

双目立体视觉技术的实现及其进展

北京理工大学信息科学技术学院光电工程系(100081) 隋 婧 金伟其

摘 要: 阐述了双目立体视觉技术在国内外应用的最新动态及其优越性。指出双目体视技术的实现分为图像获取、摄像机标定、特征提取、立体匹配和三维重建几个步骤,详细分析了各个步骤的技术特点、存在的问题和解决方案,并对双目体视技术的发展做了展望。

关键词: 双目立体视觉 计算机视觉 立体匹配 摄像机标定 特征提取

双目立体视觉是计算机视觉的一个重要分支,即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,通过计算空间点在两幅图像中的视差,获得该点的三维坐标值。80年代美国麻省理工学院人工智能实验室的 Marr 提出了一种视觉计算理论并应用在双眼匹配上,使两张有视差的平面图产生有深度的立体图形,奠定了双目立体视觉发展的理论基础。相比其他类的体视方法,如透镜板三维成像、投影式三维显示、全息照相术等,双目体视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,如微操作系统的位姿检测与控制、机器人导航与航测、三维测量学及虚拟现实等。

1 双目体视的技术特点

双目体视技术的实现可分为以下步骤:图像获取、摄像机标定、特征提取、图像匹配和三维重建,下面依次介绍各个步骤的实现方法和技术特点。

1.1 图像获取

双目体视的图像获取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,获取立体图像对。其针孔模型如图1。假定摄像机 C_1 与 C_2 的角距和内部参数都相等,两摄像机的光轴互相平行,二

维成像平面 $X_1O_1Y_1$ 和 $X_2O_2Y_2$ 重合, P_1 与 P_2 分别是空间点 P 在 C_1 与 C_2 上的成像点。但一般情况下,针孔模型两个摄像机的内部参数不可能完全相同,摄像机安装时无法看到光轴和成像平面,故实际中难以应用。

上海交大在理论上对会聚式双目体视系统的测量精度与系统结构参数之间的关系作了详尽分析,并通过试验指出,对某一特定点进行三角测量。该点测量误差与两 CCD 光轴夹角是一复杂的函数关系;若两摄像头光轴夹角一定,则被测坐标系与摄像头坐标系之间距离越大,测量得到点距离的误差就越大。在满足测量范围的前提下,应选择两 CCD 之间夹角在 $50^\circ \sim 80^\circ$ 之间。

1.2 摄像机的标定

对双目体视而言,CCD 摄像机、数码相机是利用计算机技术对物理世界进行重建前的基本测量工具,对它们的标定是实现立体视觉基本而又关键的一步。通常先采用单摄像机的标定方法,分别得到两个摄像机的内、外参数;再通过同一世界坐标中的一组定标点来建立两个摄像机之间的位置关系。目前常用的单摄像机标定方法主要有:

(1)摄影测量学的传统设备标定法。利用至少 17 个参数描述摄像机与三维物体空间的约束关系,计算量非常大。

(2)直接线性变换法。涉及的参数少、便于计算。

(3)透视变换矩阵法。从透视变换的角度来建立摄像机的成像模型,无需初始值,可进行实时计算。

(4)相机标定的两步法。首先采用透视矩阵变换的方法求解线性系统的摄像机参数,再以求得的参数为初始值,考虑畸变因素,利用最优化方法求得非线性解,标定精度较高。

(5)双平面标定法。

在双摄像机标定中,需要精确的外部参数。由于结构配置很难准确,两个摄像机的距离和视角受到限制,一般都需要至少 6 个以上(建议取 10 个以上)的已知世界坐标点,才能得到比较满意的参数矩阵,所以实际测量过程不但复杂,而且效果并不一定理想,大大地限制

