

# 立体视觉技术的应用与发展

周 星, 高志军

(上海航天技术研究院第805研究所, 上海 201108)

**摘 要:** 对双目视觉的理论基础、基本原理及其核心技术-立体匹配技术进行了系统的研究与总结, 简要介绍了双目视觉技术在 DARPA2005 挑战赛、美国勇气与机遇号火星车及上海航天局 805 所研制的攻关样机 MR-3 上的应用, 最后根据目前立体视觉的发展现状和存在的问题提出今后的发展方向。

**关 键 词:** 计算机应用; 双目视觉技术; 立体匹配; 视差

**中图分类号:** TP 391

**文献标识码:** A      **文 章 编 号:** 1003-0158(2010)04-0050-06

## Application and Future Development of Stereo Vision Technology

ZHOU Xing, GAO Zhi-jun

(NO805 Institute, Shanghai Academy of Spaceflight Technology, Shanghai 201108, China)

**Abstract:** The theory and principle of binocular vision is studied, as well as its crucial technology-stereo matching. Also, the applications of binocular vision on Stanley and Spirit and Opportunity, the Mars Exploration Rover developed by NASA as well as the research platform-MR-3, developed by NO805 Institute, Shanghai Academy of Spaceflight Technology, are presented. Finally, the application and future development of stereo vision is proposed according to the current level and technical bottleneck of stereo vision.

**Key words:** computer application; binocular vision technology; stereo matching; disparity

计算机立体视觉是被动式测距方法中最重要的距离感知技术, 它直接模拟了人类视觉处理景物的方式, 可以在多种条件下灵活地测量景物的立体信息。其作用是其它计算机视觉方法所不能取代的, 对它的研究无论从视觉生理的角度还是在工程应用中都具有十分重要的意义。计算机立体视觉的开创性工作是从 20 世纪 60 年代中期

开始的, 美国麻省理工学院的 Robert 把 2 维图像分析推广到 3 维景物分析, 标志着计算机立体视觉技术的诞生, 并在随后的 20 年中迅速发展成一门新的学科。特别是 20 世纪 70 年代末, Marr 等创立的视觉计算理论<sup>[1-2]</sup>对立体视觉的发展产生了巨大影响, 现已形成了从图像获取到最终的景物可视表面重建的比较完整的体系<sup>[3]</sup>。

收稿日期: 2008-12-28

基金项目: 上海市科学技术委员会资助项目(06dz2105)

作者简介: 周 星 (1986-), 男, 江苏南通人, 硕士研究生, 主要研究方向为视觉导航技术、路径规划技术等。

本文立足于双目视觉导航技术, 简要阐述其基本原理, 着重论述视觉导航技术的核心——立体匹配技术的研究现状与进展<sup>[4]</sup>, 最后总结分析目前双目视觉导航技术研究存在的问题和今后的发展方向。

## 1 双目视觉导航技术

### 1.1 双目视觉的基本原理<sup>[5]</sup>

人眼在观察物体时存在视差和频差, 这一发现奠定了双目立体视觉的理论基础。双目立体视觉是由不同位置的两台或者一台摄像机(CCD)经过移动或旋转拍摄同一幅场景, 通过各种算法匹配出相应像点, 从而计算出视差, 然后基于三角测量原理恢复深度(距离)信息。80年代美国麻省理工学院人工智能实验室的 Marr 创建了视觉计算理论并应用于双眼匹配, 双目视觉直接模拟人类双眼处理景物的方式, 可靠简便, 在许多领域均极具应用价值。

