

双目立体匹配算法的研究与实现

黄松梅 毕远伟 许 晓

(烟台大学 计算机与控制工程学院 山东 烟台 264005)

摘要: 立体匹配是计算机立体视觉的研究核心. 根据约束条件的不同, 将立体匹配算法分为三类: 全局立体匹配算法、半全局立体匹配算法和局部立体匹配算法. 介绍了双目立体匹配的原理和立体匹配算法, 深入研究了全局匹配算法中的 GC 算法、局部匹配算法中的 BM 算法和半全局匹配 SGM 算法. 结合实验结果, 分析了三种算法的匹配效果及匹配速度, 并对各自优缺点进行了比较和总结.

关键词: 计算机视觉; 立体匹配; 图割法; 块匹配; 半全局匹配算法

中图分类号: TP391 文献标志码: A 文章编号: 1673-8020(2018)01-0025-06

立体视觉是计算机视觉的一个重要分支, 在机器人视觉、目标识别、医学诊断、工业检测、三维测量等领域发挥着重要的作用. 立体视觉分为图像的采集与预处理、摄像机标定、立体校正、立体匹配、深度信息提取等几部分^[1], 其中立体匹配是双目立体视觉中最关键也是最难实现的一步. 匹配的精度和速度直接影响后续三维重建的效果, 因此匹配算法具有重要的研究意义, 也一直是研究的热点之一. 立体匹配的目的^[2]是从不同视角拍摄同一景物得到图像对, 并与像素之间形成一种对应关系, 也就是在两幅图像中找出同名像点(相同特征的点), 从而计算出视差, 获取图像的三维信息.

计算机视觉从上世纪 50 年代发展以来, 研究重点在于提高匹配的精度和速度, 并已取得大量的研究成果. 文献[3]提出了一种基于交互式图像分割的立体匹配方法, 设置种子点, 并采用快速图割算法完成区域分割提取; 与现有的图割法相比, 该方法匹配准确且运算量小. 文献[4]提出了一种改进的跨尺度代价聚合立体匹配算法, 用补丁匹配近似算法代替传统的胜者为王算法, 提高了匹配速度和精度. 文献[5]基于自适应窗口及对特征融合的立体匹配算法, 改进了弱纹理及斜面等情况的匹配效果. 文献[6]利用双目立体视

觉构建 3D 模型, 将立体视觉应用于移动机器人感知系统. 本文系统地介绍了双目立体视觉的模型及原理. 根据深度优化所采用的不同策略将立体匹配算法分为基于窗口的局部立体匹配算法、基于能量函数的全局匹配算法和半全局匹配算法, 并详细介绍了算法的原理. 结合实验结果对三类算法进行对比, 并计算出误匹配率, 定量评测出算法的优缺点.

1 双目立体视觉模型

物体在视网膜上的成像信息经大脑匹配, 可重建出物体的三维几何结构. 双目立体视觉是根据人眼成像的特点, 采用双目立体相机在同一时刻对同一场景进行拍摄; 并基于视差原理^[2], 对两幅图像进行三维重建. 图 1 是一种理想情况下双目立体视觉模型.

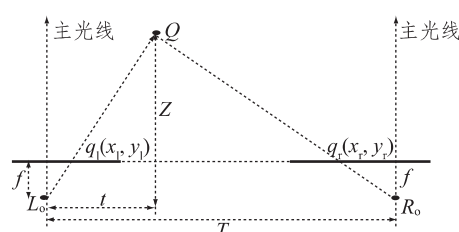


图 1 理想双目立体视觉模型

收稿日期: 2017-05-15; 修回日期: 2017-10-23

基金项目: 国家自然科学基金(61602399)

第一作者简介: 黄松梅(1989—), 女, 山东济宁人, 硕士研究生, 研究方向为计算机视觉测量. E-mail: Hsongmei@126.com

通信作者: 毕远伟(1972—), 男, 山东荣成人, 副教授, 硕士研究生导师, 硕士, 研究方向为计算机视觉测量和嵌入式系统. E-mail:

biyuanwei@126.com

理想的双目立体模型满足以下条件:相机的光轴相互平行,在无穷远处相交;两个相机的像平面精确位于同一平面,而且图像中心的连线与两相机投影中心的连线平行;两相机的成像行对准,即两图像的像素行平行且对应行在同一水平线上;所有的极点均位于无穷远处,而且对应极线严格相互对准;两相机的焦距相等;相机成像的主点与图像中心重合。