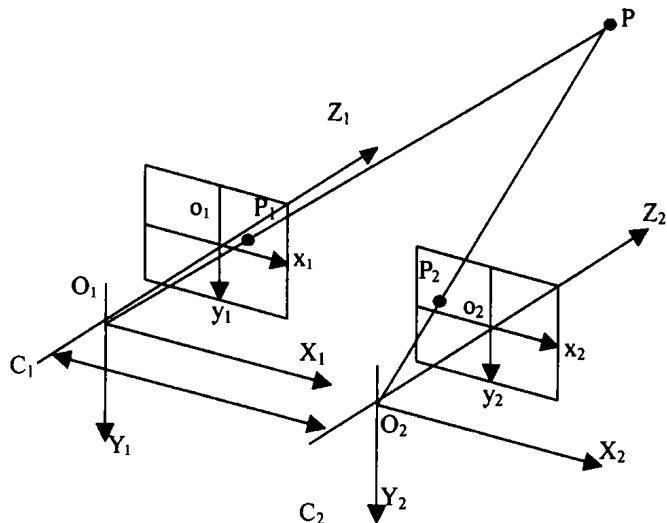


图1 双摄像机针孔模型

计算机应用

了其应用范围。此外双摄像机标定还需考虑镜头的非线性校正、测量范围和精度的问题,目前户外的应用还很少。

上海大学通信与信息工程学院提出了基于神经网络的双目立体视觉摄像机标定方法。首先对摄像机进行线性标定,然后通过网络训练建立起三维空间点位置补偿的多层前馈神经网络模型。此方法对双目立体视觉摄像机的标定具有较好的通用性,但是精确测量控制点的世界坐标和图像坐标是一项严格的工作。因此神经网络中训练样本集的获得非常困难。

1.3 特征点提取

立体像对中需要提取的特征点应满足以下要求:与传感器类型及抽取特征所用技术等相适应;有足够的鲁棒性和一致性。需要说明的是:在进行特征点像的坐标提取前,需对获取的图像进行预处理。因为在图像获取过程中,存在一系列的噪声源,通过此处理可显著改进图像质量,使图像中特征点更加突出。

1.4 立体匹配

立体匹配是双目体视中最关键、困难的一步。与普通的图像配准不同,立体像对之间的差异是由摄像时观察点的不同引起的,而不是由其它如景物本身的变化、运动所引起的。根据匹配基元的不同,立体匹配可分为区域匹配、特征匹配和相位匹配三大类。

区域匹配算法的实质是利用局部窗口之间灰度信息的相关程度,它在变化平缓且细节丰富的地方可以达到较高的精度。但该算法的匹配窗大小难以选择,通常借助于窗口形状技术来改善视差不连续处的匹配;其次是计算量大、速度慢,采取由粗至精分级匹配策略能大大减少搜索空间的大小,与匹配窗大小无关的互相关运算能显著提高运算速度。

特征匹配不直接依赖于灰度,具有较强的抗干扰性,计算量小,速度快。但也同样存在一些不足:特征在图像中的稀疏性决定特征匹配只能得到稀疏的视差场;特征的提取和定位过程直接影响匹配结果的精确度。改善办法是将特征匹配的鲁棒性和区域匹配的致密性充分结合,利用对高频噪声不敏感模型来提取和定位特征。

相位匹配是近二十年才发展起来的一类匹配算法。相位作为匹配基元,本身反映信号的结构信息,对图像的高频噪声有很好的抑制作用,适于并行处理,能获得亚像素级精度的致密视差。但存在相位奇点和相位卷绕的问题,需加入自适应滤波器解决。

1.5 三维重建

在得到空间任一点在两个图像中的对应坐标和两摄像机参数矩阵的条件下,即可进行空间点的重建。通过建立以该点的世界坐标为未知数的4个线性方程,可以用最小二乘法求解得该点的世界坐标。实际重建通常采用外极线约束法。空间点、两摄像机的光心这三点组

成的平面分别与两个成像平面的交线称为该空间点在这两个成像平面中的极线。一旦两摄像机的内外参数确定,就可通过两个成像平面上的极线的约束关系建立对应点之间的关系,并由此联立方程,求得图像点的世界坐标值。对图像的全像素的三维重建目前仅能针对某一具体目标,计算量大且效果不明显。

