**基于双目视觉的三维书写研究综述**

## 摘要

双目立体视觉是机器视觉的重要形式，具有效率高、精度合适、系统简单、成本低等优点，在识别手势时，能瞬间完成图像获取，从而高效的进行测量。本文对基于双目视觉的三维书写的研究进展进行分析。

## 关键词:计算机视觉、双目视觉、三维书写

## 1.研究背景与意义

人机交互技术发展到今天，研究中心已经由计算机转移到了人，进入了自然和谐的人机交互阶段。自然和谐的人机交互技术把机器看成一个类似于人的多功能的感知体，它通过模拟人的各种感觉器官来与人进行各种交互，比如基于视觉的肢体动作、手势、人脸、表情、视线的识别，以及基于听觉的语音识别，还有基于触觉、嗅觉、味觉等等，是一种多媒介、多模式的交互技术。

伴随着硬件设备的进步以及计算机视觉技术的发展，很多基于视觉的人机交互技术从理论变成了现实，目前这些人机交互技术多通过单个摄像头获取的二维平面图像来实现。在三维空间内，利用双目摄像头，通过立体视觉技术来实现更为自然的人机交互，是近来非常热门的研究课题。

## 2 本课题的研究进展

### 2.1人机交互

人机交互是关于设计、评价和实现供人们使用的交互式计算机系统，并围绕相关的主要现象进行研究的学科。简单的说，人机交互技术的主要研究对象是如何进行人与计算机之间的信息交换。随着多媒体、多通道和虚拟现实技术的发展，它正经历着从单通道交互向多通道交互、从二维交互向三维交互的转变。

人机交互也是一门交叉学科，如图1所示。

#### 2.1.1人机交互接口

过去最主要的人机交互接口包括鼠标、键盘、显示器、打印机和音箱等，人们每天通过这些设备与计算机进行交互。近年来随着科学技术的发展，在传统的图形用户接口之外出现了一些新型的可解释交互接口和数字交互接口，比如操纵杆、压力笔、数据手套、手写板、虚拟头盔、Kinect、Google眼镜等。图2显示了三种人机交互接口的交互方式，除了过去传统的图形用户接口之外，另外还有两类可解释交互接口和数字交互接口。

可解释交互接口可以是能够与计算机进行通信的任何物体。用户对交互设备进行位置移动时，计算机会把虚拟物体的逻辑位置也相应的进行移动，它直接对应虚拟物体，而不是像传统的图形用户接口一样对应屏幕上的光标或鼠标。它的优点是可以快速访问，并支持同时操作多个设备。

数字交互接口直接使用手作为物理设备，来对计算机中的逻辑设备进行控制。相比可解释交互接口，它不需要借助额外的物理接口，可以实现更加快速的交互，并且同时支持多个设备。常见的数字交互接口比如手机的电容屏，可以支持多个手指直接在上面滑动和点击操作。