双目立体视觉技术<sup>[6]</sup>的实现可分为以下步骤: 图像采集、图像处理、立体匹配、三维重建。

典型的双目立体视觉系统如图1所示。

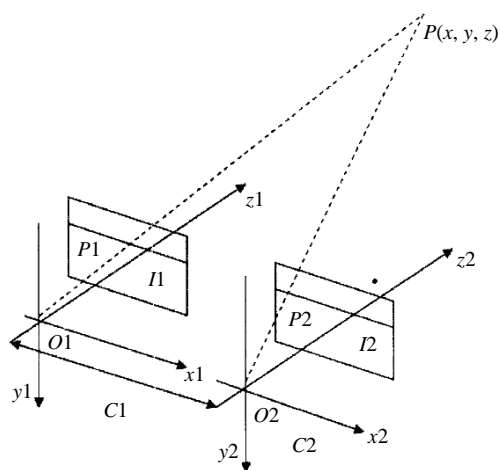


图1 双目立体视觉系统原理示意图

图1中  $C1$ ,  $C2$  是两个光轴平行放置的摄像机, 它们的  $x$  轴重合,  $y$  轴平行,  $I1$ ,  $I2$  分别是它们的成像平面,  $O1$ ,  $O2$  分别是它们的光心。两个摄像机坐标系只相差  $x$  轴方向的平移距离  $b$ , 称为基线长度。任意物点  $P$  在  $I1$ ,  $I2$  上的成像点分别为  $P1$ ,  $P2$ , 通过外极线校正使它们具有相同的  $y$  坐标, 它们的图像坐标之间的差值  $d$ , 就是视差。

当摄像机焦距  $f$  已知, 任意物点  $P$  的深度信息, 即  $P$  点在其中一个摄像机坐标系  $z$  轴上的坐标可由式(1)计算得到

$$z = \frac{b \cdot f}{d} \quad (1)$$

在得到深度信息后, 可以根据摄像机线性成像几何模型及三角测量原理求得物点  $P$  在摄像机坐标系下的  $X_c$ ,  $Y_c$  分量, 如式(2)、式(3)所示。

$$X_c = \frac{b(U - U_0)}{d} \quad (2)$$

$$Y_c = \frac{b(V - V_0)}{d} \quad (3)$$

这样就得到了视差图里各点在摄像机坐标系下的三维坐标。其中, 图像中心坐标  $U_0$ ,  $V_0$  和焦距  $f$  都属于摄像机的内部参数, 可以通过摄像机标定获得。

### 1.2 立体匹配技术<sup>[7]</sup>

立体匹配是立体视觉中最困难的一步。客观景物被投影为二维图像时, 景物的丰富信息(如光照背景, 几何形状, 环境特征, 畸变等)最终以像素的灰度值反映, 立体匹配是典型的视觉计算“病态”问题, 因此对图像无歧义的匹配是相当困难的。立体匹配方法必须解决3个问题:

- (1) 正确选择图像的匹配特征;
- (2) 寻找特征间的本质属性;
- (3) 建立正确的匹配策略。

目前, 立体匹配的研究基本上分为两个方向: ① 从理解人类视觉的立体融合机制出发, 试图建立一种通用的人类双眼视觉计算模型; ② 从实际应用和要求出发, 建立实用的立体视觉系统。

#### 1.2.1 匹配约束准则

辨认两幅图像中的对应位置之所以很难, 其原因在于存在假目标问题。这时就需要某些附加的信息, 以使用某种方法来得到正确的匹配。这里的附加信息就是匹配约束准则。对于一对立体图对, 首先必须考虑的3条约束准则是:

- (1) 相容性 两匹配元素必须由同一物理标记产生;
- (2) 唯一性 一个特征点最多只能和另一幅图上的一个特征点相匹配;
- (3) 连续性 匹配得到的视差值在图像平面上几乎处处平滑。

在实际应用中,还会分析特定的应用环境,加入完备的先验知识,增加约束条件使匹配问题得到简化,如外极线约束就把一个二维空间搜索问题降为一维搜索问题,大大减小了运算量;又如在立体相机安装的时候,由安装基线长度可以估算视差的范围,从而进一步简化一维搜索问题。实践证明,增加必要的约束条件是可行的,而且很多匹配算法的优化也是基于对实际环境约束条件的提取。

### 1.2.2 匹配算法<sup>[8]</sup>

在目前的立体视觉研究中,绝大多数都是针对各自的应用目的和要求建立实用的立体视觉系统。根据匹配基元的不同,立体匹配技术可以分为 3 大类:区域匹配、特征匹配、相位匹配。

#### (1) 区域匹配

区域匹配以基准图的待匹配点为中心创建一个窗口,用邻域像素的灰度值分布来表征该像素,然后在对准图中搜索这么一个像素,以其为中心创建同样的一个窗口,并将其邻域像素的灰度值分布来表征它,两者间的相似性必须满足一定的阈值条件。