2 双目体视的最新应用

2.1 国外研究动态

双目体视目前主要应用于四个领域:机器人导航、微操作系统的参数检测、三维测量和虚拟现实。

日本大阪大学自适应机械系统研究院研制了一种自适应双目视觉伺服系统^[1],利用双目体视的原理,以每幅图像中相对静止的三个标志为参考,实时计算目标图像的雅可比矩阵,从而预测出目标下一步运动方向,实现了对运动方式未知的目标的自适应跟踪。该系统仅要求两幅图像中都有静止的参考标志,无需摄像机参数。而传统的视觉跟踪伺服系统需事先知道摄像机的运动、光学等参数和目标的运动方式。

日本奈良科技大学信息科学学院提出了一种基于双目立体视觉的增强现实系统(AR)注册方法^[2],通过动态修正特征点的位置提高注册精度。该系统将单摄像机注册(MR)与立体视觉注册(SR)相结合,利用MR和三个标志点算出特征点在每个图像上的二维坐标和误差,利用SR和图像对计算出特征点的三维位置总误差,反复修正特征点在图像对上的二维坐标,直至三维总误差小于某个阈值。该方法比仅使用MR或SR方法大大提高了AR系统注册深度和精度。实验结果如图2,白板上三角形的三顶点被作为单摄像机标定的特征点,三个三角形上的模型为虚拟场景,乌龟是真实场景,可见基本上难以区分出虚拟场景(恐龙)和现实场景(乌龟)。



图2 基于双目视觉的AR系统虚实场景结合后的图像

日本东京大学将实时双目立体视觉和机器人整体姿态信息集成,开发了仿真机器人动态行走导航系统^[3]。该系统实现分两个步骤:首先,利用平面分割算法分离所拍摄图像对中的地面与障碍物,再结合机器人躯体姿态的信息,将图像从摄像机的二维平面坐标系转换到描述

计算机应用

躯体姿态的世界坐标系,建立机器人周围区域的地图;其次根据实时建立的地图进行障碍物检测,从而确定机器人的行走方向。

日本冈山大学使用立体显微镜、两个 CCD 摄像头、微操作器等研制了使用立体显微镜控制微操作器的视觉反馈系统,用于对细胞进行操作、对种子进行基因注射和微装配等。

麻省理工学院计算机系提出了一种新的用于智能交通工具的传感器融合方法^[4],由雷达系统提供目标深度的大致范围,利用双目立体视觉提供粗略的目标深度信息,结合改进的图像分割算法,能够在高速环境下对视频图像中的目标位置进行分割,而传统的目标分割算法难以在高速实时环境中得到令人满意的结果,系统框图如图 3。

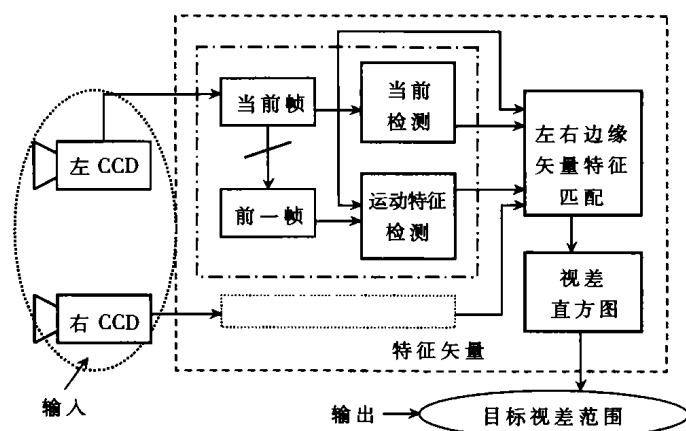


图 3 基于双目帧图像和运动融合的目标深度探测系统

华盛顿大学与微软公司合作为火星卫星“探测器”号研制了宽基线立体视觉系统^[5],使“探测器”号能够在火星上对其即将跨越的几千米内的地形进行精确的定位和导航。系统使用同一个摄像机在“探测器”的不同位置上拍摄图像对,拍摄间距越大,基线越宽,能观测到越远的地貌。系统采用非线性优化得到两次拍摄图像时摄像机的相对准确的位置,利用鲁棒性强的最大似然概率法结合高效的立体搜索进行图像匹配,得到亚像素精度的视差,并根据此视差计算图像对中各点的三维坐标。相比传统的体视系统,能够更精确地绘制“探测器”号周围的地貌和以更高的精度观测到更远的地形。

