

重庆大学

博士学位论文

基于计算机立体视觉的双目立体成像研究

姓名：刘然

申请学位级别：博士

专业：计算机应用技术

指导教师：朱庆生

20070420

摘 要

双目立体成像(Binocular Stereo Imaging)是指根据双目立体视觉(Binocular Stereo Vision)的基本原理,获取、产生和传输一个场景(Scene),并将这个场景展现成具有立体感的景象。基于计算机立体视觉(Computer Stereo Vision)的双目立体成像不仅是虚拟现实、机器视觉的关键技术,而且在多媒体教学、虚拟植物可视化、立体展现、数字娱乐、产品外观设计、雕刻与建筑等领域都有着广泛的应用。

本文研究的目的,就是通过对基于计算机立体视觉的双目立体成像的原理、技术和方法进行研究,为它的应用提供理论和技术支撑。

论文首先对基于计算机立体视觉的双目立体成像相关技术及理论进行了研究,然后在此基础上给出了一个基于监视器的双目立体成像模型,详细讨论了这个模型所具有的成像性质。最后针对该模型结合实验给出了三种立体图像对(Stereo Pair)获取的方法:基于侯原理的立体图像对的生成方法、已建三维模型的情况下立体图像对的生成方法和基于单目视频图像序列的立体图像对的生成方法。

论文的主要研究成果包括:

在成像模型方面,提出了一个基于监视器的双目立体成像模型,详细研究了 this 模型所具有的成像性质,给出了立体图像对坐标的计算公式、视差性质、像点立体成像规律以及直线段立体成像规律,并设计实验进行了验证。

针对遮挡问题,提出了利用空间关系的定性表示与推理来表示物体之间的位置关系以及让计算机理解这种位置关系的方法。在空间关系的定性表示方面,给出了 *SR* 表示模型(*SR Representation Model*),提出了距离关系的定性表示模型和方向关系的定性表示模型。在空间关系的定性推理方面,定义了基于 *SR* 表示模型的组合运算,给出了相应组合运算表(*Composition Operation Table*);构造了定性三角推理所需要的 *QTR* 组合运算表、距离定性减法表和角度定性减法表。与定量的方法相比,采用定性的方法大大降低了时间复杂度和空间复杂度。

论文提出了基于侯原理的立体图像对的生成方法,给出了评价立体效果的定量指标(交叉熵和均方根误差),以及产生符合舒适 Panum 融合区要求的随机数算法 *ProduceKmn*;构造了平移像素法、均匀分布、三角分布、分段均匀分布、拟合灰度值分布、正态分布等六种满足融像条件的分布,结合实验详细讨论了各种分布对立体效果的影响;研究了图像子块个数对立体效果的影响。

论文提出了已建三维模型的情况下立体图像对的生成的方法,详细研究了影响立体效果的目标摄像机与三维模型的位置关系、基线长度、成像位置、成像变形等重要因素,给出了相应的成像规律。

论文研究了视差对深度感知的影响、对象的提取、时序与帧间隔等关键问题，提出了一种单目视频图像序列立体化的方法，并通过实验对该方法进行了验证。

作为一项具有创新性意义的工作，论文在研究方法与思路上力求有所突破，其研究成果将为双目立体成像的具体应用提供具有实用价值的新方法。

关键词：立体视觉，计算机立体视觉，双目立体成像，立体图像对，视差，立体视频

ABSTRACT

Binocular stereo imaging means capturing, producing and transmitting a scene, and then displaying it with stereoscopic perception based on the basic principles of *binocular stereo vision*. Binocular stereo imaging based on *computer stereo vision* is not only the key technique to realize the virtual reality, machine vision, but also widely used in many fields such as multimedia education, stereoscopic display of the virtual plant, digital entertainment, appearance design of industrial products, sculpture and architecture etc.

The purpose of this dissertation is to probe into binocular stereo imaging based on computer stereo vision, which will provide the theories and techniques support for its applications.

In this dissertation, the theories and techniques related to binocular stereo imaging based on computer stereo vision are firstly studied. Based on this foundational work, a binocular stereo imaging model based on monitor is proposed and the imaging features of this model are thoroughly discussed. Finally, with the help of the model and subsequent experiments, three kinds of methods for acquiring *stereo pair* for the model are elaborated: stereo pair creation based on *Hough principle*, stereo pair creation when 3D models are given, and stereo pair creation based on *monocular video frequency image sequence*.

The main research results of this dissertation are summarized as follows.

As for imaging model, the dissertation introduces a binocular stereo imaging model based on monitor. Then it focuses on the imaging features of the model. The formula for computing the coordinates of stereo pairs is given; the properties of parallax are presented; we also give the rules of stereo imaging of point and line segment, and experiments have been carried out to prove the conclusions.

With respect to occlusion, the dissertation argued that, the theory of *qualitative representation and reasoning of spatial relations* can be used for the *representation* of the positional relations among objects in computer and the comprehension of this kind of relations by computer. As for qualitative representation of spatial relations, a *SR representation model* is presented and further provided is the qualitative representation models of distance and directional relations. As for qualitative reasoning of spatial

relations, the composition operation is defined based on *SR* representation model and the corresponding *composition operation table* is also presented. Furthermore, we create qualitative subtraction table of distance, qualitative subtraction table of angle and the QTR composition operation table which is needed for qualitative trigonometry reasoning. Compared with the quantitative method, the qualitative method can considerably reduce the time complexity and space complexity.

The dissertation presents a stereo pair creation method based on *Hou principle*. Two quantitative criterions, *cross-entropy* and *root-mean-square error*, are given to evaluate the stereo effect for this method. Besides, an algorithm named *ProduceKmn*, which is used to produce random numbers that accord with the requirement of *Panum's Fusional Area*, is also presented. Moreover, we constructed six kinds of distributions which accord with the requirement of fusion for the method, such as Moving Pixels without Stretch, Uniform Distribution, Triangular Distribution, Piecewise Uniform Distribution, Fitting Gray Value Distribution, and Normal Distribution. Meanwhile, on the basis of experiments, the impacts that the above-mentioned six distributions may have upon stereo effect have been discussed. In addition, the interrelation between the stereo effect and the number of sub-blocks has been studied.

The dissertation introduces a stereo pair creation method when 3D models are given. Meanwhile, detailed discussion is focused on the important factors that exert great influence on stereo effect, such as positional relations between the target cameras and the 3D Model, length of the baseline, position of the virtual image and imaging distortion, etc. Moreover, corresponding rules of imaging are also presented.

Last but not least, the dissertation also touches on the influence that parallax has upon depth perception, object extraction, time sequence & frame interval, etc. Accordingly, we presented a method to convert a monocular video frequency image sequence to a stereo one. Experiments have been carried out to test it.

As an innovative work, this thesis tried it's best to break through on research methods and research thoughts, and the research results will provide new methods for the practical applications of binocular stereo imaging.

Keywords: Stereo Vision, Computer Stereo Vision, Binocular Stereo Imaging, Stereo Pair, Parallax, Stereo Video

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得重庆大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：刘然

签字日期：2007年6月8日

学位论文版权使用授权书

本学位论文作者完全了解重庆大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权重庆大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

保密（☐），在 年解密后适用本授权书。

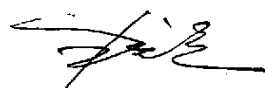
本学位论文属于

不保密（☒）。

（请只在上述一个括号内打“√”）

学位论文作者签名：刘然

导师签名：



签字日期：2007年6月8日

签字日期：2007年6月8日

1 绪 论

1.1 选题背景

“当猕猴桃树出现在电脑屏幕上时,吸引了全场的观众:芽发出来了,开始长大了,叶片展开了,花儿盛开了,果实开始成长了——一整年的生命周期被缩短至不到1分钟的时间。这一镜头实在令人难忘。”这是十多年前Tim Thwaites为我们描述的**虚拟植物可视化技术**^[1]。经过十多年的发展,伴随着虚拟现实技术的迅猛发展,虚拟植物可视化技术取得了超乎人们想象的进步^[2]。到目前为止,已有不少模拟植物生长的软件系统^[2-10]。这些软件系统能在计算机监视器上直观地展现植物的生长过程。然而,这种展现多数是平面的,并不具有立体逼真性^[5-9]。要做到立体的展现,需要运用**计算机立体视觉技术**。

计算机视觉是指用计算机实现人的视觉功能——对客观世界的三维场景的感知、识别和理解^[11]。**计算机立体视觉**是计算机视觉的一个重要分支,它直接模拟了人类视觉处理景物的方式^[12]。由于客观世界在空间上是三维的,所以计算机视觉的研究和应用从根本上来说是三维的。因此立体视觉已成为计算机视觉研究的重点和热点之一,在虚拟现实、多媒体教学、机器人视觉、医学整形和美容、工业产品的外观设计、三维轮廓测量、艺术雕塑、建筑等领域有着广泛的应用^[6, 12-14]。

计算机立体视觉技术有多种^[13, 14],目前研究较多的是**双目立体视觉(Binocular Stereo Vision)**技术。人们通常用双眼同时观看物体,由于两只眼睛视轴的间距(瞳距)的存在,左眼和右眼在观看一定距离的物体时,所接收到的视觉图像是不同的。大脑通过眼球的运动、调整,综合了这两幅图像的信息,感知到生理深度暗示,从而产生立体感。这便是双目立体视觉的基本原理^[15, 16]。利用双目立体视觉技术,人们可以实现**双目立体成像(Binocular Stereo Imaging)**^[17]。简单地说,双目立体成像就是根据双目立体视觉的基本原理,获取、产生和传输一个空间的**场景(Scene)**,并将这个场景展现成具有**立体感(Stereoscopic Perception)**的景象^[18]。基于计算机立体视觉的双目立体成像已在虚拟植物的可视化展现、机器视觉、多媒体教学、数字娱乐、产品外观设计、雕刻与建筑、检验医学^[19-25]等领域得到了应用。

基于计算机立体视觉的双目立体成像是虚拟植物可视化展现的关键技术。从2004年开始,在朱庆生教授的指导下,课题组对该课题开始了系统地研究,几年中已取得了初步的成果。这些研究背景为论文的研究工作打下了良好的基础。

利用计算机立体视觉技术进行**立体成像(Stereo Imaging)**最早是由 Waters 等人提出来的^[15, 18]。随着计算机技术的不断发展,双目立体成像逐渐成为最具应用前景的立体成像技术之一,也是人们研究最多的热点之一^[26-31]。当前双目立体成像

的研究主要集中在以下几个方向:

① 生理深度暗示的产生

基于计算机立体视觉的双目立体成像需要在显示设备上产生深度暗示,特别是生理深度暗示。一般认为,人类可以有多种方法感知对象之间的相对距离,这些有关距离的指示称为深度暗示(Depth Cue)^[32]。通常,这些深度暗示可以分为两类:心理深度暗示(Psychological Depth Cue)和生理深度暗示(Physiological Depth Cue)^[15, 33]。心理深度暗示主要由视觉经验和视觉记忆构成。生理深度暗示则是基于双目立体视觉的基本原理产生的,包括人眼晶状体调节、双眼会聚和双目视差等因素。生理深度暗示比其他深度暗示强,当与其他深度暗示发生冲突时,生理深度暗示通常要占优势, Bela Julesz 演示的随机点立体图像对就是很好的证明^[34]。目前,根据左眼视图(Left Image)和右眼视图(Right Image)是否同时显示在显示设备上的不同,将双目立体成像系统分为时间并行双目立体成像系统(Time-paralleling Binocular Stereo Imaging System)和分时双目立体成像系统(Time-sharing Binocular Stereo Imaging System)^[15, 18]。当前实际应用较多的是分时双目立体成像系统^[16, 35-39]。较早的分时双目立体成像系统有 SGI 公司的 Tessma, StereoGraphics 公司的 Lipton^[15, 18]。当前, NVIDIA 公司开发的基于 NVIDIA 系列显卡的 3D 立体驱动在 PC 机上有较广泛的应用^[38]。国内一些科研单位和高等院校也在分时双目立体成像系统的硬件方面进行了研究^[16, 36, 39]。代表未来发展方向的自由立体显示器是当前研究的一个热点。利用自由立体显示器,无需佩戴任何眼镜就能产生立体效果。目前在国外已经有一些研究机构研制出自由立体显示器的原型机,例如 Dimension Technologies 公司生产的 DTI 1100C“虚拟窗口”^[40], 德国 Dresden3D GmbH 公司生产的 Dresden3D D4D 显示器等。

② 立体效果(Stereo Effect)的评价

立体成像是—门艺术,然而较少有文献对人们通过立体成像系统所能看到的立体效果进行研究和评价^[37]。当前已有公司和科研机构意识到了这个问题的重要性。NVIDIA 公司针对大多数电脑游戏,根据实际经验定制了相应的配置,力图使游戏的显示达到较为满意的立体效果^[38]。然而不同的用户对立体效果的感受也不同,定制的配置不一定适用于每一个用户。国内的豪杰计算机技术有限公司在其产品《豪杰立体影院》中提供了立体参数(Stereo Parameter)设置功能,供用户在实际观察时调节立体效果。然而繁杂的参数调节往往使用户感到力不从心。Frank L.Kooi 和 Alexander Toet 在大量实验的基础上对双目立体成像的视觉舒适(Visual Comfort)问题作了较为详尽的讨论^[41],他们的研究有助于对立体效果进行评价。国内国防科技大学等单位的学者也从不同侧面对立体效果进行了一定的探讨^[42, 43]。但总体来看目前对立体效果的研究还需要进一步深入。

③ 立体图像对(Stereo Pair)的获取

左眼视图和右眼视图一起构成立体图像对。要实现双目立体成像,首先要获得立体图像对。有两类方式获得立体图像对:

第一类方式是使用立体相机拍摄场景得到立体图像,这类方式获得的立体图像效果好。美国 Cornell 大学的 Eaglesham, B.S. 等人采用 CLSM (Confocal Laser Scanning Microscope) 获取立体图像对进行立体成像,以方便细菌观察^[44]。美国 Minnesota 大学的 Erlandsen S.L. 等人利用免疫胶体金技术和立体成像技术对 P 选择因子 (CD62P), GPI-IX (CD42A/CD42B), GPIIb/IIIa (CD41/CD61) 等人血小板的单体细胞辅助因子进行高分辨率 CryoFESEM 检测^[45]。然而,这类方式需要至少两台以上经过光学特性、机械特性和电子特性严格校准的摄像机,并且这两台摄像机的聚焦系统、变焦系统、几何失真、增益控制、光圈控制、会聚控制和视差控制等都要求非常精确一致。这样就增加了拍摄立体图像的难度^[15]。

第二类方式是通过计算机绘制得到立体图像对。这类方式引起了国内外学者的广泛兴趣。

在已建好场景的三维模型的情况下生成立体图像对要相对容易一些。奥地利 Habbware 开发了一个组件用于已建模的情况下生成立体图像或立体视频^[46],但是用户需要自己调整参数,并且用户也不清楚什么样的参数才能使立体效果最好,有些情况下还会有图像的变形。国内的北方工业大学 CAD 研究中心在“863”计划的资助下在这一领域取得了一定的成果^[47]。

在没有场景的三维模型的情况下,生成立体图像对要困难一些。人们尝试着如何从平面图像获得立体图像。1969 年, L. Weiner 首先提出了倾斜投影屏立体成像法,后来 Shanks 对 Weiner 的方法进行了改良,他用一个向后倾斜的马鞍型曲面代替了 Weiner 的倾斜投影屏。天津大学的侯春萍等人在平面图像的立体化方面做了开创性工作。他们提出了一种平面图像立体化方法。该方法利用了心理深度暗示和生理深度暗示的原理,将一幅图像划分成 $M \times N$ 个图像子块,然后使每一个子块在水平方向上随机的偏离它们原来所处的位置,这样得到的图像和原图像一起构成立体图像对。近年来,基于图像的绘制 (Image-Based Rendering, IBR) 技术得到迅速发展,这种技术不仅应用于真实感场景的绘制、估计几何参数^[48]和摄像机参数^[49]、产生立体视觉^[50]、稳定视频^[51]、合成高分辨率的立体图像序列^[23],也可被用来生成立体图像对。1993 年, Chen 等提出了图像内插的方法^[52],第一次提出了基于图像绘制的思想,并于 1995 年将 IBR 思想用于“Quicktime VR”^[53]。从那以后,在图像图形领域,IBR 成为十分重要的研究方向。这种全新的产生真实感图形的方法与传统技术的主要区别在于它是直接从图像得到图像,不需要经过复杂的建模过程^[54]。典型的基于图像的绘制方法有:

1) 先从图像序列中恢复和建立场景的三维模型, 再用经典的计算机图形学工具绘制出新的图像^[34, 55-57]。

加州大学伯克利分校开发的 **Façade** 系统利用了这种绘制方法。它能够从数码相片中恢复和建立建筑物场景的三维模型, 绘制新的图像。它利用许多建筑物具有相对简单的整体几何学, 简化了场景结构和摄像机运动估计。它还使用了简单而强大的基于模型立体视觉思想给粗略的建筑物外形增加细节^[34]。

2) 视图插值法^[54]

视图插值法根据从给定相互之间密集光流的两幅输入图像来重建任意视点的图像。它能在相邻采样点图像之间建立光滑自然的过渡, 从而真实展现了各相邻采样点间场景透视变换的变化。由于相邻两幅场景画面之间存在很大的连贯性, 视图插值法基于相邻画面的摄像机参数及画面上各像素代表的可见点处的深度值来自动决定它们之间的像素对应。在预处理阶段预先计算存储好这种对应关系, 绘制时只需在对应像素间进行插值即可获得中间过渡画面。由于计算机生成的深度图像实际上记录了每一像素可见点的三维信息, 若摄像机的有关参数已知, 则像素对应关系可基于可见点一致性来建立。对实际拍摄的图像来说, 有两种方法可用来获得有关信息: 采用深度摄像机, 并记录其摄像机的有关参数; 或者采用计算机视觉理论来恢复深度信息及其摄像机参数。

3) 纹理映射

在 ACMSIGGRAPH'2000 上, Oliveira 等人将三维变换的图像绘制技术与传统的二维纹理映射相结合, 提出了一种称为“浮雕纹理映射”(relief texture mapping) 的新视点图像生成算法, 实现了用传统纹理映射方法难以实现的三维凹凸细节效果。该算法首先将三维物体正投影到它的包围盒表面上, 以得到带有深度的浮雕纹理 (relieftexture, 简称 RT); 然后通过一个可分离为两次一维变换的二维“预变换”(pre-warp) 过程, 将 RT 图像变换为一个分辨率大小固定且与新视点相关的中间纹理图像; 最后用传统的纹理映射技术将此中间纹理图像变换成新视点的图像。中国科学院的郑新等人则用“层次纹理映射”替代了“浮雕纹理映射”, 也生成了效果较好的新视点图像^[58]。

4) 视图变形^[54]

Seitz S M 等人较早阐述了视图变形的思想^[59]。如果有两幅输入图像, 则通过视图变形人们能将位于摄像机光心连线上的任何视点重构出来。如果摄像机的运动方向垂直于视图方向, 那么中间视图恰好是两幅输入图像的线性组合, 对应的扫描线也是平行的。因此, 在后处理阶段, 视图变形能用于反校正中间图像。Seitz S M 等人利用视图变形技术, 根据《蒙娜丽莎》(左眼视图) 和它的反向映射图(右眼视图)生成了一幅蒙娜丽莎正面照片。

上海交通大学的郭英凯等人在只有一幅原始图像的情况下, 从这幅图像抽取有用信息, 再改变成像时的几何条件生成新图像^[60]。

除了基于图像绘制的方法, 我们还可以从一段平滑的视频序列中提取立体图像对, 论文第六章将详细论述我们在这方面的研究工作。

当前, 立体图像对的获取正成为双目立体成像的一个研究热点。其中第二类方式, 即通过计算机绘制得到立体图像对的方式在虚拟植物可视化展现、计算机游戏、电视广告、电影特技等领域有着广阔的应用前景。

1.2 研究意义与目的

计算机立体视觉能够运用两个或多个摄像机对同一景物从不同位置成像并进而从视差中恢复深度信息^[61]。它直接模拟了人类视觉处理景物的方式, 在虚拟植物可视化展现、多媒体教学、机器人视觉、医学整形和美容、工业产品的外观设计、三维轮廓测量、艺术雕塑、建筑等领域有着广泛的应用。计算机立体视觉的开创性工作始于 20 世纪 60 年代中期。20 世纪 70 年代末, Marr 等创立了视觉计算理论, 对立体视觉的发展产生了巨大的影响^[12, 61]。目前, 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^[61]。从整体上来看, 国内对计算机立体视觉的学术研究和实际运用相对落后。

双目立体视觉技术是人们研究较多和最具实用价值的计算机立体视觉技术, 它可用于实现双目立体成像。双目立体成像涵盖生理深度暗示的产生、立体效果的评价、立体图像对的获取等许多方面, 涉及的具体技术和方法则包括了计算机图形学、摄像机的几何模型的建立、摄像机的几何标定、模式识别、基于图像的绘制等。因此, 开展双目立体成像的研究对提高我国在这一领域的研究水平具有重要的学术意义。

在实际运用方面, 双目立体成像在虚拟植物的可视化展现、机器视觉、多媒体教学、数字娱乐、产品外观设计、雕刻与建筑等领域都有着广阔的应用前景。

由上可知, 研究、分析、探讨基于计算机立体视觉的双目立体成像, 对于提高国内在计算机立体视觉领域的学术水平, 促进双目立体成像在虚拟现实、机器视觉、多媒体教学、数字娱乐、产品外观设计、雕刻与建筑等领域的应用, 都有着重要的学术和实用意义。

基于计算机立体视觉的双目立体成像是虚拟植物可视化展现的关键技术, 对于虚拟植物可视化展现具有重要价值。2003 年 8 月, 国家“863”计划开始实施“数字农业技术与示范”重大专项课题, 将虚拟植物可视化技术研究作为重要研究内容之一。2005 年重庆大学计算机学院申报成功教育部高等学校博士学科点专项科研基金项目: 虚拟作物生长可视化关键技术研究(项目编号: 20050611027);

2006 年申报成功国家高技术研究发展计划(“863”计划): 虚拟作物可视化引擎关键技术研究(项目编号: 2006AA10Z233); 2007 年申报成功重庆大学研究生科技创新基金(项目编号: 200701Y1A0080194)。本学位论文的研究工作是这些项目的研究工作的重要组成部分。其研究目的, 就是希望在对基于计算机立体视觉的双目立体成像技术进行全面、系统研究的基础上, 重点探讨实现立体成像的一些关键问题, 并在此基础上给出我们课题组的一些研究结论, 据此开发出组件, 最终将双目立体成像技术应用于虚拟植物可视化立体展现中。

1.3 研究内容

基于计算机立体视觉的双目立体成像涉及计算机视觉、模式识别、人工智能、认知心理学、计算机图形学等领域中许多具有挑战性的研究问题。已有一些文献对相关问题进行了探讨。主要存在的问题是所成的立体图像仍不够逼真和自然, 人们对人眼的功能以及双目立体成像模型的了解还不够彻底, 立体图像对的获取还有一些难题没有得到较好的解决。本文在对双目立体成像相关技术及理论进行了研究的基础上, 给出了一个基于监视器的双目立体成像模型, 详细讨论了这个双目立体成像模型所具有的成像性质, 然后针对该模型提出了三种立体图像对获取的方法, 这些工作把基于计算机立体视觉的双目立体成像的研究推向深入, 也为双目立体成像在虚拟植物可视化立体展现中的应用提供了理论和技术上的支撑。

根据上述研究内容, 本论文做如下章节安排:

第一章提出本文的选题背景、研究意义与目的, 以及研究内容。

第二章介绍了双目立体成像相关技术及理论。首先论述了双目立体成像相关技术, 着重介绍了双目立体成像系统, 包括双目立体成像系统的发展现状、分时双目立体成像系统以及立体图像文件类型, 等等。其次, 论述了双目立体成像相关理论, 包括会聚角变化值、Panum 融合区、生理深度暗示和心理深度暗示以及遮挡问题, 重点研究了遮挡问题。

第三章提出并研究了基于监视器的双目立体成像模型。首先给出了一个基于监视器的双目立体成像模型; 然后研究了该模型的成像性质, 包括立体图像对坐标的计算、视差及其性质、像点立体成像分析以及直线段立体成像的定量分析; 最后设计了实验对相关的结论进行了验证。

第四章提出了基于侯原理的立体图像对的生成方法, 给出了评价转换后的立体效果的定量指标(以交叉熵和均方根误差两个定量指标对转换效果进行评价), 以及产生一组符合融像要求的随机数的算法 **ProduceKmn**, 构造了平移像素法、均匀分布、三角分布、分段均匀分布、拟合灰度值分布、正态分布等六种分布, 通

过实验详细研究了各种分布对立体效果的影响。最后,研究了图像子块个数对立体效果的影响。

第五章提出了一种已建三维模型的情况下立体图像对生成方法。首先介绍了坐标变换,给出了相应的坐标变换公式。其次介绍了 3DS MAX 中如何利用摄像机对象(Camera Object)生成立体图像对。最后研究了影响立体效果的几个重要因素,包括目标摄像机与三维模型的位置关系、基线长度、成像位置的控制、成像变形分析等内容。

第六章提出了一种基于单目视频图像序列的立体图像对的生成方法,在研究视差对深度感知的影响、对象的提取、时序与帧间隔等关键问题的基础上,提出了一种单目视频图像序列立体化的方法,并通过实验对该方法进行了讨论。

第七章为总结和展望,对全文的研究工作进行了总结,归纳了论文中的创新点,并对下一步的研究工作进行了展望。

论文涉及的基本概念之间的关系如图 1.1 所示。

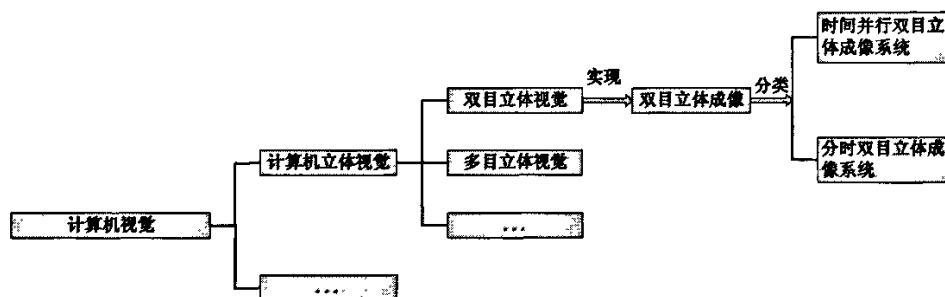


图 1.1 本论文涉及的基本概念之间的关系

Fig1.1 Basic relations between the concepts in this dissertation

论文的逻辑结构如图 1.2 所示。

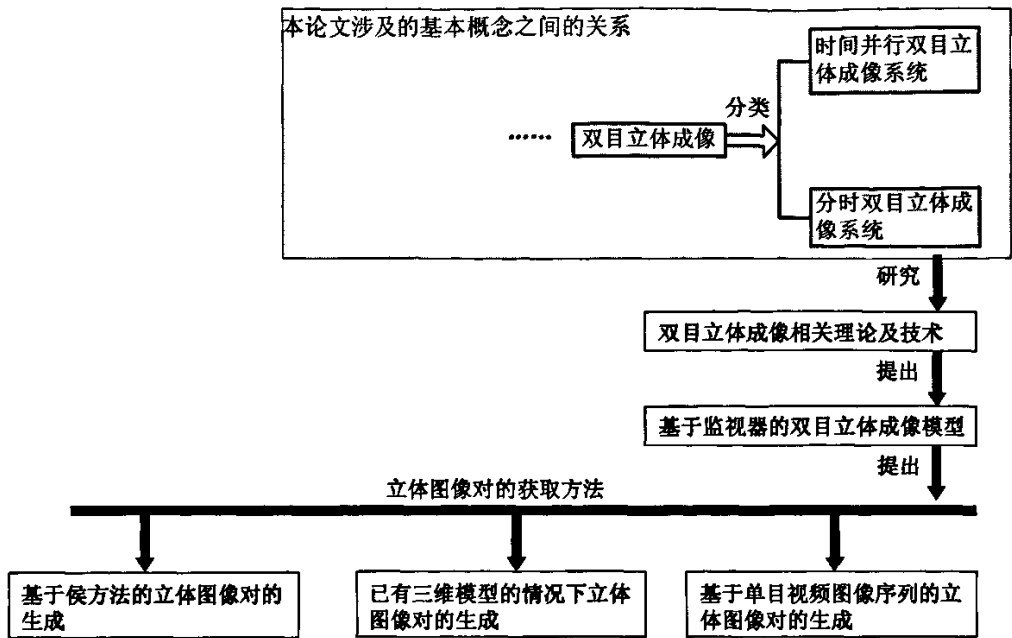


图 1.2 本论文的逻辑结构

Fig1.2 Logical structure of this dissertation

2 双目立体成像相关技术及理论

2.1 双目立体成像系统

2.1.1 发展现状

如前所述,由于人眼两只眼睛存在瞳距,我们在观看一定距离的物体时,左眼和右眼所接收到的视觉图像是不同的。大脑能够很巧妙地将两眼细微的差别融合,产生有空间感的立体景物。如果我们能够为左右眼分别提供同一场景的不同照片(立体图像对),就能够利用双目立体视觉的基本原理看到立体图像。双目立体成像系统就是采用一定的策略让人的左眼只能看到左眼视图,右眼只能看到右眼视图。

双目立体成像系统主要由显示芯片、驱动程序和观察显示设备三部分组成。其中观察显示设备有多种,不同的观察显示设备采用的立体图像对显示方式也不同。目前常用的观察显示设备有双色眼镜^[33, 62]、液晶光阀眼镜(LCD shutter glasses)^[16, 37, 38, 63]、头盔显示器、自由立体(Autostereo)显示器^[64-67]等。

