, then the spatial information is used to correct the images before matching. The matching points the corrected images inherited are used as the seeds of regional growth. The regional growth method optimized by graph cuts is employed to obtain dense parallax and completes the reconstruction of the virtual perspective according to the principle of binocular vision. Finally, the virtual perspective is projected based on the dense parallax according to the principle of binocular vision. Experiments show that the algorithm has certain adaptability to the complex environment in the experimental scene. And when the initial displacement is known, based on the images' matching information, a virtual camera image with a small scene distortion and little missing scene at the preset location of the experimental scene is obtained.

Key words: monocular stereo vision; regional growth; dense match; graph cuts; view reconstruction

3D 显示技术, 因与传统二维视频图像相比能够给出目标的深度信息更加符合人的视觉特性, 而备受关注。当前常见三维图像源的获取技术主要有双目^[1]与单目^[2]之分, 双目立体图像获取技术通过立体标定过的两架相机即可得到立体观察图像, 实现简单; 相对单目立体图像获取技术就需要添加特殊制定的器材或算法才能得到双目观察图像。但当三维图像获取技术在一些特定环境如水下机器人, 小尺寸内窥镜等中应用时, 单目的实现方式在尺寸、重量、便携等方面较之双目有很大优势。

在单目立体视觉方面, Deepu 等^[3]通过建立相机的离焦模型, 从单相机图像中自动生成立体视觉图像中的第二幅图像。Logothetis 等^[4]提出新的半校准近场光度立体技术, 在光源位置已知, 亮度未知的环境下重构了目标三维表面。Zhou 等^[5]通过根据内窥镜序列图像间成功匹配的特征点间的视差, 对未匹配点进行线性插值的方式获得稠密视差, 并在三维重建后, 重投影得到了立体视觉图像。杨军等^[6]使用单目序列图像间的最佳匹配点集, 通过扩展稀疏点云的方式对目标物体进行三维表面

重建。在稠密匹配方面,李彬彬等^[7]用图割法优化基于置信度传播的稠密匹配策略,提高了底纹理区域和遮挡区域匹配。程思培^[8]等提出一种基于金字塔分层双向动态规划的立体匹配算法,降低了金典动态匹配算法中横向条纹的产生,同时提高了边缘和弱纹理区域的匹配性能。Tatsunori Tanai等^[9]结合马尔可夫随机场对中倾斜平面匹配算法^[10]和曲率正则方法^[11]提出图割法^[12-13]优化的区域扩展运动(Local Expansion Moves, LocalExp)立体匹配方法,能够得到精确的平滑分区视差图。

本文假设平滑运动中相机获得的相邻图像之间的视场与视角均存在一定的变化,但差别不大,此时可近似将相邻或相近的具有一定差别的图像对看作双目视觉中的立体图像对,经过相应立体矫正后即成为双目立体视图,但此时双目立体图像对的基线长度会随相机的运动状态的改变而改变,不能直接用以人眼观察。为此本文结合已知初始环境下的同步定位与建图(Simultaneous Localization And Mapping, SLAM)^[14-15]技术定位出相机的运动轨迹,同时根据校正后图像对的稠密匹配结果生成在物理空间中相对目标视图固定位置的虚拟相机图像。

基于上述构想,本文的主要内容可分为相机位置求取、稠密匹配和视图重建三个部分。相机位置求取部分,本文运用尺度不变特征转换匹配算法(Scale-invariant feature transform, SIFT)^[16]将每幅图像的最优 SIFT 特征匹配集及其对应物点作为相

机运动中的标志。在稠密匹配部分,本文采用三维平面参数约束的视差平面区域扩展算法,提高匹配算法的不同场景的适应性,本文将 SIFT 特征匹配点作为视差平面区域增长的种子点,并以此作为视差优化的限制条件。在视图重建部分,本文应用双目视觉模型,重投影出假定的虚拟相机图像。

1 基于 SIFT 的相机位置定位原理

由于 SIFT 特征匹配具有旋转不变、放缩不变、较强的稳定性,且生成的匹配对多,对光照变化不敏感,特征点的利用率高等特性,能够为摄像机位置的求取提供更佳稳定数据基础,更利于视角重建算法得出稳定、准确的结果^[17]。单目 SIFT-SLAM 相机定位算法的流程如图 1 所示。

单目 SIFT-SLAM 相机定位算法的具体流程为:

(1) 相机内部参数的预求取;通过张正友教授的棋盘标定法对摄像头的内部参数进行标定,其中相机的畸变参数用以对每一帧图像进行畸变校正,相机的特征参数则为三维重建及视角重建使用。对图像进行必要的高斯滤波,减小噪声影响。

(2) 特征点提取与匹配;采用 SIFT 算法进行特征提取与描述,即从高斯尺度空间中具有尺度、旋转不变特征点,根据其相应领域梯度信息生成归一化的 128 维特征描述向量^[18];进行当前帧特征点与特征点库中的特征点匹配,得到特征匹配对。

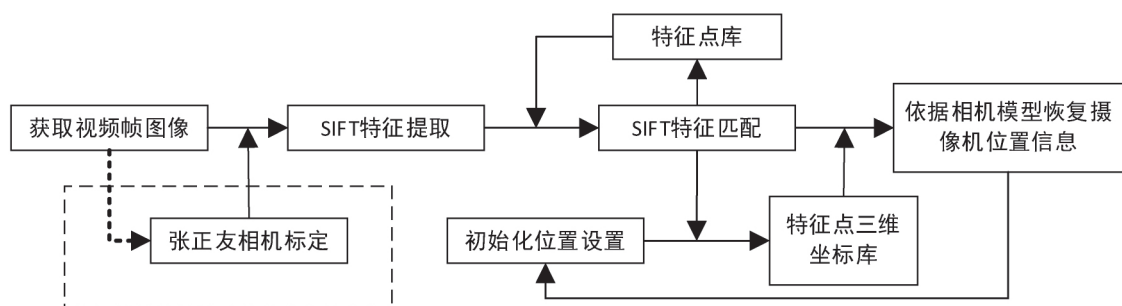


图 1 单目 SIFT-SLAM 相机定位算法

(3) 摄像机位置恢复;通过特征匹配对,实现了当前帧图像特征点与特征点三维坐标库中三维坐标的对应关系,再依据相机投影模型,恢复摄像机的位置、姿态。

(4) 更新特征点参数;运用三维重建技术对已经得到的信息进行重建工作,即可得到新的特征点三

维坐标信息,并更新特征三维坐标库,同时用匹配对更新特征点库。

(5) 对每一帧重复上述步骤,实现对摄像机三维空间位置的不断更新。

该相机定位算法实现了摄像机的三维坐标的求取,为视角重建提供物理尺寸上的参照。

2 种子区域扩展的稠密匹配

2.1 三维视差平面的全局能量函数

文献[10]中指出: 在一个经过校正的立体视觉系统中, 三维空间平面上的点对应的像素间的视差亦是受三维平面约束的, 即每一个像素 p 的视差均用三维平面 $d_p = a_p \mu + b_p \nu + c_p$ 来描述, 由此可知图像中具有相近灰度特征的区域块可共享视差平面。如此, 只需为左右图中的每个像素 p 所在区块寻找其视差平面 $f_p = (a_p, b_p, c_p)$ 限定的映射 $f_p = f(p): \Omega \rightarrow \mathcal{L}$, 并使用全局能量函数(1)最小。

$$E(f) = \sum_{p \in \Omega} \varphi_p(f_p) + \lambda \sum_{(p, q) \in \Omega} \psi_{pq}(f_p, f_q) \quad (1)$$

式中, 第一项为数据项, 用来衡量匹配像素间的图像一致性, 视差平面 f_p 定义了一个从像素 p 到另一图像中对应点的非线性映射。第二项为平滑项, 用以描述相邻像素对 $(p, q) \in \Omega$ 间视差信息的约束关系。

能量函数中的数据项可定义为:

$$\varphi_p(f_p) = \sum_{s \in W_p} \omega_{ps} \rho(s | f_p) \quad (2)$$

其中, W_p 像素 p 为中心的方形窗; ω_{ps} 适应窗口权重, 该权重本文采用引导图像滤波器, 用式(3)描述。

$$\omega_{ps} = \frac{1}{\|W'\|^2} \sum_{k: (p, s) \in W'_K} (1 + (I_p - \mu_K)^T (\sum_K + e)^{-1} (I_s - \mu_K)) \quad (3)$$

式(3)中, $I_p = I_L(p) / 255$, 是一个归一化的色彩向量; μ_K 和 Σ_K 分别是局部回归窗口 W'_K 中灰度矩阵 I_p 的均值和协方差; e 是为避免过度拟合的单位阵。