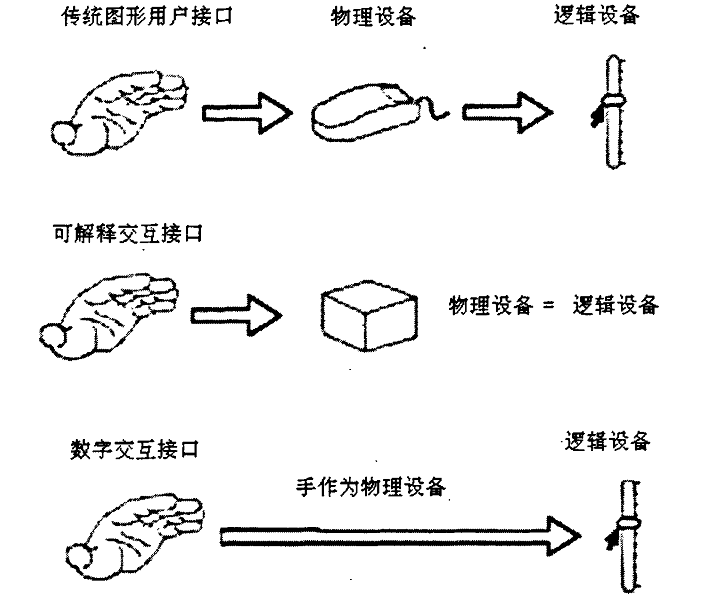


图1 人机交互与其他学科的关系 图2 人机交互的三种方式

#### 2.1.2实时人机交互

实时人机交互[37]是指人和计算机的交互过程中需要保持同步，这个过程应当是紧密耦合的。而实际上真实世界和虚拟世界在交互过程中不可能做到完全紧密的耦合，总会存在一些延时，当延时足够小到可以忽略不计时，我们可以认为这种交互是实时的。比如每秒30帧和35帧的视频，虽然在播放时每帧图像的延时不同，但其差别细微，并且这两种视频的延时人眼都无法观察到，因此可以认为它们是实时播放的。

### 2.2计算机视觉

计算机视觉 [2]是使用计算机及相关设备对生物视觉的一种模拟，其研究目标[38]是使计算机能够通过二维图像来获取物体在三维空间内的几何形状、位置、运动以及姿态等信息，并能将这些信息进行存储、识别、理解和描述。计算机视觉是工程领域和科学领域中的一个富有挑战性重要研究领域；它还是一门包括人工智能、机器学习、信号处理、应用数学、物理学、神经生物学等的综合性交叉学科。

### 2.3基于视觉的人机交互

在人们日常生活与人交际过程中，肢体语言能够用来表达丰富的意图，是除自然语言之外最为常用的一种表达方式。在大部分情况下，它可以不受文化和地域的限制，在很大程度上弥补自然语言表达的不足。

基于视觉技术的人机交互相比其他人机交互技术具有很多优势：随着硬件技术的发展，现在视频采集设备成本低廉且非常普及，很多智能手机和平板等移动终端设备都配备了高像素的摄像头甚至3D双摄像头。基于视觉技术的人机交互，也为这些设备实现更加丰富的功能、更加友好的用户体验提供了技术支持。

基于视觉的人机交互通常有两种实现方法：通过深度传感器（3D Depth Sensors）结合单个摄像头的方式；利用两个或多个摄像头，采用立体视觉技术来实现。微软的Kinect[3]就是前一种方式的代表。

### 2.4三维手写技术

手写技术作为人工智能与模式识别的一个传统的研究领域，吸引了众多研究者对其开展了深入的研究和探索。手写技术根据字符书写的空间不同，可以分为平面手写技术与三维手写技术。自上世纪50年代开始至今，平面手写技术的研究已经非常成熟，识别率和识别速度都完全能够满足实际应用的要求，市场上已经有很多实用的相关产品。近年来随着硬件技术和人机交互技术的不断发展，以及人们对于手写技术用户体验的要求不断提高，推动了三维手写技术的出现。

三维手写技术可以在三维空间随意的进行书写，从而摆脱了传统书写板和书写平面的束缚，逐渐成为手写领域的发展趋势。三维手写信息的输入方式主要有两种：三维加速度传感器输入和计算机图像输入。

### 2.5双目立体视觉

双目立体视觉利用双目或者多目摄像头实现三维重建的方法，是计算机视觉的一个重要研究分支。它的实现是利用两台处于不同位置的摄像机（或者一台摄像机经过移动或旋转）来拍摄同一场景，这样该场景中的同一空间点在两台摄像机中会分别得到的两个二维投影，然后通过计算该点的两个二维投影的视差，最终求得该点的世界坐标系坐标值，从而达到三维重建的目的。

#### 2.5.1发展历史

上世纪60年代中期美国MIT的Roberts把过去的简单二维图像分析推广到了复杂的三维场景[40]，他将立方体、棱柱体和楔形体等简单规则多面体的三维结构从数字图像中提取出来并描述了它们的形状和空间关系，标志着立体视觉技术的诞生。Roberts研究的范围非常广泛，包括提取边缘和角点特征，分析各种几何要素，直到研究图像的明暗以及纹理，摄像机的运动和成像几何等等。上世纪80年代初，美国麻省理工学院人工智能实验室的Marr[4]提出了一种基于双目匹配的视觉计算理论。Marr还创立了视觉计算理论框架。双目视觉直接模拟人类双眼处理景物的方式，可靠简便，极具应用价值，经过几十年来的发展，在很多领域中得到了广泛运用，例如医学成像、航空测绘、工业检测、虚拟现实和增强现实等。

