

3D Reconstruction Based on a Fast Area-based Stereo Matching Algorithm

Hongwei Gao^{1,2}, Chengdong Wu³, Bin Li¹

¹⁾ Robotics Laboratory, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016

(E-mail: ghw1978@sia.cn)

²⁾ Graduate School of the Chinese Academy of Sciences, Beijing 100039

³⁾ School of Information Science & Engineering, Northeast University, Shenyang 110004

(E-mail: wuhaoli@online.ln.cn)

Abstract—According to the parallel binocular vision system, a connotative constraint was proved by geometry theory in this paper. That is for any space point projected into two images, along the horizontal direction in image coordinate, there always exists $u_1 > u_2$. Based on this constraint, a fast stereo cross-correlation algorithm and correlative 3D reconstruction process were introduced. Epipolar rectification, 3D sparse and dense reconstruction results show the validity of the algorithm.

Keywords—Parallel binocular vision, $u_1 > u_2$ constraint, Image rectification, Gray-level cross-correlation, 3D reconstruction

基于一种快速区域匹配算法的三维重建方案研究

高宏伟^{1,2} 吴成东³ 李斌¹

¹⁾ 中国科学院沈阳自动化研究所机器人学重点实验室 沈阳 110016

²⁾ 中国科学院研究生院 北京 100039

³⁾ 东北大学信息科学与工程学院 沈阳 110004

摘要 本文针对平行双目视觉系统的成像特点,从摄像机成像几何上证明了对于空间任一点在两个图像坐标系中的水平方向坐标均存在 $u_1 > u_2$ 这一隐含的约束条件。论述了以该约束条件下的灰度互相关匹配算法为核心的三维重建过程,基于真实环境图像的极线校正和三维稀疏与稠密重建的实验结果证明了所提出算法的有效性。

关键词 平行双目视觉, $u_1 > u_2$ 约束, 图像校正, 灰度交叉互相关, 三维重建

1. 引言

传统的立体匹配算法可以分为基于特征^[1]、基于区域^[2]、基于相位的匹配算法三类^[3]。基于特征的算法主要检测景物自身结构的特性,如图像中的边缘点、角点等灰度不连续点,它们一般是空间物体棱边上的点或是不同曲面上的交点。在特征点匹配上以后,可以通过插值进行特征点间的非特征点的匹配,此方法只能得到精确的稀疏匹配点。基于区域的匹配主要是根据图像的区域相关性进行匹配,可得到稠密的匹配点。基于相位的匹配是利用具有局域频率特征的相位信号作为匹配基元进行匹配,其匹配过程具

有由粗到精的多分辨率特性,亦可得到稠密的匹配点。基于区域的稠密匹配算法可分为两种,一种是局部窗匹配法,另一种是全局优化算法。局部窗匹配法通过比较图像间一个窗口之中相邻像素的灰度相似性来确定该窗口的中心点是否匹配。在该算法中,选择适合的窗口大小对得到一个光滑而清晰的视差图是至关重要的。匹配区域在统计意义上的相似性以及算法的计算速度均与相关窗口的大小有关。窗口小计算速度快,但统计相似性不够,容易出现误匹配;窗口大统计相似性增强,不易出现误匹配,但算法速度降低,在应用中需要折中考虑。全局优化算法通过优化某一视差函数来求解匹配问题,视差函数中通常都包含一个匹配值项和一个平滑约束项,寻求这两项之间正确的

国家“863”计划项目(资助号:2002AA422130)

平衡关系是全局优化算法的一个关键问题^[4]。本文首先从摄像机成像几何上证明了在平行双目视觉系统中的图像坐标下存在 $u_1 > u_2$ 这一隐含的约束条件，然后基于该约束条件设计了一种局部窗稠密匹配算法。与非平行双目视觉系统下的匹配算法相比，由于平行双目视觉系统存在这一隐含约束条件，该算法可以减少搜索空间，加快匹配速度。同时讨论了以该匹配算法为核心的三维重建过程。最后基于真实环境图像的极线校正和三维稀疏与稠密重建的实验结果证明了所设计的算法的有效性。