2.2 国内研究动态

浙江大学机械系完全利用透视成像原理,采用双目体视方法实现了对多自由度机械装置的动态、精确位姿检测,仅需从两幅对应图像中抽取必要的特征点的三维坐标,信息量少,处理速度快,尤其适于动态情况。与手眼系统相比,被测物的运动对摄像机没有影响,且不需知道被测物的运动先验知识和限制条件,有利于提高检测精度。

东南大学电子工程系基于双目立体视觉,提出了一种灰度相关多峰值视差绝对值极小化立体匹配新方法,

可对三维不规则物体(偏转线圈)的三维空间坐标进行非接触精密测量。

哈工大采用异构双目活动视觉系统实现了全自主足球机器人导航^[6]。将一个固定摄像机和一个可以水平旋转的摄像机,分别安装在机器人的顶部和中下部,可以同时监视不同方位视点,体现出比人类视觉优越的一面。通过合理的资源分配及协调机制,使机器人在视野范围、测距精度及处理速度方面达到最佳匹配。双目协调技术可使机器人同时捕捉多个有效目标,观测相同目标时通过数据融合,也可提高测量精度。在实际比赛中其他传感器失效的情况下,仅仅依靠双目协调仍然可以实现全自主足球机器人导航。

火星 863 计划课题“人体三维尺寸的非接触测量”,采用“双视点投影光栅三维测量”原理,由双摄像机获取图像对,通过计算机进行图像数据处理,不仅可以获取服装设计所需的特征尺寸,还可根据需要获取人体图像上任意一点的三维坐标。该系统已通过中国人民解放军总后勤部军需部鉴定。可达到的技术指标为:数据采集时间小于 5s/人;提供身高、胸围、腰围、臀围等围度的测量精度不低于 1.0cm。

3 双目体视的发展方向

就双目立体视觉技术的发展现状而言,要构造出类似于人眼的通用双目立体视觉系统,还有很长的路要走,进一步的研究方向可归纳如下:

(1)如何建立更有效的双目体视模型,能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)探索新的适用于全面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(透视、旋转、缩放等)、噪声干扰、特殊结构(平坦区域、重复相似结构等)及遮掩景物的匹配问题;

(3)算法向并行化发展,提高速度,减少运算量,增强系统的实用性;

(4)强调场景与任务的约束,针对不同的应用目的,建立有目的的和面向任务的体视系统。

双目体视这一有着广阔应用前景的学科,随着光学、电子学以及计算机技术的发展,将不断进步,逐渐实用化,不仅将成为工业检测、生物医学、虚拟现实等领域的关键技术,还有可能应用于航天遥测、军事侦察等领域。目前在国外,双目体视技术已广泛应用于生产、生活中,而我国正处于初始阶段,尚需广大科技工作者共同努力,为其发展做出贡献。

参考文献

- 1 Minoru Asada, Takamaro Tanaka. Visual Tracking Of Unknown Moving Object By Adaptive Binocular Visual Servoing.