① 双色眼镜

客观世界中人们观察到的各种颜色可由红、绿、蓝三种基本颜色(加色基色)按不同比例混合而成,计算机正是通过这种方法来产生不同的色彩,这就是所谓的加色法;同样,一种颜色的光通过另一种颜色的镜片的过滤,其中一些波长的光会被吸收,变成第三种颜色的光,这就是减色法^[68]。减色法的基本颜色是青色、洋红色及黄色。它们被称为减色基色,分别用于吸收过滤红色、绿色及蓝色的光。根据减色法的原理,可采用如下方法使左眼只能看到左眼视图、右眼只能看到右眼视图:将屏幕底色设计成黑色,左眼视图和右眼视图分别采用红、蓝两种颜色显示;观察者戴上双色眼镜,左眼用黄色镜片,右眼用青色镜片。这样,由于镜片的过滤作用,左眼只能看到红色的视图,右眼只能看到蓝色的视图。左右双眼看到的图像通过大脑的综合,就产生了深度感知。

使用双色眼镜来观看立体图像是一种经济快速的方法。该方法成本低,不需要在计算机上安装专门的立体驱动,甚至可以用此方法观看纸质图片。但用双色眼镜观看黑白图片的效果较好,观看彩色图片的效果较差。目前这种观察设备有被淘汰的趋势。

② 液晶光阀眼镜

液晶光阀眼镜所采用的显示方式是页交换(Page Flipped)的显示方式^[38, 69]。

在液晶上加一定的电压会改变其分子排列,从而可以控制液晶镜片的开关状态,允许或阻止光线通过液晶镜片。液晶光阀眼镜的页交换的显示方式是指:在

某一时刻,关闭右眼液晶镜片,打开左眼液晶镜片,同时让计算机 CRT 显示器显示左眼视图;在下一时刻,关闭左眼液晶镜片,打开右眼液晶镜片,同时让计算机 CRT 显示器显示右眼视图。在此工作方式中改变液晶的开关状态,使左右眼镜片及左、右眼视图以较快的速度切换,大脑会认为左右眼同时看到了各自的图像,并综合左、右眼视图产生深度感知,这就达到了“左眼看左眼视图,右眼看右眼视图”的目的。

计算机的 CRT 显示器存在一定的刷新率,必须采用外部辅助电路连接 CRT 显示器与液晶镜片,使 CRT 显示器的刷新率与液晶镜片的开关速率同步,才能达到以上的目的。由于在刷新率不变的情况下左、右眼视图需要交替显示,因此每只眼睛看到的图像刷新率实际上是 CRT 显示器刷新率的一半。若 CRT 显示器刷新速率低的话,眼睛看到的图像刷新率更低,会出现画面闪烁感。可见, CRT 显示器的刷新率越高越好。

液晶光阀眼镜的价格适中,并可用于观看彩色图像。它虽然对 CRT 显示器刷新率要求较高,但立体效果(色彩、沉浸(Immersion)感等)较双色眼镜好^[33]。因此,它在游戏娱乐、虚拟现实等领域里有广泛的应用。本文实验采用的分时双目立体成像系统就使用了液晶光阀眼镜。

③ 头盔显示器

头盔显示器(Head-Mounted Display, HMD)是 3D 显示技术中起源最早、发展得最完善、应用也较广泛的技术,是专为用户提供虚拟现实场景中景物彩色立体显示的显示器。它通常固定在用户的头部,用两个 LCD 或 CRT 显示器分别向两只眼睛显示计算机立体驱动生成的立体图像对。大脑将综合左、右眼视图产生深度感知(如图 2.1 所示)。头盔显示器的种类较多,根据需求的不同,有全投入式的和半投入式的,等等^[33]。



图 2.1 头盔显示器

Fig 2.1 Head-Mounted Display

头部位置跟踪设备是头盔显示器上的主要部件。通过跟踪头部位置,虚拟现实用户的运动感觉和视觉系统能够得以重新匹配跟踪,计算机随时可以知道用户头部的位置及运动方向。因此,计算机就可以随着用户头部的运动,相应地改变呈现在用户视野中的图像,从而提高了用户对虚拟系统知觉的可信度。头部位置跟踪还可以增加双眼视差和运动视差,这些视觉线索能改善用户的深度感知。

头盔显示器的分辨率较低,屏幕成像小,必须放大以达到和人的视野相一致,失真大。人眼在近距离内聚焦会感到疲劳,佩戴它观察无法得到舒适自然的体验。而且头盔显示器的造价较高。但头盔显示器具有较好的临场感和沉浸感,在许多特定的场合具备特殊的优势,因此被广泛应用于军事、CAD/CAM、工业生产、模拟和训练、3D 显示与电子游戏、显微技术和医疗等领域。

④ 自由立体显示器

自由立体显示器是立体显示技术研究的一个重要方向。利用自由立体显示器,无需佩戴任何眼镜就能产生立体效果^[18]。美国、日本、德国、加拿大等发达国家早在 20 世纪 60 年代末就开始进行这个领域的研究,并相继提出了一些实现原理和方法。TFT LCD 技术成熟后,掀起了自由立体显示器研究的又一轮热潮^[66, 70]。

自由立体显示按实现方法分主要有透镜法和光栅法两种^[71]。两种方法都通过控制活动的 LCD 彩色阵列,显示一种合成的图像。该图像包含竖直交替排列的图像条纹,这些条纹构成了具有双目视差的左、右眼视图像。左、右眼视图以一定的速率切换,分别显示在奇数列和偶数列上,交替的并行显示使得每只眼睛只看到其对应列,亦即其对应的左眼视图或右眼视图,因而产生了深度感知。

目前在国外已经有一些研究机构研制出自由立体显示器的原型机,例如 Dimension Technologies 公司生产的 DTI 1100C“虚拟窗口”就是这种显示器^[40]。它的每个 LCD 像素点由许多窄光源从背部照亮,背部光源的距离 d 与 LCD 到使用者的距离 p_d 之间的关系决定了立体观察锥。使用者必须在此观察锥内才能看到立体效果,否则无立体效果。此外,德国的 Dresden3D GmbH 公司生产的 Dresden3D D4D 显示器,已经在医疗方面得到了应用^[71];美国纽约大学也研制出自由立体显示器的原型机。但是现在的自由立体显示器造价相当高,而且有许多技术问题尚待解决,因此没有得到广泛的应用。

2.1.2 分时双目立体成像系统

分时双目立体成像是模拟人眼观察物体的方式,将左右两幅视图交替显示来产生立体图像的一种常用的系统^[15, 18, 38]。本文实验采用的是基于 NVIDIA 显卡的分时双目立体成像系统。目前市面上显卡所采用的显示卡芯片中,支持立体

显示功能且提供立体驱动程序的主要是 NVIDIA 公司生产的 TNT 系列、MX 系列、Geforce 系列的显示卡芯片,而其他公司生产的大部分显示卡芯片尚不能支持立体显示功能。因此,本论文中使用的显示卡芯片是 NVIDIA 公司生产的 Geforce 系列显示卡芯片,其中芯片类型为 GeForce4 MX 440 with AGP8X。CRT 显示器分辨率为 800×600 时刷新率最大能达到 100Hz。观察设备是液晶光阀眼镜,采用的实现方式为页交换(Page Flipped)显示方式。图像信号输出部分的硬件线路连接示意图如图 2.2 所示。

在图 2.2 所示的硬件线路连接示意图中,同步信号线接头的作用是连接显示卡输出接口和 CRT 显示器的信号线接头,并且引出一条显示信号的同步信号线,经过控制盒,最终与液晶光阀眼镜相连接,这样同步信号就能控制左右液晶光阀分时、交替开启和关闭,以保证左眼只能看到左画面,右眼只能看到右画面。

此外,CRT 显示器的信号线接头为 VGA 插头,控制盒和液晶光阀眼镜的插头为 3.5mm 的立体声耳机插头。

在图像输出过程中,很有可能出现左、右眼视图反转(即左眼看到了右画面,右眼看到了左画面)的情况,这样将无法使观看者产生正确的生理深度暗示。因此控制盒上装有一个左右眼交换开关。当左、右眼视图反转的时候,按一下这个开关就会使这两幅图像再次反转,从而恢复正常。

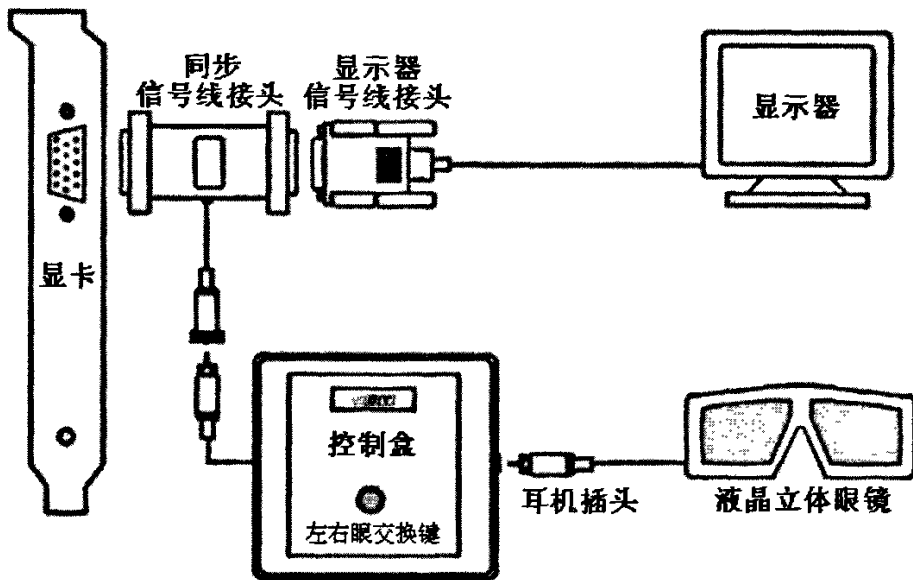


图 2.2 分时双目立体成像系统的硬件

Fig 2.2 Hardware of the Time-sharing Binocular Stereo Imaging System

该分时双目立体成像系统的工作原理是：计算机将左眼视图和右眼视图交替迅速地显示在监视器上。当左眼视图显示时，液晶光阀眼镜使右眼的液晶镜片遮断，使得只有左眼才能看到监视器；同理，当右眼视图显示时，液晶光阀眼镜使左眼的液晶镜片遮断，使得只有右眼才能看到监视器，这样便产生了立体感^[36, 69]。NVIDIA 系列显卡能够将 JPS 文件中的左右眼视图(立体图像对)交替显示在监视器上^[38]。因此，只要生成含有立体图像对的 JPS 文件，就能通过这个系统看到立体图像。

2.1.3 立体图像文件类型

常用的立体图像文件类型有 H3D(.h3d)文件类型和 JPS(.jps)文件类型等等^[38]。本论文采用了 JPS 文件格式存储立体图像。JPS(JPEG Stereo)文件格式的立体图像是由同一空间场景中两幅同为 JPG 格式的左右眼视图拼凑成的，因此 JPS 格式与 JPG 格式没有本质上的区别，把后缀名为 jps 的图像文件改为 jpg 就可以作为 JPG 文件使用。

标准的 JPS 文件是由左眼视图与右眼视图交错拼凑成的，图 2.3 是 JPS 文件中图像的一种排列情况。

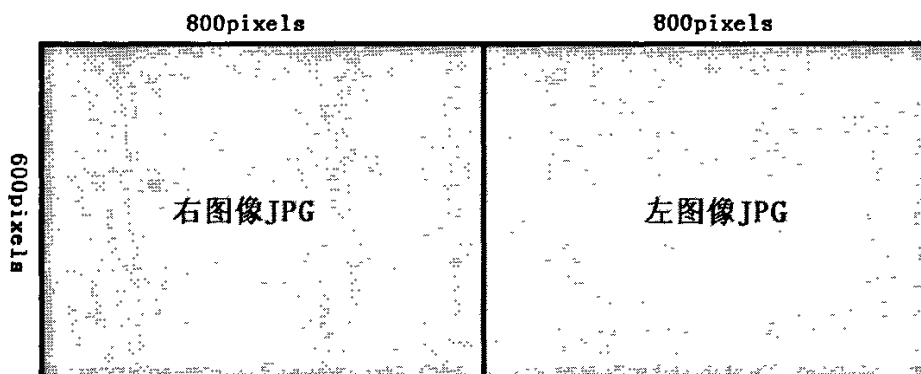


图 2.3 JPS 文件类型

Fig 2.3 JPS file type

打开 JPS 文件进行观看时，立体驱动程序会自动将文件中的图像一分为二，作为左右眼视图分别进行加载。

从 JPS 文件类型可以看出，存储或传输立体图像对是一件非常消耗空间和时间的事，因此有必要研究比二维图像的压缩更高效的压缩方案。立体图像对压缩方法的研究开始于 20 世纪 80 年代末^[72]，经过十几年的发展，现在已经有了几种

较为成熟的算法。但从实际应用的角度看，目前还没有形成统一的编码标准。目前使用比较广泛的是基于块(block-based)匹配的压缩方法和基于物体(object-based)的压缩方法。其中，基于块匹配的视差估计和补偿的方法研究较多[73]。



图 2.4 立体图像

Fig 2.4 Stereo image

图 2.4 是用立体相机的两个摄像头模拟人的双眼对物体从两个不同角度获取的立体图像。从图 2.4 可以看出，立体图像对之间存在极大的相似性。因此，通常对于立体图像使用视差估计和补偿的方法以去除两图像之间的相关性，从而达到压缩的目的。

视差估计和补偿技术类似于视频压缩中的运动补偿技术。在该方法中，一幅图像进行单独编码，并作为参考图像，另一幅是由参考图像来对其进行预测的目标图像。而作为原始图像和预测图像之差的残余图像还要进行单独的压缩编码，以修正预测误差。视差估计过程是确定每个像素的位移矢量。由于对每个像素单位单独处理将使运算复杂化，所以常用 $k \times k$ 的图像子块来进行处理。通常取 $k = 8$ 或 $k = 16$ 。

把目标图像分成大小相等的图像子块后，需要根据一定的误差匹配准则，对每一块在参考图像中确定的搜索窗内搜索其最佳匹配块。一般匹配误差最小的一

块被选中。两个常用的误差匹配准则是平均绝对误差(MAE)和均方误差(MSE)准则。

定义 2.1: 设右眼视图 $rImg$ 为参考图像, 左眼视图 $lImg$ 为目标图像, 且左眼视图被分割成 $M \times N$ 个图像子块, 则

$$MAE = \frac{1}{M \times N} \left(\sum_{m=1}^M \sum_{n=1}^N |lImg(m, n) - rImg(m+h, n+v)| \right) \quad (2.1)$$

$$MSE = \frac{1}{M \times N} \left(\sum_{m=1}^M \sum_{n=1}^N |lImg(m, n) - rImg(m+h, n+v)|^2 \right) \quad (2.2)$$

式中 h 和 v 分别表示水平视差和垂直视差。

尽管 MSE 比 MAE 有更好的性能, 一般选择 MAE 而不是 MSE 作为误差匹配准则。因为 MAE 在硬件实现时更简单有效。

视差矢量图通常用差分脉冲编码调制(DPCM)进行无损编码, 接着进行哈夫曼编码进一步压缩数据量。在给出图像的视差估计后, 残余图像由原始图像减去估计值来得到。它主要包含边缘和低频信息, 然后对残余图像和参考图像分别进行编码。

我们已初步实现了基于视差估计和补偿方法的立体图像对压缩, 这使得系统性能在存储和传输方面有所提高, 但更多的改进工作仍需要进一步深化。

2.2 双目立体成像相关理论

2.2.1 会聚角变化值

经验表明, 人们习惯于调整焦点和转动双眼来观察不同深度的物体。在进行双目立体成像时, 虽然人们感兴趣的像点的深度不同于屏幕, 但是仍要求我们聚焦于屏幕。在这种情况下, Valys发现大多数人能容忍的会聚角变化值为 1.6° ^[42]。如果这个值过大, 会引起“调节/汇聚”矛盾, 如图2.5所示。

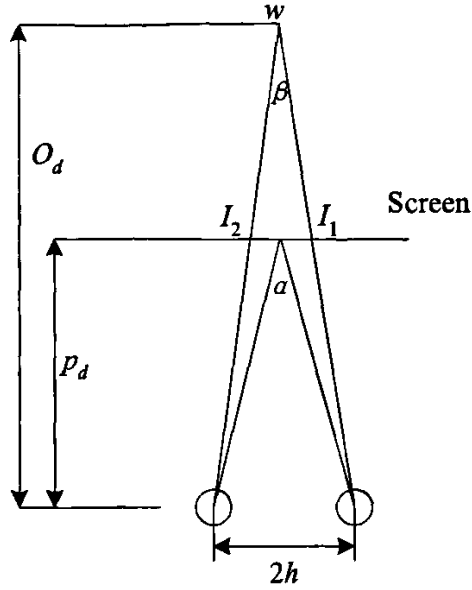


图 2.5 调节与汇聚

Fig 2.5 Adjustment and focus

Valyus指出

$$D_{\max} = 0.03p_d, \quad (2.3)$$

其中 D_{\max} 为最大允许视差。Valyus 的理论与论文后面提到的 Panum 融合区 (Panum's fusional area) 的理论本质上是一致的。通常取两者中更为严格的限制条件。

公式(2.3)限制了立体图像的景深(在摄影机镜头或其他成像器前沿着能够取得清晰图像的成像器轴线所测定的物体距离范围)。从图 2.5 可得

$$\alpha = 2\arctg \frac{h}{p_d} \quad (2.4)$$

$$\beta = 2\arctg \frac{h}{O_d} \quad (2.5)$$

因此会聚角变化值为

$$\Delta\theta = \alpha - \beta = 2(\arctg \frac{h}{p_d} - \arctg \frac{h}{o_d}) \quad (2.6)$$

由于 $\Delta\theta$ 很小时, 有 $tg\theta \approx \theta$ (θ 用弧度表示), 则由(2.4)、(2.5)、(2.6)可得

$$o_d = \frac{h}{tg \frac{\beta}{2}} \approx \frac{h}{\frac{\beta}{2}} = \frac{h}{\frac{\alpha}{2} - \frac{\Delta\theta}{2}} \approx \frac{h}{tg \frac{\alpha}{2} - \frac{\Delta\theta}{2}} = \frac{2h}{\frac{2h}{p_d} - \Delta\theta} \quad (2.7)$$

因为 $-\frac{1.6\pi}{180} \leq \Delta\theta \leq \frac{1.6\pi}{180}$, 所以可以根据公式(2.7)计算出立体图像的景深。由此可见, 生成的立体图像对的景深不是无限制的。我们不能将虚像成像在任意的位置。

2.2.2 Panum 融合区

另一个在立体成像中经常遇到的问题是 Panum 融合区问题, 是 Panum 于 1858 年发现并提出的, 主要解决双象单视问题^[15, 74, 75]。

早在 1858 年, Panum 就指出^[76, 77], 当人眼的双眼注视一个外界物体时, 另外一个远于或近于双眼单视界的物体, 虽然没有刺激到左、右眼视网膜的对应点上, 但只要刺激在两个视网膜对应点附近的一定范围内, 也可以在大脑中产生单一的立体视觉。视网膜对应点附近这个很小的范围被叫做 **Panum 融合区** (Panum's Fsional Area)。如果几个物体刺激两个视网膜形成的像的视差过大, 便会出现复视, 即不能把该物体在左、右眼视网膜上形成的像在大脑中融合成单一的立体视觉。Panum 融合区说明并不是所有在两个视网膜上形成含有双目视差的图像都能在大脑中形成单一的立体视觉, 只有满足一定双目视差的立体图像对才能被融合成单一的立体图像。如果立体图像对的左、右眼视图的视差过大, 那么在一定的距离下观看立体图像时, 人的眼睛和大脑就无法协调晶状体调节、双眼会聚与双目视差等多种生理立体视觉因素所提供的深度信息之间的关系, 将左、右眼视图融合成单一的立体图像, 其结果会给观察者造成很大的不舒适感。图 2.6 说明了晶状体调节、双眼会聚与双目视差是如何相互影响的。

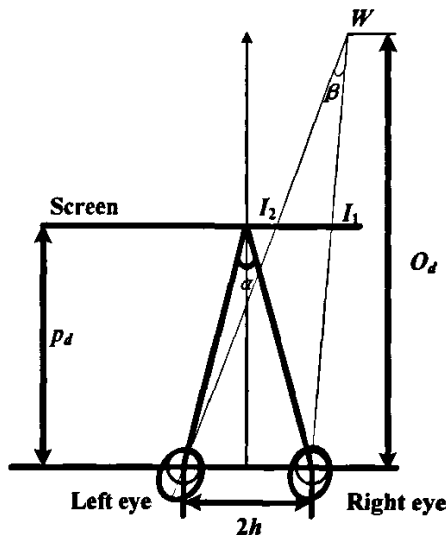


图 2.6 视差与 Panum 融合区

Fig 2.6 Parallax and Panum's Fsional Area

观察者在观看立体图像时，设 p_d 为观察者与屏幕之间的距离， W 是在观察者大脑中形成的双眼单视的像点，当 p_d 与 W 的深度之间的差超过一定限制时，晶状体调节所提供的深度暗示将起主导作用，而双目视差所提供的深度暗示无法被融合成单一的立体图像。超出Panum融合区限制的立体图像，会使观察者感到疲劳、头痛、甚至引起复视。实验表明， o_d 不能超过的限制为^[78-80]

$$\frac{Ep_d}{E+\eta p_d} < o_d < \frac{Ep_d}{E-\eta p_d} \quad (2.8)$$

其中 E 是人眼的瞳孔直径，典型值为0.4 cm， η 为视锐度，一般取 $\eta = 1' \approx 2.907 \times 10^{-4} \text{ rad}$ 。为了获得舒适的立体图像，显示器屏幕前后的像点深度值都不应超过公式(2.8)的限制。落入公式(2.8)范围内像点集合所覆盖的区域称为**舒适Panum融合区**。一种成功的双目立体成像系统展示给观察者的立体图像，应该是不使观察者产生视疲劳的图像。因此水平视差的取值范围不能超出舒适Panum融合区的限制。

2.2.3 生理深度暗示和心理深度暗示

由于立体图像是在人的大脑视区中形成的，因此它与人类的立体视觉是密不可分。人类的立体视觉与深度暗示是紧密联系在一起。通常，这些深度暗示可

以分为两类：生理深度暗示(Physiological Depth Cue)和心理深度暗示(Psychological Depth Cue)。

生理深度暗示是基于双目立体视觉的基本原理产生的，包括人眼晶状体调节、双眼会聚和双目视差等因素。其中，双目视差是人眼最强烈的生理深度暗示因素。Bela Julesz 利用随机点图已经证明，在排除一切心理深度暗示之后，一组完全无意义的视觉刺激，只要具备视差条件，就能经双眼产生深度上的感觉。这说明双目视差可以与任何视觉经验无关。由于双目视差对立体视觉的贡献最大，是最强烈的生理深度暗示，所以也是双目立体成像考虑的主要因素^[15]。

心理深度暗示主要由视觉经验和视觉记忆构成。人们在观看一张平面彩色图片时，可以根据图片上的内容判断其中物体、人物之间的距离关系，而这种判断通常十分准确。这说明平面图像中尽管不存在能用人的双目视差等生理深度暗示因素来识别的深度信息，却存在着其它的深度暗示。这些暗示信息是人类对自然景物长期观察而得到的一种立体的视觉记忆和视觉经验，依靠这种视觉记忆和经验，观察者能够从平面图像中准确地提取出物体间的相对位置和深度信息，这种深度暗示被称为心理深度暗示。

通过大量的实验人们发现：当我们观察自然界的景物时，在观察者每只眼睛的视网膜上都形成了反映被观察景物深度关系的像，当不考虑这些像在两眼之间形成的视差时，这种反映被观察景物深度关系的像就构成了心理深度暗示；当考虑在两只眼睛视网膜上形成的像的视差时，它们构成了另一种深度暗示，即生理深度暗示^[15]。由于被观察的景物是同一个景物，其深度关系是同一个深度关系，因此生理深度暗示和心理深度暗示所反映的被观察景物的深度关系应该是一致的。这种一致性作为一种恒常的视觉经验被记忆，以至于任何违反这种恒常性的深度信息都会受到强烈的抑制。因此，人类的立体视觉是综合作用的结果，生理深度暗示和心理深度暗示在不同的情况下有相互增强或相互抑制减弱的作用。

图 2.7 是为说明两者之间的关系所设计的一个试验的结果。在试验中我们请受试者指出观看到图像的深度感知。图中的横坐标 X 表示屏幕的水平方向，纵坐标 E 表示显示在屏幕上的图像给观察者造成的深度感知程度。“.”表示观察者对屏幕上一个图像单元的心理深度感知，“■”表示观察者对屏幕上一个图像单元的生理深度感知，“△”表示屏幕上一个已经含有心理深度信息的图像单元，在加入了生理深度暗示之后给观察者造成的综合的深度感知。

例如图 2.7 中的第一个图像单元给观察者造成的心理深度感知是 Ex_1 ，随机加入生理深度暗示给观察者造成的生理深度感知是 Es_1 ，而在这个单元中加入的生理深度暗示与原有的心理深度暗示不矛盾，生理深度感知和心理深度感知共同形成的综合深度感知为 Ez_1 ，这个图像单元的深度感知得到加强。

再看图 2.7 中的第二个图像单元，其心理深度感知是 Ex_2 ，随机加入生理深度暗示给观察者造成的生理深度感知是 Es_2 ，由于加入的生理深度暗示与原有的心理深度暗示矛盾，便形成综合深度感知 Ez_2 。这时心理深度暗示对生理深度暗示产生了抑制减弱的作用。用同样的方法可以分析得出所有图像单元的综合深度感知，如图 2.7 所示。这样得到的综合深度感知符合心理深度暗示深度的变化规律。由于心理深度暗示反映的也是客观景物的深度变化，因此从这样构成的立体图像中，观察者可以获得符合客观景物深度变化规律的含有双目视差的深度感知。

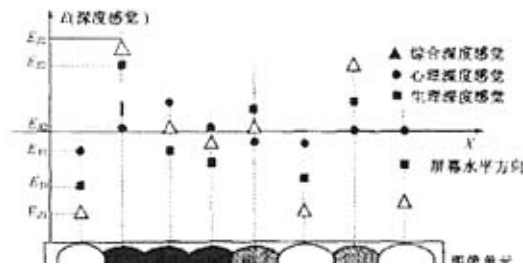


图 2.7 生理深度暗示和心理深度暗示的关系

Fig 2.7 Relation between Physiological Depth Cue and Psychological Depth Cue

由此得出结论：当生理深度暗示与心理深度暗示一致时，综合的深度感知增强；而当生理深度暗示与心理深度暗示矛盾时，生理深度暗示会被心理深度暗示抑制减弱^[15]。产生这一立体视觉现象的原因是，心理深度暗示尽管不能提供最强的深度暗示，但对违反视觉经验的深度暗示将产生极强的抑制作用，这是单眼视觉模式识别的重要特征，也是人类视觉提供给自身的一种重要的自我保护功能。

侯等人指出^[81]，在每个图像单元中随机加入双目视差，在心理深度暗示的作用下，会使整个立体图像的综合深度感觉得到统计规律上的加强。当观看含有随机的双目视差图像时，心理深度暗示会自动地对加入的双目视差信息进行增强或抑制减弱，从而使观察者获得具有一定深度感知的立体图像。这对于处理那些几乎没有运动信息或运动变化十分复杂的平面序列图像更加有效。

2.2.3 遮挡问题

双目立体成像有时候会涉及遮挡问题。例如，图 2.8(a)是观察者在某一位置看到的立体图像，此时他不一定能正确地分辨物体之间的位置关系；但只要他轻微地转动头部，就可以看到物体之间实际的位置关系，如图 2.8(b)所示。

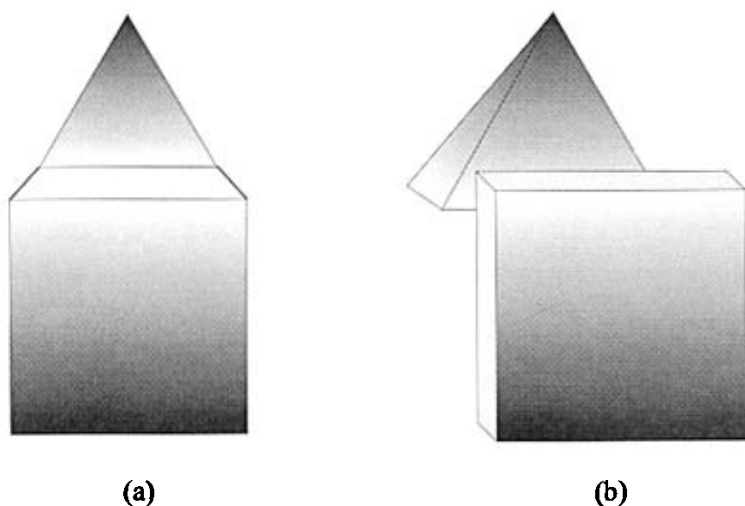


图 2.8 物体的位置关系

Fig 2.8 Position relation between objects

在计算机立体视觉领域，遮挡问题是比透视几何畸变更复杂的问题^[61]。遮挡问题首先要解决的是如何检测遮挡区域。有一种算法是根据 Marr 和 Poggio 提出的影像匹配的唯一性和连续性的条件，采用协同匹配算法检测遮挡区域^[82, 83]；另一种算法采用自适应窗口和“双向匹配方法”的算法检测遮挡区域^[84-87]。Scharstein 和 Szelski 提出了一个基于贝叶斯估计的影像匹配模型^[88]，该模型视差分布是非线性的。该方法取得了与自适应窗口算法类似的效果。还有几种方法采用了动态规划思想，在算法中应用了顺序约束来同时获取视差和检测遮挡区域^[89-91]。另外一种方法是应用于城市三维模型提取中遮挡区域检测的，该方法建立在三维物体模型假设的基础上，将影像特征和三维物体模型假设相结合提取三维物体模型的真实数据^[92-94]。

其次要解决的问题是，在检测出遮挡区域以后，如何表示物体之间的位置关系以及让计算机理解这种位置关系，并且大大节省计算量。我们认为，可以用空间关系的定性表示与推理来解决这个问题。

① 空间关系的定性表示

空间关系是指物体之间存在的一些具有空间特性的关系^[95]。空间关系主要有拓扑关系、方向关系、度量关系三大类。其中拓扑关系是指在拓扑变换(旋转、平移、缩放等)下保持不变的空间关系，即拓扑不变量，如物体之间的相离和相交关系；方向关系描述了物体在空间中的某种排序，如前后、东西等等；度量关系是用某