#### (2) 特征匹配<sup>[9-10]</sup>

特征匹配是为使匹配过程满足一定抗噪能力且减少歧义性问题而提出。一般地,用于匹配的特征应满足以下特性:唯一性、再现性、具有物理意义。

基于特征的匹配一般含有特征提取和定位、特征描述、特征匹配三个步骤。作为匹配基元的特征分为局部特征和全局特征两大类。面向点、线、面和局部能量等局部特征提取的算法很多。全局特征的定义和提取和具体应用有关,如用于人脸匹配的三角形网格结构就是全局特征的一种。

#### (3) 相位匹配

继区域匹配和特征匹配出现之后,Kuglin 和 Hines<sup>[11]</sup>等提出了第三种立体视觉匹配算法——相位匹配。相位匹配基于这么一个假定,即认为像中对对应点的局部相位是相等的。考虑傅立叶变换的空间支撑为无限,一般地,相位匹配方法对带通滤波信号的相位信息进行处理而得到像对间的视差。最常用的相位匹配方法有相位相关法和相位差一频率法<sup>[12-13]</sup>。

### 1.3 双目视觉导航技术的技术难点

虽然双目视觉导航技术随着计算机科学技术

的发展,逐渐成为主流的导航技术,但就其自身而言存在一些目前还未能很好解决的问题,归纳起来有以下几点:

#### (1) 摄像机对环境的自适应能力

摄像机作为整个导航系统的数据来源,能否获得高质量的图像决定了整个导航系统的成败。

#### (2) 运算量过大

在进行图像预处理及之后的立体匹配时,如果不对图像进行压缩会导致图像数据量过大而影响实时性要求,目前普遍采用的方法是首先对图像进行降采样,但这种做法只是在特定的假设条件下才是可行了,并没有理论依据,只是在算法精度与实时性之间获得折中。

#### (3) 立体匹配精度

所谓立体匹配就是在左右摄像机所拍摄的两幅图中寻找对应点。通过外极线矫正可以使匹配空间由 2 维降至 1 维,大大减少了运算量。但目前的技术瓶颈是如何建立左右匹配对的相似性准则。对于不同的特征,相似性准则的选取决定着匹配精度。

#### (4) 科学视觉模型的建立

目前,双目视觉模型是建立在 Marr 的视觉理论基础之上的,要能够完全达到人眼视觉模型还有很长的路要走。也许有一天,通过建立更为科学的双目视觉模型,以上问题都会迎刃而解。

## 2 双目视觉导航技术的应用

### 2.1 DARPA2005 挑战赛上的应用

DARPA(Defense Advanced Research Projects Agency)由美国国防部主办,旨在开发出适合在未知地形里行驶的自主车。DARPA2005 挑战赛斯坦福大学的自主车 Stanley<sup>[14]</sup>以 6 小时 53 分钟 58 秒自动驾驶 142 英里最终获得冠军。

Stanley 的成功取决于很多方面,如多传感器融合,路径规划策略,速度选择等等。而双目视觉技术也在其中得到应用并发挥着至关重要的作用,Stanley 利用雷达进行障碍物检测,最大检测距离为 22m,由此对它造成的影响是车速最高只能达到 25 英里/每小时,否则难以保证安全行驶。在 2005 挑战赛中,开发小组估计要想成功的完成比赛,速度必须达到 35 英里/每小时,而由于受到雷达的局限,开发小组最终采用彩色相机来探测更大范围内的可行使区域。图 2 比较了

雷达成像地图和视觉成像地图。左图为雷达地图，其障碍物检测距离是 22m，右图是视觉地图，其障碍物检测距离是 70m。

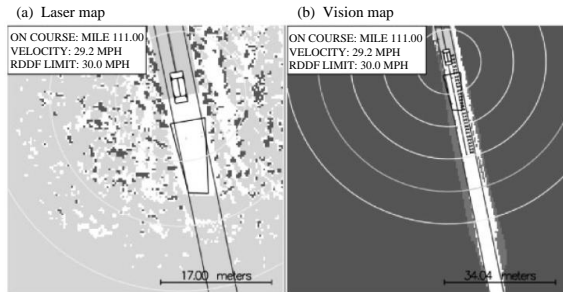


图 2 雷达地图与视觉地图比较

## 2.2 勇气号、机遇号火星车上的应用

双目立体视觉的成功运用典范当属在美国 NASA 勇气号与机遇号火星探测车的应用，火星与地球的距离是  $5.57 \sim 40.13 \times 10^7 \text{ km}$ ，从地面发射电磁波需要将近 20 分钟才能到达火星，这么大