2. 隐含约束条件的证明

在针孔模型下的双目视觉系统的成像原理图如图1所示。摄像机光轴互相平行，两个摄像机坐标系只差 x 轴方向上的一个平移量 b ，也称作基线距离。这样的摄像机配置使得空间中一点在两个相机中成像点的图像坐标在 v 方向相同， u 方向不同， $u_1 - u_2$ 又称为“视差”。相应的对极几何特点是极线与图像扫描线(垂直于 v 轴的直线)平行，如图1所示的 E_{1i} 与 E_{2i} 为两条极线，空间点 $P(x_{1i}, y_{1i}, z_{1i})$ 在两个相机中的成像点为 p_{1i} 和 p_{2i} 。理论上两条极线应该共线，这样使寻找匹配点的空间由二维降至一维。但由于相机实际摆放不准确而使得两者接近共线。

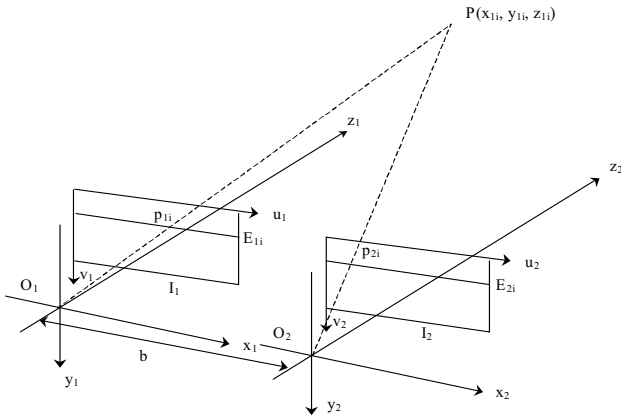


图1 平行双目视觉系统成像原理图

引理：在如图2摆放的平行双目视觉系统中，空间中任一点在两个图像坐标系中均成像且坐标分别为 (u_1, v_1) 和 (u_2, v_2) ，则恒有 $u_1 > u_2$ 成立。

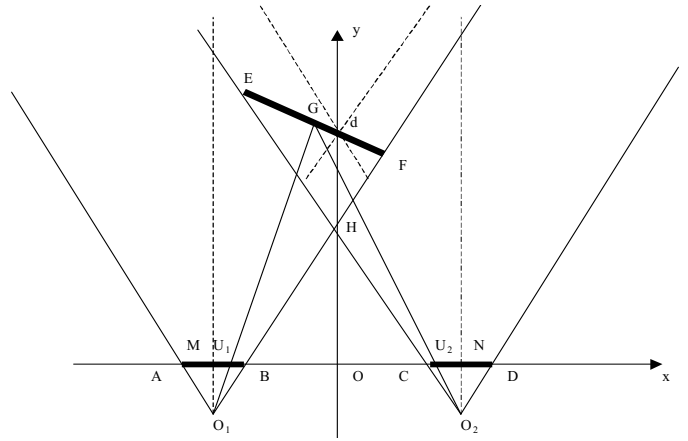


图2 平行双目视觉系统成像原理图

证明：假设两个相机参数一致。如图所示 x 轴上的两条粗线段代表 CCD 成像平面，与 y 轴相交的粗线段代表特征点所在的投影直线，分别与两个相机视场边缘交于 E 和 F ，相机 1、2 的视场角分别为 AO_1B 和 CO_2D ，交点为 H ，在公共视场角 EHF 内的所有特征点点在两个 CCD 上均成像， O_1 and O_2 到 AB 与 CD 的距离为相机焦距 f ，射线 O_1M and O_2M 代表两个相机的光轴方向， AB 与 CD 的长度相等，均为 CCD 横向 (u 方向) 宽度，并采用 $width$ 表示， O_1 and O_2 间距为基线长度并用 b 表示。各关键点的坐标可表示为： $O_1(-b/2, -f)$ ， $O_2(b/2, -f)$ ， $A(-width/2 - b/2, 0)$ ， $B(width/2 - b/2, 0)$ ， $C(-width/2 + b/2, 0)$ ， $D(width/2 + b/2, 0)$ 。根据

相似三角形关系 $\frac{OH}{O_1M} = \frac{OB}{BM}$ 可求出 OH 的长度

$$L = \frac{f(b - width)}{width}$$

H 点的坐标表示为 $(0, L)$ 。

根据两点式可写出直线 O_1H 的方程为：

$$y = L + \frac{2(f+L)}{b}x$$

同理 O_2H 的方程为：

$$y = L - \frac{2(f+L)}{b}x$$

假设特征点在过 y 轴上 $(0, d)$ 且斜率为 k 的直线 EF 上，其中 $d > L$ 以保证特征点在两相机中均可见。那么该直线方程为 $y = kx + d$ 。为了保证在两个相机中都看到观测物体表面的特征点以及满足序的一致