给定视差平面 $f_p = (a_p, b_p, c_p)$, 式(2)中的 $\rho(s | f_p)$ 是衡量窗口 W 中支撑像素 $s = (s_\mu, s_\nu)$ 与其在右图中匹配点 s' 之间的差异, 且有:

$$s' = s - (a_p s_\mu + b_p s_\nu + c_p) \rho \quad (4)$$

$$\rho(s | f_p) = (1 - \alpha) \min(\|I_L(s) - I_R(s')\|_1, \tau_{col}) + \alpha \min(\|\nabla_x I_L(s) - \nabla_x I_R(s')\|, \tau_{grad}) \quad (5)$$

式(5)中, $\nabla_x I$ 是图像 I 在 x 轴方向的灰度梯度; α 是对颜色项和梯度项的平衡系数; 并且两项分别被 τ_{col} 和 τ_{grad} 截断, 以提高算法的鲁棒性。计算右图的灰度 $I_R(s')$ 时采用线性插值方法, 并用 sobel 算子 $[-0.5 \ 0 \ 0.5]$ 来计算 x 轴方向梯度 ∇_x 。当数据项为

右视图中像素能量函数项时, 交换式(3)、(5)中的 I_L 和 I_R , 并将式(4)改为加视差值。

能量函数中的平滑项定义为式(6):

$$\psi_{pq}(f_p, f_q) = \max(\omega_{pq}, \varepsilon) \min(\bar{\psi}_{pq}(f_p, f_q), \pi_{dis}) \quad (6)$$

式(6)中, ω_{pq} 是对比度敏感度权重, 定义为式(7):

$$\omega_{pq} = e^{-\|I_L(p) - I_L(q)\|_1 / \gamma} \quad (7)$$

其中, γ 是用户定义变量; ε 是小的常数项, 用以限制权重 ω_{pq} , 以增强算法对噪声的鲁棒性。函数 $\bar{\psi}_{pq}(f_p, f_q)$ 以视差的形式代表 f_p 和 f_q 之间的不连续性。

$$\bar{\psi}_{pq}(f_p, f_q) = |d_p(f_p) - d_p(f_q)| + |d_q(f_q) - d_q(f_p)| \quad (8)$$

式(8)中, $d_p(f_p) = a_p p_\mu + b_p p_\nu + c_p$ 。式中, 第一项用 f_p 和 f_q 在 p 点视差之差衡量它们之间的差异, 而第二项表示它们之间的相似性。

2.2 图割法优化的种子区域增长原则

本文中的区域分割采用的固定窗口大小的全图分割, 首先将全图划分为若干 5×5 的单元格, 在对每个单元格建立 3×3 八邻域单元格区域的视差共享区域。对于单元格其视差平面的优化过程一共有三个:

过程1: 初始化视差平面: 根据单元格中已知的视差数据结合扰动量, 随机产生视差平面, 并经行块内图割法优化赋值;

过程2: 单元格内视差优化: 根据循环次数和视差范围产生修正量, 确定新的视差平面, 使用图割法优化块内视差平面的更新范围;

过程3: 共享区域内视差共享: 从单元格内随机选取视差平面, 依据图割法优化平面视差更新范围。

上述三个过程中, 只有具有特征匹配对的单元格才会执行过程1, 即初始化视差平面的过程, 并在该过程结束时, 将随机选取视差平面传递给视差共享区域, 对于不具有特征匹配点的单元格, 只能根据从其他单元格共享的视差平面进行视差优化的后续过程2、3。本文通过上述三个视差优化过程, 将全图能量函数的优化求解转化为每个单元与其视差共享区域内的能量函数最优化问题。

本文基于种子区域扩展的稠密匹配算法的具体步骤如下:

- (1) 对做图像进行单元格分割;
- (2) 具有特征匹配对的单元格执行过程 1, 初始化视差平面;
- (3) 循环, 每个单元格顺序执行过程 2, 3, 求解全局能量函数最优的视差组合;
- (4) 根据全图视差平面, 求解稠密视差。