(下转第 12 页)

计算机应用

```

THREAD_IN_PDD,
0,
NULL,
(PHW_VTBL) &CTRLVTbl
};
PHWOBJ rgpHWObjects[] = { /* 包含两个 PDD 实例
                             的数组 */
    &BluetoothObj,
    &BluetoothCTRLObj
};
DEVICE_LIST SerDL = { /* 存储设备驱动程序中所有
                        串口设备 */
    "CESerial.dll", /* 串口驱动程序的名字 */
    sizeof(rgpHWObjects)/sizeof(PHWOBJ), /* 串口设备的数目 */
    rgpHWObjects
};

PDEVICE_LIST GetSerialObject(VOID)
{
    return (&SerDL);
}

```

在串口驱动程序中注册了两个串口设备后,就要实现这两个 PDD 实例对应的流接口函数。微软为要实现的串口 PDD 模块提供了一个 HWOBJ(Hardware Object)类型的串行对象表,这个结构列出了实现串口驱动的所有接口函数指针。见下面的描述:

```

typedef struct __HW_VTBL {
    PVOID (*HWInit)(ULONG Identifier, PVOID pMDDCon-
        text, PHWOBJ pHWObj);
    ...,
    ...,
    BOOL (*HWIoctl)(PVOID pHead, DWORD dwCode,
        PBYTE pBufIn, DWORD dwLenIn, PBYTE pBufOut, DWORD
        dwLenOut, PDWORD pdwActualOut);
} HW_VTBL, *PHW_VTBL;

```

(上接第 6 页)

- Proceeding of the 1999 IEEE International Conference on Multisensor Fusion and Intelligent Systems;
- 2 Vallerand Steve, Kanbara Masayuki, Yokoya Naokazu. Binocular Vision-Based Augmented Reality System With An Increased Registration Depth Using Dynamic Correction Of Feature Positions. Proceedings Of the 2003 IEEE, Virtual Reality. March 2003; Vol.22~26, p271~272
 - 3 Kei Okada, Masayuki Inaba. Integration Of Real-Time Binocular Stereo Vision And Whole Body Information For Dynamic Walking Navigation Of Humanoid Robot. IEEE Conference on Multisensor Fusion and Intergration for Intelligent Systems

用户通过修改串行口 PDD 的串行对象表改变函数集或函数名。下面的结构 BluetoothVTbl 定义了蓝牙仿真串口 DDSI 函数的名称。

```

const HW_VTBL BluetoothVTbl = {
    BluetoothInitSerial,
    ...,
    ...,
    BluetoothIoctl
};

```

同样用结构变量 CTRLVTbl 定义了控制端口的 DDSI 函数名列表。由于这个串口设备用作内部实现特殊的功能,下面只列出了需要关心的主要函数名。

```

const HW_VTBL CTRLVTbl = {
    CTRLInitSerial,
    CTRLDeinit,
    CTRLOpen,
    CTRLClose,
    ...,
    ...,
    CTRLIoctl
};

```

蓝牙作为一个驱动新经济的引擎,将会对以嵌入式系统为主的“后 PC 机”时代的到来产生巨大的推动作用,同时它还将面临来自象 IrDA、802.11、HomeRF 等无线通信技术的挑战。而将 OBEX 成功引入蓝牙,使得在嵌入式系统下工作的蓝牙通信设备上传文件、同步、打印等应用成为可能,因此在蓝牙中起着举足轻重的作用。

参考文献

- 1 Draft Foundation. Specification of the Bluetooth System. v1.0' July, 1999
- 2 ATI. Nucleus Plus Internals. 1999
- 3 isi. PSOS+ V2.3 User's Reference Manual. 1999
- 4 金纯,许光辰,孙睿. 蓝牙技术. 北京:电子工业出版社 (收稿日期:2004-06-30)

2003

- 4 Yajun Fang, Ichiro Masaki, and Berthold Horn. Depth-Based Target Segmentation for Intelligent Vehicles: Fusion of Radar and Binocular Stereo. IEEE Transactions On Intelligent Transportation Systems, 2002;3(3)
- 5 Clark F. Olson, Habib Abi-Rached, Ming Ye, Jonathan P. Hendrich. Wide-Baseline Stereo Vision For Mars Rovers. Proceedings Of The 2003 IEEE/RSJ Intl. Conference On Intelligent Robots And Systems. October 2003
- 6 高庆吉,洪炳熔,阮玉峰. 基于异构双目视觉的全自主足球机器人导航. 哈尔滨工业大学学报, 2003;35(9)

(收稿日期:2004-05-30)