性约束, 要求 k 不能介于 O_1H 与 O_2H 的斜率之间, 即

$$k \notin \left[-\frac{2(f+L)}{b}, \frac{2(f+L)}{b} \right].$$

进一步可求得 EF 与 O_2H 的交点 $E \left(\frac{b(L-d)}{kb+2(f+L)}, \frac{kbl+2d(f+L)}{kb+2(f+L)} \right)$, EF 与 O_1H 的交点

$$F \left(\frac{b(L-d)}{kb-2(f+L)}, \frac{kbl-2d(f+L)}{kb-2(f+L)} \right).$$

在 EF 上任取一点 $G(x_0, kx_0 + d)$, 由两点式写出直线 O_1G 的方程为

$$\frac{x-x_0}{-b/2-x_0} = \frac{y-kx_0-d}{-f-kx_0-d}, \text{ 令 } y=0 \text{ 求得与 } x \text{ 轴交点}$$

$$x_1 = \frac{2fx_0 - b k x_0 - bd}{2(f + kx_0 + d)}.$$

$$\frac{x-x_0}{b/2-x_0} = \frac{y-kx_0-d}{-f-kx_0-d}, \text{ 令 } y=0 \text{ 求得与 } x \text{ 轴交点}$$

$$x_2 = \frac{2fx_0 + b k x_0 + bd}{2(f + kx_0 + d)}.$$

设 A 、 C 两点为相机 1、2 的图像坐标原点, U 为成像点在 CCD 平面上的横向物理距离, 则有:

$$U_1 = x_1 - x_A = \frac{2fx_0 + bf + fwidth + dwidth + kwidthx_0}{2(f + kx_0 + d)} \quad (1)$$

$$U_2 = x_2 - x_C = \frac{2fx_0 - bf + fwidth + dwidth + kwidthx_0}{2(f + kx_0 + d)} \quad (2)$$

$$\text{可推得 } U_1 - U_2 = \frac{bf}{f + kx_0 + d},$$

由于 $d > L$ 且在 EF 上均有 $y > 0$ 成立, 所以 $kx_0 + d > 0$, 所以 $U_1 - U_2 > 0$, $U_1 > U_2$. 不等式两边同时除以 CCD 平面的横向尺度因子 d_u , 即可得到在图像坐标系下有 $u_1 > u_2$ 成立。

也就是说在平行双目视觉系统中, 只要特征点同时在两个相机内成像, 那么在图像坐标系下总有 $u_1 > u_2$ 成立。这一结论为接下来讨论的互相关搜索算法提供了理论基础。为了在实际应用中, 满足上述引理, 需要在固定双摄像机时进行有针对性的调节。基本思想是取一标定板与光轴垂直置于双摄像机前方, 测得图像上同一直线段的 v 坐标看是否相差小于两个像素, 若大于两个像素, 则调节固定螺丝松紧并加垫必要的纸片, 如此反复调节直到 u 、 v 坐标相差均小于两个像素为止, 即有如下约束成立 $u_A - u_B < 2$, $v_A - v_C < 2$, $u_{A'} - u_{B'} < 2$, $v_{A'} - v_{C'} < 2$ 。至此系统已基本贴近理想的平行双目系统, 成像点的坐标也满足上述推导的结论, 如图 3 所示。

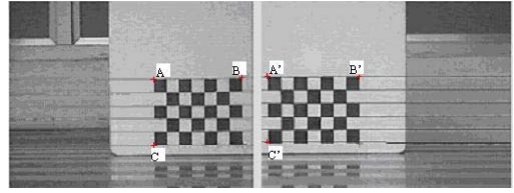


图 3 调整结果

3. 三维重建算法

3.1 摄像机标定

采用基于径向畸变的 TSAI 经典两步方法在高精度的标定台上进行标定, 为了提高标定与测量精度, 利用直接光学方法^[5]确定摄像机的光心坐标, 同时采用精确到子像素级的角点检测算法。这样可以保证标定后的立体视觉系统平均测量误差在 xyz 三个方向上均在 1mm 内。

3.2 匹配

3.2.1 金字塔法降低图像分辨率

对待匹配的场景图像和模板图像中 2×2 邻域内的像素灰度值取平均值, 得到分辨率低一级的图像。依此类推, 可得到一组呈金字塔式的图像, 用 k 表示级数, K 为最高级, $k=0$ 级图像为原图像。从最低分辨率图像开始匹配搜索, 找到粗匹配位置后, 再在第 $K-1$ 级图像上粗匹配位置附近进行搜索, 依此类推, 直到 $k=0$ 的那一级图像(即原图像)为止。由于从第 $K-1$ 级开始仅限于在一个或几个粗匹配点的附近进行搜索, 所以计算量并不大。金字塔法可以提高匹配速度, 但随着分辨率的越低, 匹配的精度也降低。综合考虑速度与精度要求, 本文采用两级金字塔。