在双目立体视觉系统中,一般要将两成像面校正为图1所示的理想情况,此时,左右匹配点处于同一水平线上。一幅图像上的特征点在另一幅图像上的匹配点一定在对应的极线上。立体匹配时,只需在另一幅图像的对极线上匹配即可,而无需对另一幅图像上的所有特征点进行一一搜索。如图1所示,点 Q 在左右平面上的成像点为 q_l, q_r ,左图像上的点 q_l 在右图像上的匹配点 q_r 一定在极线上^[2],因此,通过极线约束,二维搜索空间也就转化为一维。

空间点的深度是通过三角测量原理计算得到。由图1可知 q_l, q_r 纵坐标值相等,横坐标值分别为 x_l, x_r ,图中 L_o, R_o 分别为左右相机的投影中心,相机焦距为 f ,点 Q, q_l, q_r, L_o, R_o 位于同一平面内。视差定义为同一空间点在两图像物理坐标系 x 轴坐标之差,即 $d = x_l - x_r$,结合三角相似原理得到

$$Z = \frac{fT}{x_l - x_r} \quad (1)$$

由式(1)可知,空间点深度与视差成反比,与焦距、基线距离成正比。双相机标定时可计算出 f 和 T ,视差 d 通过立体匹配计算,所以立体匹配是双目视觉中的关键一步。

2 立体匹配算法

立体匹配的步骤包括:图像预处理、匹配代价计算、视差计算和视差优化。立体匹配的匹配基元可分为三种:点、线和块。本文根据匹配策略将立体匹配算法分为全局匹配算法、半全局匹配算法和局部匹配算法。全局匹配算法的重点是最小化全局能量函数,半全局匹配算法是使用互信息计算匹配代价,而局部匹配算法的重点是构造匹配代价函数及聚合视差^[7]。

在不同的应用领域,对匹配精度的要求不同,因而需要采用不同的匹配算法。例如目标识别、人

体跟踪等领域,不需要稠密的视差图,但对算法的实时性要求较高;而工业检测和医学诊断等就需要精密的视差图。不同的匹配算法各具特点,适合于不同的场合。

局部立体匹配算法以像素及其周围的小区域为约束,构造匹配代价函数,可分为:区域匹配、特征匹配和相位匹配^[8]。特征匹配基于几何不变性原理,匹配时选取图像的拐点、边缘、线段等为匹配基元。由于几何特征的不连续性和稀疏性,特征匹配只能得到稀疏的视差图。相位匹配基于傅里叶平移定理,寻找局部相位相等的对应点。区域匹配以邻域窗的灰度为匹配基元,以相似性测度为依据构造视差图。由于区域匹配算法能得到相对密集的视差图,鲁棒性较强,使得它在局域匹配算法中占据重要的位置。

基于全局的立体匹配算法实质是最小化全局能量函数^[9],该算法分为三类:动态规划算法、图割法和置信传播算法。动态规划算法是在左右视图对应扫描线上,寻找最小匹配代价路径,将全局问题划分为小问题,并分阶段求解。置信度传播算法引入置信度和信号思想,通常与其它算法结合使用。图割法将图像分割引入立体匹配,通过迭代优化寻找全局最优解。

半全局匹配算法基于互信息计算匹配代价,采用平滑约束,引用动态规划算法搜索最优路径,求解视差。

全局匹配算法中常用的是图割法,相对而言它是处理效果最好的一种匹配算法。局部匹配算法中使用较多的是区域匹配算法,这种方法匹配速度快,处理简单。半全局匹配算法无论从效果还是速度分析,都处于全局匹配算法和局部匹配算法之间,具有一定的参照作用。本文将重点介绍区域匹配算法、图割法和半全局匹配算法三种算法。

2.1 区域匹配算法

区域匹配算法又称为块匹配算法或基于窗口的匹配算法。该算法引入支撑窗的概念,并假设支撑窗内的像素具有相同的视差^[10]。首先需要设定小窗口,根据窗口提取的特征向量在待匹配图像上遍历搜索,计算该窗口与遍历过程中每一个窗口间的相似程度,以相似最大的窗口为最终结果。匹配特征可以是像素或者像素块,也可以根据图像特征自适应调节窗口大小^[10]。窗口的选择影响匹配的准确性,窗口过小不能包含灰度上的足够

变化;窗口过大影响计算效率且容易受摄像机投影畸变的干扰.文献[11]提出了比较经典的非自适应窗口大小的算法,认为支撑窗内视差遵循高斯分布,视差估计基于初始视差;在窗口四邻域分别进行一个像素的扩展,再计算新窗口的视差不确定度,最后比较扩展前后的视差不确定度.Yoon和Kweon^[12]提出根据窗口内像素关系,在不改变窗口形状和大小的情况下,适当调节像素的权重,可提高算法的灵活性.文献[13]提出了自适应窗口,针对每个像素计算最优支撑窗.