的时延以及火星与地球之间直接通信的局限，很难做到通过遥控操作来控制火星车，这就对火星车的自主能力提出了很高的要求。

路径规划和位置检测需要高精度和高分辨率的地图做基础，而卫星地图不能满足要求，为了解决上述问题，Athena 火星车<sup>[15]</sup>共装备有四对立体相机，相机参数如表 1 所示。

表 1 立体相机的技术参数

	分辨率	视场角	视角	基线长
Navcam	1024×	0.82	450	15cm
Pancam	1024×	0.27	160	25cm

立体相机获取图像，经立体匹配、路径规划找到合适的导航路径，具体流程图及导航结果如图 3 所示<sup>[16]</sup>。

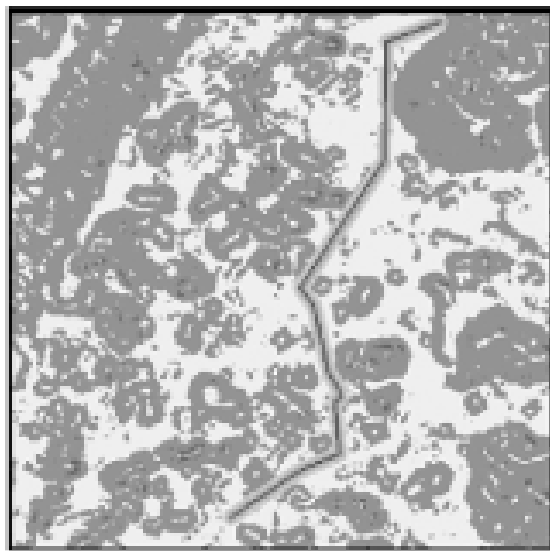
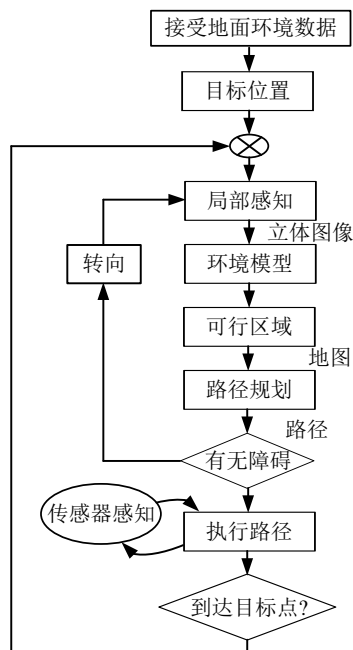


图 3 自主导航流程及结果

## 2.3 在研月球车上的应用

我国的嫦娥工程规划为三期，简称为“绕、落、回”三步走。二期工程“落”即发射月球软着陆器，并携带月球巡视勘察器（俗称月球车），在着陆区附近进行就位探测。

目前，全国高校及各科研单位都对月球车的

研制有浓厚的兴趣并且投入了相当大的精力。上海航天局 805 所自 1995 年成立月球探测小组以来，完成了原理样机、试验样机和攻关样机的研制，在月球车探测技术领域取得了领先的优势。这里简要介绍双目视觉导航技术在攻关样机（简称 MR-3，如图 4 所示）上的应用。

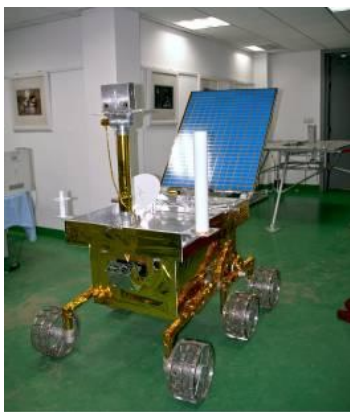


图4 MR-3 攻关样机

### 2.3.1 立体相机

立体视觉系统采用了符合工业标准的高分辨率(1024\*768)1394 接口的数字 CCD 摄像机, 采样率最大为 10 帧, 两个摄像机固定安装在一个紧凑的立方体金属盒中, 如图 5 所示。系统外形尺寸为: 长 157mm, 宽 47.4mm, 高 36mm, 通过 1394 图像采集卡捕捉图像。镜头焦距固定, 水平视角为 70°。系统在应用之前经过精确的标定, 标定信息包括: 图像原点的位置, 基线长度, 最大视差范围和镜头径向失真校正。