3.2.2 图像极线校正

极线校正的目的是使立体像对中的对极线与扫描线重合, 即新图像对中的对极线与 u 轴重合, 两幅图像仅存在 u 轴方向的视差, 这就将图像对的匹配问题从二维降到了一维, 即在寻找匹配点时, 只需沿 u 轴方向搜索, 这样有效的缩小了搜索范围, 提高了匹配算法的效率。根据校正后图像 v 方向坐标相等的原则, 根据标定获得的投影矩阵求出变换矩阵(一般定义为 H 阵), 采用变换矩阵对原图像进行校正, 通过重采样和插值等技术得到整数坐标的图像坐标值及相应的灰度值, 并进行显示^[6]。

3.2.3 交叉互相关计算

读取两幅图像, 以存储左图数据数组的第一个元素为基准, 在算法设定的搜索空间内在右图中通过灰度相关法搜索候选匹配点, 记录分数值最大的点。以刚才记录的右

图中最大分数值的点为基准，反过来在左图中进行相关分数值最大的候选匹配点的搜索并记录，判断该点与左图中的基准点是否相同，相同则认为匹配正确并保存匹配点对。依次类推完成剩余的交叉互相关匹配工作。

$$\begin{cases} u_2 = u_1 + sd(u_1, v_1) \\ v_2 = v_1 \end{cases} \quad (3)$$

其中 $s = \pm 1$ ，以保证视差总为正值。 (u_1, v_1) 和 (u_2, v_2) 分别为空间任一点在左右两幅图中的图像坐标。

一般的视差搜索方法是以左图中某一待匹配点的坐标 (u, v) 为基准，在右图中点 $(u - d, v)$ 和点 $(u + d, v)$ 之间进行搜索候选匹配点。根据我们的定理则只需在点 $(u - d, v)$ 和 (u, v) 之间进行搜索候选匹配点就行了，反过来以右图中某一待匹配点的坐标 (u, v) 为基准，在左图中点 (u, v) 和点 $(u + d, v)$ 之间进行搜索候选匹配点就行了，这样便可以减少一半的搜索空间，匹配运行速度也将提高一倍。由于通过手工调整双摄像机光轴近似于平行，使得成像过程中图像变形较小，从而减少了极线校正算法引入的误差，也间接减少了三维测量误差。折中考虑，在实验中选择 7×7 大小的匹配窗口。

3.2.4 误匹配点的剔除

误匹配点的主要来源有三个方面。一种是一对多情况下引起的误匹配，另一种是匹配错误但相关分数值高而收入引起的误匹配，再有就是由于视差原因不对而收入引起的误匹配。对于稠密匹配，上述交叉互相关技术可以有效的剔除部分误匹配点对，除此之外还可以采用点到极线距离设定阈值、峰值滤波等方法进一步剔除误匹配点对^[7]。每次剔除部分误匹配点对后要重新计算对极几何（即基础矩阵），然后再次重复上述过程，几个循环以后可以将误匹配点对数大幅度降低。对于稀疏匹配，可采用松弛迭代法和最小中值法（LMedS）剔除部分误匹配点对。

3.3 三维重建

对匹配点对的 u 、 v 乘以变换矩阵 H 的逆矩阵，求出校正前的匹配点对的图像坐标。由于标定时用到的是未校正的图像，所以在重建时要采用已确定的匹配点对的校正前的图像坐标。根据三角测量原理进行测量^[8]，转化为世界坐标系下的三维坐标 (X, Y, Z) ，或以视差值的形式表示。已知摄像机成像模型和匹配关系后，三维重建相对比较容易一些。以开放图形库 OpenGL 为工具，根据数据的

相关性进行虚拟物体重建。

4. 实验结果与分析

采用两台美国 IPX-2M30 黑白 CCD 摄像机，分辨率为 1600×1200 。 μ tron12.5 毫米镜头，视场角为 53° 。双摄像机用螺丝固定于方盒内，光轴近似平行放置，在试验中平行双目系统的基线长度为 200 毫米，在高精度的标定台上进行摄像机的标定与测量精度的验证工作，如图 4 所示。系统的硬件配置为奔腾 IV3.0GHz 处理器，软件平台由 VC++6.0 编程语言实现。