窗口形状和大小的选择难以在线性时间内求解,搜索最佳窗口时有大量的优化算法,包括赢者通吃、非线性扩展、最小比例圆、动态规划等.最常见的是赢者通吃方法,较直观且方便实现,其原理是先计算不同匹配窗口下每个像素点对应的匹配代价,再选择最小匹配代价窗口作为最终的匹配窗口^[13].

区域匹配算法首先计算匹配代价,视差 $d \in (d_{\min}, d_{\max})$ 在左视图中选取以 $p_l(x, y)$ 为中心的矩形窗口,在右视图中寻找以 $(x + d, y)$ 为中心的矩形窗口,用匹配代价函数计算两个窗口的相似性.匹配代价函数通常有平方差之和(SSD)、归一化互相关(ZSSD)、绝对差之和(SAD)等^[14],函数表达式分别为:

$$C(x, y, d) = \sum_{(x, y) \in U_p} (I_l(x, y) - I_r(x + d, y))^2, \quad (2)$$

$$C(x, y, d) = \sum_{(x, y) \in U_p} ((I_l(x, y) - I_l) - (I_r(x + d, y) - I_r))^2, \quad (3)$$

$$C(x, y, d) = \sum_{(x, y) \in U_p} |I_l(x, y) - I_r(x + d, y)|. \quad (4)$$

式(2)~(4)中, I_l 和 I_r 表示左图和右图的像素灰度, U_p 表示以 P 为中心的邻域窗口.本文采用SAD查找左右图对中的最优匹配点,该算法引入对极约束查找匹配点.对极约束就是立体图像对中的极线相应对准,即立体校正已完成的工作,使得图像对共面且行对准.BM算法的主要步骤^[2]如下:

1) 预过滤,使得图像对亮度归一化并加强图像纹理程度;

2) 沿极线采用SAD窗口进行匹配搜索;

3) 再过滤,去除误匹配点.

2.2 基于能量函数的图割算法

Tao等^[15]将图像分割与立体匹配相结合,并将基于像素的匹配问题转化为基于图像分割的匹配.Boykov等^[9]提出基于图割(Graph Cut)的立体匹配技术,将能量最小化问题转化成图像最小割问题.为了使能量最小化,需求解一个标记组合,这是一个迭代求解的过程.图割法的提出,使马尔科夫随机场能量最小值问题得以解决^[16].Graph Cut模型的核心问题是构造数据约束项和平滑约束项,使得到的分割结果更加准确.

Graph Cut模型中能量函数的表达式为:

$$E(f) = \sum_{\{p, q\} \in N} V(f_p, f_q) + \sum_{p \in P} D_p(f_p), \quad (5)$$

式中: P 表示所有像素, p 表示某个像素位置; N 表示邻接关系; D 是数据约束, f 是标记, $D_p(f_p)$ 表示像素 p 的视差为 f_p 时的匹配代价; V 称为光滑约束,表示相邻像素间在视差上相互影响的关系.能量函数最小化问题等价于最小割求解问题,而最大流是解决最小割问题的方法^[16],所以,Graph Cut引用网络流算法来解决能量函数最小化问题.图割法首先依据每个因子在能量函数中的取值构造赋权图,然后使用赋权图中的最小割极小化目标能量函数,将全局能量最小化问题等价为赋权图中的最小割问题,最后采用最大流算法求解.

图割法分为以下几个步骤:

1) 使用某种分割方法,对图像进行一定程度的分割;

2) 通过基于区域的局部立体匹配算法快速求出各点的初始视差;

3) 将各个区域进行结合视差信息的平面化建模,得到平面集合^[17];

4) 对步骤3)得到的平面集合进行修正;

5) 以区域为单位进行全局匹配,得到各个区域最终视差.

2.3 半全局立体匹配算法

半全局匹配算法以互信息作为相似性测度,匹配点只能搜索周围路径上的像素,不能基于整幅图像,从而导致在弱纹理区域产生误匹配^[18].

半全局匹配算法的主要步骤如下:

1) 左右扫描线上逐像素匹配计算;

2) 聚合匹配代价,采用8条规划路径进行匹配代价聚合;

- 3) 计算视差,采用迭代的方法计算视差;
- 4) 去除误匹配的像素点,采用左右图像同名像点的一致性来去除无效匹配点.