3 视图重建

双目立体视觉模型中两相机的像平面平行, 右相机 x_r 轴位于左相机 x_l 轴延长线上, 使得两幅图像的对极线恰好在同一水平线上, 由图 2 看出, 空间点在两相机中投影点 x' 和 x'' 只在水平 x 轴方向存在视差 d 。很容易得到:

$$d = \frac{f_l B}{z} \quad (9)$$

其中 B 是立体相机基线长度 $f_l = f_r$ 是左右相机的焦距。

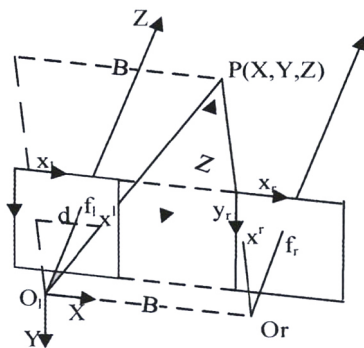


图 2 双目视觉模型

为应用双目视觉模型, 获得假定中虚拟相机的图像, 本文利用稠密视差数据, 使用三角测量法获得目标相机图像所有像素对应的空间点坐标 $P(x, y, z)$, 利用双目立体视觉的视差公式 (9), 可计算得到目标相机 (本文设为左相机) 像素 x' 在虚拟相机成像平面上的对应点 x'' , 此时 x'' 并未正好落在像素点上, 虚拟相机像素点 u' 的具体灰度值难以确定。为此, 本文根据 $x' \rightarrow x''$ 的一一映射, 建立反向 $u' \rightarrow x'$ 的一对多映射, 取 u' 对应多点 x' 的平均值。

4 实验结果及分析

为验证本文算法, 本文使用焦距为 6 mm, 视场角为 60° , 分辨率为 640×480 的摄像拍摄自行搭建的真实实验环境, 如图 3 所示。本文算法的运行环境为: Windows7 64 系统, Intel Core i5 主频 2.6 GHz; 在 Visual Studio2017 编译平台下, 基于 Opencv3.1.

0, 使用 C++ 语言实现。

在虚拟相机视图的投影试验中, 使用三张相机在特定位置时拍摄的图像模拟相机移动过程中的拍摄图像, 图 3 中 a, b, c 三张图像分别为相机顺序左移移动 5 mm 时拍摄的图像。本文的验证策略为: 设图像 a 和 b 间的位移幅值 $T_0 = 5$ mm, 算法通过 T_0 计算图像 b, c 间的位移 T_1 , 并通过对 b, c 进行稠密匹配后的视差数据, 对图像 c 进行虚拟相机投影。试验中为方便对比, 将虚拟相机位置设置为图像 b 对应的相机位置, 即图像 c 右侧 5 mm 处。在本文场景中参数的设置为: 数据项中 $e = 2.5$, $\alpha = 0.85$, $\tau_{col} = 7.0$, $\tau_{grad} = 2.0$; 平滑项中 $\lambda = 10$, $\gamma = 15$, $\epsilon = 0.01$, $\tau_{dis} = 0.5$ 。

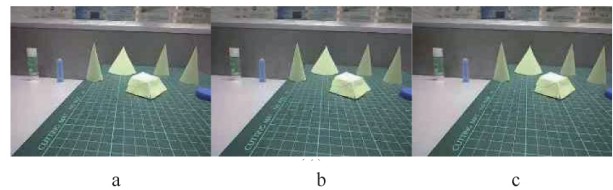


图 3 实验场景

通过算法计算得到的 T_1 的模值为 4.59 mm, 由于实验并未在工作台上进行, 该误差在目测误差范围之内。为加速匹配与算法的鲁棒性, 稠密匹配前需先对图像 b, c 进行立体校正, 图 5 中图像 a, c 分别对应 b, c 的校正结果。图 4 为稠密匹配算法的稠密匹配结果。

从图中可以看出, 对强纹理部分具有很强鲁棒性, 表面接触区域视差传递良好, 但是算法仍存在倾斜面匹配中的阶梯视差情况, 且对于跳跃视差区域存在明显的过平滑。图 5 中图像 b 为最终重建的虚拟相机图像, 与原图 a 相比, 存在明显的视场变化, 对比相应实际视图可以看到两图中的直线的倾斜角度几乎一致, 表明虚拟视角已经很接近实际校正图像; 另外虚拟视图像素有实际视图像素间存在一定的整体偏移, 初步分析是由于实验中位移测量、计算中的误差及部分投影误差所致。