图 4 设备与高精度标定台

实验中的两组场景分别取自室内和室外。从表 1 可以看出，稀疏匹配点对经过松弛迭代法后可以去除一部分误匹配点对，稠密匹配点对经过几轮过滤算法后也剔除一些误匹配点对。通常情况下，剔除的点数越多，匹配的正确率越高，但对于稠密匹配而言，剩下的点数也不能太少否则重建出来的图像就会缺少细节。另外，匹配计算耗时的长短与视差有关，视差越大匹配耗时越多。在平行双目视觉系统中，对于同一空间点而言，视差与基线长度有关，基线越长视差越大，计算速度越慢，但三维重建精度越高，在实验中对基线长度的选择也要折中考虑。从图 5 (b) 和图 6 (b) 可以看出，经过极线校正后的左右两幅图像相应的对极线基本共线，误差小于一个像素。图像校正后变形不大，说明本文第二部分论述的手工调节方法起了一定的作用。重建图像中有毛刺存在，说明此处过渡不平滑，曲率突变，个别空间点的重建误差较大，这与误匹配点的存在和标定误差有关。图 5 (e)、6 (e) 所示的三维网格图是未进行纹理粘贴的三维空间点构成的图，在 OpenGL 环境下，可以从不同的视角进行观察。从图 5 (d)、(f) 和图 6 (d)、(f) 的对比中可以看出，稀疏重建的效果要好于稠密重建的效果，这是因为稀疏重建中用到的匹配点数较少，因此含有误差较大的三维点的几率要比稠密重建的小。

表 1 不同图像匹配结果比较

场景名称	视差范围	匹配计算 耗时(秒)	稀疏匹配点对数 松弛前-松弛后	稠密匹配点对数 过滤前-过滤后
办公室	13-60	22	842-752	84209-75211
岩石	10-30	14	723-638	77723-68638



图 5 (a)



原始采集的



图 5 (b)



极线校正以



图 5 (c) 视差图



图 5 (d) 稀疏重建图



图 5 (e) 三维网格图



图 5 (f) 稠密重建图

图 5 室内景物三维重建结果



图 6 (a)



原始采集的



图 6 (b)



极线校正以

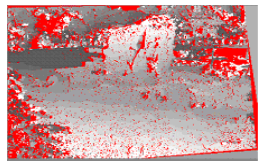


图 6 (c) 视差图



图 6 (d) 稀疏重建图

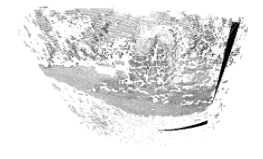


图 6 (e) 三维网格图



图 6 (f) 稠密重建图

图 6 室外景物三维重建结果

5. 结论

本文论述了一种基于隐含约束条件的局部窗稠密匹配算法和以该算法为核心的三维重建过程,给出了基于真实环境图像的极线校正和三维重建的实验结果。由于三维重建的精度与 CCD 量化误差、计算截断误差、摄像机标定精度及特征检测与匹配精度等有关,所以要设计一个高精度的立体视觉系统,必须综合考虑各方面的因素,保证各个环节都具有较高的精度。这些方面的问题有待于进一步的研究,以使立体视觉系统的三维重建精度更高。

参考文献

- [1] W. Grimson and L. Eric, "Computational experiments with a feature based stereo algorithm," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 7, no. 1, pp. 17-34, 1985.
- [2] C.M. Sun, "A fast stereo matching method," in *Digital Image Computing: Techniques and Applications*, pp. 95-100, Auckland, Nov, 1997.
- [3] B. Crespi and A.G. Cozzi, "Analog computation for phase-based disparity estimation: continuous and discrete models," *Machine Vision and Applications*, vol. 11, no. 2, pp. 83-95, 1998.
- [4] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1, pp. 7-42, 2002.
- [5] R.K. Lenz and R.Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 713-720, 1988.
- [6] F. Andrea, T. Emanuele and V. Alessandro, "A compact algorithm for rectification of stereo pairs," *Machine Vision and Applications*, vol. 12, no. 1, pp. 16-22, 2000.
- [7] S.B. Goldberg, M.W. Maimone, L. Matthies, "Stereo Vision and Rover Navigation Software for Planetary exploration," in *Proc. IEEE Aerospace. Conf.*, vol. 5, Montana, March 2002, pp. 2025-2036.
- [8] 马颂德,张正友.计算机视觉—计算理论与算法基础.北京:科学出版社,2003.