3 实验结果及算法比较

实验选用 Qt Creator 开发平台,基于 OpenCV 进行开发,图像由 Middlebury 资料库提供^[19]. 立体匹配算法的结果是得到稠密准确的视差图. 现在仍未出现一种速度快、匹配精确并适合所有场景的匹配算法,不同的匹配算法有各自的优缺点. 本文对块匹配算法(BM)、半全局匹配算法(SGM)、图割法(GC)进行实验,实验结果如图2~4所示. 图5是真实视差图.

立体匹配公认的定量评价标准为误匹配率,误匹配率公式如下:

$$B = \frac{1}{N} \sum_{(x,y)} |d_c(x,y) - d_t(x,y)| > \delta_d, \quad (6)$$

式中: d_c 表示计算所得视差值, d_t 表示真实视差值, δ_d 表示阈值.

由式(6)可分别算出 BM、SGM 和 GC 三种算法的误匹配率,如表1所示.



图3 SGM 视差图

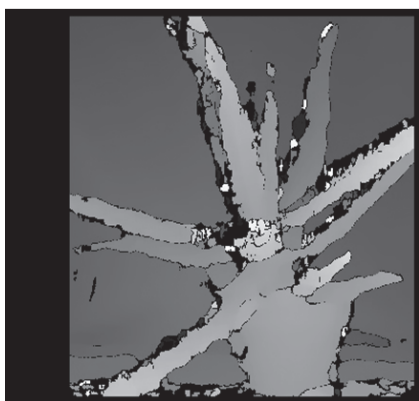


图2 BM 视差图

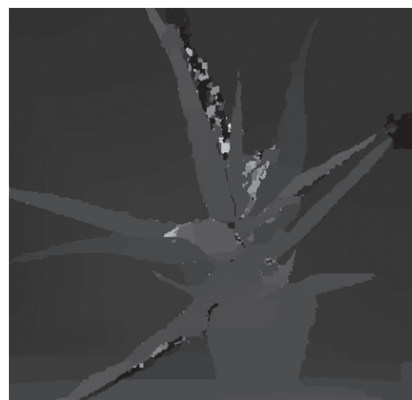


图4 GC 视差图



图 5 真实视差图

表 1 BM,SGM,GC 算法的误匹配率 (%)

算法	aloe	baby	平均
BM	49.20	47.33	48.27
SGM	46.13	41.65	43.89
GC	34.57	32.34	33.46

为比较三种算法的运行速度,计算出三种算法的运行时间,如表 2 所示。

表 2 BM,SGM,GC 算法的运行时间 ms

算法	aloe	baby	平均
BM	50.44	41.84	46.14
SGM	188.69	229.08	208.89
GC	60861.27	62759.30	61810.29

1) 匹配效果分析. 图割法具有平滑项,客观上满足了 Marr 的平滑性约束的要求^[20],对于缺少纹理的图像匹配效果较好;BM 算法以邻域窗的灰度为匹配基元,以相似性测度为依据,求解局部最优解,无法在全局图像范围内优化视差,对于遮挡、弱纹理等区域比较敏感,易造成误匹配,从而导致匹配精度较低,只能得到稀疏的视差图;SGM 算法引用动态规划算法搜索最优路径,采用平滑约束保证视差的连续性,能够得到较为稠密

的视差图. 根据图 2~5 比较三种算法的视差图与真实视差图,根据表 1 比较三种算法的误匹配率,可以得到:GC 算法的匹配效果最精确,SGM 次之,BM 算法的匹配效果精确度最差。

2) 匹配速度分析. GC 算法通过迭代优化寻求全局最优解,过分依赖分割结果,导致匹配速度较慢;BM 算法基于窗口寻求局部最优解,无需平滑项约束,因此匹配速度快;SGM 算法虽运用平滑项约束提高匹配精确度,增加了计算量,使得匹配速度慢与 BM 算法,但匹配过程引用动态规划算法,相比于 GC 算法无需全局迭代,节省了全局迭代所需的时间,因此匹配速度比 GC 算法快. 由实验得出的三种算法的运行时间如表 2,可以得出三种算法的运行速度关系:BM > SGM > GC。

通过上述分析与比较,可以得出三种算法各自的优缺点,如表 3 所示。

表 3 BM/SGM/GC 算法比较

算法	类别	优点	缺点
BM	局部匹配	效率高,计算量少	对遮挡和弱纹理处理效果不好
SGM	半全局	速度快,鲁棒性强,对光照不敏感	对弱纹理处理效果不好
GC	全局匹配	对于遮挡边界、和弱纹理处理效果好	算法复杂度高,导致匹配速度慢