种度量空间中的度量来描述的物体之间的关系，如物体之间的距离。

空间关系表示的基本任务是以数学或逻辑的方法区分不同的空间关系，给出形式化的表示。其意义在于澄清不同用户关于空间关系的语义，为构造空间查询语言和空间分析提供形式化工具^[95]。由于空间问题固有的复杂性和不确定性，空间关系的表示普遍采用定性的方法^[96]。

拓扑关系的定性表示与推理主要有两种基本的方法：基于点集拓扑学的“**n-交集**”模型和基于区域连接的 **RCC 方法**^[97]。

方向关系的定性表示主要有**锥形法(Cone)**和**投影法(Projection)**。锥形法是将空间物体及其周围的区域分成带有方向性的几个区域，通过各物体本身与这些区域之间的“交”的结果来描述方向关系；投影法是将空间物体投影到特定的坐标轴上，通过各物体投影间的关系来描述方向关系。由于两种表示各有优缺点，我们综合了投影法和锥形法两种方向关系表示方法，提出了 **DR 方向关系表示模型**。

为了定义 **DR 方向关系表示模型**，我们首先给出如下几个概念的形式化定义。

定义 2.2: 基于投影法的方向关系

设 $V = \{<, >, m, mi, o, oi, s, si, di, d, fi, f, =\}$ 表示 **JEPD 关系**^[97]的符号集合，则物体 **A** 与 **B**(参照物体)之间的方向关系可定义为二元组 $(\chi(A, B), \phi(A, B))$ ，其中 $\chi(A, B) \in V$ ， $\phi(A, B) \in V$ ， $\chi(A, B)$ 表示物体 **A**、**B** 在 **X**轴上投影的 **JEPD 关系**， $\phi(A, B)$ 表示物体 **A**、**B** 在 **Y**轴上投影的 **JEPD 关系**。

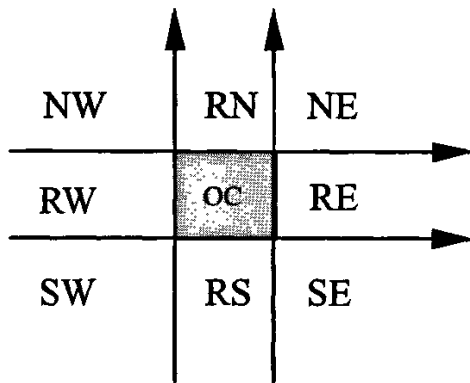


图 2.9 井字空间

Fig 2.9 Space with Neutral Zone

定义 2.3: 井字空间中的方向关系

设 $D = \{NE, RN, NW, RW, SW, RS, SE, RE, 0\}$ 表示方向关系的符号集合，其中 **0** 是表示“同位”方向关系的符号(当两个物体的最小边界矩形^[97, 98]的内部有交点时，

用“同位”表示其方向)。设 2^D 表示集合 D 的幂集, 则物体 A (目标物体)与 B (参照物体)的方向关系可表示为二元函数 $\text{dir}(A, B)$, $\text{dir}(A, B) \in 2^D$ 。其中 $\{NE\}$ 、 $\{RN\}$ 、 $\{NW\}$ 、 $\{RW\}$ 、 $\{SW\}$ 、 $\{RS\}$ 、 $\{SE\}$ 、 $\{RE\}$ 、 $\{0\}$ 称为强方向关系。空关系 \emptyset 表示不能确定或存在矛盾的方向关系。

有了上述概念, 我们就可以给出如下定义:

定义 2.4: DR 方向关系表示模型

二维空间物体 A 、 B 的空间方向关系 $DR(A, B)$ 可表示为:

$$DR(A, B) = \begin{cases} (\chi(A, B), \varphi(A, B)), & \text{dir}(A, B) = \{0\}, \\ \text{dir}(A, B), & \text{dir}(A, B) \neq \{0\}. \end{cases}$$

但是, 当 $\text{dir}(x, y) = \{0\}$ 时, 仅仅采用定义 2.4 来描述两个物体之间的方向关系不能保证正确地推理^[97, 99], 因此有必要将它与拓扑关系结合起来。为此, 我们提出了 SR 表示模型。

定义 2.5: SR 表示模型

设 $\delta(A, B)$ 表示物体 A 、 B 之间的拓扑关系(RCC8, RCC15 或 RCC23[97, 100]), $DR(A, B)$ 表示 A 、 B 间的方向关系, 称

$$SR(A, B) = (\delta(A, B), DR(A, B))$$

为空间物体 A 、 B 间的空间关系, 或 SR 表示模型。

在有些情况下根据定义 2.5 所定义的 SR 表示模型来描述空间关系可能仍然满足不了对精度的要求。这种情况下我们对 SR 表示模型的定义稍稍做了修改, 增加了拓扑关系度量, 提出了度量型 SR 表示模型。

定义 2.6: 度量型 SR 表示模型:

$$SR(A, B) = \begin{cases} (\delta(A, B), \delta_m(A, B), (\chi(A, B), \varphi(A, B))), & \text{dir}(A, B) = \{0\}, \delta(A, B) \in T, \\ (\delta(A, B), \delta_m(A, B), \text{dir}(A, B)), & \text{dir}(A, B) \neq \{0\}, \delta(A, B) \in T. \end{cases}$$

其中 $\delta_m(A, B)$ 表示 $\delta(A, B)$ 的度量。

表 2.1 拓扑关系度量

Table 2.1 Measurement of topological relations

$\delta(A, B)$	$\delta_m(A, B)$	备注
$DC(A, B)$	ds/dc	
$EC(A, B)$	0	
$PO(A, B)$	$\text{area}(A \cap B) / (\text{area}A + \text{area}B)$	ds: A 、 B 间的最短距离; dc: A 、 B 中心间的距离; $\text{area}A$: A 的面积; $\text{area}B$: B 的面积; $\text{area}(A \cap B)$: A 与 B 相交的面积
$TPP(A, B)$	$\text{area}A / \text{area}B$	
$NTTP(A, B)$	$\text{area}A / \text{area}B$	
$TPPi(A, B)$	$\text{area}B / \text{area}A$	
$NTPPi(A, B)$	$\text{area}B / \text{area}A$	
$EQ(A, B)$	0	

对 RCC8 的情况, $\delta_m(A, B)$ 可按表 2.1 所示进行计算; 而在 RCC15、RCC23 情况下如何计算仍需进一步研究。

上述所有定义都可以推广到三维物体的情况。

度量关系主要是距离关系, 其定性表示的研究也处在起步阶段。在综合前人研究成果的基础上^[97, 101, 102], 我们给出了一个距离关系的定性表示模型:

定义 2.7: 设 $Const$ 为一个参考常量, 任意一个距离量 x 与 $Const$ 的比值记为 $[x/Const]$, 参考点取 $[x/Const]=2/3$ 、 1 、 $3/2$, 我们据此用 7 个标签对距离的大小进行定性划分:

$$\begin{aligned} l &= \{x \mid x \in R, 0 < [x/const] < 2/3\} \\ \alpha &= \{x \mid x \in R, [x/const] = 2/3\} \\ sl &= \{x \mid x \in R, 2/3 < [x/const] < 1\} \\ e &= \{x \mid x \in R, [x/const] = 1\} \\ sg &= \{x \mid x \in R, 1 < [x/const] < 3/2\} \\ \beta &= \{x \mid x \in R, [x/const] = 3/2\} \\ g &= \{x \mid x \in R, 3/2 < [x/const] < \infty\} \end{aligned}$$

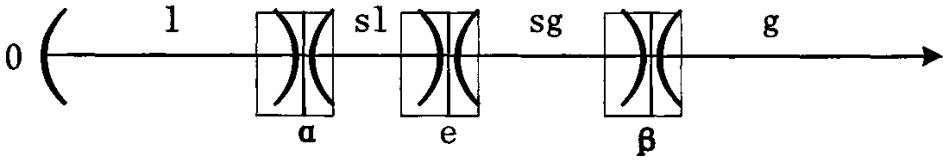


图 2.10 距离的定性划分

Fig 2.10 Qualitative labeling of distance

采用这种距离关系的定性表示模型时, 需要重新定义一个方向关系的定性表示模型以配合完成定性三角推理^[97, 101-103]。

定义 2.8: 在二维平面中, 我们用 7 个标签对角度进行定性划分:

$$\begin{aligned} a &= \{\theta \mid 0 < \theta < \pi/3\} \\ \delta &= \{\theta \mid \theta = \pi/3\} \\ sa &= \{\theta \mid \pi/3 < \theta < \pi/2\} \\ r &= \{\theta \mid \theta = \pi/2\} \\ so &= \{\theta \mid \pi/2 < \theta < 2\pi/3\} \\ \varphi &= \{\theta \mid \theta = 2\pi/3\} \\ o &= \{\theta \mid 2\pi/3 < \theta \leq \pi\} \end{aligned}$$

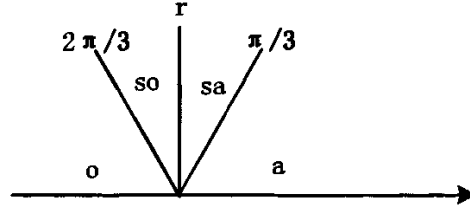


图 2.11 角度的定性划分

Fig 2.11 Qualitative labeling of angle

当前, 将拓扑关系、方向关系和度量关系结合起来对空间关系进行定性表示的研究还在进行当中。

② 空间关系的定性推理

空间推理是指利用空间理论和人工智能 AI(Artificial Intelligence)技术对空间物体之间的空间关系(方向关系、拓扑关系、度量关系)进行定性或定量分析和处理的过程^[104]。目前, 空间推理也采用定性的方法。

空间关系的定性推理一般都要借助组合运算的形式。我们重点研究了 SR 表示模型的组合运算和定性三角推理。

1) SR 表示模型的组合运算

定义 2.9: SR 表示模型的组合运算

设 ∞ 是建立在函数 $SR(x, y)$ 上的运算, 且对任意物体 A, B 和 $C, SR(A, B) \infty SR(B, C) =_{\text{def}} SR(A, C)$, ∞ 称为 SR 表示模型的组合运算。组合运算 ∞ 的运算规则是函数 SR 表示模型中的分量分别进行组合运算, 即

$$SR(A, C) = SR(A, B) \infty SR(B, C) = SR(\delta(A, B) \infty \delta(B, C), DR(A, B) \infty DR(B, C))$$

如果分量 $\delta(x, y)$ 采用的是 RCC8, 则其组合运算表如表 2.2 所示^[97]。

从表 2.2 可以看出, 对于基于 RCC8 的一阶推理的实现是极其困难的。而构造基于 RCC23 的组合运算表, 并且实现基于 RCC23 的一阶推理就更为困难了。为了解决这个问题, Bennett 提出了将 RCC 的一阶逻辑转换为零阶逻辑的思想^[97]。如果采用经典的零阶逻辑对 RCC8 的关系进行编码, 则这种编码并没有足够充分的表达力, 因此人们引入了一种叫做 I_0 的零阶逻辑^[105]。之后人们完善了 I_0 , 提出了所谓的 I_0^+ 。但是, 基于 I_0^+ 的推理依然有许多要克服的困难。目前, 这一问题已成为国际学术界研究的前沿问题, 也是我们课题组今后研究的一个重要内容。

分量 $DR(x, y)$ 的组合运算表的构造也较复杂。假如我们严格按照定义 2.3 来构造组合运算表的话, 那么得出的组合运算表不但庞大复杂, 而且也不符合人们日常生活常识, 偏离了定性表示的初衷。为简化问题的复杂性, 真正体现定性表示的特点, 在较大尺度空间中, 我们认为任意两个物体 A, B 的大小相差不大, 这样

当 $\text{dir}(A, B)=\{0\}$ 时, 可以认为它们完全相等。在上述约定下, 我们构造出了近似的强方向关系组合运算表, 如表 2.3 所示。

表 2.2 RCC8 的组合运算表($\delta(A, B) \in \delta(B, C)$)

Table 2.2 Composition operation table of RCC8($\delta(A, B) \in \delta(B, C)$)

$\delta(A, B) \backslash \delta(B, C)$	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ
DC	no.info	DR,P O,PP	DR,P O,PP	DR,P O,PP	DR,P O,PP	DC	DC	DC
EC	DR,P O,PPi	DR,P O,TPP	DR,P O,PP	EC,PO ,PP	PO,PP	DR	DC	EC
PO	DR,P O,PPi	DR,P O,PPi	no.info	PO,PP	PO,PP	DR,P O,PPi	DR,P O,PPi	PO
TPP	DC	DR	DR,P O,PP	PP	NTPP	DR,P O,TPP, TPPi	DR,P O,PPi	TPP
NTPP	DC	DC	DR,P O,PP	NTPP	NTPP	DR,P O,PP	no.info	NTPP
TPPi	DR,P O,PPi	EC,PO ,PPi	PO,PP i	PO,TP P,PPi	PO,PP	PPi	NTPPi	TPPi
NTPPi	DR,P O,PPi	PO,PP i	PO,PP i	PO,PP i	O	NTPPi	NTPPi	NTPPi
EQ	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ

表 2.3 近似的强方向关系组合运算表($\text{dir}(A, B) \in \text{dir}(B, C)$)

Table 2.3 Approximate composition operation table of Strong Directional Relations ($\text{dir}(A, B) \in \text{dir}(B, C)$)

$\text{dir}(A, B) \backslash \text{dir}(B, C)$	{RN}	{NE}	{RE}	{SE}	{RS}	{SW}	{RW}	{NW}	{0}
{RN}	{RN}	{NE}	{NE}	{E}	{SH}	{W}	{NW}	{NW}	{RN}
{NE}	{NE}	{NE}	{NE}	{E}	{E}	Φ	{N}	{N}	{NE}
{RE}	{NE}	{NE}	{RE}	{SE}	{SE}	{S}	{SL}	{N}	{RE}
{SE}	{E}	{E}	{SE}	{SE}	{SE}	{S}	{S}	Φ	{SE}
{RS}	{SH}	{E}	{SE}	{SE}	{RS}	{SW}	{SW}	{W}	{RS}
{SW}	{W}	Φ	{S}	{S}	{SW}	{SW}	{SW}	{W}	{SW}
{RW}	{NW}	{N}	{SL}	{S}	{SW}	{SW}	{RW}	{NW}	{RW}
{NW}	{NW}	{N}	{N}	Φ	{W}	{W}	{NW}	{NW}	{NW}
{0}	{RN}	{NE}	{RE}	{SE}	{RS}	{SW}	{RW}	{NW}	{0}

2) 定性三角推理(Qualitative Trigonometry Reasoning, QTR)

以前的研究工作主要侧重于拓扑关系^[100, 103, 106]和方向关系^[104, 107], 而在考虑空间问题的时候加入距离的因素能提高定性描述的准确性和推理结果的有效性。定性三角推理就是综合了距离和方向两类空间关系的推理方法。它基于外部参考框架而建立, 利用三角形的边-角关系来推导两点间的距离和方向。

根据定义 2.7 和定义 2.8, 我们可以构造出定性三角推理的组合运算表, 如表 2.4 所示。

表 2.4 定性三角推理组合运算表($Const = AB$)

Table 2.4 Composition operation table of QTR ($Const = AB$)

		$\angle A$						
		a	δ	sa	r	so	φ	o
(AB, AC)	(e, l)	(l \sim sl, a)	(sl, a)	(sl \sim sg, a)	(sg, a)	(sg, a)	(sg, a)	(sg \sim g, a)
	(e, α)	(l \sim sl, a)	<u>(sl, a)</u>	(sl \sim sg, a)	<u>(sg, a)</u>	(sg, a)	(sg, a)	(sg \sim g, a)
	(e, sl)	(l \sim sl, a \sim sa)	(sl, a)	(sl \sim sg, a)	(sg, a)	(sg \sim g, a)	(sg \sim g, a)	(sg \sim g, a)
	(e, e)	(l \sim sl, sa)	<u>(e, δ)</u>	(sg, a)	<u>(sg, a)</u>	(sg \sim g, a)	<u>(g, a)</u>	(g, a)
	(e, sg)	(l \sim sg, sa \sim o)	(sg, sa)	(sg \sim g, a \sim sa)	(sg \sim g, a)	(sg \sim g, a)	(g, a)	(g, a)
	(e, β)	(l \sim sg, sa \sim o)	<u>(sg, sa)</u>	(sg \sim g, a \sim sa)	<u>(g, a)</u>	(g, a)	<u>(g, a)</u>	(g, a)
	(e, g)	(l \sim g, sa \sim o)	(sg \sim g, sa \sim so)	(sg \sim g, a \sim sa)	(g, a \sim sa)	(g, a \sim sa)	(g, a)	(g, a)

注: 表中加下划线的项表示结果是一个确定的值而不是一段范围。

在定性三角推理当中, 由于可以在点与点之间任意构造三角形, 所以其中间步骤免不了会涉及到边-边和角-角的定性加、减操作(加法和减法本质上是一致的)。基于定义 2.7 和定义 2.8 中距离与方向的定性划分, 我们令 $G = \{l, \alpha, sl, e, sg, \beta, g\}$ 表示定性距离的值域, $\omega = \{0^\circ, a, \delta, sa, r, so, \varphi, o, \pi\}$ 表示定性角度的值域, 则距离的定性减法可以表示为

$$b_3 = \{b \mid b_1, b_2 \in G, b_1 > b_2, b_3 = b_1 - b_2\} \quad (2.9)$$

角度的定性减法可以表示为

$$\theta_3 = \{\theta \mid \theta_1, \theta_2 \in \omega, \theta_1 > \theta_2, \theta_3 = \theta_1 - \theta_2\} \quad (2.10)$$

由此得到关于距离的定性减法表和关于角度的定性减法表, 分别如表 2.5 和表 2.6 所示。类似的我们也可以构造出加法表。

表 2.5 和表 2.6 在具体的推理过程中常常用到。例如, 在已知三角形的两个内

角的情况下，根据内角和为 180° 可以用减法求得第三个角的定性范围。在使用关于距离的定性减法表时，还可依据三角形的两边之和大于第三边、两边之差小于第三边的原则检验推理结果的正确性和合理性，同时尽可能地缩小结果的范围。

利用表 2.4、表 2.5 和表 2.6，就可以实现定性三角推理。但是，上述推理是针对二维情况而言的，如何把这些推理规则推广到三维，还需要进一步研究。

表 2.5 距离定性减法表

Table 2.5 Qualitative subtraction table of Distance

被减数 \ 减数	l	α	sl	e	sg	β	g
l	0~l						
α	l	0					
sl	l~sl	l	0~l				
e	l~sl	$l(1/3)$	l	0			
sg	l~sl	l~sl	l~sl	l	0~l		
β	sl~sg	$sl(5/6)$	l~sl	$l(1/2)$	l	0	
g	sl~g	sl~g	l~g	l~g	l~g	l~g	0~g

表 2.6 角度定性减法表

Table 2.6 Qualitative subtraction table of Angle

被减数 \ 减数	0°	a	δ	sa	r	so	φ	o	π
0°	0°								
a	a	$0^\circ \sim a$							
δ	δ	a	0°						
sa	sa	$a \sim sa$	a	$0^\circ \sim a$					
r	r	$a \sim sa$	<u>$a(\pi/6)$</u>	a	0°				
so	so	$a \sim so$	a	a	a	$0^\circ \sim a$			
φ	φ	$sa \sim so$	<u>$\delta(\pi/3)$</u>	a	<u>$a(\pi/6)$</u>	a	0°		
o	o	$sa \sim o$	<u>$sa \sim so$</u>	$a \sim so$	$a \sim sa$	$a \sim sa$	a	$0^\circ \sim a$	
π	π	o	<u>$\varphi(2\pi/3)$</u>	so	<u>$r(\pi/2)$</u>	sa	<u>$\delta(\pi/3)$</u>	a	0°

注：表中加下划线的项表示结果是一个确定的值而不是一段范围。

综上，空间关系的定性表示与推理为我们解决遮挡问题提供了新的思路和新的方法。与定量的方法相比，采用定性的方法大大降低了时间复杂度和空间复杂度^[108]。我们对空间关系的定性表示与推理作了一定基础性的研究，下一步的工作就是要把已有的成果完善，并应用到遮挡问题中去。

2.3 小结

本章主要论述了双目立体成像相关技术及理论。首先,论述了双目立体成像相关技术,着重介绍了双目立体成像系统,包括双目立体成像系统的发展现状、分时双目立体成像系统以及立体图像文件类型,等等。

当前流行的双目立体成像系统主要由显示芯片、驱动程序和观察显示设备三部分组成。其中观察显示设备有多种,不同的观察显示设备采用的立体图像对显示方式也不同。目前常用的观察显示设备有双色眼镜、液晶光阀眼镜(LCD shutter glasses)、头盔显示器、自由立体(Autostereo)显示器等。其中,液晶光阀眼镜的价格适中,并可用于观看彩色图像,因此在游戏娱乐、虚拟现实等领域里得到了广泛的应用。本文实验采用的观察设备就是液晶光阀眼镜。

分时双目立体成像是模拟人眼观察物体的方式,将左右两幅视图交替显示来产生立体图像的一种常用的系统。

以液晶光阀眼镜作为观察设备的分时双目立体成像系统的工作原理是:计算机将左眼视图和右眼视图交替迅速地显示在监视器上。当左眼视图显示时,液晶光阀眼镜使右眼的液晶镜片遮断,使得只有左眼才能看到监视器;同理,当右眼视图显示时,液晶光阀眼镜使左眼的液晶镜片遮断,使得只有右眼才能看到监视器,这样便产生了立体感。

立体图像文件是双目立体成像系统的重要组成部分。常用的立体图像文件类型有 JPS(.jps)文件等类型。JPS(JPEG Stereo)文件格式的立体图像是由同一空间场景中两幅同为 JPG 格式的左右眼视图拼凑成的,因此 JPS 格式与 JPG 格式没有本质上的区别。

其次,本章论述了双目立体成像相关理论。

经验表明,在进行双目立体成像时,虽然人们感兴趣的像点的深度不同于屏幕,但是仍要求我们聚焦于屏幕。Valyus 发现大多数人能容忍的会聚角变化值为 1.6° 。Valyus 的理论限制了双目立体成像的景深,我们不可能将虚像成像在任意的位置。

Panum 发现,当人眼的双眼注视一个外界物体时,另外一个远于或近于双眼单视界的物体,虽然没有刺激到左右眼视网膜的对应点,但只要刺激在两个视网膜对应点附近的一定范围内(Panum 融合区),也可以在大脑中产生单一的立体视觉。Panum 融合区理论告诉我们并不是所有在两个视网膜上形成含有双目视差的图像都能在大脑中形成单一的立体视觉。一种成功的双目立体成像系统展示给观察者的立体图像,其水平视差的取值范围不能超出舒适 Panum 融合区的限制。

侯等人通过实验发现,当生理深度暗示与心理深度暗示一致时,综合的深度感觉增强;而当生理深度暗示与心理深度暗示矛盾时,生理深度暗示会被心理深

度暗示抑制减弱。

针对遮挡问题,本文提出利用空间关系的定性表示与推理来表示物体之间的位置关系以及让计算机理解这种位置关系的方法。这种定性的方法能大大节省计算量。在空间关系的定性表示方面,本文给出了一个 *SR* 表示模型,提出了距离关系的定性表示模型和方向关系的定性表示模型,以方便进行定性三角推理。在空间关系的定性推理方面,本文定义了基于 *SR* 表示模型的组合运算,给出了相应组合运算表;构造了定性三角推理所需要的组合运算表、距离定性减法表和角度定性减法表。

上述双目立体成像相关技术和理论,构成了本论文的研究基础。在此基础上,我们给出了一个基于监视器的双目立体成像模型,并在第三章详细地研究这个模型。

3 双目立体成像模型

3.1 基于监视器的双目立体成像模型

利用双目立体视觉的基本原理, 我们可以通过一定的观察显示设备让观察者看到立体图像。有不少学者针对具体的应用讨论了各种成像模型及其涉及的关键问题^[109-114]。例如, 在检验医学领域, Jiang, H. 等人给出了一个基于CCD的单翼数字立体X射线成像系统模型, 基于概率理论推导了该模型立体成像平均值和定位误差的标准差^[115]。在虚拟植物研究领域, 虚拟植物软件系统常采用标准的计算机监视器(又称为fishtank^[33])作为显示设备, 因此有必要讨论基于监视器的双目立体成像模型。

本文给出的基于监视器的双目立体成像模型如图3.1所示。图中采用的坐标系是世界坐标系^[11, 116]。其中, α (监视器屏幕)是平行于 XOY 面的投影平面; β 是平行于 XOY 面的像平面, W 为 β 中的一点。设 $A_1(h, 0, 0)$ 、 $A_2(-h, 0, 0)$ ($h>0$)分别是右眼和左眼的坐标, 因此原点处在中央眼^[11]的位置, XOY 面即为观察者双眼所在平面。设像平面与 XOY 面的距离为 o_d ($o_d>0$), 投影平面与 XOY 面的距离为 p_d ($p_d>0$)。 $MO'N$ 为监视器屏幕的图像坐标系^[11, 117], $O'M//OX$, $O'N//OY$, O' 在坐标系 XYZ 中的坐标为 (x_0, y_0, p_d) 。

通过 A_1 和 A_2 观看 W , 我们可以分别得到 W 在 α 内的立体图像对^[16, 69] I_1 和 I_2 (图3.1所示为正视差^[16]的情形)。当 A_1 和 A_2 分别被放置成两台相同摄像机的像平面时, 便是很多文献讨论过的立体摄像机几何模型^[12, 14, 16, 114, 116-118]。

这些文献主要研究图像的获取和深度信息的确定, 没有深入讨论成像性质和立体效果。然而在使用监视器作为视觉显示设备时, 观察者会发现当观察点发生变化时, 图像的大小和成像的位置将会发生变化。这将对虚拟植物软件系统生成与实际植物大小和尺寸一样的立体图像产生影响。因此, 有必要研究图3.1所示的模型的成像性质和立体效果。

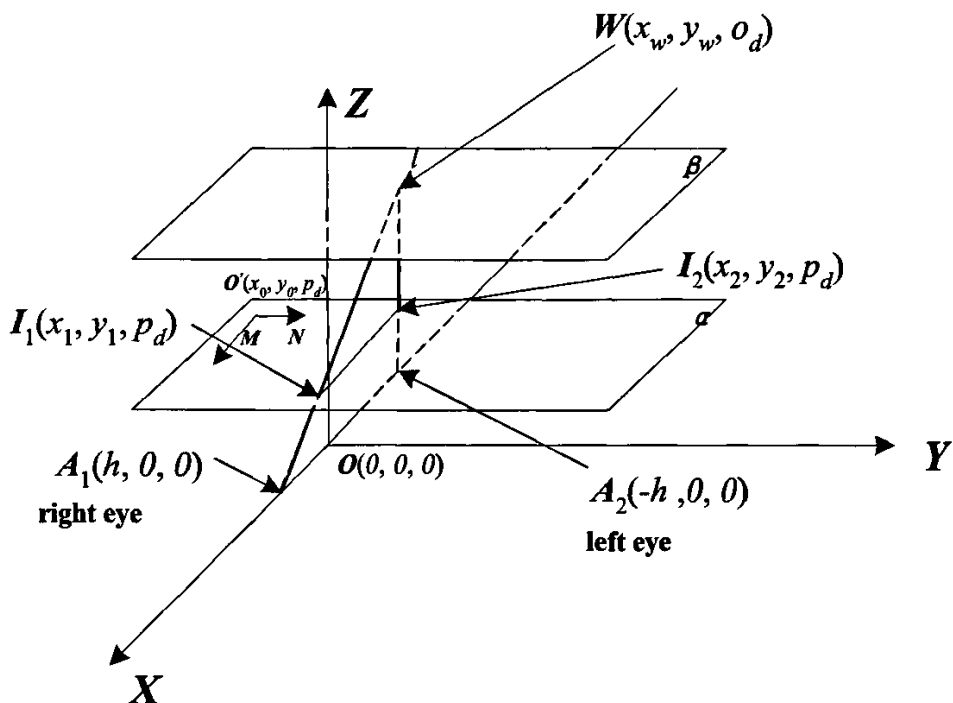


图 3.1 基于监视器的双目立体成像模型

Fig 3.1 Binocular stereo imaging model based on monitor

3.2 双目立体成像模型的成像性质

3.2.1 立体图像对坐标的计算

如图 3.1 所示, 设人的左、右眼分别处在 A_2 、 A_1 位置。通过 A_1 和 A_2 观看 W , 分别得到 W 在 α 内的立体图像对 I_1 和 I_2 , 根据透视投影原理计算可得 I_1 、 I_2 和 W 的坐标满足如下关系:

公式(3.1)

$$I_1\left(\frac{(x_w - h) \times p_d}{o_d} + h, \frac{y_w \times p_d}{o_d}, p_d\right), I_2\left(\frac{(x_w + h) \times p_d}{o_d} - h, \frac{y_w \times p_d}{o_d}, p_d\right) \quad (3.1)$$

公式(3.1)虽然是由正视差的情形推理得到的, 但是可以证明它也适用于负视差的情形。

根据公式(3.1)和坐标轴的平移公式(这里为讨论方便, 设 $MO'N$ 和 XYZ 两坐标系的单位相同), 可得 I_1 和 I_2 在监视器屏幕上的坐标分别为:

公式(3.2)

$$I_1'(\frac{(x_w-h) \times p_d}{o_d} + h - x_0, \frac{y_w \times p_d}{o_d} - y_0), I_2'(\frac{(x_w+h) \times p_d}{o_d} - h - x_0, \frac{y_w \times p_d}{o_d} - y_0) \quad (3.2)$$

公式(3.2)实现了从世界坐标系到监视器屏幕的图像坐标系的变换。 W 的深度信息经过变换隐含在立体图像对 I_1' 和 I_2' 的位置关系之中, 具体说就是水平视差 D 之中^[11, 16, 41, 42]。因此 I_1' 和 I_2' 没有竖坐标也能表达深度信息。

3.2.2 视差及其性质

由公式(3.2)可得:

性质 1

$$y_1' = y_2' = \frac{y_w \times p_d}{o_d} - y_0$$

设 α 为监视器屏幕。性质 1 表明, 在图 3.1 的模型中如果用计算机来生成立体图像对 I_1' 和 I_2' , 则应使 I_1' 和 I_2' 处在屏幕的同一水平位置, 才能保证不会在图像中人为地产生垂直视差^[41, 42, 119]。

性质 2

水平视差

$$D = x_1 - x_2 = x_1' - x_2' = 2h \frac{o_d - p_d}{o_d}$$

性质 2 中的 $2h$ 为瞳距, 一般为 65mm。性质 2 表明, 水平视差只与 p_d 和 o_d 相关, 只要平面 α 、 β 一定, 则 β 中任一点所对应的水平视差都相等。今后除非特别声明, 视差是指水平视差。

由性质 2 可知:

$$\textcircled{1} \lim_{o_d \rightarrow +\infty} D = 2h$$

这表明对于远处的物体 W , 水平视差趋于定值^[42]。因此, 双目立体视觉模型对近距离(5m 以内)的物体特别有效^[33], 对于远处的物体双眼难以分辨出它们之间的深度信息, 因而立体感较弱。

② 设 α 、 β 固定, 则 $o_d - p_d$ 为定值 d_p 。当观察者远离屏幕时, 有

$$\lim_{p_d \rightarrow +\infty} D = \frac{2hd_p}{p_d + d_p} = 0$$

如图 3.2 所示, 设平面 $A_1A_2W \cap \beta = l$ (显然 $W \in l$)。当观察者双眼沿着 X 轴平移到 A'_1 、 A'_2 , 再观看计算机已生成的立体图像对 I_1 和 I_2 , 会看到怎样的像点立体成像现象呢? 为了研究这个问题, 先要证明下面的性质 3。

$$I // I_1 I_2 // A_1 A_2$$

设 $W' \in l$, 在平面 A_1A_2W' 内分别作直线 $W'I_1$ 和 $W'I_2$, 与 X 轴分别相交于 A'_1 和 A'_2 , 且:

对于图 3.2 中的 A_2, A_1 , 有 $A_2, A_1 = 2h$ 。

证明：如图 3.2 所示，过 W 作 $WP \perp OX$ 交 OX 于 P ，交 I_1I_2 于 Q ；过 W' 作 $W'R \perp OX$ 交 OX 于 R ，交 I_1I_2 于 S 。 \because 由性质 3 有 $I_1I_2 \parallel A_1A_2$ ， \therefore 根据平行线分线段成比例定理，有 $\frac{WQ}{WP} = \frac{W'S}{W'R}$ 。又 $\because I_1I_2 \parallel A_1A_2$ ， $\therefore \frac{WQ}{WP} = \frac{I_2I_1}{A_2A_1}$ ， $\frac{W'S}{W'R} = \frac{I_2I_1}{A_2'A_1'}$ 。 $\therefore \frac{I_2I_1}{A_2'A_1'} = \frac{W'S}{W'R} = \frac{WQ}{WP} = \frac{I_2I_1}{A_2A_1}$ ， $\therefore A_2'A_1' = A_2A_1$ 。 $\because A_1A_2 = 2h$ ， $\therefore A_2'A_1' = 2h$ 。

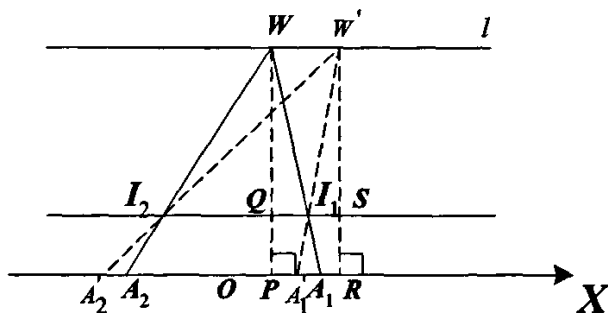


Fig 3.2 Lines and their relations in plane A, A, W

由性质 4 可知, $A_2'A_1'$ 等于瞳距。因此, 可以反过来考虑: 当双眼沿着 X 轴平移到 A_1' 、 A_2' , 再观看计算机已生成的立体图像对 I_1 和 I_2 时, W 移至 W' 。这表明, 双眼沿着 X 轴平移时 W 始终只在 l 上朝反方向移动, 即所看到的像点的深度不会发生改变。

由图 3.2 还可以直观地看出, 如果双眼沿着 X 轴朝一个方向向无穷远处平移, W 会向反方向移动, 直至看不见。

类似于性质 4 的证明, 我们可以证明图 3.3 中更一般的情况。

性质 5

如图 3.3 所示, 设 $W' \in \beta$, 作直线 $W'I_1$ 和 $W'I_2$, 与 XOY 面分别相交于 A_1' 和 A_2' , 则有 $A_2'A_1' = 2h$ 。

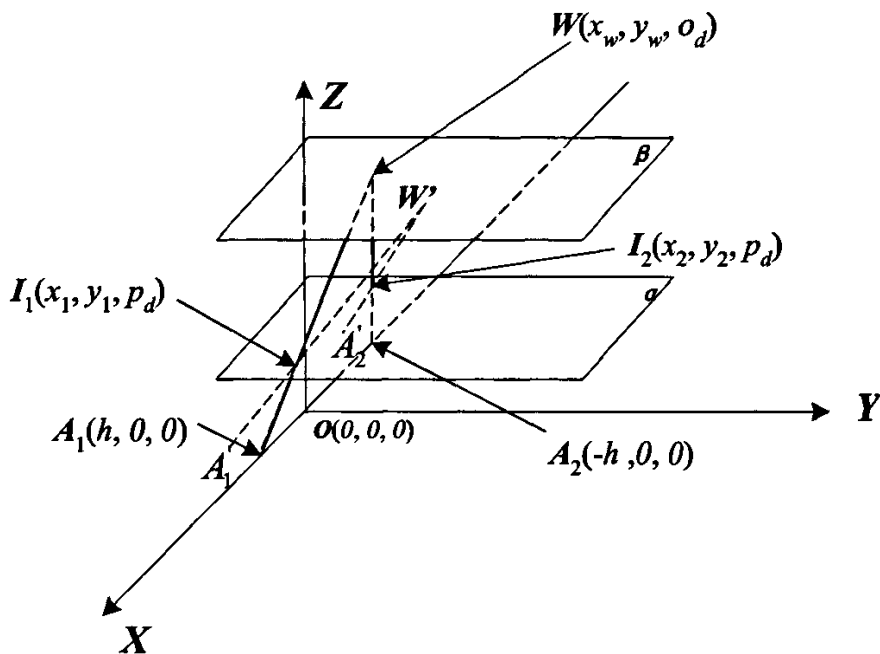


图 3.3 平面 XOY 内双眼平移的一般情形

Fig 3.3 A general case when eyes moving in parallel in the plane XOY

性质 5 表明, 观察者双眼在平面 XOY 内平行于 X 轴平移时, 所看到的像点的深度不会发生改变。也就是说, 只要观察者保持与屏幕的距离不变, 他左右上下摆头不会觉得像点和他的距离有变化; 但是感觉看到的图像在平移。

3.2.4 直线段立体成像的定量分析

虚拟植物的立体图像通常都是复杂的图像。不过复杂图像可以看成由许多条直线段组成。因此,我们以立体直线段为例考察在图 3.1 所示的模型中,当双眼位置变化时直线段立体成像的变化规律。

由性质 2 可得: $p_d = \frac{o_d(2h-D)}{2h}$ 。此表达式说明如果保持 D 不变,则 p_d 增大时 o_d 也会增大。这表明正视差的情况下,在观看计算机已生成的立体图像对 I_1 和 I_2 时,如果观察者保持双眼所在直线平行于 X 轴,沿着 Z 轴远离或靠近屏幕,他会感觉到 W 沿相反方向远离或靠近屏幕;而负视差情况下,他会感觉到 W 沿相同方向远离或靠近屏幕。

更深入的定量计算如图 3.4 所示(为了方便做实验,我们讨论负视差的情形,并且使直线段的位置有一定的特殊性)。

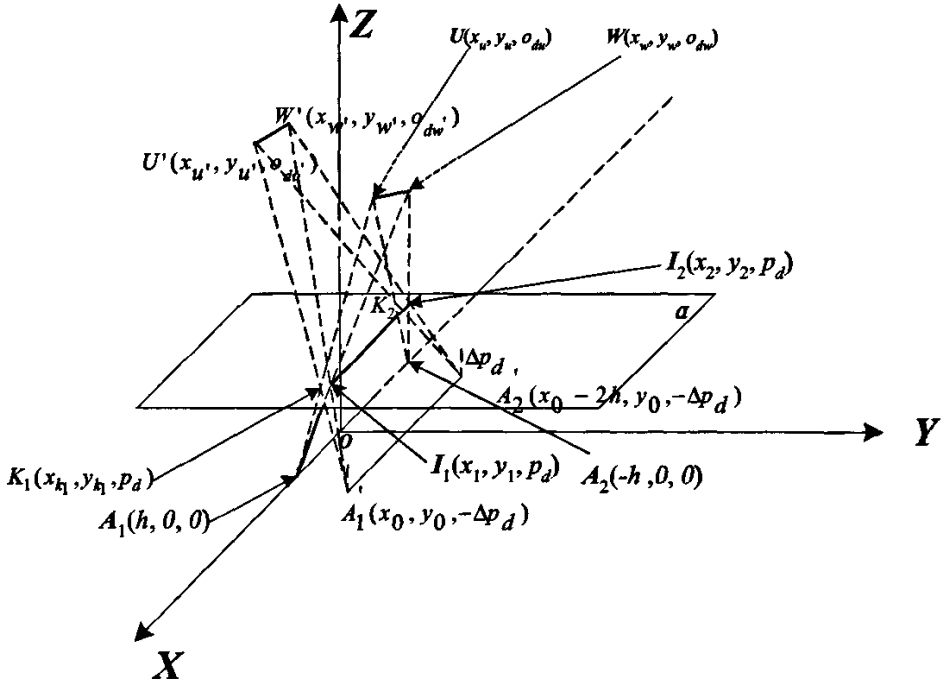


图 3.4 直线段立体成像(水平视差 D 保持不变)

Fig 3.4 Stereo imaging of line segment (horizontal parallax D is a constant)

如图 3.4 所示,设观察者通过 A_1 和 A_2 观看 U , 分别得到 U 在 α 内的立体图像对 K_1 和 K_2 。根据公式(3.1)可得 K_1 和 K_2 的坐标分别为

$$K_1\left(\frac{(x_u - h) \times p_d}{o_{du}} + h, \frac{y_u \times p_d}{o_{du}}, p_d\right), K_2\left(\frac{(x_u + h) \times p_d}{o_{du}} - h, \frac{y_u \times p_d}{o_{du}}, p_d\right) \quad (3.3)$$

将 A_1A_2 平移至 $A'_1A'_2$ 。设 A'_1 到 α 的距离为 $p_d + \Delta p_d$ ($p_d + \Delta p_d > 0$)，因此 A'_1 的坐标为 $(x_0, y_0, -\Delta p_d)$ 。∵ $A'_1A'_2 \parallel A_1A_2$ ， $A'_2A'_1 = A_2A_1 = 2h$ ，∴ A'_2 的坐标为 $(x_0 - 2h, y_0, -\Delta p_d)$ 。此时再通过 A'_1 和 A'_2 观看立体图像对 I_1 和 I_2 时， W 移至 W' ；观看 K_1 和 K_2 时， U 移至 U' 。直线 A'_1W' 的方向向量可取为 $s = \overrightarrow{A'_1I_1} = \{x_1 - x_0, y_1 - y_0, p_d + \Delta p_d\}$ ，故直线 A'_1W' 的对称式方程为

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z + \Delta p_d}{p_d + \Delta p_d} \quad (3.4)$$

又∵ W' 和 W 对应的水平视差都为 I_2I_1 ，∴ 由性质 2 得

$$2h \frac{o_{dw} - p_d}{o_{dw}} = I_2I_1 = 2h \frac{(o_{dw} + \Delta p_d) - (p_d + \Delta p_d)}{o_{dw} + \Delta p_d} \quad (3.5)$$

其中 o_{dw} 为 W' 的竖坐标。

由公式(3.5)解出 o_{dw} 得

$$o_{dw} = \frac{o_{dw}(p_d + \Delta p_d)}{p_d} - \Delta p_d \quad (3.6)$$

将公式(3.6)代入公式(3.4)得 W' 的坐标为

$$\left(\frac{o_{dw}}{p_d}(x_1 - x_0) + x_0, \frac{o_{dw}}{p_d}(y_1 - y_0) + y_0, o_{dw}\right) \quad (3.7)$$

其中根据公式(3.1)有

$$\begin{cases} x_1 = \frac{(x_w - h) \times p_d}{o_{dw}} + h, \\ y_1 = \frac{y_w \times p_d}{o_{dw}}. \end{cases}$$

同理可得 U' 的坐标为

$$\left(\frac{o_{du}}{p_d}(x_{k_1} - x_0) + x_0, \frac{o_{du}}{p_d}(y_{k_1} - y_0) + y_0, o_{du} \right) \quad (3.8)$$

其中

$$o_{du} = \frac{o_{du}(p_d + \Delta p_d)}{p_d} - \Delta p_d$$

并且根据公式(3.3)有

$$\begin{cases} x_{k_1} = \frac{(x_u - h) \times p_d}{o_{du}} + h, \\ y_{k_1} = \frac{y_u \times p_d}{o_{du}}. \end{cases}$$

由公式(3.7)和公式(3.8)可知, $W'U'$ 满足

$$\begin{cases} x_{u'} - x_{w'} = \frac{1}{p_d} [o_{du}(x_{k_1} - x_0) - o_{dw}(x_1 - x_0)] = x_u - x_w + \frac{o_{du} - o_{dw}}{p_d} (h - x_0), \\ y_{u'} - y_{w'} = \frac{1}{p_d} [o_{du}(y_{k_1} - y_0) - o_{dw}(y_1 - y_0)] = y_u - y_w - \frac{o_{du} - o_{dw}}{p_d} y_0, \\ o_{du'} - o_{dw'} = \frac{(p_d + \Delta p_d)}{p_d} (o_{du} - o_{dw}) \end{cases} \quad (3.9)$$

因此, 像和原像的拉伸比例为

$$\frac{|W'U'|}{|WU|} = \frac{\sqrt{(x_{u'} - x_{w'})^2 + (y_{u'} - y_{w'})^2 + (o_{du'} - o_{dw'})^2}}{\sqrt{(x_u - x_w)^2 + (y_u - y_w)^2 + (o_{du} - o_{dw})^2}} \quad (3.10)$$

现在来计算 \overrightarrow{WU} 和 $\overrightarrow{W'U'}$ 之间的夹角。

\overrightarrow{WU} 的方向向量可取为 $s_1 = \{m_1, n_1, p_1\} = \{x_u - x_w, y_u - y_w, o_{du} - o_{dw}\}$, $\overrightarrow{W'U'}$ 的方向向量可取为 $s_2 = \{m_2, n_2, p_2\} = \{x_{u'} - x_{w'}, y_{u'} - y_{w'}, o_{du'} - o_{dw'}\}$, 则 s_1 与 s_2 的夹角 θ 满足

$$\cos \theta = \frac{(x_u - x_w)(x_{u'} - x_{w'}) + (y_u - y_w)(y_{u'} - y_{w'}) + (o_{du} - o_{dw})(o_{du'} - o_{dw'})}{\sqrt{(x_u - x_w)^2 + (y_u - y_w)^2 + (o_{du} - o_{dw})^2} \sqrt{(x_{u'} - x_{w'})^2 + (y_{u'} - y_{w'})^2 + (o_{du'} - o_{dw'})^2}} \quad (3.11)$$

讨论:

情形 1: 线段平行于眼睛所在的观察平面, 且观察者保持头部无偏转平行移动。

设线段 WU 平行于 XOY 面, 即 $o_{dw} = o_{du}$, 则公式(3.9)可化简为

$$\begin{cases} x_{u'} - x_{w'} = x_u - x_w, \\ y_{u'} - y_{w'} = y_u - y_w, \\ o_{du'} = o_{dw'}. \end{cases} \quad (3.12)$$

将公式(3.12)代入公式(3.10), 可得 $|W'U'| = |WU|$; 代入公式(3.11), 可得 $\theta = 0$, 即 $\overrightarrow{WU} \parallel \overrightarrow{W'U'}$ 。由此可知, 如果 W 和 U 在 Z 轴上的坐标相等, 当观察者保持双眼所在直线平行于 X 轴, 远离(靠近)屏幕时, 他看到的图像虽然会产生移动, 但是在平行于 X 轴方向不会产生形变。

如果线段 WU 垂直于 YOZ 面, 或者垂直于 XOZ 面, 则可归结为 WU 平行于 XOY 面的情况。

情形 2: 线段垂直于观察平面, 且观察者保持头部无偏转前后移动。

设线段 WU 垂直于 XOY 面, 即 $x_u = x_w$, $y_u = y_w$, 则公式(3.9)可化简为

$$\begin{cases} x_{u'} - x_{w'} = \frac{o_{du} - o_{dw}}{p_d} (h - x_0), \\ y_{u'} - y_{w'} = -\frac{o_{du} - o_{dw}}{p_d} y_0, \\ o_{du'} - o_{dw'} = \frac{(p_d + \Delta p_d)}{p_d} (o_{du} - o_{dw}). \end{cases} \quad (3.13)$$

将公式(3.13)代入公式(3.10), 得

$$\frac{|W'U'|}{|WU|} = \frac{\sqrt{(h - x_0)^2 + y_0^2 + (p_d + \Delta p_d)^2}}{p_d}$$

将公式(3.13)代入公式(3.11), 得

$$\cos \theta = \frac{p_d + \Delta p_d}{\sqrt{(h - x_0)^2 + y_0^2 + (p_d + \Delta p_d)^2}}$$

进一步讨论特殊情况, 当 $x_0 = h$ 且 $y_0 = 0$ 时, 即 A_1A_2 在平移时保持 X 轴和 Y 轴坐标不变, 有

$$\begin{cases} \frac{|W'U'|}{|WU|} = \frac{p_d + \Delta p_d}{p_d} \\ \cos \theta = 1 \end{cases}$$

由此可知, 若 WU 垂直于 XOY 面, 双眼沿 Z 轴方向远离(或靠近)屏幕并保持其 X 轴和 Y 轴坐标不变时, 他看到的图像与原图像平行, 长度会发生变化, 拉伸比例为 $\frac{p_d + \Delta p_d}{p_d}$ 。当 Δp_d 为负, 即双目靠近屏幕时, 图像以比例 $\frac{p_d + \Delta p_d}{p_d}$ 缩小;

当 Δp_d 为正, 即双目远离屏幕时, 图像以比例 $\frac{p_d + \Delta p_d}{p_d}$ 放大。

情形 3: 线段平行于 XOZ 面, 且观察者保持头部无偏转前后移动。

设线段 WU 平行于 XOZ 面, 即 $y_u = y_w$ 。考虑特殊情况: 当 $h = x_0$, $y_0 = 0$ 时(即 A_1A_2 在平移时保持 X 轴和 Y 轴坐标不变), 则公式(3.9)可化简为

$$\begin{cases} x_{u'} - x_{w'} = x_u - x_w, \\ y_{u'} - y_{w'} = 0, \\ o_{du'} - o_{dw'} = \frac{(p_d + \Delta p_d)}{p_d} (o_{du} - o_{dw}) \end{cases} \quad (3.14)$$

将公式(3.14)代入公式(3.10), 得

$$\frac{|W'U'|}{|WU|} = \sqrt{1 + \frac{\frac{2p_d\Delta p_d + \Delta p_d^2}{p_d^2}}{\frac{(x_u - x_w)^2}{(o_{du} - o_{dw})^2} + 1}}$$

将公式(3.14)代入公式(3.11), 得

$$\cos\theta = \frac{\frac{(x_u - x_w)^2}{(o_{du} - o_{dw})^2} + \frac{(p_d + \Delta p_d)}{p_d}}{\sqrt{\frac{(x_u - x_w)^2}{(o_{du} - o_{dw})^2} + 1} \sqrt{\frac{(x_u - x_w)^2}{(o_{du} - o_{dw})^2} + \frac{(p_d + \Delta p_d)^2}{p_d^2}}}$$

令 $\frac{x_u - x_w}{o_{du} - o_{dw}} = d$ (d 为任意常数), 可将上两式化简为

$$\begin{aligned} \frac{|W'U'|}{|WU|} &= \sqrt{1 + \frac{\frac{2p_d\Delta p_d + \Delta p_d^2}{p_d^2}}{d^2 + 1}} \\ \cos\theta &= \frac{d^2 + \frac{(p_d + \Delta p_d)}{p_d}}{\sqrt{d^2 + 1} \sqrt{d^2 + \frac{(p_d + \Delta p_d)^2}{p_d^2}}} \end{aligned}$$

由此可以看出, 当 d 一定时, $\frac{|W'U'|}{|WU|}$ 和 $\cos\theta$ 只与 p_d 、 Δp_d 有关, 也就是说, 当线段 WU 在 X 轴上的投影与在 Z 轴上投影的比值为一定值时, 拉伸比例和夹角只与眼睛离屏幕的距离(p_d)和移动的改变量(Δp_d)有关, 而与 WU 本身的位置、长度无关。

情形 4: 线段平行于 YOZ 面, 且观察者保持头部无偏转前后移动。

设线段 WU 平行于 YOZ 面, 即 $x_u = x_w$, 考虑特殊情况: $h = x_0$, $y_0 = 0$ 时(即 A_1A_2 在平移时保持 X 轴和 Y 轴坐标不变), 则公式(3.9)可化简为

$$\begin{cases} x_{u'} - x_{w'} = 0, \\ y_{u'} - y_{w'} = y_u - y_w \\ o_{du'} - o_{dw'} = \frac{(p_d + \Delta p_d)}{p_d} (o_{du} - o_{dw}) \end{cases} \quad (3.15)$$

比较公式(3.14)与公式(3.15), 不难看出, 情形 4 与情形 3 完全类似。按同样的方法分析, 得到的结论为: 若线段平行于 YOZ 面, A_1A_2 在平移时保持 X 轴和 Y 轴坐标不变, 则当线段 WU 在 Y 轴上的投影与在 Z 轴上投影的比值为一定值时, 拉伸比例和夹角只与眼睛离屏幕的距离(p_d)和移动的改变量(Δp_d)有关, 而与 WU 本身的位置、长度无关。

3.3 实验

从上面的理论计算可知,对于情形 1,只要观察者保持双眼平行地远离或靠近监视器屏幕,尽管图像会产生平移,但是他看到的各点深度都相同的图像的大小和角度都不会变。这样,对于由深度相同的点所构成的立体图像,我们可以利用公式(3.2)生成我们需要的尺寸的图像。对于特定参数(特别是随机参数)下经过复杂的演算过程生成的立体图像,我们也可以通过虚拟对象进行实际测量的方法来确定对象在某一方向上的尺寸。为了验证上述结论并验证测量的可操作性,我们设计了下面的实验。

实验采用了分时双目立体成像系统。其中计算机的显卡为 NVIDIA 系列显卡,芯片类型为 GeForce4 MX 440 with AGP8X。当 CRT 显示器分辨率达到 800*600 时刷新率可达 100Hz^[36]。系统的组成及其连接方式如图 2.2 所示。立体图像可以利用公式(3.2)编写程序通过计算机绘制得到。图 3.5 是按这种方式生成的立体直线段 WU (为了让观察者更容易的分辨正负视差以便于实际测量,我们在直线段 WU 的基础上补画了一个立方体)。它是采用如下参数绘制的: $2h = 65\text{mm}$, $p_d = 40\text{cm}$, $1\text{cm} = 28.35$ 像素(监视器的有效显示区域约为 $28.2\text{cm} \times 21.2\text{cm}$), O 在坐标系 XYZ 中的坐标为 $(-14.11\text{cm}, -10.58\text{cm}, 40\text{cm})$ 。图 3.6 说明了图 3.5 中的图像是按照怎样的坐标绘制的。显然,这是负视差的情形。绘制完成后,保存为 JPS 立体图像格式,然后通过 JPS 浏览器配合立体液晶光阀眼镜进行观察,便可看到飘浮在监视器和眼睛之间的一个立方体和一条直线段。

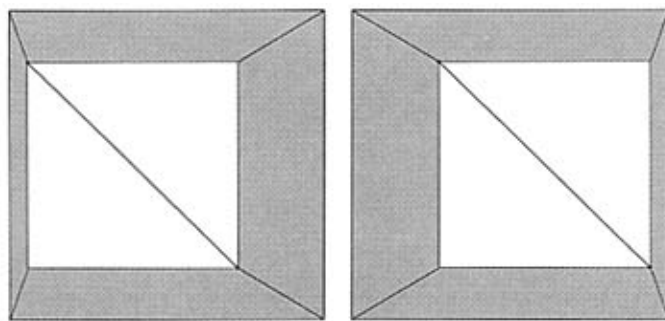


图 3.5 根据公式(3.2)由计算机程序绘制的立体图像(左/右型)

Fig 3.5 A stereo image drawn by computer program according to Formula (3.2) (left/right)

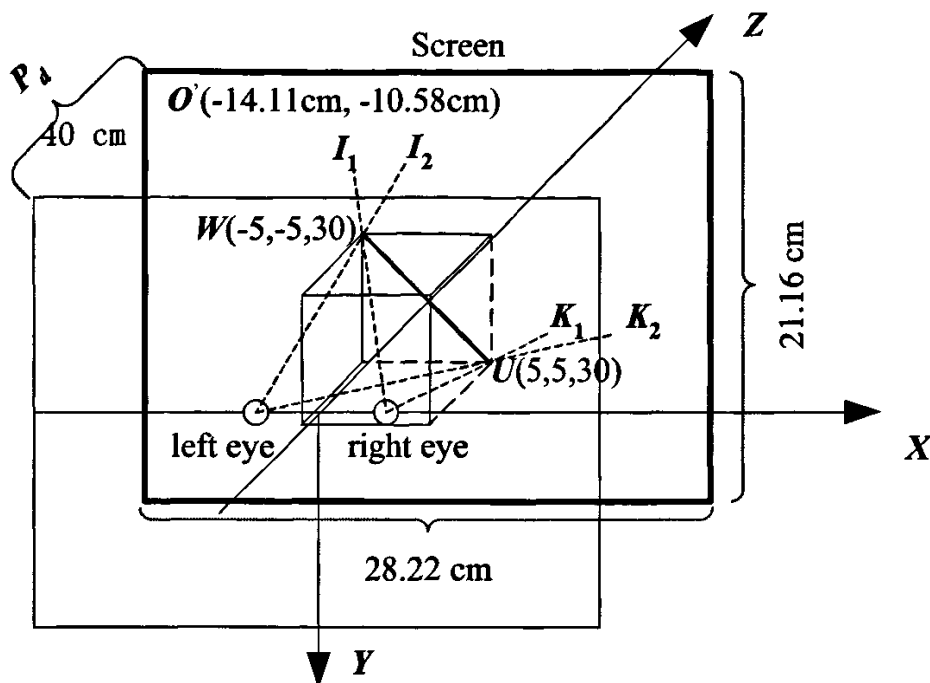


图 3.6 图 3.5 中生成的立体图像与坐标系之间的关系

Fig 3.6 Relation between the coordinate systems and the stereo image drawn in Fig 3.5

如果按图 3.6 所示的坐标绘制, 观察者应该看到立体直线段位于监视器和眼睛的中间位置, 距离监视器屏幕 10cm。由于是负视差, 我们可以使用光具座和尺实际测量出直线与监视器的实际距离, 实验装置(部分)如图 3.7 所示, 实验主要步骤如下:

- ① 通过光具座上的望远镜调整监视器的位置使监视器屏幕与光具座保持垂直;
- ② 利用游标确定观察平面的位置, 将立体液晶光阀眼镜固定在光具座游标处;
- ③ 观察者通过立体液晶光阀眼镜观察监视器屏幕, 慢慢移动另一游标到虚像——直线段 WU 处;
- ④ 记下游标读数, 算出直线段 WU 和屏幕的距离。

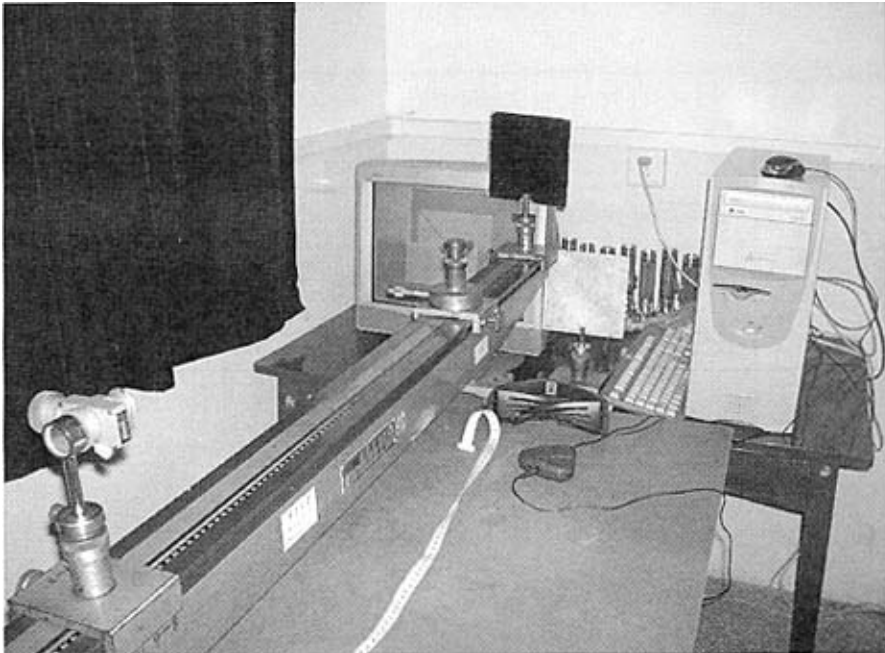


图 3.7 实验装置图。其中立体液晶光阀眼镜距监视器屏幕 40cm，
双眼处于屏幕中央位置的正前方。

Fig 3.7 Apparatus used in the experiment. The distance between LCD shutter glasses and the screen is 40cm, and the eyes are kept in front of the center of the screen during the experiment.