#### 2.5.2 基本原理

双目立体视觉系统由两台摄像机组成，其基本原理如图7所示。

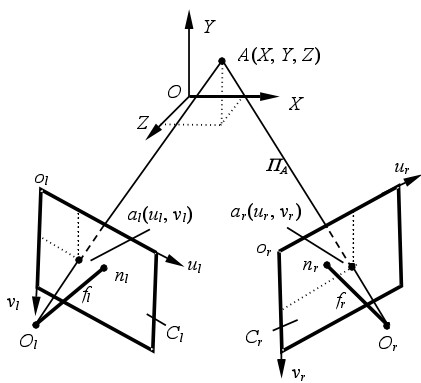


图3 双目立体视觉的基本原理

#### 2.5.3 国内外研究动态

双目立体视觉目前主要应用于机器人导航、三维测量、虚拟现实和增强现实等领域。

在机器人导航领域，双目立体视觉技术已经得到了广泛的应用。

日本东京大学研发的仿真机器人动态行走导航系统，继承了机器人的姿态信息技术和双目立体视觉技术，它首先建立机器人所在区域的地图信息，再根据实时的地图信息来检测障碍物，通过调整机器人姿态，变换它的行走方向，使其绕过障碍物。

双目立体视觉还可以用于三维测量，由于它具备效率高、系统结构简单、精度合适、成本低廉等优点，适合于制造现场的在线、非接触产品检测和质量控制。由

东南大学电子工程系研究了一种立体匹配的新方法，基于双目立体视觉，通过灰度相关多峰值视差绝对值极小化，可以对三维不规则物体进行非接触的三维空间坐标精密测量。

浙江大学开发了一种利用透视成像原理，采用双目立体视觉方法实现的机械系统。该系统只需从双目摄像机拍摄的两幅图像中提取特征点得三维坐标，就可以检测多自由度机械装置的动态和精确位置。

日本冈山大学研制了一套视觉反馈系统。它使用立体显微镜、双目摄像头等设备，来控制微操作器，对种子的细胞进行基因注射等操作。

#### 2.5.4 发展方向

双目立体视觉技术的进一步的研究方向可归纳如下：

1. 建立更为有效的双目立体视觉模型，使得立体视觉不确定性的本质属性得到更充分的反映，从而获得更多的匹配约束信息，使立体匹配的难度得到降低；
2. 解决透视、旋转、缩放等几何畸变，以及灰度失真、噪声干扰、特殊结构和遮掩景物等情况下的匹配问题。需要进一步探索有效的算法结构和匹配策略；
3. 研究并行化算法以提高运算速度，增强系统的实用性；
4. 建立面向任务的双目立体视觉系统，针对不同的应用场景进行设计，增强系统的适用性。

## 3.总结

双目立体视觉技术随着计算机技术的发展以及光学、电子学的不断进步，将不仅在工业检测、生物医学、航天遥测等领域得到应用，在军事侦察和虚拟现实等领域也有着广阔的应用前景。目前双目立体视觉技术在国外已经广泛应用于生产和生活中，但在我国还处于初始阶段，具有极大的发展潜力。

## 参 考 文 献

1. 百度百科: 计算机视觉http://baike.baidu.com/view/155265.htm.
2. http://baike.baidu.com/view/3766855.htm.
3. M Liévin, F Luthon. Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of face features in video[J]. IEEE Transactions on Image Processing, 2004,1 3(1): 63~71.
4. T Gevers, AWM Smeulders. Color-based object recognition[J]. Pattern Recognition, 1999, 32(3): 453~464.
5. 游亚平,李明,袁保宗.可变光照下的人脸检测[J]. 信号处理, 2004, (2): 101~107.
6. 石伟栋. 精确制导中目标识别与跟踪的应用研究[D]. 北京交通大学, 2011.
7. 孟祥旭. 人机交互基础教程[M]. 清华大学出版社, 2010.
8. 隋婧, 金伟其. 双目立体视觉技术的实现及其进展[J]. 电子技术应用, 2004.