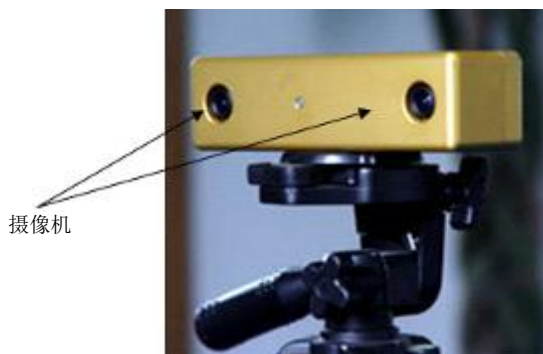


图5 立体视觉系统

经过实测, 相机双目公共区水平视角为 55°, 垂直视角为 40°。

### 2.3.2 实现算法

由于月球地表在局域范围内的纹理色彩并不是很丰富, 采用基于特征提取的方法来进行立体匹配获得的正确率不会很高。所以所采用的算法并不致力于特征点、线或面的提取, 而是直接采用图像对进行特征匹配。同时月球表面由各种各样的矿物质组成, 其表面粗糙度比环境光的波长大得多, 因此在地表面的反射主要由漫散射组

成。当两个相距不是很远的摄像机拍摄同一目标时, 目标特征的灰度值及其邻域中各点的灰度值在两个摄像机中具有一定的相似性。具体的算法流程如图 6 所示。

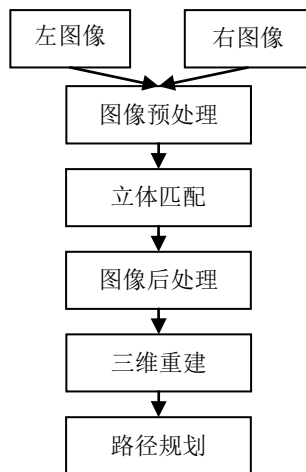


图6 立体视觉算法流程

### 2.3.3 测试结果

用开发的“立体分析.exe”软件可查看左右相机拍摄到的图像, 以及特征点图和深度信息图, 如图 7 所示。

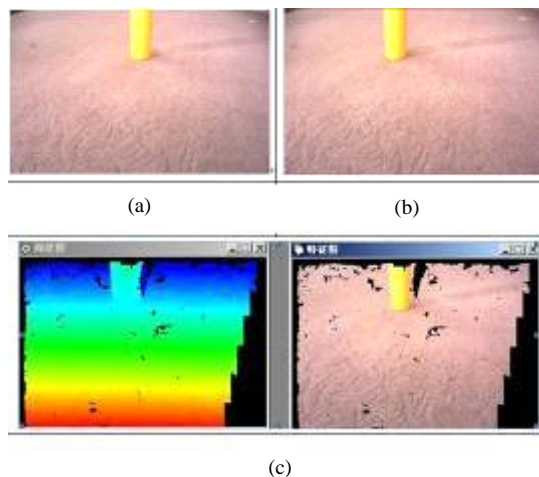


图7 由立体视觉系统获得的左右两幅图像(a), (b)及由他们重建出的深度图和特征图(c)

由图 7 可以发现目前该系统已经能够很可靠的匹配出两幅图像中的特征, 进而恢复出目标的深度信息, 为视觉导航提供了比较可靠的数据源。

### 2.3.4 存在的问题及初步解决方法

在调试中发现在深度图右斜侧和左下脚区域有时会出现不稳定或错误数据, 这主要是因为

这两个区域为左右相机公共视野区的两侧边界, 处理起来会有些问题, 如图 8 所示。目前解决方法是在提取三维数据时, 丢弃了这两块区域的数据, 以消除其影响, 当然这对视野大小有一定影响, 但这影响很小。