实验中，NVIDIA 3D 立体驱动的基本立体参数^[38]设置如表 3.1 所示，测量结果参见表 3.2。

从表 3.2 可以看出，实际测量得到的距离值与理论值是比较接近的。实验的误差主要来自眼睛与眼镜镜片之间的距离，这一误差目前实验条件下不容易减小。类似于上述方法，我们将双眼移到距监视器屏幕 60cm 的位置，再通过立体液晶光阀眼镜观察图 3.5 所示的立体图像，测量结果如表 3.3 所示，可以看出直线段 WU 到监视器屏幕的距离变为 14.8cm，即 WU 远离了屏幕。参与实验的观察者未感到直线段有明显形变。这些与 3.2.4 中分析得到的结论相符。

表 3.1 NVIDIA 3D 立体驱动的基本立体参数

Table 3.1 Basic stereo parameters of the NVIDIA 3D stereo driver

参数	设置值
Monitor Size	19.0"
Bit Depth	16 bits per pixel
Resolution	800*600
Stereo Refresh Rate	100Hz
Stereo Gamma	1.00
Stereo Separation	20%

表 3.2 立体直线段 WU 到监视器屏幕的距离($p_d=40\text{cm}$)Table 3.2 Distance between the line segment and the screen ($p_d=40\text{cm}$)

实验次数	距离读数(cm)	平均值(cm)	理论值(cm)
1	9.8	10.0	10.0
2	10.0		
3	9.9		
4	10.2		
5	10.3		
6	10.3		
7	10.4		
8	9.9		
9	9.8		
10	9.8		

注：测量重复了 10 次。每次距离读数都是由不同观察者估读得到的。

表 3.3 立体直线段 WU 到监视器屏幕的距离($p_d=60\text{cm}$)Table 3.3 Distance between line segment and the screen ($p_d=60\text{cm}$)

实验次数	距离读数(cm)	平均值(cm)	理论值(cm)
1	14.7	14.8	15.0
2	14.7		
3	15.2		
4	14.9		
5	15.1		
6	14.6		
7	14.7		
8	15.0		
9	14.7		
10	14.9		

注：测量重复了 10 次。每次距离读数都是由不同观察者估读得到的。

类似的，也可以设计实验验证情形 2、情形 3 和情形 4。只是这些情形较为复杂。

上述实验证明，公式(3.1)和公式(3.2)是正确的，实验中的虚拟对象是可以测量的。图 3.8 是根据公式(3.1)和公式(3.2)利用 3DS MAX 生成的一株兰花苗的立体图像对。兰花主干的中心点被设计成离屏幕 10 cm，即中心点的水平视差为 60 像素。实际测量发现，中心点距离屏幕确为 10 cm 左右。显然，这种立体可视化展现技术，要比目前多数虚拟植物软件系统采用的二维平面可视化展现形象、逼真得多。

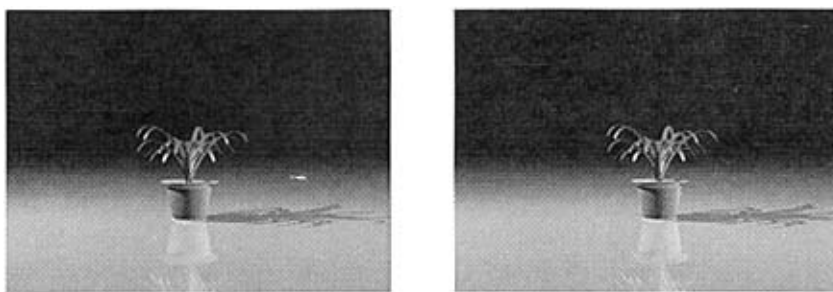


图 3.8 兰花的立体图像对(中心点距屏幕 10 cm)

Fig 3.8 Stereo pair of an orchid (the distance between the central point and the screen is 10 cm)

3.4 小结

立体成像是虚拟植物可视化立体展现的关键技术。然而较少有文献对基于监视器的双目立体成像的成像性质及立体效果进行深入研究。本章通过理论和实验研究了通过基于监视器的双目立体成像模型所能看到的立体效果，并重点对直线段立体成像进行了定量分析，得出了相关的成像规律。

当实验中的立体图像换成由虚拟植物软件系统生成的虚拟植物时，成像规律也是一样的。定量分析和实验表明，当观察者保持双眼平行地远离或靠近监视器屏幕，他看到的各点深度都相同的图像的不会产生形变；但是，各点深度不同的图像则会产生形变。这就是说，假如我们利用虚拟植物软件系统在特定的参数下生成了一颗秧苗的立体图像，我们可以通过上述实测的方法测量这颗秧苗在垂直方向上能长多高。不过，较难测量出它在其他方向上的尺寸。

在现实中，观察者的观察位置是随意的。有关各点深度不同的图像产生形变的规律，以及任意情形下通过模型看到的图像变化有什么规律还需要进一步考察。像与原像的定量关系也需要更深入研究。实验中我们也发现，虽然所有观察者看到的基本现象是一样的，但是每个人对立体视觉效果的感受还是有细微的差别，并且有时候还会对虚像的位置产生错觉。这种差别可能是由于视力、色觉等方面存在个体差异。有关这方面的研究还需要进一步深入。

4 基于侯原理的立体图像对的生成

4.1 侯原理简介

如绪论所述,在已有立体照片(立体图像对)时双目立体成像是容易实现的。但实际生活中,立体照片并不多。比如,人类在漫长的历史长河中积累了大量珍贵的绘画和摄影作品,这些作品本身是二维的平面图像。然而客观世界在空间上是三维的,如果能在普通的计算机 CRT 监视器上将这些平面图像转换成立体图像,则会给人们带来更加真实的视觉享受。要在监视器上实现平面图像立体化,就要用到计算机立体视觉技术^[120]。已有一些文献对相关的问题进行了探讨^[15, 17, 120],但由于视觉计算本身是病态的,因此在这一领域目前仍然没有完美的解决方法。本章主要研究一种新颖的利用心理深度暗示和生理深度暗示相互作用的原理生成立体图像对的方法——基于侯原理的立体图像对的生成方法。

侯等人指出,在每个图像单元中随机加入双目视差,在心理深度暗示和生理深度暗示相互作用下,会使观察者获得具有一定深度感知的立体图像^[15, 33]。我们把这样一种方法称为侯原理。根据侯原理的基本原理,我们将一幅图像划分成如图 4.1 所示的 $M \times N$ 个图像子块(M 是行数, N 是列数),然后使每一个子块在水平方向上随机地偏离它们原来所处的位置,这样得到的图像和原图像合一起就能构成立体图像对。

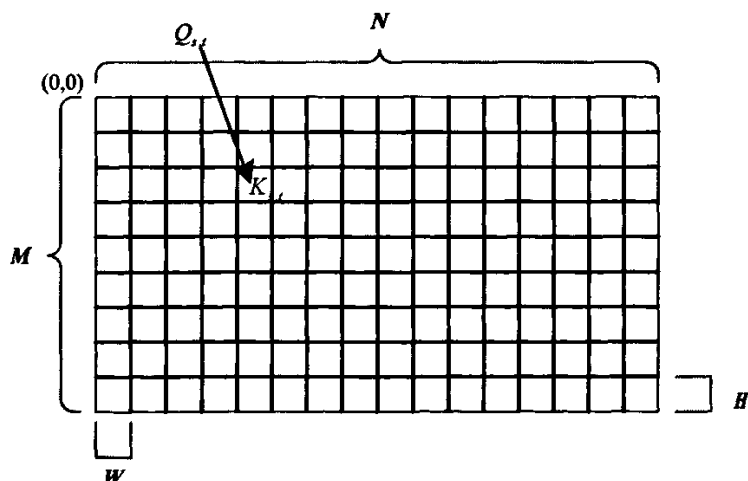


图 4.1 一幅图像被分割成 $M \times N$ 个图像子块,水平视差被随机地加入每一个子块当中

Fig 4.1 An image is segmented into $M \times N$ sub-blocks, horizontal parallax is added to each sub-block randomly

基于侯原理的平面图像立体化流程图如图 4.2 所示。

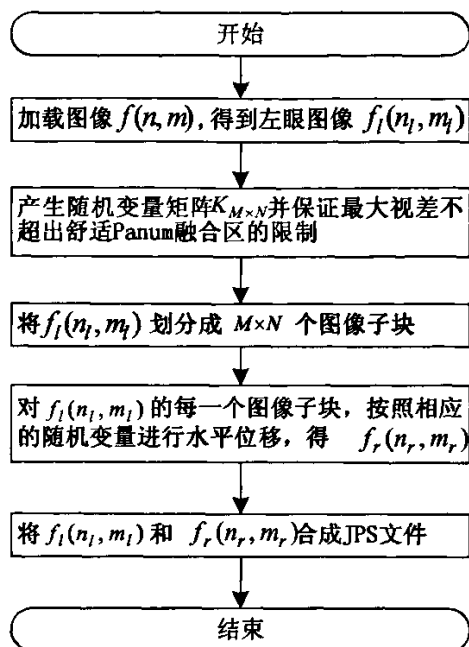


图 4.2 平面图像立体化流程图

Fig 4.2 Procedure of image conversion from planar to stereo

然而侯等人没有对平面图像立体化的立体效果给出定量的评价指标, 也没有深入讨论立体化的实现、立体化参数对转换后的立体效果的影响。本章利用侯原理在标准的计算机监视器上对普通的平面图像进行了立体化, 较为深入地讨论了这些问题。

4.2 立体化效果的评价指标

立体化效果是带有主观色彩的, 要想给出一个定量的评价指标是一件困难的事情。不过根据侯原理的特点, 我们认为下面这样一个评价方法是符合逻辑的、可行的: 在事先已有立体图像对的情况下, 根据其中一幅图像来立体化生成另一幅图像, 如果生成的图像和立体图像对中真实的图像差异越小, 这说明立体化的效果越好。

我们以下面的定量指标来评价转换效果的好坏(设立体相机拍摄的或根据双目立体成像的基本原理由软件生成的左、右眼视图分别为 $lImg$ 和 $rImg$ 。立体化算法对 $lImg$ 进行立体化, 得到右视图 $lImg'$, 大小与 $rImg$ 相同):

① 交叉熵(Cross-entropy)

定义 4.1: 设 X 是具有有限 n 个状态值的随机变量, $p_i = P\{X = x_i\}; i = 0, 1, \dots, n-1$ 。在图像分析领域, p_i 常常是对灰度值 i 的直方图统计, n 表示灰度级数。交叉熵 $I_{[P,Q]}$ 用来度量两种概率分布 $P = \{p_0, p_1, \dots, p_{n-1}\}$ 和 $Q = \{q_0, q_1, \dots, q_{n-1}\}$ 之间的信息量的差异:

$$I_{[P,Q]} = \sum_{i=0}^{n-1} p_i \ln \frac{p_i}{q_i} (q_i \neq 0) \quad (4.1)$$

交叉熵是评价两幅图像差别的关键指标, 它直接反映了两幅图像像素的信息量差异。

从公式(4.1)可以看出, 由于 p_i 是对灰度值 i 的直方图统计, 因此交叉熵反映的只是整体上一幅图像与另一幅图像之间的信息量差异。有可能出现这种情况: 两幅图像 P 和 Q 画面差别很大, 但是交叉熵 $I_{[P,Q]}$ 却比较小, 所以还有必要参考另外一个指标——均方根误差。

② 均方根误差(Root-mean-square error)

定义 4.2: 设图像的高和宽分别为 m 和 n (单位为像素), $r\text{Img}(i, j)$ 表示图像 $r\text{Img}$ 在像素点 (i, j) 处的灰度值, 则均方根误差

$$E_{RMS}(r\text{Img}', r\text{Img}) = \left\{ \frac{1}{mn} \left[\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [r\text{Img}'(i, j) - r\text{Img}(i, j)]^2 \right] \right\}^{1/2} \quad (4.2)$$

由公式(4.1)和公式(4.2)可知, 如果 $I_{[r\text{Img}', r\text{Img}]}$ 和 $E_{RMS}(r\text{Img}', r\text{Img})$ 的绝对值都比较小, 则说明转换后所得到的右眼视图与真正拍摄得到的右眼视图接近, 较符合人们心理和生理立体视觉的观察习惯, 转换后的立体效果就好。大量的实验和统计结果表明事实的确如此。

4.3 随机变量

在利用侯原理进行平面图像立体化过程中, 为了使图像子块在水平方向的移动具有随机性, 我们对每一个图像子块 $Q_{s,t}$ 都设置了一个随机变量 $K_{s,t}$, 这样整幅图像就对应一个随机变量矩阵 $K_{M \times N}$ ^[81], 如图 4.1 所示。下面讨论服从同一种均匀分布的随机变量取不同的值时对转换效果的影响。实验使用的原始图像如图 4.3(每幅图像的大小为 800×600)所示, 程序对右眼视图进行立体化。实验需要重复多次。要使各次实验的结果具有可比性, 必须让每次实验所使用的系统参数相同。而且在下面的实验中, 我们尽力让 $0.5 \leq K_{s,t} < 1.5$, 并且对于需要将图像分割成图像子

块的情况固定取 $M = 150$ ， $N = 200$ ，即将每幅图像划分成 150 行、200 列，每个图像子块的大小是 4×4 (像素)。



图 4.3 原始图像(左眼视图 $lImg$ + 右眼视图 $rImg$)

Fig 4.3 Original image (left image of $lImg$ / right image of $rImg$)

实验中，程序产生的随机变量 $K_{s,t}$ 应满足均值 $\alpha = 1$ ，以保证转换后图像大小总体上不变。同时，无论服从何种分布的随机变量 $K_{s,t}$ ，都应该使 $Q_{s,t}$ 在水平方向上的最大位移不超出舒适 Panum 融合区^[15, 80]的限制。侯等人提出可以采用方差实现对各图像子块的水平视差的控制，以使图像子块随机移动产生的水平视差不超出舒适 Panum 融合区的限制。但这种方法没有直观的物理意义，计算量也比较大。本文采用文献^[80]的计算方法。该方法要求转换后原图像的每一个像素对应的水平视差 $\Delta q_{t,k}$ 满足不等式(4.3)^[5]：

$$-\frac{2h\eta p_d}{EP_\sigma} < \Delta q_{t,k} < \frac{2h\eta p_d}{EP_\sigma} \quad (4.3)$$

其中 $2h$ 是瞳距，一般为 6.5 cm； η 为视锐度，一般取 $\eta = 1' \approx 2.907 \times 10^{-4}$ rad；监视器屏幕与观察者的距离为 p_d ($p_d > 0$)，这里取 100 cm， E 是瞳孔直径，一般为 0.4 cm， P_σ 是监视器的像素间隔，这里取 $P_\sigma = 3.53 \times 10^{-2}$ cm^[17]；因此可得， $-13 < \Delta q_{t,k} < 13$ 。由于立体化的转换过程是一张原始图像通过一定方法的平移而产生立体图像对完成的，因此可以认为这样的立体图像对是由平行双目立体成像系统^[80]得到的，可以不考虑文献[80]中提到的理论限制。产生一组符合要求的随机数的算法如下：

Algorithm ProduceKmn: produce a random variable matrix $K_{M \times N}$ which satisfies the constraint of Panum's Fusional Area.

Input: a random variable matrix $K_{M \times N}$ which contains M rows and N columns ($M \geq 1, N \geq 1$). Each element $K_{[s, t]}$ in the matrix corresponds to a sub-block $Q_{s, t}$.

Output: a random variable matrix $K_{M \times N}$ in which each element satisfies the constraint of Panum's Fusional Area

Method:

for each row K_s in $K_{M \times N}$ do

begin

repeat

for each element $K_{[s, t]}$ in row K_s do

produce a random number which has certain distributions, $K_{s, t} \in [0.5, 1.5]$;

if $K_{s, t}$ satisfies the constraint of Panum's Fusional Area then

$K_{[s, t]} := K_{s, t}$;

else

repeat producing $K_{s, t}$; if $K_{s, t}$ still not satisfies the constraint after three times, clear all elements in the front of $K_{[s, t]}$

until all elements in row K_s have random numbers which satisfy the constraint;

end.

在最好的情况下，算法 ProduceKmn 的时间复杂度是 $O(M \times N)$ 。

从侯等人提出的平面图像立体化方法原理可以看出^[81]，将一幅图像划分成 $M \times N$ 个图像子块之后，行与行之间的图像子块对应的随机数没有关系，而每一行内部各图像子块之间对应的随机数是有依赖关系的。因此，为了降低时间复杂度，算法 ProduceKmn 采取的方法是如果某一图像子块对应的随机数不符合要求，则再重复产生。如果重复三次还是不符合要求，那么这一行对应的随机数就全部重新产生，并且每产生一个随机数，就立刻检查它是否符合要求。

在下面的实验中，我们讨论了6种不同情形的参数对立体效果的影响。

4.3.1 平移像素法

首先让我们考虑最简单的情形。设 $M = 1, N = 1$ ，并且随机变量 $K_{1,1}$ 的取值为 1。此时得到的 $limg$ 和 $rimg$ 是同一幅图像。如果我们将 $limg$ 的场景平移一段距离 d (单位为像素)，得到 $limg'$ ， $limg'$ 再和 $rimg$ 构成立体图像对 (如图 4.4 所示)，这样是否也能产生较好的立体效果呢？我们通过与 3.3 节相同的分时双目立体成像系统来验证这个问题。为了便于比较，实验所采用的参数也与 3.3 节相同，并且保证 $-13 \leq \Delta q_{t,k} \leq 13$ ，即两幅图像对应像素的水平视差的绝对值不超过 13 个像素。我们采用平移法重复实验 10 次，实验得到的数据如表 4.1 和图 4.5 所示。从表 4.1 和图 4.5 可以看出，当 $3 \leq d \leq 5$ 时，交叉熵和均方根误差都比较小，立体效果相对要好一些。

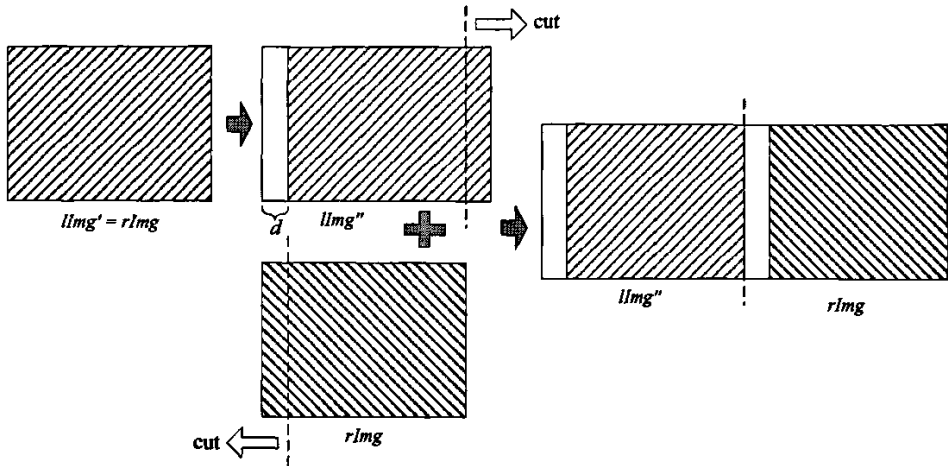


图 4.4 通过平移像素法生成立体图像对

Fig 4.4 By moving pixels without stretch, a stereo pair is created

表4.1 交叉熵和均方根误差(平移像素法)

Table 4.1 Cross-entropies and root-mean-square errors (Moving pixels without stretch method)

d	$I[limg', limg]$	E_{RMS}
1	0.002834	38.33027
2	0.001599	38.27087
3	0.000372	38.22086
4	-0.000854	38.19509
5	-0.002081	38.22177
6	-0.003305	38.28793
7	-0.004531	38.36434
8	-0.005758	38.43292
9	-0.006992	38.47522
10	-0.009653	39.48690
11	-0.010899	39.54044
12	-0.012141	39.59407
13	-0.013382	39.64767
Mean value	-0.004984	38.69757
Standard deviation	0.005361	0.609734

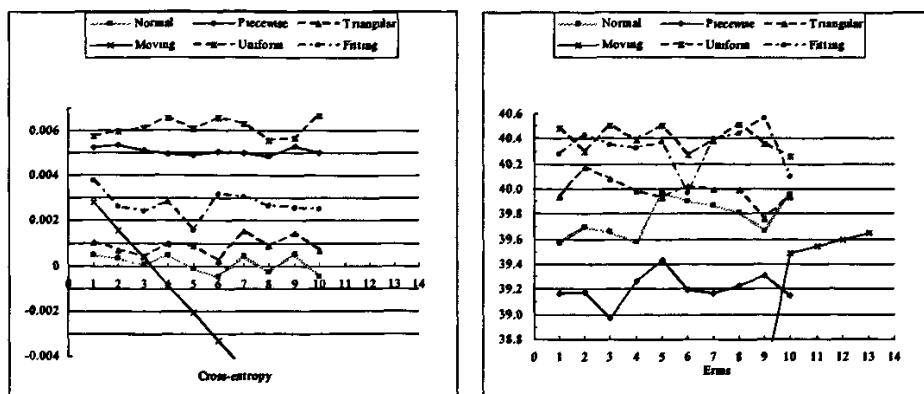


图 4.5 根据实验所获得的数据，我们可以将曲线画在一张图中做比较

Fig 4.5 According to the data get from experiments, we
can draw lines in one figure for comparison

注:

Moving: Moving pixels without stretch

Uniform: Uniform distribution

Triangular: Triangular distribution

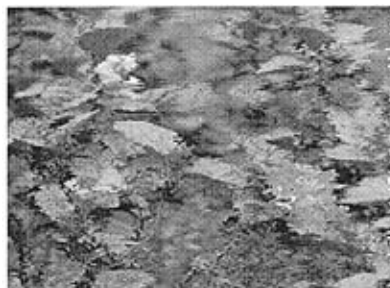
Piecewise: Piecewise uniform distribution

Fitting: Fitting the distribution of gray value

Normal: Normal distribution

4.3.2 均匀分布

设程序产生的 $K_{s,t}$ 服从均匀分布，且满足 $0.5 \leq K_{s,t} \leq 1.5$ 。之所以要求 $0.5 \leq K_{s,t} \leq 1.5$ ，是因为要让程序产生的随机变量的均值 $\alpha=1$ ，以保证转换后图像大小总体上不变。同时， $K_{s,t}$ 还应该使 $Q_{s,t}$ 在水平方向上的最大位移不超出舒适 Panum 融合区的限制。



(a) Uniform

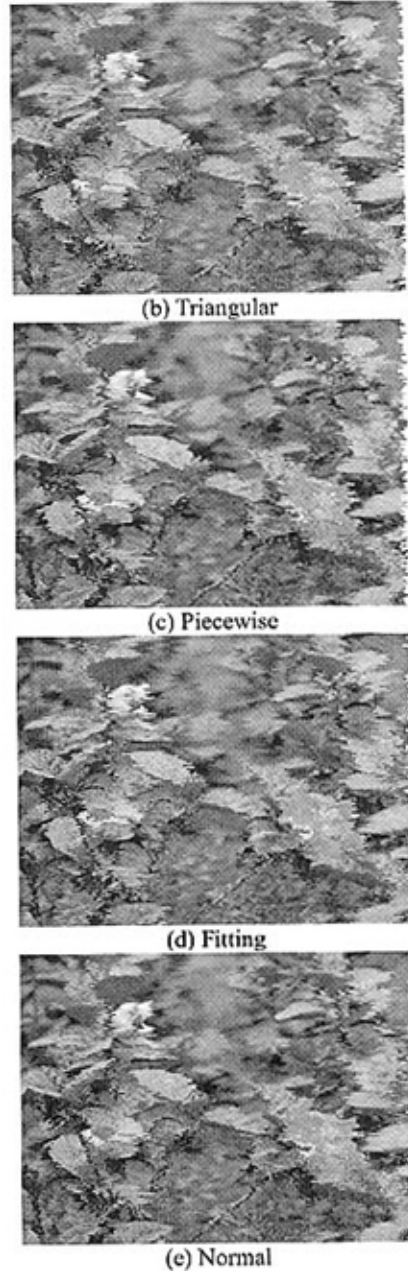


图 4.6 左眼视图 lmg' 可以根据表 4.2 中不同的随机数集合生成；它们有一些不同

Fig 4.6 The left image of lmg' can be created according to different sets of random numbers in Table 4.2. They are slightly different

表 4.2 中的“Uniform”列是一组符合要求的服从均匀分布的随机变量取值。根据这组取值得到的左眼视图如图 4.6(a)所示。将它和原始图像中的左眼视图进行对比，可得到相应的交叉熵和均方根误差。重复进行实验 10 次，可得到实验数据

如表 4.3 中“Uniform”列所示。根据这些数据所画的曲线图见图 4.5。

表 4.2 服从不同分布的随机数数据集；所有这些数都由算法 ProduceKmn 产生

Table 4.2 Data sets of random numbers that have different distributions; all the numbers are produced by algorithm ProduceKmn

Random variables	Value				
	Uniform	Triangular	Piecewise	Fitting	Normal
$K_{1,1}$	0.9854	0.5000	1.1131	1.4412	0.7294
$K_{1,2}$	1.4960	0.6134	0.8892	0.6862	1.1507
...
$K_{150,200}$	1.3390	0.8848	1.2481	1.2761	0.6891

表 4.3 交叉熵和灰度均方根误差

Table 4.3 Cross-entropies and root-mean-square errors of gray scale

No.	Uniform		Triangular		Piecewise	
	$I(I_{img'}, I_{img})$	E_{RMS}	$I(I_{img'}, I_{img})$	E_{RMS}	$I(I_{img'}, I_{img})$	E_{RMS}
1	0.005744	40.48661	0.001033	39.93968	0.005240	39.16967
2	0.005968	40.29314	0.000666	40.17308	0.005372	39.17799
3	0.006120	40.50277	0.000421	40.07825	0.005132	38.97137
4	0.006564	40.38166	0.000965	39.98112	0.004953	39.26402
5	0.006051	40.50212	0.000887	39.92865	0.004892	39.43780
6	0.006559	40.26641	0.000243	40.01958	0.005070	39.19563
7	0.006322	40.37273	0.001516	39.99385	0.004991	39.16959
8	0.005552	40.51129	0.000860	39.99393	0.004883	39.23213
9	0.005656	40.35854	0.001437	39.76041	0.005306	39.30676
10	0.006666	40.26092	0.000679	39.94126	0.005008	39.15154
Mean value	0.006120	40.39362	0.000871	39.98098	0.005085	39.20765
Standard deviation	0.000326	0.085663	0.001033	39.93968	0.000172	0.11994

表 4.3 交叉熵和灰度均方根误差(续表)

Table 4.3 Cross-entropies and root-mean-square errors of gray scale (Continued)

No.	Fitting		Normal	
	$I(I_{img'}, I_{img})$	E_{RMS}	$I(I_{img'}, I_{img})$	E_{RMS}
1	0.003782	40.2754	0.000491	39.57265
2	0.002635	40.4307	0.000344	39.69419
3	0.002396	40.3512	0.000061	39.66012
4	0.002808	40.3192	0.000482	39.58124
5	0.001578	40.3682	-0.000098	39.97126
6	0.003185	39.9659	-0.000456	39.89830
7	0.003069	40.3905	0.000414	39.86347
8	0.002649	40.4378	-0.000287	39.81272
9	0.002567	40.5623	0.000537	39.67035
10	0.002520	40.0976	-0.000452	39.95245
Mean value	0.002719	40.3199	0.000104	39.76767
Standard deviation	0.0005733	0.1732	0.000350	0.131965

4.3.3 三角分布

我们通过程序模拟了服从三角分布的随机变量 $K_{s,t}$ ，以研究不同分布对立体化效果的影响。

定义 4.3: 三角分布

我们设

$$p(x) = \begin{cases} -\frac{\pi}{2} \cos(\pi x) & 0.5 \leq x < 1.5, \\ 0 & \text{otherwise.} \end{cases} \quad (4.4)$$

其中, $p(x)$ 是 $K_{s,t}$ 的概率密度函数, 我们称随机变量 $K_{s,t}$ 服从三角分布。

因此,

$$\alpha = E(K_{s,t}) = \int_{-\infty}^{\infty} xp(x)dx = \int_{0.5}^{1.5} x \cdot \left(-\frac{\pi}{2}\right) \cos(\pi x) dx = 1 \quad (4.5)$$

等式(4.5)表明 $K_{s,t}$ 的均值等于 1。表 4.2 中的第三列 “Triangular” 是一组服从三角分布且满足要求的随机变量取值。根据这组取值得到的左眼视图如图 4.5(b) 所示。同样, 我们采用该分布重复实验 10 次, 将实验数据分别填入表 4.3 相应的栏中。根据这些数据所画的曲线图参见图 4.5。

4.3.4 分段均匀分布

设 $K_{s,t}$ 服从均匀分布, 限制 $K_{s,t} \in [0.5, 1.5)$, 同时, 保证 $K_{s,t}$ 使 $Q_{s,t}$ 在水平方向上的最大位移不超出舒适 Panum 融合区的限制。从图 4.3 中可以看出, 在同一时刻, 左眼看到的场景相对右眼要向右偏一点, 因此, 当根据右眼视图 $rlmg$ 生成左眼视图 $llmg$ 时, 可将右眼视图划分成均等的左右两部分, 左半部分整体上多拉伸一些, 右半部分整体上多压缩一些, 以减少 $llmg$ 和 $llmg'$ 的不同。根据这个思想, 我们给出分段均匀分布的定义。

定义 4.4: 分段均匀分布

设图像对应的随机变量矩阵 $K_{M \times N}$ 的左半部分的随机变量服从 $[1.5 - \beta, 1.5]$ 上的均匀分布, 右半部分服从 $[0.5, 0.5 + \beta]$ 上的均匀分布, 其中 $\beta \in (0, 1]$ 。我们称这种均匀为分段均匀分布(Piecewise Uniform Distribution)。