4 结语

本文对全局匹配算法、半全局匹配算法和局部匹配算法进行了分析,并对三种算法进行了实现,得到三种算法的匹配效果及匹配效率. 由实验分析可知,现有立体匹配算法并不能兼顾准确性及实时性,局部匹配算法中具有代表性的算法是块匹配算法,该算法的实时性较高,但匹配的准确性较低,只能应用于对精确度要求较低的实时系统;全局匹配算法中具有代表性的算法是图割法,该算法的准确性较高,但处理速度较慢,不能应用于实时性系统;半全局匹配算法是匹配速度和精确度都处于两者之间,不能充分使用全局像素信息,并不适用于虚拟现实这类对实时性和准确性要求较高的系统. 经过几十年的发展,立体匹配不断改进,给机器视觉、医学诊断等研究带来新的进展,但仍未研究出一种快速、精确又通用的匹配算法,立体匹配仍是立体视觉中最困难的环节,需要更加深入的研究。

参考文献:

- [1] 于乃功,秦永钢,阮晓钢. 立体匹配算法进展[J]. 计算机测量与控制, 2009, 17(5): 817-819.
- [2] BRADSKI G, KAEHLER A. 学习 OpenCV (中文版) [M]. 于仕琪, 刘瑞祯, 译. 北京: 清华大学出版社, 2009.
- [3] 王雅宁, 梁新刚. 基于交互式图像分割的立体匹配方法[J]. 计算机技术与发展, 2016, 26(9): 164-166.
- [4] 汤春明, 蒋昂. 基于跨尺度代价聚合的改进立体匹配算法[J]. 计算机工程, 2016, 42(11): 272-276.
- [5] 时华, 朱虹. 基于自适应匹配窗口及多特征融合的立体匹配[J]. 模式识别与人工智能, 2016, 29(3): 193-203.
- [6] TAMAS L, GORON L C. 3D semantic interpretation for robot perception inside office environments [J]. Engineering Applications of Artificial Intelligence, 2014, 32(6): 76-87.
- [7] 曹之乐, 严中红, 王洪. 双目立体视觉匹配技术综述[J]. 重庆理工大学学报(自然科学版), 2015, 29(2): 70-75.
- [8] 白明, 庄严, 王伟. 双目立体匹配算法的研究与进展[J]. 控制与决策, 2008, 23(7): 721-729.
- [9] BOYKOV Y, VEKSLER O, ZABIH R. Fast approximate energy minimization via Graph Cuts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1222-1239.
- [10] 周秀芝, 文贡坚, 王润生. 自适应窗口快速立体匹配[J]. 计算机学报, 2006, 29(3): 473-479.
- [11] HOSNI A, BLEYER M, GELAUTZ M, et al. Local stereo matching using geodesic support weights [C] // Proceedings of the International Conference on Image Processing, 2009: 2093-2096.
- [12] YOON K J, KWEON I S. Adaptive support-weight approach for correspondences search [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 650-656.
- [13] KANDE T, OKUTOMI M. A stereo matching algorithm with an adaptive window: theory and experiment [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(9): 920-932.
- [14] 肖艳青, 刘党辉, 孙朋. 图像立体匹配研究进展[J]. 测控技术, 2009, 28(8): 1-5.
- [15] TAO H, SAWHNEY H S, KUMAR R. A global matching framework for stereo computation [C] // Proceedings of Eighth IEEE International Conference on Computer Vision, 2002: 532-539.
- [16] 梁芳. 基于图像分割的立体匹配算法研究[D]. 北京: 北京邮电大学, 2014.
- [17] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167-181.
- [18] HIRSCHMULLER H. Stereoprocessing by semi-global matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341.
- [19] INTILLE S S, BOBICK A F. Disparity-space images and large occlusion stereo [M] // Computer Vision - ECCV 94. Berlin: Springer, 1994: 179-186.
- [20] 徐欢欢. 基于能量函数的图像分割方法的研究[D]. 合肥: 中国科学技术大学, 2009.

Research and Implementation of Binocular Stereo Matching Algorithms

HUANG Songmei, BI Yuanwei, XU Xiao

(School of Computer and Control Engineering, Yantai University, Yantai 264005, China)

Abstract: Stereo matching was the core of computer vision. According to the different constraint conditions, the stereo matching algorithms were divided into three categories: global stereo matching algorithm, semi global stereo matching algorithm and local stereo matching algorithm. The principle of binocular stereo matching and the stereo matching algorithms were introduced. The GC algorithm of the global matching algorithms, the BM algorithm of the local matching algorithms and the semi global matching SGM algorithm were studied intensively. Based on the experimental results, the matching accuracy and matching speed of the three algorithms were analyzed, and their advantages and disadvantages were compared and summarized.

Keywords: computer vision; stereo matching; Graph Cut method; block matching; semi-global matching
(责任编辑 李秀芳)