图 4 图像 b, c 的稠密视差图

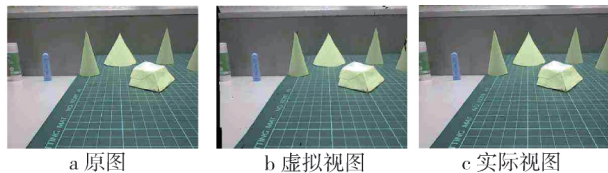


图5 校正后视图和重建虚拟视图

5 小结

基于相机运动过程中视角与视场小幅度变化假设,本文提出了结合 SLAM 与图割法优化的种子区域扩展的稠密匹配技术获取单目立体视图的算法。实验验证表明运用基于 SIFT 的 SLAM 位置求取定位模块能够在初始位置已知的情况下给出相机的位移信息,同时基于种子区域扩张的稠密匹配算法,在高纹理和视差连续区域具有较强鲁棒性,并成功从单目相机的序列图像重投影得到了假定位置处具有较小影像畸形和缺失的虚拟相机视图。

参考文献:

- [1] 高艳艳. 双目内窥镜三维重建方法的研究[D]. 天津: 天津工业大学, 2017.
- [2] GENG J, XIE J. Review of 3 - D Endoscopic Surface Imaging Techniques [J]. IEEE Sensors Journal, 2014, 14 (4): 945 - 960.
- [3] DEEPU R, HONNARAJU B, MURALI S. Automatic generation of second image from an image for stereovision [C]. International Conference on Advances in Computing, Communications and Informatics. IEEE, 2014: 920 - 927.
- [4] LOGOTHETIS F, MECCA R, CIPOLLA R. Semi - Calibrated Near Field Photometric Stereo [C]. Computer Vision and Pattern Recognition. IEEE, 2017: 4521 - 4530.
- [5] ZHOU J, ZHANG Q, Li B, et al. Synthesis of stereoscopic views from monocular endoscopic videos [C]. Computer Vision and Pattern Recognition Workshops. IEEE, 2010: 55 - 62.
- [6] 杨军, 石传奎, 党建武. 基于序列图像的鲁棒三维重建方法[J]. 计算机应用, 2011, 31(6): 1566 - 1568.
- [7] 李彬彬, 王敬东, 李鹏. 基于图像分割的置信传播立体匹配算法研究[J]. 红外技术, 2011, 33(3): 167 - 172.
- [8] 程思培, 达飞鹏. 基于动态规划的分层立体匹配算法研究[J]. 仪器仪表学报, 2016, 37(7): 1665 - 1672.
- [9] TANAI T, MATSUSHITA Y, SATO Y, et al. Continuous 3D Label Stereo Matching using Local Expansion Moves [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016(99): 1.
- [10] BLEYER M, RHEMANN C, ROTHER C. PatchMatch Stereo - Stereo Matching with Slanted Support Windows [C]. British Machine Vision Conference. 2011: 14. 1 - 14. 11.
- [11] LI G, ZUCHER S W. Differential Geometric Inference in Surface Stereo [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2010, 32(1): 72.
- [12] 钟德海, 姚剑敏, 郭太良. 一种基于图像分割的立体匹配算法[J]. 电视技术, 2014, 38(19): 5 - 7 + 12.
- [13] 段军, 何明一, 戴玉超. 基于线段分割的图割立体匹配[J]. 电视技术, 2012, 36(11): 15 - 18.
- [14] 吕丹娅, 姚剑敏, 郭太良. 改进 ORB 算法在单目视觉 SLAM 特征匹配中的应用[J]. 电视技术, 2016, 40(11): 107 - 111 + 115.
- [15] 权美香, 朴松昊, 李国. 视觉 SLAM 综述[J]. 智能系统学报, 2016, 11(6): 768 - 776.
- [16] 李校林, 李银, 王志锋. 基于 SIFT 算法在遥感图像匹配中的研究[J]. 电视技术, 2016, 40(9): 108 - 111.
- [17] LUO J, GWUN O. A Comparison of SIFT, PCA - SIFT and SURF [J]. International Journal of Image Processing, 2013, 3(4): 143 - 152.
- [18] 胡海青, 谭建龙, 朱亚涛, 等. 改进 SIFT 算法在文字图像匹配中的应用[J]. 计算机工程, 2013, 39(1): 239 - 243.



作者简介:

王连兵(1992—), 硕士研究生, 主研: 机器视觉;
叶兵(1982—), 博士, 主研: 激光技术应用、电子技术应用及光学精密测量;
马伟东(1992—), 硕士研究生, 主研: 机器视觉。
责任编辑: 辛美玉 收稿日期: 2018 - 05 - 28