这里取 $M = 150$, $N = 200$, $\beta = 0.82$, 从而有

$$\begin{aligned} E(\Delta q_{t,k}) &= E(W \times [(t-1) - (K_{s,1} + K_{s,2} + \dots + K_{s,t-1})] + kW(1 - K_{s,t}) / K) \\ &= \begin{cases} W(\beta-1)(t-1+k/K)/2, & 1 < t \leq N/2 \\ W(1-\beta)(t-1-N+k/K)/2, & t > N/2 \end{cases} \end{aligned} \quad (4.6)$$

其中 $k \in [0, K-1]$ 是图像子块中水平像素的序号。从公式(4.6)可以看出, 当 $t < N$ 时, $E(\Delta q_{t,k}) < 0$; 当 $t = N$, $k = K$ 时, $E(\Delta q_{N,K}) = 0$ 。这说明, 右眼视图最后一列

像素对应的水平视差的均值为 0，生成的左眼视图没有丢失右眼视图的内容。由这两幅图像构成立体图像对是负视差的情形。

表 4.2 中的第四列“Piecewise”是一组服从分段均匀分布且满足限制条件的随机变量取值($\beta = 0.82$)。根据这组取值得到的左眼视图如图 4.6(c)所示。相应的 $I[limg', limg]$ 和 $E_{RMS}(limg', limg)$ 可以根据图 6(c)和图 4.3 中的左眼视图得到。重复实验 10 次，将实验得到的数据填入表 4.3 相应的栏中，并根据这些数据在图 4.5 中画出相应的曲线。

分别取 β 为 0.80、0.82、0.85、0.87 和 0.89 等值进行实验，可以发现 β 为 0.82 时效果较好，但耗时也较长。

4.3.5 拟合灰度值分布

我们可以对原始图像 $rimg$ 的灰度值进行统计，拟合灰度直方图得到一个函数 $p(k)$ 作为随机变量 $K_{s,t}$ 的概率密度函数，通过程序模拟产生服从该分布的随机数，最终可求得 $limg'$ 。这是拟合灰度值分布方法的思想。

$rimg$ 的灰度值的分布如图 4.7 所示。我们将其进行归一化，其中横坐标为 $k/255$ ，纵坐标为 $n_k/\max(n_k)$ ， n_k 是图像中灰度值为 k 的像素点个数， $k = 0, 1, 2, \dots, 255$ 。得到的归一化灰度值分布如图 4.8 所示。

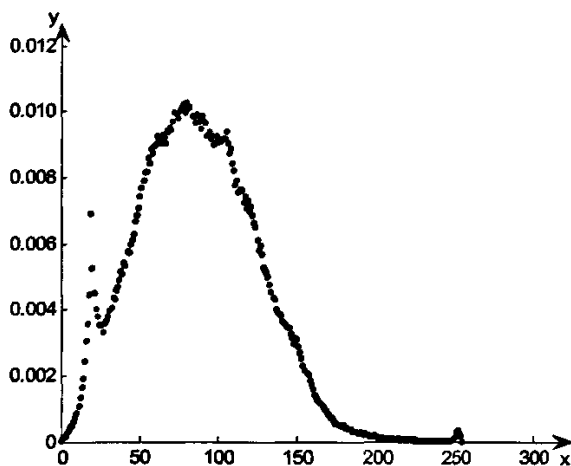


图 4.7 $rimg$ 的灰度值的分布是不规则的，它取决于图像本身

Fig 4.7 Distribution of the gray value of $rimg$ is not regular, it depends the image itself

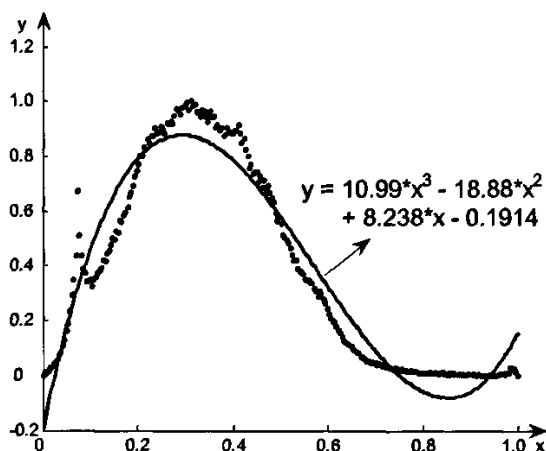

 图 4.8 归一化灰度值分布及 $p(k)$ 对应的曲线

 Fig 4.8 The normalized distribution of the gray value and the curve of $p(k)$

我们采用最小二乘法对图 4.8 中的离散点进行曲线拟合。设 $p(k) = p_1 * k^3 + p_2 * k^2 + p_3 * k + p_4$ 。根据 $rImg$ 的灰度值的分布，当置信界限(confidence bound)为 95% 时，可求得： $p(k) \approx 10.993 * k^3 - 18.882 * k^2 + 8.2377 * k - 0.1914$ 。它所对应的曲线如图 4.8 所示。

我们希望转换后的图像大小不变，这就要求 $E(K_{s,t}) = 1$ ，即

$$\int_{0.5}^{1.5} kp(k)dk = 1 \quad (4.7)$$

根据概率密度函数的性质， $p(k)$ 还应满足等式(4.8)。

$$\int_{-\infty}^{\infty} p(k)dk = \int_{0.5}^{1.5} p(k)dk = 1 \quad (4.8)$$

为了使得 $p(k)$ 同时满足等式(4.7)和等式(4.8)，我们须将曲线 $p(k)$ 进行一定的平移。设 $p(k)$ 平移后的曲线为 $f(k) = p(k - k_0) + c_0$ ，那么 $f(k)$ 必须同时满足

$$\int_{k_0}^{k_0+1} f(k)dk = \int_{k_0}^{k_0+1} [p(k - k_0) + c_0]dk = 1$$

和

$$\int_{k_0}^{k_0+1} kf(k)dk = \int_{k_0}^{k_0+1} k[p(k - k_0) + c_0]dk = 1$$

因此可求得 $k_0 = 0.5614$ ， $c_0 = 0.6196$ ，故

$$f(k) = \begin{cases} 10.993k^3 - 37.4360k^2 + 39.9222k - 12.1417 & 0.5626 \leq k \leq 1.5626 \\ 0 & \text{else} \end{cases} \quad (4.9)$$

其中 $f(k)$ 是 $K_{s,l}$ 的概率密度函数。

然后我们可用如下方法获得服从概率密度函数为 $f(k)$ 的随机数(Random number):

① 求 $f(k)$ 的分布函数 $F(k)$:

$$F(k) = \int_{-\infty}^k f(k)dk = \int_{0.2630}^k (10.993k^3 - 37.4360k^2 + 39.9222k - 12.1417)dk \quad (4.10)$$

$$= 2.7483k^4 - 12.4787k^3 + 19.9611k^2 - 12.1417k + 2.4596$$

② 求 $F(k)$ 的反函数 $F^{-1}(k)$:

可以用一元四次方程求根公式求解, 也可以用数值分析方法近似求解。

得到 $F^{-1}(k)$ 以后, 让程序产生 $[0, 1)$ 区间上服从均匀分布的随机数, 再代入公式 (4.10), 我们就可计算得到服从概率密度函数为 $f(k)$ 的随机数。

注意我们同样要求程序产生的随机数满足舒适 Panum 融合区的限制。

表 4.2 中的第五列是一组拟合灰度值分布且满足限制条件的随机变量取值。根据这组取值得到的左眼视图如图 4.6(d) 所示。与前面几节相似, 我们采用拟合灰度值的分布的方法重复实验 10 次, 将实验得到的数据填入表 4.3 相应的栏中。根据这些数据所得到的图形参见图 4.5。

4.3.6 正态分布

设 $K_{s,l}$ 服从正态分布, 且它的均值 $\alpha = 1$ 。为了便于比较, 我们要让 $K_{s,l}$ 的取值绝大部分落在区间 $[0.5, 1.5]$, 并且做到当一个随机数产生时, 就立即检查它是否落在区间 $[0.5, 1.5]$, 如果落在这个区间就保留, 否则舍弃。根据这个要求及 “ 3σ ” 原则, 有 $P = P\{\alpha - 3\sigma \leq K_{s,l} < \alpha + 3\sigma\} = 0.9974$, 从中求得 $\sigma = 1/6$, $D(K_{s,l}) = \sigma^2 = 1/36$ ($\sigma > 0$)。

表 4.2 中的第 6 列 “Normal” 是一组服从正态分布且满足限制条件的随机变量取值 (随机数)。根据这组取值得到的左眼视图如图 4.6(e) 所示。与前面几节相似, 我们重复实验 10 次, 将实验得到的数据填入表 4.3 相应的栏中, 对应的曲线画在图 4.5 中。

4.3.7 实验结果

从前面 6 种不同的情形我们发现:

① 对平移像素法自身而言, 当 $3 \leq d \leq 5$ 时, 交叉熵 $I[lmg', lmg]$ 和均方根误差 $E_{RMS}(lmg', lmg)$ 都比较小, 立体效果相对要好一些。不过, 实际观察发现, 虽然此时平移像素法的交叉熵和均方根误差比起其他分布都要小, 但得到的图像的立体感并没有明显的增强。造成这种现象的原因是, 在使用该方法时场景的信息会有丢失。

② 从图 4.5 可以看出, 分段均匀分布和正态分布的交叉熵和均方根误差都比较理想。就这两种分布而言, 分段均匀分布的交叉熵较大, 而均方根误差较小。实际观察显示, 两种分布对应的立体效果差不多, 都比其他分布要较好一些。对其他的图像进行实验, 可以发现正态分布的效果要比分段均匀分布的效果更稳定

一些,并且多数情况下比其他分布要好。这可能是由于正态分布是自然界中最常见的分布的原因。

③ 理论上, $I_{[limg', limg]}$ 和 $E_{RMS}(limg', limg)$ 的最佳值应该都是 0, 这时表明 $limg'$ 和 $limg$ 几乎没有差别。但从表 4.3 和图 4.5 可以看出, 在每次实验中随机变量取不同的值, $I_{[limg', limg]}$ 和 $E_{RMS}(limg', limg)$ 都没有衰减为 0 的趋势, 而不管随机变量服从何种分布。相反地, 它们基本上都稳定在平均值附近。重复更多次实验我们发现, $I_{[limg', limg]}$ 和 $E_{RMS}(limg', limg)$ 本身也是随机变量, 并且相互是独立的。这表明当随机变量服从同一种分布时, 随机变量的取值对立体效果没有多大的影响。当随机变量服从不同分布时, 不同分布的交叉熵和均方根误差会有所不同, 表明不同的分布会对立体效果产生一定的影响。一般说来, 分段均匀分布和正态分布的立体效果较好, 并且多数情况下正态分布表现更稳定一些。实际观察也证实了这一点。

4.4 图像子块个数

利用侯原理将平面图像立体化需要将一副原始图像划分成 $M \times N$ 个图像子块。下面讨论参数 M 和 N 对立体效果的影响。实验使用的原始图像同图 4.3, 随机变量 $K_{s,i}$ 服从均匀分布, 且满足 $0.5 \leq K_{s,i} \leq 1.5$ 。程序对左眼视图进行立体化。实验过程和评价指标的计算类似于第 4.3 节, 改变 M 和 N 的值多次进行实验, 实验结果如表 4.4 所示。

表 4.4 M 和 N 取不同值时的交叉熵和均方根误差

Table 4.4 Cross-entropy and root-mean-square error of gray scale (M and N are set different pairs)

实验次数	M	N	交叉熵 $I_{[rlmg', rlmgl]}$	灰度均方根误差 $E_{RMS}(rlmg', rlmgl)$
1	3	4	0.0751694366335869	41.8372383117676
2	6	8	0.0772337689995766	42.7408256530762
3	12	16	0.0743808373808861	42.0463600158691
4	24	32	0.0709674656391144	42.3715324401855
5	60	80	0.076013058423996	42.3922271728516
6	120	160	0.0784920901060104	42.3936805725098
7	150	200	0.075246050953865	42.1790199279785
8	300	400	0.072163350880146	41.9058036804199
9	600	800	0.051229078322649	41.0751113891602
10	1	4	0.0500383973121643	40.0455017089844
11	1	8	0.119796976447105	43.7675170898438
12	1	16	0.0608777217566967	42.0727272033691
13	1	25	0.0559485889971256	40.6947746276855
14	1	32	0.0760729908943176	41.9188499450684
15	1	80	0.0581528469920158	41.2985877990723
16	1	100	0.0559812597930431	41.2457847595215
17	1	160	0.0754962936043739	42.1061859130859
18	1	200	0.0532086342573166	40.9138870239258
19	1	400	0.0592886358499527	41.0359001159668
20	1	800	0.05742496997118	41.4469299316406
...

从表 4.4 可以看出:

① 从交叉熵和均方根误差这两个评价指标来看, M 取 1(对原始图像只做垂直划分)和 M 取其它值之间的差别不是很大。并且, 两个指标也没有衰减为 0 的趋势。从理论上分析这种现象是可以理解的: 因为随机变量 $K_{s,l}$ 服从均匀分布, 并且都是对同一副图像 $l\text{Img}$ 立体化。可以计算出 $I_{[l\text{Img}, r\text{Img}]} = 0.005149$, 对应的 $E_{RMS}(l\text{Img}, r\text{Img}) = 38.92347$ 。根据上述立体化方法可知, $E_{RMS}(r\text{Img}', r\text{Img})$ 不应偏离 $E_{RMS}(l\text{Img}, r\text{Img})$ 太远, $I_{[r\text{Img}', r\text{Img}]}$ 也类似。因此这两个指标没有衰减为 0 的趋势。

② 实验中发现, 当 $M=1$ 时, 相当于对原始图像只做垂直划分, 这样得到的图像较为平滑, 所以从实验的效果来看, 看到的立体图像要平滑一些, 不过立体感要弱一些。在 $M=1$ 的情况下, 当 N 趋于 $r\text{Img}$ 的宽度时, 看到的立体图像变得越来越平滑; 但是立体感变得越来越弱, 直到最后没有立体感。从理论上分析这种现象是可以理解的: 当 N 趋于 $r\text{Img}$ 的宽度时, 差不多几个像素就构成了一个图像子块, 由于像素的坐标需要取整数值, 这就造成不少像素没有移动, 因此得到的 $r\text{Img}'$ 与 $l\text{Img}$ 相差不大, 立体感就会减弱。

根据①和②, 可以得出结论: 在 p_d 为一个合适的值的前提下, 图像子块的个数取不同的值对立体化效果的综合影响不大, 只对立体图像的平滑和立体感产生微弱的影响。当图像变得平滑时立体感就变弱, 反之, 立体感就变强。平滑与立体感是互相矛盾的, 人们只能取一个折中值。

当随机变量 $K_{s,l}$ 服从其他分布时, 也有类似的结论。

4.5 小结

本章主要论述了基于侯原理的立体图像对的生成。

将平面图像转换成立体图像很有意义。获取同一场景的立体图像对是实现平面图像立体化的一个关键问题。侯等人根据生理深度暗示和心理深度暗示的关系理论, 在每个图像单元中随机加入双目视差, 在心理深度暗示和生理深度暗示相互作用下, 会使观察者获得具有一定深度感知的立体图像, 我们称之为侯原理。利用侯原理将平面图像立体化的核心思想是: 将一幅图像划分成若干个图像子块, 在图像子块中随机地加入水平视差。本章利用侯原理在标准的计算机监视器上对普通的平面图像进行了立体化, 给出了评价转换后的立体效果的定量指标(交叉熵和均方根误差), 并讨论了各参数对立体效果的影响。实验结果表明:

① 当随机变量矩阵中的每一个随机变量都服从同一种分布时, 随机变量的取值对立体效果没有多大的影响。

② 当随机变量矩阵中的随机变量每次采用不同的分布时, 立体效果化的效果会有一些的不同。一般地, 采用分段均匀分布和正态分布立体效果较好, 并且多数情况下正态分布表现更稳定一些。

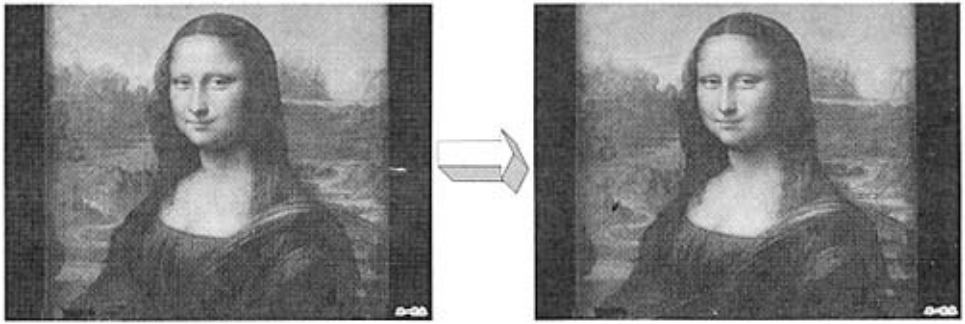
③ 当监视器屏幕与观察者的距离为一个合适的值(例如 1 m)时, 图像子块的个数取不同的值对立体效果的综合影响不大, 只对立体图像的平滑和立体感产生微弱的影响。平滑与立体感是互相矛盾的, 人们只能取一个折中值。

有几个问题值得注意:

① 平面图像立体化的转换效果的评价指标——交叉熵反映的只是整体上一幅图像与另一幅图像之间的信息量差异, 实验表明, 单独使用交叉熵来评价转换效果是不准确的。我们首先可以考察交叉熵, 在交叉熵比较小的情况下, 再考察灰度均方根误差, 一般地, 交叉熵和灰度均方根误差的值越小效果越好。可以用这两个指标一起控制平面图像立体化的过程。

② 在讨论随机变量对转换效果的影响时, 我们注意到算法 `ProduceKmn` 需要对产生的随机数进行检查。由于涉及对随机数的操作, 这个算法在很坏的情况下有可能变得不稳定。在最好的情况下, 算法 `ProduceKmn` 的时间复杂度是 $O(M \times N)$ 。在一定的情况下, 特别是 Δq_{jk} 的取值范围变得较小时, 计算量急剧增大, 甚至不收敛。因此, 需要进一步研究如何保证算法的收敛性, 以及加快算法的收敛速度。

③ 由于平滑与立体感是互相矛盾的, 人们只能取一个折中值, 因此由侯原理生成的立体图像的立体感是有限的, 其立体效果是不可能达到尽善尽美的, 它适合于那些版画等主要含有心理深度暗示的平面图像的立体化。例如图 4.9 是达·芬奇的作品《蒙娜丽莎》。要将这幅图像立体化采用其他方法是复杂而困难的。采用侯原理简单实用且效果也明显。但是要达到非常逼真的立体效果侯原理是无能为力的。通过实验我们发现, 采用正态分布已经接近侯原理所能达到的最佳立体效果的极限, 必须寻求另外的办法来生成立体效果更好的立体图像对。第五章和第六章分别阐述了另外的立体图像对获取方法。



左眼视图

右眼视图

图 4.9 采用侯原理立体化《蒙娜丽莎》

Fig 4.9 Conversion of "Mona Lisa" from planar to stereo with Hou principle

5 已建三维模型的情况下立体图像对的生成

5.1 引言

能在标准的计算机 CRT 监视器上看到植物生长过程的立体图像是一件令人兴奋的事情。它可以给人们带来更加真实直观的视觉效果，提供较好的沉浸 (Immersion)。如前所述，目前人们较多的使用基于监视器的分时双目立体成像系统 (Time-sharing Binocular Stereo Imaging System) 来实现立体成像。对于这样的系统，只要得到立体图像对，就能通过系统看到立体图像。因此，获得立体图像对是一个非常关键的问题。

我们知道，有两类方式获得立体图像对^[15]：第一类方式是使用立体相机拍摄场景得到立体图像对，这种方式有一个很大的缺点：不适合计算机游戏、影片特技等场合；第二类方式是通过计算机绘制得到立体图像对。这类方式引起了国内外学者的广泛兴趣。侯原理就属于这一类。还有一些学者讨论了场景的三维模型已建好的情况下立体图像对的生成^[34, 35, 121-124]。但是较少深入讨论这个问题。本章给出了一种在已建好三维模型的情况下生成立体图像对的方法，并对相关问题作了较为深入的分析，指出了在何种情形下成像能够达到较好的立体效果（立体感逼真程度）。

5.2 坐标变换

在立体图像对的生成过程中，经常遇到坐标变换的问题。我们给出了一个基于监视器的双目立体成像模型。该模型如图5.1所示。所标参数的具体含义参见文献^[17]。针对该模型我们可以求出从景物坐标系^[68, 117]到世界坐标系的坐标变换矩阵。

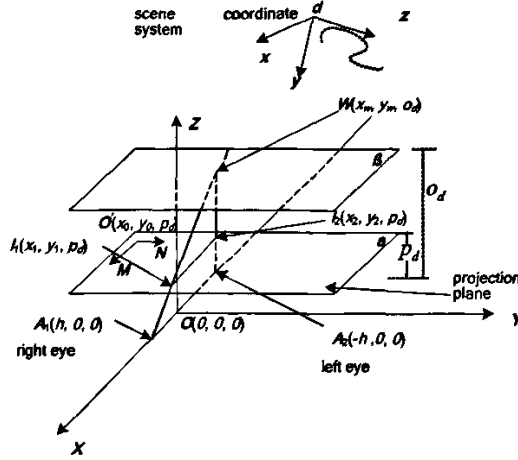


图 5.1 双目立体成像模型及景物坐标系

Fig 5.1 Binocular stereo imaging model and scene coordinate system

① 从景物坐标系到世界坐标系

设景物坐标系原点 d 在世界坐标系中的坐标为 (d_x, d_y, d_z) ，且两坐标系的单位长度相同。若已知坐标轴间的欧拉角分别为 ε_x (俯仰角, pitch)、 ε_y (偏航角, yaw) 和 ε_z (滚动角, roll)，则根据布尔莎公式，景物坐标系中任意一点 (x_j, y_j, z_j) 在世界坐标系中的坐标为^[68]

$$[x \ y \ z \ 1] = [x_j \ y_j \ z_j \ 1] T \quad (5.1)$$

其中 T 为变换矩阵，

$$T = \begin{bmatrix} \cos \varepsilon_y \cos \varepsilon_z & \sin \varepsilon_x \sin \varepsilon_y \cos \varepsilon_z - \cos \varepsilon_x \sin \varepsilon_z & \cos \varepsilon_x \sin \varepsilon_y \cos \varepsilon_z - \sin \varepsilon_x \cos \varepsilon_z & 0 \\ \cos \varepsilon_y \sin \varepsilon_z & \sin \varepsilon_x \sin \varepsilon_y \sin \varepsilon_z + \cos \varepsilon_x \cos \varepsilon_z & \cos \varepsilon_x \sin \varepsilon_y \sin \varepsilon_z - \sin \varepsilon_x \cos \varepsilon_z & 0 \\ \sin \varepsilon_y & \sin \varepsilon_x \cos \varepsilon_y & \cos \varepsilon_x \cos \varepsilon_y & 0 \\ d_x & d_y & d_z & 1 \end{bmatrix}$$

若已知 z 轴和 X 轴、 Y 轴、 Z 轴的夹角分别为 α 、 β 和 γ ，则 T 为^[125]：

$$\begin{bmatrix} \frac{\cos \alpha \cos \beta}{(\cos^2 \alpha + \cos^2 \beta)^{1/2}} & \frac{\cos \beta \cos \gamma}{(\cos^2 \alpha + \cos^2 \beta)^{1/2}} & -(\cos^2 \alpha + \cos^2 \beta)^{1/2} & 0 \\ -\frac{\cos \beta}{(\cos^2 \alpha + \cos^2 \beta)^{1/2}} & \frac{\cos \alpha}{(\cos^2 \alpha + \cos^2 \beta)^{1/2}} & 0 & 0 \\ \cos \alpha & \cos \beta & \cos \gamma & 0 \\ d_x & d_y & d_z & 1 \end{bmatrix}$$

若已知景物坐标系三个基本单位向量在世界坐标系中的坐标分别为 (x_u, y_u, z_u) 、 (x_v, y_v, z_v) 和 (x_w, y_w, z_w) ，则 T 为^[126]：

$$\begin{bmatrix} x_u & y_u & z_u & 0 \\ x_v & y_v & z_v & 0 \\ x_w & y_w & z_w & 0 \\ d_x & d_y & d_z & 1 \end{bmatrix}$$

② 从世界坐标系到图像坐标系

如图5.2所示, 从世界坐标系到计算机屏幕的图像坐标系(又叫设备坐标系, Device Coordinate System)的公式为^[17]:

$$I_1(\frac{(x_w - h) \times p_d}{o_d} + h - x_0, \frac{y_w \times p_d}{o_d} - y_0), I_2(\frac{(x_w + h) \times p_d}{o_d} - h - x_0, \frac{y_w \times p_d}{o_d} - y_0) \quad 5.2$$

其中, $MO'N$ 和 XYZ 两坐标系的单位长度相同。这个公式已在3.2.1节介绍过。

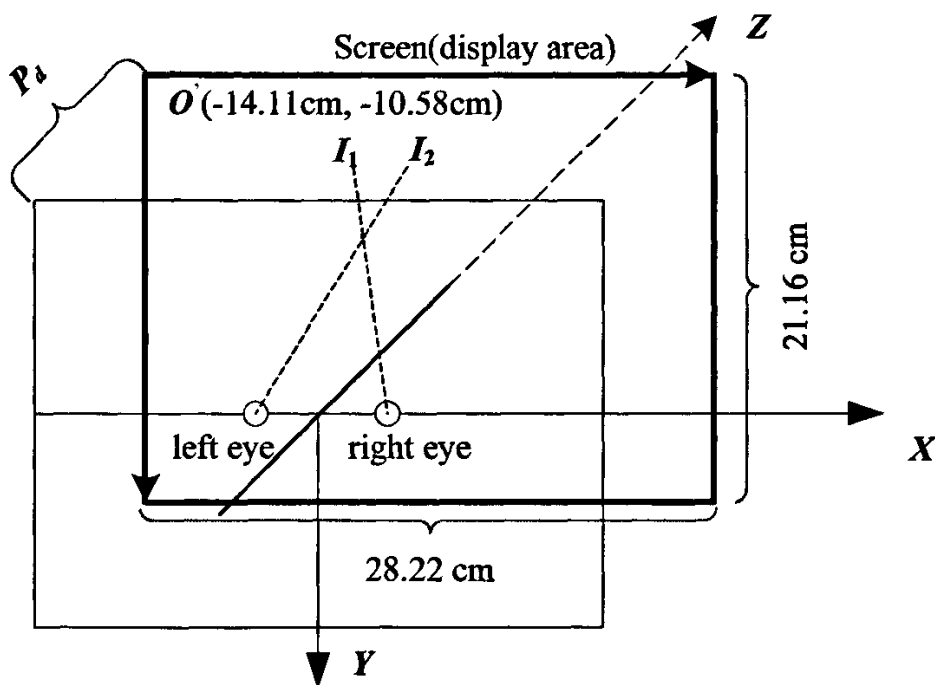


图 5.2 世界坐标系与图像坐标系

Fig 5.2 World coordinate system and image coordinate system

在我们的实验中, 将监视器的显示区域(display area)调整为 $28.2 \text{ cm} \times 21.2 \text{ cm}$, 分辨率调整为 800×600 , 因而有 $1 \text{ cm} = 28.35 \text{ pixels}$ 。设 O' 在坐标系 XYZ 中的坐标为 $(-14.11, -10.58, p_d)$, 即 $x_0 = -14.11 \text{ cm}$, $y_0 = -10.58 \text{ cm}$, 如图 5.2 所示, 则公式(5.2)可改写为

$$\begin{aligned}
& I_1'((\frac{(x_w - h) \times p_d}{o_d} + h + 14.11) \times 28.35, (\frac{y_w \times p_d}{o_d} + 10.58) \times 28.35), \\
& I_2'((\frac{(x_w + h) \times p_d}{o_d} - h + 14.11) \times 28.35, (\frac{y_w \times p_d}{o_d} + 10.58) \times 28.35)
\end{aligned} \tag{5.3}$$

其中，坐标的单位为像素。

根据上述讨论，可以采用如下方法在计算机监视器上生成立体图像对：首先在特定的景物坐标系中绘制场景，然后根据公式(5.1)和公式(5.3)将该场景分别投影成图像坐标系中的左眼视图和右眼视图。通过基于监视器的分时双目立体成像系统便可以观察到立体图像。

5.3 3DS MAX 中摄像机的使用

在三维场景已经建好的情况下，需要根据公式(5.1)和公式(5.3)将该场景分别投影成图像坐标系中的左眼视图和右眼视图以得到立体图像对。有些绘图软件(例如 3DS MAX)提供了摄像机对象，它能够简化这种投影过程。就像人的眼睛观察场景中的物体一样，摄像机记录图像的工作原理是：先对场景中的物体进行坐标变换，然后再进行透视投影。物体经过摄像机的投影就是拍摄到的底片(film)，而监视器上摄像机视口(viewport)的图像则是底片冲洗出来的照片。如果把两个完全相同的摄像机对象分别放置在不同位置，使它们之间的距离等于瞳距，再对同一场景进行拍照，就能生成立体图像对。本节研究如何利用摄像机对象生成立体图像对。

5.3.1 度量单位

在 3DS MAX 中，默认情况下，数值使用的度量单位是通用单位(generic unit)，它表示在 3DS MAX 的工作区域中实际显示的单位。使用者可以定义每个通用单位代表任何实际单位。例如，每个通用单位可以代表 1 英寸、5 米、或者 100 海里等。在 3DS MAX 中进行正确的单位设置十分重要。因为 3DS MAX 中高级光照特性使用真实世界的尺寸进行计算，它要求建立的模型与真实世界的尺寸一致。

5.3.2 摄像机对象

在 3DS MAX 中可以生成摄像机对象(目标摄像机和自由摄像机)。它们的作用与真实世界的摄像机一样，用于记录静态或动态的图像。生成静态的立体图像对主要用到目标摄像机对象，它包含了视点(viewpoint)、目标点(target)。目标点是一个非渲染对象，它用来确定摄像机的观察方向，摄像机的视点总是盯着目标点。视点目标点的连线称为视线。