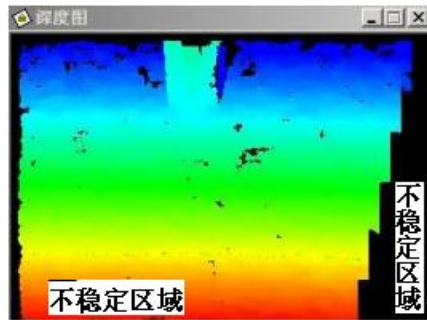


图 8 不稳定区域

目前立体系统对车轮在沙地上留下的齿轮印识别存在的问题, 有时会识别出错误的数数据, 经过查看错误三维数据, 发现大部分数据的  $y$  (高度) 都很大或很小 (一般是 1 米多, 或两米多), 而目前相机安装倾角和高度使得它最高只能看到离地 0.5 米, 最低约深 0.6 米的地方, 因此可在程序中判断  $y$  高度是否在  $(-0.6 \sim 0.5)$  米之间, 可滤除一部分错误数据, 经过测试效果有明显改善。当然还正在努力想从根本上解决误识别的问题。

### 3 双目视觉导航技术的发展方向

综上所述, 要精确模拟人眼立体视觉系统还存在很多困难, 但就其今后的发展方向有如下几个方面:

(1) 深入研究人眼视觉机理, 建立更有效的双目视觉模型, 更充分地反映立体视觉不确定性的本质属性, 为匹配提供更多的约束信息, 降低立体匹配的难度;

(2) 利用对人眼选择性注意机制的研究和主动视觉的方法解决视觉计算病态结构问题, 选择性注意机制将不适定问题转化为适定问题并在已知摄像机的外部参数情况下, 将非线性问题转化为线性问题;

(3) 探索新的适用于立体视觉的计算理论和匹配策略, 选择有效的匹配准则和算法结构, 以解决存在灰度失真、几何畸变、噪声干扰、特

殊结构的匹配问题;

(4) 算法向并行化发展, 提高速度, 减少运算量, 增强系统的实时性。

### 参 考 文 献

- [1] Marr D. Vision [M]. W H Freeman and Company, 1982. 25-40.
- [2] [美] Marr D. 视觉计算理论[M]. 姚国正等译. 北京: 科学出版社, 1998. 65-90.
- [3] 罗 翔, 席文明, 颜景平. 一种双目主动立体视觉系统的目标定位算法[J]. 东南大学学报, 2002, 32(1): 60-63.
- [4] 游素亚, 徐光祐. 立体视觉研究的现状与进展[J]. 中国图形图像学报, 1997, 2(1): 17-21.
- [5] 贾云德. 机器视觉[M]. 北京: 科学出版社, 2001. 243-322.
- [6] 李 奇, 冯华君. 计算机立体视觉技术综述[J]. 光学技术, 1999, (5): 71-73.
- [7] 徐 奕, 周 军, 周源华. 立体视觉匹配技术[J]. 计算机工程与应用, 2003, (1): 1-6.
- [8] 潘 华, 郭 戈. 立体视觉研究的进展[J]. 计算机测量与控制, 2004, 12(12): 1121-1124.
- [9] Loaiza H, Triboulet J, Lelandais S. Matching segments in stereoscopic vision [J]. IEEE Instrumentation & Measurement Magazine, 2001, (3): 37-42.
- [10] Kuijper A, Florack L M J. On the creation of critical points in scale space with application to medical image analysis [R]. UU-CS-2001-25 TR.
- [11] Kuglin C D, Hines D C. The phase correlation image alignment method [C]//Proc. of IEEE Int. Conf. on Cybernetics and Society, 1975: 163-165.
- [12] 徐彦君, 杜利民, 侯自强. 基于相位的尺度自适应立体匹配方法[J]. 电子学报, 1999, 27(7): 38-41.
- [13] 游素亚, 柳 建, 徐光祐. 利用视觉相位鉴别能力求解立体视觉匹配[J]. 电子学报, 1996, 24(10): 23-27.
- [14] Sebastian Thrun, Mike Montemerlo. Stanley: the robot that won the DARPA grand challenge [J]. Journal of Field Robotics, 2006, 23(9): 661-692.
- [15] 顾 征, 苏显渝, 李美菊. “勇气”号火星车的全景照相系统分析[J]. 激光杂志, 2005, 26(4): 48-50.
- [16] Maurette M. Mars rover autonomous navigation [J]. Autonomous Robots, 2003, (14): 199-208.