摄像机在创建后就被指定了默认的参数。这些参数包括：

- Lens(镜头)和 FOV(视野)：镜头和视野是相关的，改变镜头的长度自然也会改变摄像机的视野。真正的摄像机，其镜头长度和视野是被约束在一起的，但是不

同的摄像机镜头配置有不同的视野和镜头长度比。影响视野的另外一个因素是图像的纵横比,一般用 X 轴方向的数值比 Y 轴方向的数值来表示。例如,如果镜头长度是20 mm,图像纵横比是2.35,那么视野将是 94° ;如果镜头长度是20 mm,图像纵横比是1.33,那么视野将是 62° 。人眼所能聚焦的角度是 48° ,因此为了符合人们的习惯,摄像机镜头一般使用48 mm左右^[122]。

- **目标距离(Target Distance):** 目标距离是摄像机到目标点的距离,可以通过改变这个距离来使目标点在视线(视点 to 目标点的连线)方向上靠近或者远离摄像机。当使用景深(depth of field)时,这个距离非常有用。

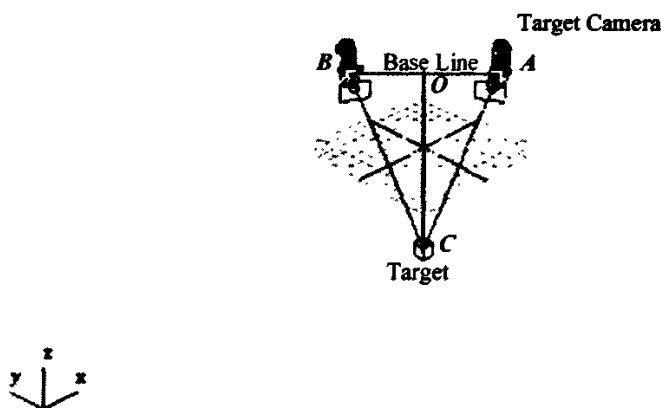


图 5.3 目标摄像机

Fig 5.3 Target Camera

5.4 立体效果

理论和实验表明,两个目标摄像机与三维模型的位置关系、两个目标摄像机之间的距离(这里等于基线^[11, 127]的长度)、会聚角变化值^[42]、舒适 Panum 融合区的限制,成像变形等等,是影响立体效果的重要因素,下面就定性地讨论这些因素。

5.4.1 目标摄像机与三维模型的位置关系

在3DS MAX中利用两个目标摄像机生成立体图像对时,这两个摄像机的镜头、视野和目标距离都要相同,它们的视点和目标点必须都在同一平面内。这样可以避免产生垂直视差^[42]。

根据目标摄像机与三维模型的位置关系, 视差可分为正视差、负视差和正负视差同时存在等三种情形。

① 当两个目标摄像机的两条视线相互平行或相交于目标点, 且三维模型在目标点的内侧时(如图5.4所示), 生成负视差的立体图像对。此时人眼感觉物体在监视器之前。直线 l 上的点具有零视差, 即显示时刚好重合, 因此这些点就成像在屏幕上。

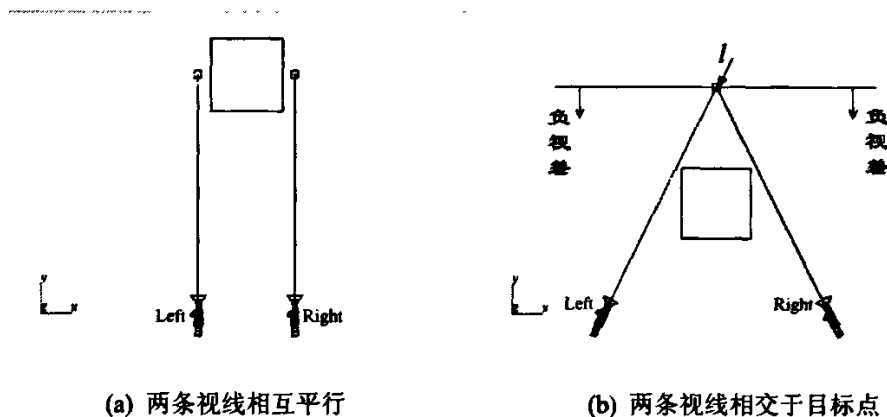


图 5.4 负视差的情形

Fig 5.4 A case of negative disparity

② 当两个目标摄像机的两条视线相交于目标点, 且三维模型在目标点的外侧时(如图5.5所示), 生成正视差的立体图像对。此时人眼感觉物体在监视器里面。

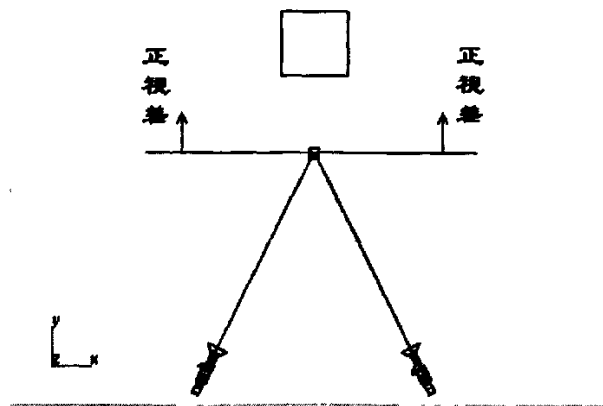


图 5.5 正视差的情形

Fig 5.5 A case of positive disparity

③ 当两个目标摄像机的两条视线相交于目标点，且目标点内外侧都存在三维模型时(如图5.6所示)，生成的立体图像对正负视差同时存在。此时人眼感觉物体镶嵌在监视器的表面。

通过实验可以验证，上述结论是正确的。

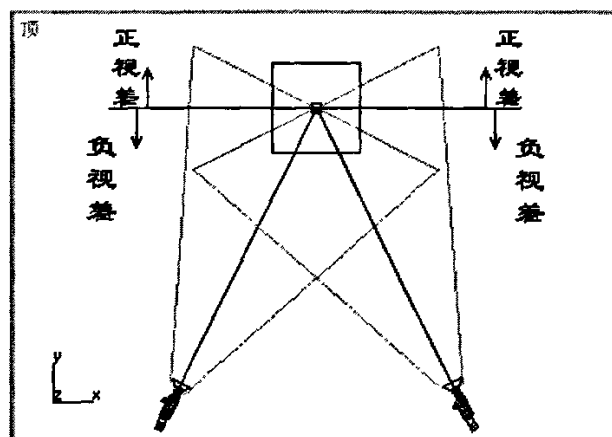


图 5.6 正负视差同时存在的情形

Fig 5.6 A case which negative disparity positive disparity are both exist

5.4.2 基线长度

基线长度(两个目标摄像机之间的距离)是影响立体效果的另一个重要因素。这是因为两个目标摄像机代表了双眼，而它们之间的距离就相当于瞳距。瞳距过大会使人眼聚焦困难从而降低甚至丧失立体效果。瞳距只有在一定的范围内才能使人产生视觉舒适的立体图像^[122]。因此需要仔细调整基线长度。

最佳基线长度与输出图像的监视器大小有关，也与人眼到监视器之间的距离有关。也就是说，生成的立体图像对的立体效果，与它们在显示设备上的偏差和观察距离有关。对于某一特定的立体图像对，在观察距离一定的情况下，在15英寸监视器上可能观看到的立体效果极佳，而在19英寸监视器上可能会丧失立体效果；同样，在给定监视器的情况下，在1 m的距离观看可得到极佳的立体效果，而在5 m的距离观看时也可能没有立体效果。而所成虚像的物理位置，可以根据公式(5.2)或公式(5.3)计算得到。

此外，在目标摄像机的目标点重合的情况下，最佳基线长度也与目标点到基线的距离和摄像机本身的镜头长度有关。如图5.3所示，摄像机和目标点围成 $\triangle ABC$ ， CO 为基线 AB 上的高，则 AB 的最佳长度与 CO 的长度和摄像机本身的镜头长度都有关系。

由此可见,基线长度与立体效果的关系较为复杂。在特定的监视器和观察距离的情况下,保证两摄像机的镜头长度不发生变化,采用多次渲染、多次试验的方法是确定摄像机最佳基线长度的一个可行方法。

本节采用这种方法讨论了在表5.1所示参数的情况下,基线长度的最佳值。

首先我们确定在目标点固定不变的情况下,最佳的基线长度。

实验步骤如下:

- ① 在模型的场景中加入两个目标摄像机,两个目标点的坐标均为(0, 0, 120),两个视点的坐标分别为(5, -1000, 400)、(-5, -1000, 400);
- ② 渲染两个摄像机拍到的模型的图像,得到此时的立体图像对;
- ③ 保持目标点固定不变,保持两个摄像机视点的Y坐标为-1000、Z坐标为400不变,改变X坐标,使两个摄像机向相反的方向对称地远离(如图5.7所示),从而增加基线长度;
- ④ 重复②、③步骤,渲染得到基线长度为10、20、30、.....、90、100、150、200、250、300时的模型立体图像对;
- ⑤ 使用分时双目立体成像系统观看得到的立体图像对,根据3DS MAX中模型的实际比例,寻找出与实际比例最接近的立体图像对,记录下其渲染时的基线长度,便是此时的最佳基线长度;

实验发现,随着基线长度从10开始增加,观看到的模型在水平方向上的不会产生明显的变形,而在垂直方向上的大小会随之产生变形(第5.4.4节将做具体的成像变形分析)。当观看到的模型在水平和垂直方向上的大小比例与实际模型的比例近似相同时,可认为此时的基线长度为最佳基线长度。实验确定的最佳基线长度为50,这个结果会因观察者的不同而略有区别。

通过多人多次实验,可得出结论:保持目标点固定不变,当采用表5.1的参数时,最佳基线长度为50(单位为通用单位)左右。此时 $AB/CO = 0.05$ 。

表 5.1 实验采用的参数

Table 5.1 Parameters used in the experiment

Parameters	Lens(mm)	Monitor Size	Resolution	P_d (cm)	CO长度 (Generic Unit)	目标点是 否重合
Value	50	17"	800 * 600	40	1000	是

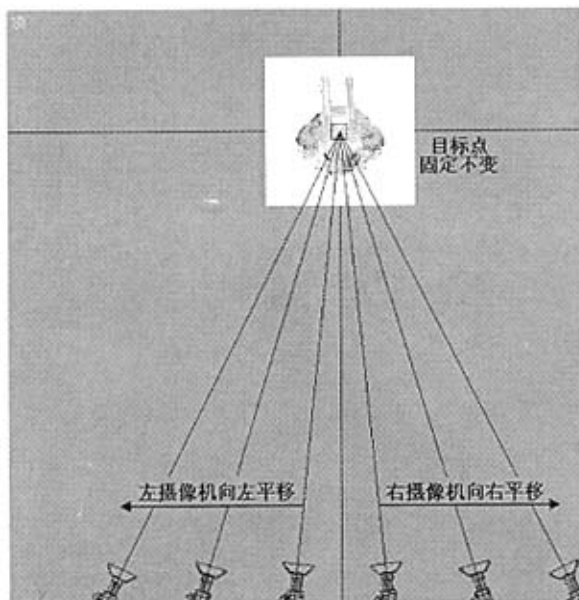


图 5.7 改变基线长度以确定最佳基线长度

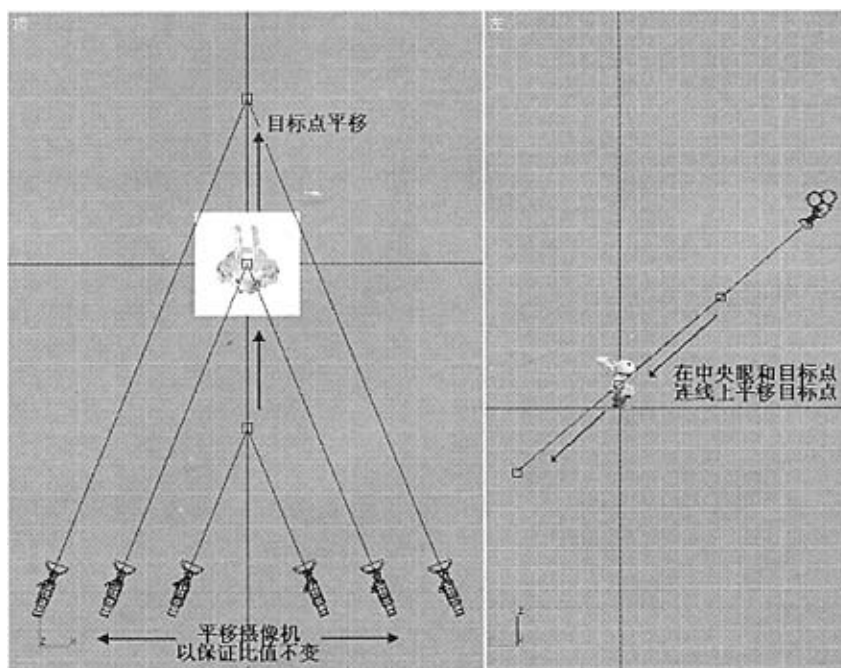
Fig 5.7 Change the length of the baseline to find the best length

接下来我们会自然而然地想到这样一个问题：在上述实验的基础上，保持 $AB/CO = 0.05$ 这一比例不变，让 AB 和 CO 同时增大，如图5.8所示，再重复渲染模型进行实验，立体感逼真程度会不会发生明显变化呢？

为此，接着上面的实验，设计具体的实验步骤如下：

- ⑥ 计算两个摄像机在距离为最佳基线长度时，基线长度 AB 和、目标点到基线的距离 CO 之比： AB/CO (这里 $AB/CO = 0.05$)；
- ⑦ 在中央眼到目标点的连线上移动两个摄像机的目标点，从而改变目标点到基线的距离；
- ⑧ 为保持 a) 步骤中的比值不变，我们保持两个摄像机的 Y 、 Z 坐标不变，只改变其 X 坐标，以调整基线长度；
- ⑨ 渲染两个摄像机的图像，得到两个摄像机此时拍到的模型左右图像对；
- ⑩ 重复⑦、⑧和⑨步骤，可以渲染得到两个摄像机的目标点分别移动到坐标为 $(0, -750, 330)$ 、 $(0, -500, 260)$ 、 $(0, 0, 120)$ 、 $(0, 500, -20)$ 、 $(0, 1000, -160)$ 的位置时(如图5.8所示)模型的立体图像对；

使用分时双目立体成像系统观看得到的立体图像对(输出图像的显示器大小固定为17英寸，人眼到显示器的观看距离固定为40cm)，观察不同图像的立体效果。

图 5.8 保持 AB/CO 的比值不变，平移摄像机和目标点Fig 5.8 Cases when the camera and the target translating while the value of AB/CO do not change

采用多人反复实验，反复评价，我们得到实验结果如表 5.2 所示。

表 5.2 实验结果

Table 5.2 Experiment results

目标点坐标	基线长度	立体感逼真程度是否发生变化
(0,-750,330)	12.5	否
(0,-500,260)	25	否
(0,0,120)	50	—
(0,500,-20)	75	否
(0,1000,-160)	100	否

从表 5.2 中可以看出，保持 $AB/CO = 0.05$ 这一比例不变，让 AB 和 CO 同时增大，立体感逼真程度不会发生明显变化；但是，所成的立体图像会经历正视差、正负视差同时存在和负视差三种情形。所成的像的大小取决于摄像机与模型的距离，而与模型的大小、模型在场景中的位置等其他因素无关。

5.4.3 成像位置的控制

在基于监视器的双目立体成像模型中,如果 p_d 和 I'_1 、 I'_2 的坐标已知,则根据公式(5.3)可以得到像 W 的坐标。也就是说,可以根据立体图像对中同一对象的视差来定量地控制对象所成虚像的位置(但要满足会聚角变化值与Panum融合区的限制),相关的实验也证实了这一点^[37]。如果是利用3DS MAX中的摄像机对象来生成立体图像对,则必须知道3DS MAX场景中从景物坐标系到图像坐标系的坐标变换,然后才能根据视差控制虚像位置。

摄像机对象简化了从景物坐标系到图像坐标系的坐标变换。如果3DS MAX的显示驱动器(display driver)选择的是OpenGL,则摄像机对象会按OpenGL的工作方式工作。OpenGL三维场景中从景物坐标系到图像坐标系主要经历如下所示的变换过程:



图 5.9 景物坐标系到图像坐标系的坐标变换

Fig 5.9 Coordinate transformation from image coordinate system to world coordinate system

其中关键的一步是投影变换。它类似于用摄像机调焦拍照。OpenGL 提供了两个透视投影函数: `glFrustum(GLdouble left, GLdouble right, GLdouble bottom, GLdouble top, GLdouble zNear, GLdouble zFar)`和 `gluPerspective(GLdouble fovy, GLdouble aspect, GLdouble zNear, GLdouble zFar)`,它们通过透视投影变换矩阵^[128]可以将眼坐标系(这里的眼坐标系是以视点为原点,以视线的方向为 Z 轴负方向的坐标系)中的坐标变换到投影平面(通常就是近裁剪面^[128])上的坐标,然后进行裁剪变换。眼坐标系可以看作是摄像机对象自身的坐标系。

如果知道了景物坐标系中点在眼坐标系中的坐标以及透视投影变换矩阵,便可求得其在图像坐标系中的坐标,从而能够计算并控制视差。但是,由第二章的双目立体成像相关理论可知,由于会聚角变化值、舒适 Panum 融合区的限制,生成的立体图像对的景深不是无限制的。我们不能将虚像成像在任意的位置。

5.4.4 成像变形分析

当两个完全相同的摄像机对象分别放置在不同位置来模拟人眼获取立体图像对时,对于图5.4(b)、图5.5和图5.6所示情况,两个摄像机的像平面互不平行;但是在成像时,这两个像平面的像将被显示在同一屏幕上。这种制作与播放的差异会

导致成像变形，严重的时候将出现复视。因此有必要分析成像变形对立体效果的影响。

设两摄像机分别放在图 5.1 中的 A_1 和 A_2 。为方便讨论，以 A_2 为原点建立左眼坐标系如图 5.10 所示，它是一种左手坐标系。

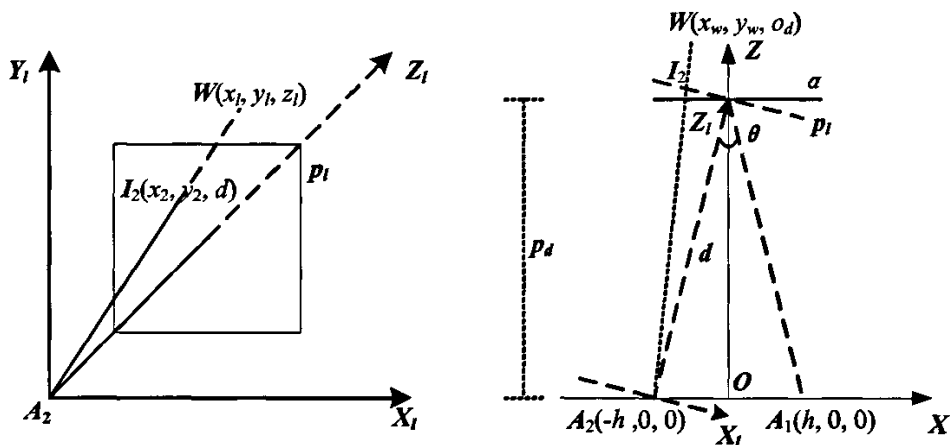


图 5.10 以 A_2 为原点的左眼坐标系

Fig 5.10 Left eye's coordinate system whose origin is A_2

成像时，左摄像机的像平面的像本来应该被显示在 p_l 的位置。因此，平面 p_l 是左眼坐标系中的投影平面。设 A_2 到 p_l 的距离为 d (视距^[69])，会聚角为 θ 。点 W 在左眼坐标系中的坐标为 (x_l, y_l, z_l) ， $I_2(x_2, y_2, d)$ 是 W 在 p_l 上的透视投影(单视点投影)。则有： $x_2 = x_l d / z_l$ ， $y_2 = y_l d / z_l$ ，对应的透视投影变换矩阵 T_l 为：

$$T_l = \begin{bmatrix} d/z_l & 0 & 0 & 0 \\ 0 & d/z_l & 0 & 0 \\ 0 & 0 & d/z_l & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

即 $[x_2 \ y_2 \ d \ 1] = [x_l \ y_l \ z_l \ 1] T_l$ 。事实上 OpenGL 等图形标准提供的函数所采用的透视投影变换矩阵要稍微复杂一些，但本质是相同的。例如调用函数 `glFrustum(l, r, b, t, n, f)` 将产生一个透视投影变换矩阵：

$$R^T = \begin{bmatrix} \frac{2n}{r-l} & 0 & 0 & 0 \\ 0 & \frac{2n}{t-b} & 0 & 0 \\ \frac{r+l}{r-l} & \frac{t+b}{t-b} & -\frac{f+n}{f-n} & -1 \\ 0 & 0 & -\frac{2fn}{f-n} & 0 \end{bmatrix}$$

该矩阵将 OpenGL 的眼坐标系从右手系变换成左手系, 最终将函数 `glFrustum()` 所定义的棱台变换成正方体视景体。为不失一般性, 这里采用 T_l 进行讨论, 它不会对结论产生较大影响。

为进行变形分析, 我们需要求出世界坐标系中任意一点 $W(x_w, y_w, o_d)$ 在左眼坐标系中的坐标。图 5.11 为世界坐标系向左眼坐标系的坐标变换过程。图 11(a) 为初始情况, 两个坐标系重合。坐标系变换经四步完成:

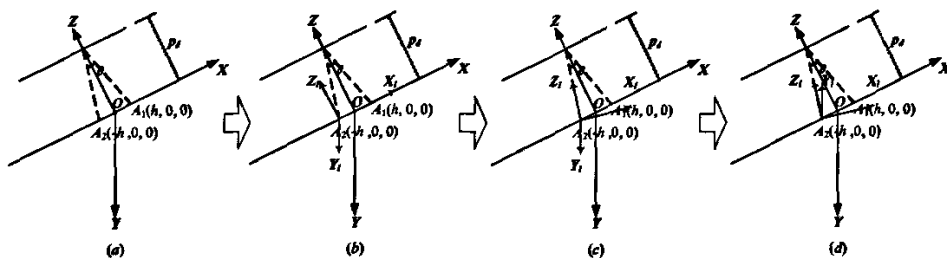


图 5.11 世界坐标系向左眼坐标系的坐标变换

Fig 5.11 Coordinate transformation from world coordinate system to left eye's coordinate system

① 将左眼坐标系原点平移至 A_2 , 如图 5.11(b) 所示。对应的变换矩阵为:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ h & 0 & 0 & 1 \end{bmatrix}$$

② 将左眼坐标系绕 Y_l 轴顺时针旋转 $\frac{\theta}{2}$, 角度的正方向规定如下: 用右手握住旋转轴, 大拇指指向旋转轴的方向, 则四指旋转握住旋转轴的方向就是角度的正方向, 如图 5.11(c) 所示。对应的变换矩阵为:

$$\begin{bmatrix} \cos \frac{\theta}{2} & 0 & \sin \frac{\theta}{2} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

③ 将左眼坐标系转换为左手系，即将 Y_l 轴反向，如图5.11(d)所示。对应的变换矩阵为：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

因此，可得世界坐标系向左眼坐标系的坐标变换矩阵：

$$T_{LView} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ h & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} & 0 & \sin \frac{\theta}{2} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta}{2} & 0 & \sin \frac{\theta}{2} & 0 \\ 0 & -1 & 0 & 0 \\ -\sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ h \cos \frac{\theta}{2} & 0 & h \sin \frac{\theta}{2} & 1 \end{bmatrix}$$

所以，

$$\begin{cases} x_l = x_w \cos \frac{\theta}{2} - o_d \sin \frac{\theta}{2} + h \cos \frac{\theta}{2} \\ y_l = -y_w \\ z_l = x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2} \end{cases} \quad (5.4)$$

进而可求出世界坐标系中的点在左眼坐标系中的透视投影的透视投影变换矩阵：

$$\begin{aligned}
 T_{LVP} = T_{LView} \bullet T_l &= \begin{bmatrix} \cos \frac{\theta}{2} & 0 & \sin \frac{\theta}{2} & 0 \\ 0 & -1 & 0 & 0 \\ -\sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ h \cos \frac{\theta}{2} & 0 & h \sin \frac{\theta}{2} & 1 \end{bmatrix} \begin{bmatrix} d/z_l & 0 & 0 & 0 \\ 0 & d/z_l & 0 & 0 \\ 0 & 0 & d/z_l & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} d \cos \frac{\theta}{2}/z_l & 0 & d \sin \frac{\theta}{2}/z_l & 0 \\ 0 & -d/z_l & 0 & 0 \\ -d \sin \frac{\theta}{2}/z_l & 0 & d \cos \frac{\theta}{2}/z_l & 0 \\ hd \cos \frac{\theta}{2}/z_l & 0 & hd \sin \frac{\theta}{2}/z_l & 1 \end{bmatrix}
 \end{aligned}$$

图 5.12 世界坐标系向右眼坐标系的坐标变换

Fig 5.12 Coordinate transformation from world coordinate system to right eye's coordinate system

类似地，可得世界坐标系向右眼坐标系变换的变换矩阵：

$$T_{RView} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -h & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} & 0 & -\sin \frac{\theta}{2} & 0 \\ 0 & 1 & 0 & 0 \\ \sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta}{2} & 0 & -\sin \frac{\theta}{2} & 0 \\ 0 & -1 & 0 & 0 \\ \sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ -h \cos \frac{\theta}{2} & 0 & h \sin \frac{\theta}{2} & 1 \end{bmatrix}$$

即点在右眼坐标系中的坐标 $[x_r \ y_r \ z_r \ 1] = [x_w \ y_w \ o_d \ 1] T_{RView}$ ，所以，

$$\begin{cases} x_r = x_w \cos \frac{\theta}{2} + o_d \sin \frac{\theta}{2} - h \cos \frac{\theta}{2} \\ y_r = -y_w \\ z_r = -x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2} \end{cases} \quad (5.5)$$

进而可求出世界坐标系中的点在右眼坐标系中的透视投影的透视投影变换矩阵:

$$\begin{aligned}
 T_{RVP} &= T_{RView} \bullet T_r = \begin{bmatrix} \cos \frac{\theta}{2} & 0 & -\sin \frac{\theta}{2} & 0 \\ 0 & -1 & 0 & 0 \\ \sin \frac{\theta}{2} & 0 & \cos \frac{\theta}{2} & 0 \\ -h \cos \frac{\theta}{2} & 0 & h \sin \frac{\theta}{2} & 1 \end{bmatrix} \begin{bmatrix} d/z_r & 0 & 0 & 0 \\ 0 & d/z_r & 0 & 0 \\ 0 & 0 & d/z_r & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} d \cos \frac{\theta}{2} / z_r & 0 & -d \sin \frac{\theta}{2} / z_r & 0 \\ 0 & -d / z_r & 0 & 0 \\ d \sin \frac{\theta}{2} / z_r & 0 & d \cos \frac{\theta}{2} / z_r & 0 \\ -hd \cos \frac{\theta}{2} / z_r & 0 & hd \sin \frac{\theta}{2} / z_r & 1 \end{bmatrix}
 \end{aligned}$$

由 T_{LVP} 和 T_{RVP} 可得世界坐标系中任意一点 $W(x_w, y_w, o_d)$ 在两个眼坐标系中的透视投影 $I_2(x_2, y_2, d)$ 和 $I_1(x_1, y_1, d)$, 其中:

$$\begin{cases} x_2 = \frac{d}{z_l} (x_w \cos \frac{\theta}{2} - o_d \sin \frac{\theta}{2} + h \cos \frac{\theta}{2}) = d \frac{x_w \cos \frac{\theta}{2} - o_d \sin \frac{\theta}{2} + h \cos \frac{\theta}{2}}{x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2}} \\ y_2 = -y_w d / z_l = -\frac{y_w d}{x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2}} \end{cases} \quad (5.6)$$

$$\begin{cases} x_1 = \frac{d}{z_r} (x_w \cos \frac{\theta}{2} + o_d \sin \frac{\theta}{2} - h \cos \frac{\theta}{2}) = d \frac{x_w \cos \frac{\theta}{2} + o_d \sin \frac{\theta}{2} - h \cos \frac{\theta}{2}}{-x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2}} \\ y_1 = -y_w d / z_r = -\frac{y_w d}{-x_w \sin \frac{\theta}{2} + o_d \cos \frac{\theta}{2} + h \sin \frac{\theta}{2}} \end{cases}$$

由于 $\tan \frac{\theta}{2} = \frac{h}{p_d}$ 以及 $\cos \frac{\theta}{2} = \frac{p_d}{d}$, 公式(5.6)可化简为:

$$\begin{cases} x_2 = \frac{d}{z_l} (x_w \cos \frac{\theta}{2} - o_d \sin \frac{\theta}{2} + h \cos \frac{\theta}{2}) = \frac{x_w p_d - o_d h + h p_d}{x_w h + o_d p_d + h^2} \sqrt{h^2 + p_d^2} \\ y_2 = -y_w d / z_l = -\frac{y_w (h^2 + p_d^2)}{x_w h + o_d p_d + h^2} \end{cases} \quad (5.7)$$

$$\begin{cases} x_1 = \frac{d}{z_r} (x_w \cos \frac{\theta}{2} + o_d \sin \frac{\theta}{2} - h \cos \frac{\theta}{2}) = \frac{x_w p_d + o_d h - h p_d}{-x_w h + o_d p_d + h^2} \sqrt{h^2 + p_d^2} \\ y_1 = -y_w d / z_r = -\frac{y_w (h^2 + p_d^2)}{-x_w h + o_d p_d + h^2} \end{cases}$$

下面进行成像变形分析:

设平面 p_r 是右眼坐标系中的投影平面。由对称性可知 p_l 、 p_r 和平面 YOZ 相交于一条直线 m , 会聚点 $A(0, 0, p_d)$ 在 n 上, $d = \sqrt{h^2 + p_d^2}$ 。屏幕 α 在世界坐标系中的方程为 $z = p_d$, 如图 5.13 所示。

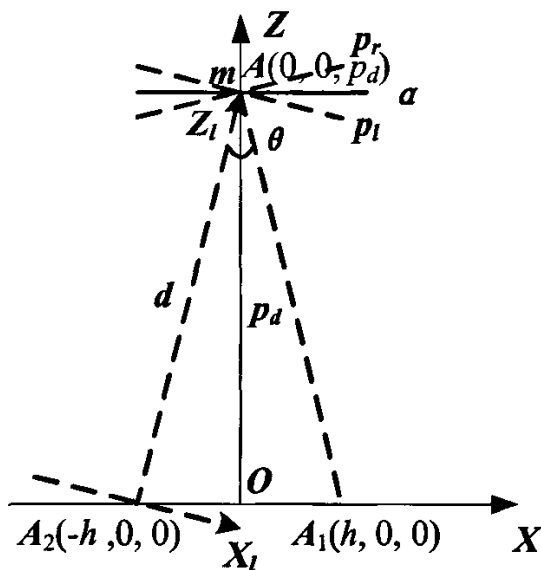


图 5.13 平面 p_l 和平面 p_r 的位置关系

Fig 5.13 Relation between plane p_l and plane p_r

通过计算平面 α 上点的透视投影, 我们可以得到成像变形的规律。

① 将会聚点 $A(0, 0, p_d)$ 的坐标代入公式(5.7), 得

$$\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}; \begin{cases} x_2 = 0 \\ y_2 = 0 \end{cases}$$

因此, 会聚点 A 的立体图像对在屏幕 α 中心重合, 水平视差和垂直视差均为 0。观看时, 其成像位置就在屏幕上, 如图 5.14 所示。

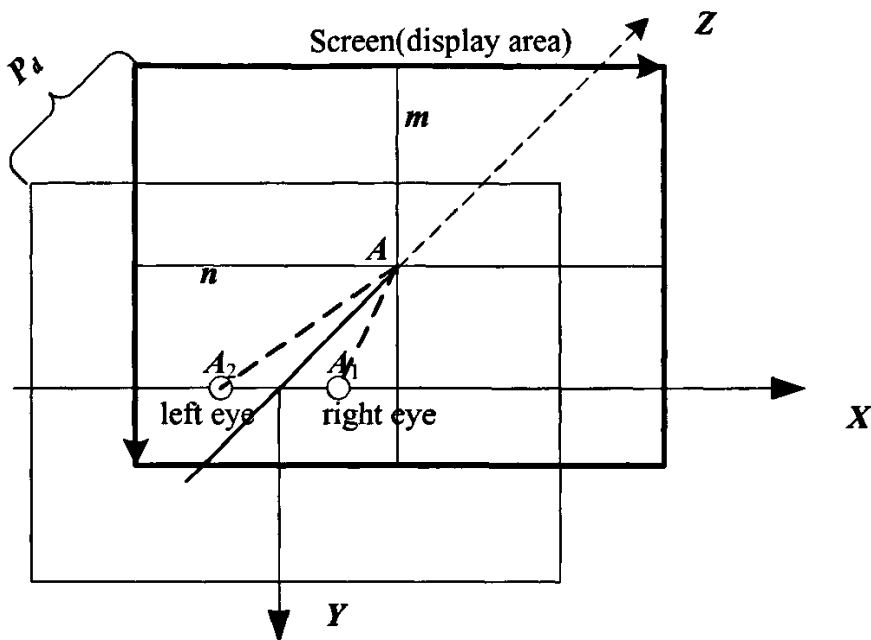


图 5.14 成像变形分析

Fig 5.14 Analysis of image distortion

② 将直线 m 上任意一点 $W(0, y_w, p_d)$ 的坐标代入公式(5.7), 得

$$\begin{cases} x_1 = 0 \\ y_1 = y_w \end{cases}; \begin{cases} x_2 = 0 \\ y_2 = y_w \end{cases}$$

因此, 直线 m 的立体图像对在屏幕 α 上也重合, 水平视差和垂直视差均为 0。观看时, 其成像位置也在屏幕上, 是一条过会聚点 A 的垂线, 这种情况与图 5.4(b) 所示一致。

③ 设直线 n 是 α 内过会聚点 A 的水平线, 将 n 上任意一点 $W(x_w, 0, p_d)$ 的坐标代入公式(5.7), 得

$$\begin{cases} x_1 = \frac{x_w p_d}{-x_w h + p_d^2 + h^2} \sqrt{h^2 + p_d^2} \\ y_1 = 0 \end{cases}; \begin{cases} x_2 = \frac{x_w p_d}{x_w h + p_d^2 + h^2} \sqrt{h^2 + p_d^2} \\ y_2 = 0 \end{cases}$$

因此, n 上 $x_w \neq 0$ 的点对应的垂直视差为 0, 但水平视差不为 0, 其值为:

$$D_H = x_1 - x_2 = \frac{2h p_d x_w^2 \sqrt{h^2 + p_d^2}}{(h^2 - x_w h + p_d^2)(h^2 + x_w h + p_d^2)} \quad (x_w \ll \frac{d}{h}).$$

因为 $d \gg h$, 所以 $\frac{d}{h} \rightarrow +\infty$, 故有 $x_w^2 < (\frac{d}{h})^2$ (d 为大于 0 的常数)。

$$\begin{aligned}
 x_w^2 &< \left(\frac{d}{h}\right)^2 \\
 \Rightarrow x_w^2 &< \left(\frac{d}{h}\right)^2 d^2 \\
 \Rightarrow x_w^2 &< \frac{(h^2 + p_d^2)^2}{h^2} \\
 \Rightarrow (h^2 + p_d^2)^2 - x_w^2 h^2 &> 0 \\
 \Rightarrow (h^2 - x_w h + p_d^2)(h^2 + x_w h + p_d^2) &> 0
 \end{aligned}$$

从而有 $D_H > 0$ ，即是正视差。也就是说，除了会聚点 A ，直线 n 上其他点成像在屏幕稍后处，如图 5.15 所示。

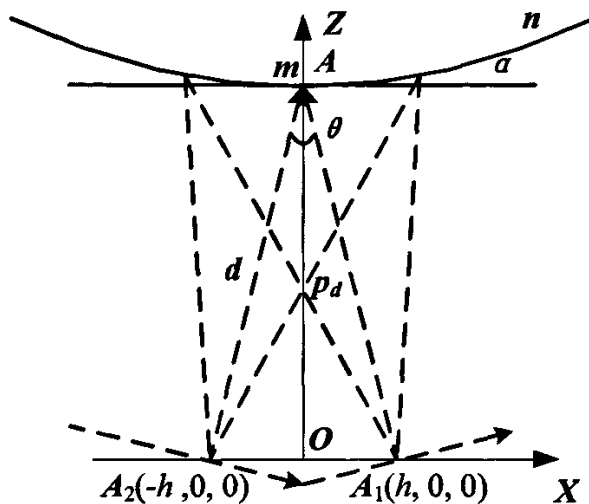


图 5.15 直线 n 的变形分析

Fig 5.15 Analysis of the distortion of line n

因为

$$D_H' = 2hp_d d \left[\frac{x_w^2}{d^4 - h^2 x_w^2} \right]' = \frac{4d^5 x_w}{(d^4 - h^2 x_w^2)^2},$$

所以当 $x_w > 0$ 时， $D_H' > 0$ ，函数 D_H 在 $[0, +\infty]$ 上单调增加；当 $x_w < 0$ 时， $D_H' < 0$ ，函数 D_H 在 $[0, -\infty]$ 上单调减少，即 $|D_H|$ 单调增加。这表明，直线 n 的立体图像将是一条从中间往屏幕两侧稍向后弯曲的曲线。

④ 对于 a 中的任意一点 $W(x_w, y_w, p_d)$ ，根据公式(5.7)，有

$$\begin{cases} x_2 = \frac{x_w p_d}{x_w h + p_d^2 + h^2} \sqrt{h^2 + p_d^2} \\ y_2 = -\frac{y_w (h^2 + p_d^2)}{x_w h + p_d^2 + h^2} \end{cases} \quad (5.8)$$

$$\begin{cases} x_1 = \frac{x_w p_d}{-x_w h + p_d^2 + h^2} \sqrt{h^2 + p_d^2} \\ y_1 = -\frac{y_w (h^2 + p_d^2)}{-x_w h + p_d^2 + h^2} \end{cases}$$

因此水平视差为

$$D_H = x_1 - x_2 = \frac{2h p_d x_w^2 \sqrt{h^2 + p_d^2}}{(h^2 - x_w h + p_d^2)(h^2 + x_w h + p_d^2)}$$

它的变化规律与③相同。

而垂直视差为

$$D_V = y_1 - y_2 = \frac{2h(h^2 + p_d^2)x_w y_w}{h^2 x_w^2 - (h^2 + p_d^2)^2}$$

人眼对垂直视差很敏感。如图 5.16 所示, 当纵向角视差 $\delta > 60'$ 时, 人眼就很难将左眼视图和右眼视图融合成单一立体图像^[43]。

因此要求 $\delta = |\arctg \frac{y_1}{d} - \arctg \frac{y_2}{d}| \leq \frac{\pi}{180}$ 。设监视器的显示区域为 282 mm×212 mm, $p_d = 400$ mm, $2h = 65$ mm。取屏幕的左上角一点 $W(-141.1, -105.8, 400)$, 代入公式(5.8), 得:

$$\begin{cases} d = 405.3 \\ y_1 = 100.2 \\ y_2 = 112.1 \end{cases}$$

所以点 $W(-141.1, -105.8, 400)$ 对应的纵向角视差 $\delta = 0.0026 \leq \pi / 180$, 点 W 能够融合成单一的立体的点。

类似的, 在上述参数下, 分析显示区域边缘上的其他关键点, 都能得到同样的结论, 也就是说, 对于 α 中的任意一点, 左眼视图和右眼视图在一定的参数下是能够融合成单一立体图像的。又由于会聚角变化值的限制, 立体图像的点不可能偏离 α 太多, 只能在 α 附近, 它们同样能够融合成单一立体图像。

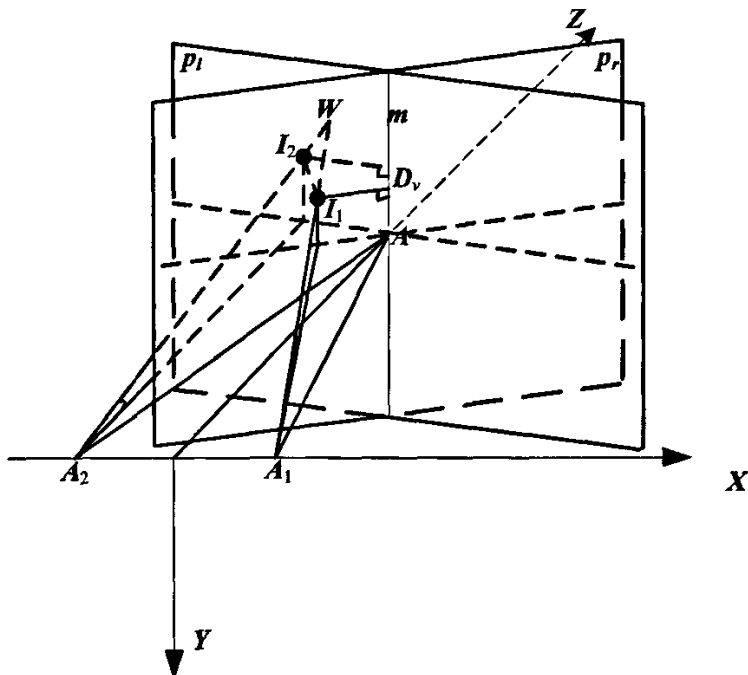


图 5.16 纵向角视差 $\delta = |\arctg \frac{y_1}{d} - \arctg \frac{y_2}{d}|$

Fig 5.16 Vertical angular parallax $\delta = |\arctg \frac{y_1}{d} - \arctg \frac{y_2}{d}|$

理论和实验表明，当两个完全相同的摄像机对象分别放置在不同位置来模拟人眼获取立体图像对时，有如下结论：

- 1) 采用分时双目立体成像系统进行立体成像时存在成像变形。
- 2) 虽然存在成像变形，但在一定的参数下，可以满足纵向角视差 δ 小于域值，从而大大减弱成像变形的影响，使人眼能够将立体图像对融合成单一的立体图像；在没有含心理深度暗示的背景时，成像变形更加不明显。
- 3) 该方法可用于测量精确度要求不高的虚拟现实。

5.5 立体视频的生成

由于动态的立体视频可看成是由多幅连续的静态立体图像对构成，因此，上述静态立体图像对的研究成果也可用于立体视频的生成。图5.17是已建三维模型的情况下立体视频生成的流程图。其中，计算左、右摄像机视点坐标是依据如下公式进行的：

$$(lx, ly, lz) = (cx, cy, cz) + \frac{dis \tan ce \times ratio \times (cy - ty, tx - cx, 0)}{2 \times \sqrt{(cy - ty)^2 + (tx - cx)^2}} \quad (5.9)$$

$$(rx, ry, rz) = (cx, cy, cz) - \frac{distance \times ratio \times (cy - ty, tx - cx, 0)}{2 \times \sqrt{(cy - ty)^2 + (tx - cx)^2}}$$

其中 (cx, cy, cz) 为中央眼处摄像机视点的坐标, (tx, ty, tz) 为中央眼处摄像机目标点的坐标, (lx, ly, lz) 、 (rx, ry, rz) 分别为左、右摄像机视点的坐标, $distance$ 为中央眼处摄像机的视点和目标点之间的距离, $ratio$ 为左、右摄像机的距离(基线长度)与 $distance$ 的比值。

公式(5.9)推导如图5.18所示。

在图 5.18 中, 直线 AB 是平面 α 的垂线, 点 A 为垂足, E_1E_2 为平面 α 上以 A 点为中点且平行于坐标面 XOY 的线段, 因此它们有如下关系:

$$AB = distance, \quad \frac{E_1E_2}{AB} = ratio$$

因 $E_1E_2 \parallel XOY$ 面, 故 $E_1E_2 \perp Z$ 轴; 又 $AB \perp \alpha$ 面, 且 $E_1E_2 \in \alpha$ 面, 故 $AB \perp E_1E_2$ 。由立体解析几何中向量乘法知识可知:

$$\overrightarrow{e_{AE_1}} = \frac{\overrightarrow{BA} \times \overrightarrow{e_z}}{|\overrightarrow{BA} \times \overrightarrow{e_z}|} = \frac{(cx - tx, cy - ty, cz - tz) \times (0, 0, 1)}{|\overrightarrow{BA} \times \overrightarrow{e_z}|} = \frac{(cy - ty, tx - cx, 0)}{\sqrt{(cy - ty)^2 + (tx - cx)^2}}$$

从而有:

$$\overrightarrow{AE_1} = |\overrightarrow{AE_1}| \times \overrightarrow{e_{AE_1}} = distance \times \frac{ratio}{2} \times \frac{(cy - ty, tx - cx, 0)}{\sqrt{(cy - ty)^2 + (tx - cx)^2}}。$$

因此:

$$\overrightarrow{OE_1} = \overrightarrow{OA} + \overrightarrow{AE_1} = (cx, cy, cz) + \frac{distance \times ratio}{2} \times \frac{(cy - ty, tx - cx, 0)}{\sqrt{(cy - ty)^2 + (tx - cx)^2}},$$

即为 E_1 点的坐标。同理可得 E_2 点的坐标为:

$$\overrightarrow{OE_2} = \overrightarrow{OA} - \overrightarrow{AE_1} = (cx, cy, cz) - \frac{distance \times ratio}{2} \times \frac{(cy - ty, tx - cx, 0)}{\sqrt{(cy - ty)^2 + (tx - cx)^2}}。$$

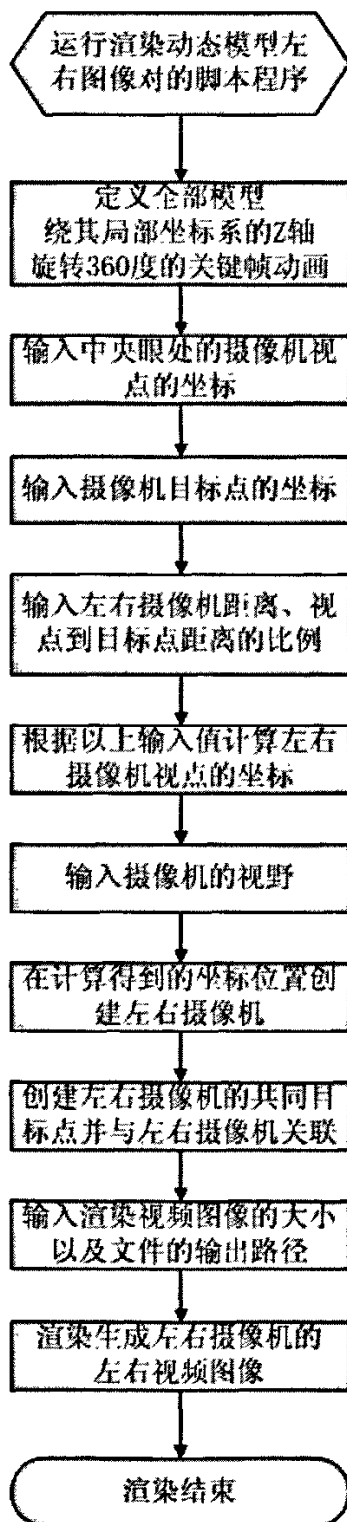


图 5.17 生成立体视频的流程图

Fig 5.17 Flow chart of stereoscopic video creation

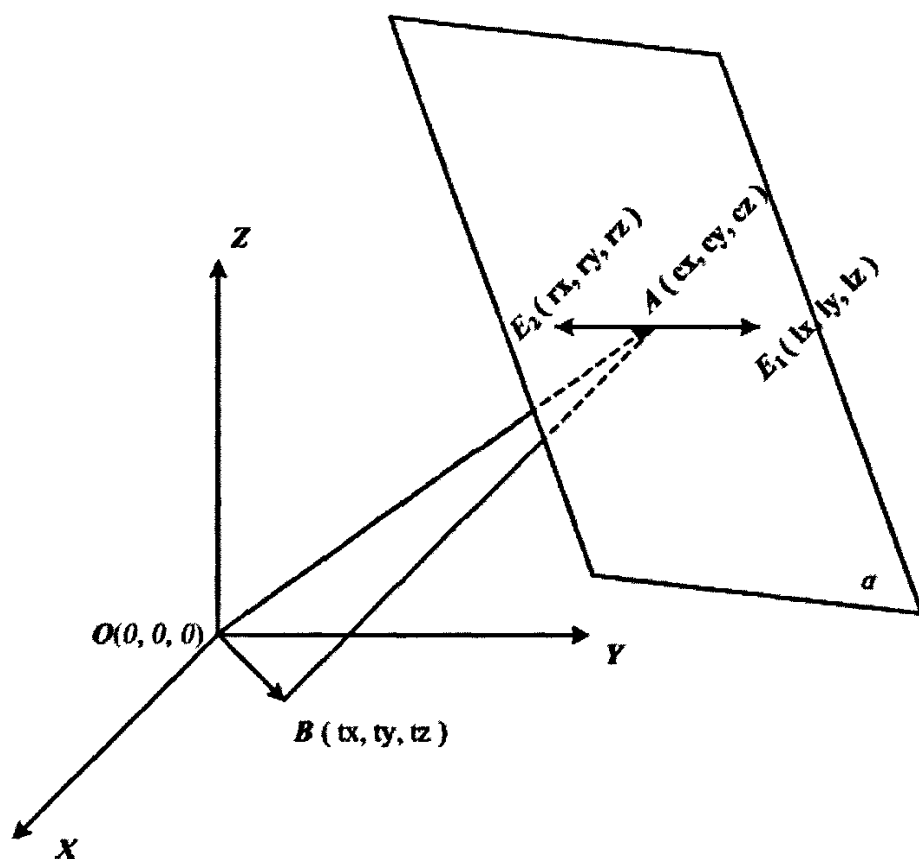


图 5.18 左右摄像机的位置关系

Fig 5.18 Positional relations between the left camera and the right camera

5.6 小结

本章介绍了已建三维模型的情况下立体图像对的生成的方法。

本章首先介绍了坐标变换,给出了景物坐标系到世界坐标系以及世界坐标系到图像坐标系的坐标变换矩阵。

其次介绍了 3DS MAX 中如何利用摄像机对象生成立体图像对。在三维场景已经建好的情况下,如果把两个完全相同的摄像机对象分别放置在不同位置,使它们之间的距离等于瞳距,再对同一场景进行拍照,就能生成立体图像对。

最后分析了影响立体效果的几个重要因素。包括目标摄像机与三维模型的位置关系、基线长度、成像位置的控制、成像变形分析等内容,得出的主要结论如下:

① 根据目标摄像机与三维模型的位置关系的不同,所成的立体图像可分为正视差、负视差和正负视差同时存在三种情形。

② 目前较少有文献对立体效果进行定量的评价,多数评价是定性的^[17,41]。在确定最佳基线长度的实验中,我们也采用了人眼观察的定性评价方法。实验发现,随着基线长度从 10 开始增加,观看到的模型在水平方向上不会产生明显的变形,而在垂直方向上的大小会随之拉伸。当观看到的模型在水平和垂直方向上的大小比例与实际模型的比例近似相同时,可认为此时的基线长度为最佳基线长度。实验确定的最佳基线长度为 50,这个结果会因观察者的不同而略有区别。在上述实验中,当基线长度达到最佳基线长度时, $AB / CO = 0.05$ 。保持这一比例不变,让 AB 和 CO 同时增大,再重复渲染模型进行实验,可以发现立体感逼真程度不会发生明显变化。

③ 如果知道了景物坐标系中点在眼坐标系中的坐标以及透视投影变换矩阵,便可求得其在图像坐标系中的坐标,从而能够计算并控制视差。但是,由于会聚角变化值、舒适 Panum 融合区的限制,生成的立体图像对的景深不是无限制的。我们不能将虚像成像在任意的位置。

④ 制作与播放的差异会导致成像变形。虽然存在成像变形,但在一定的参数下,可以满足纵向角视差 δ 小于域值,从而大大减弱成像变形的影响,使人眼能够将立体图像对融合成单一的立体图像。

本章研究了静态立体图像对的生成,详细讨论了它的立体效果。由于动态的立体视频可看成是由多幅连续的静态立体图像对构成,因此,静态立体图像对的研究成果也可用于立体视频的生成。本章给出了生成立体视频的流程图。

在没有三维模型的情况下,立体图像对的获取相对要复杂一些。第六章介绍了一种没有三维模型的情况下,从视频图像序列中提取立体图像对的方法。

6 基于单目视频图像序列的立体图像对的生成

6.1 引言

第4章介绍的改进的侯原理主要用于静态图像的立体化。随着立体成像显示设备的不断研发成功,立体展现将会成为虚拟植物可视化的主流显示方式,这种显示方式不但可以是静态的,还可以是动态的。比如,立体地展现植物的生长过程。因此,有必要研究视频的立体化。

如绪论中所述,现在可用的立体视频(stereoscopic video)大部分是通过模拟人眼观察世界的方式,采用两架精确定位的摄像机拍摄得到的。当双眼分别看到同一场景的两幅图像时,大脑能够很巧妙地将两眼细微的差别融合,产生具有空间感的立体场景。但是这要求两架摄像机具有完全一致的机械特征,并且对摄像机操作者在立体几何和摄像机定标方面有很高的技术要求。因此采集、制作与传输过程技术要求高、花费大^[129]。

自从第一张胶片发明以来,人类历史上积累了大量优秀的二维视频,因此一个比较可行的方式是将已有的二维视频转换成立体视频,已有一些文献对相关问题进行了探讨。Cassandra T. Swain通过在平面图像中添加单目暗示如模糊、阴影、明亮及遮挡来增强立体感^[130]。Yamada根据冷暖颜色理论产生深度数据^[131],文献^[132]中的作者利用单幅图与其相关深度图生成立体图像对。文献^[133]通过基于形态匹配的方法,在立体图像对库中提取适当的视差。Okino等根据运动对象的速度,采用MTD (Modified Time Difference)方法,为电视发明了一个立体图像转换系统^[134]。上述方法都各有优缺点,一般也比较复杂。本章在参考上述研究成果的基础上,提出了一种基于单目视频图像序列(video frequency image sequence)生成立体图像对的方法。该方法根据视频图像序列每帧中对象的相对位置,通过在帧间选择立体图像对的方式,将传统的单目视频图像序列转换成双目立体视频图像序列。其基本原理可用图6.1解释。

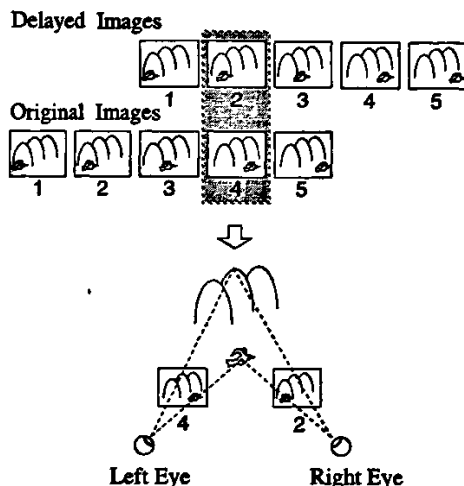


图 6.1 基于单目视频图像序列生成立体图像对的原理

Fig 6.1 Principle of stereo pairs creation based on video frequency image sequence

6.2 视差对深度感知的影响

在帧间选择立体图像对的重要依据是视差，它影响着人们的深度感知(depth perception)。因此，必须先讨论视差对深度感知的影响。

如图 6.2 所示，图中所标参数的具体含义参见 3.1 基于监视器的双目立体成像模型。由双目立体成像模型的性质 2 可知，水平视差

$$D = x_1 - x_2 = x'_1 - x'_2 = 2h \frac{o_d - p_d}{o_d}$$

设 E 是瞳距，则 $E = 2h$ 。设 $o_d - p_d = d_p$ ，则 d_p 是人眼感觉到的像相对于屏幕的深度，我们称之为视觉距离，它是深度感知的一种体现。

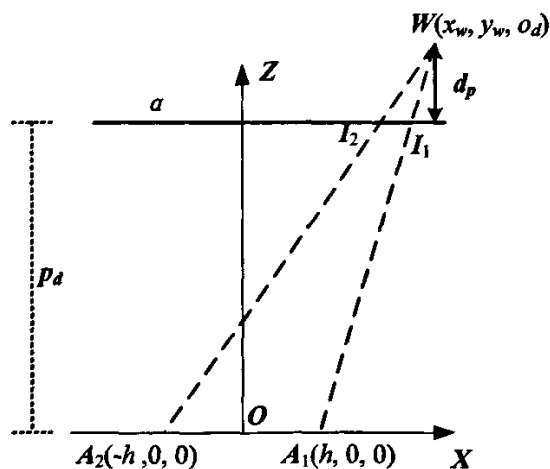


图 6.2 双目立体成像模型

Fig 6.2 Binocular stereo imaging model

因此, 视差和感知到的深度之间的关系可描述为

$$D = \frac{Ed_p}{p_d + d_p} \quad (6.1)$$

将 d_p 表示成 D 的函数形式

$$d_p = \frac{p_d D}{E - D} \quad (6.2)$$

设 p_d 和 E 是常量, 则视觉距离仅受视差变化影响。

当 $0 < D < E$ 时, 有

$$d_p = \frac{p_d D}{E - D} > 0 \quad (6.3)$$

此时是正视差的情形。虚像在屏幕之后。

当 $D > E$ 时, 有

$$d_p = \frac{p_d D}{E - D} < 0 \quad (6.4)$$

此时视线是发散的, 不能正确成像。

当 $D < 0$ 时, 有

$$d_p = \frac{p_d D}{E - D} < 0 \quad (6.5)$$

此时虚像在屏幕和观察者头部之间。公式(6.5)表明如果想要虚像成像在观察者头部和监视器之间, 视差必须小于零, 此时就是第三章所提到的负视差的情形。

考虑到

$$\lim_{D \rightarrow -\infty} d_p = -p_d \quad (6.6)$$

这说明, 负视差的绝对值越大, 所成虚像离屏幕越远, 反之则离屏幕越近, 但与屏幕最远不超过 p_d , 这是显而易见的。

为了方便实验我们重点考察负视差的情形。对负视差而言, 公式(6.5)可变为

$$|d_p| = \frac{p_d |D|}{E + |D|} \quad (6.7)$$

由公式(6.7)可得 $|D|$ 和 d_p 理论上的关系曲线图, 如图2所示。

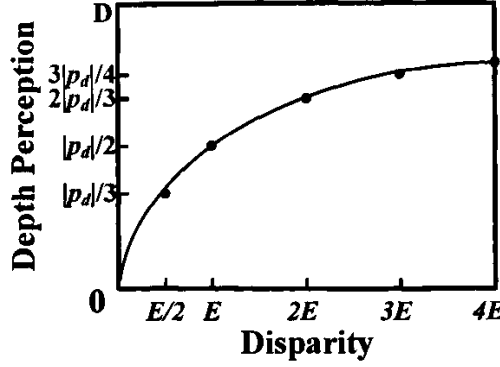


图6.3 视差和深度感知之间的关系

Fig 6.3 Relation between parallax and depth perception

可以发现,对负视差而言,当视差的绝对值较小时,曲率很大,表明此时深度感知对视差很敏感;当视差的绝对值变大时,曲线变得平滑,此时视差的变化对深度感知的影响变小。

类似地,对正视差而言,当视差较小时,曲率较小,表明此时深度感知对视差不敏感;当视差变得接近 E 时,曲率很大,此时视差的变化对深度感知的影响较大。

实际上, d_p 的值还应该满足会聚角变化值的限制。由公式(2.7)我们有

$$\Delta\theta = \alpha - \beta = 2(\arctg \frac{E}{2p_d} - \arctg \frac{E}{2(p_d + d_p)}) \quad (6.8)$$

其中, $-\frac{1.6\pi}{180} \leq \Delta\theta \leq \frac{1.6\pi}{180}$ 。

所以

$$-\frac{1.6\pi}{180} \leq 2(\arctg \frac{E}{2p_d} - \arctg \frac{E}{2(p_d + d_p)}) \leq \frac{1.6\pi}{180} \quad (6.9)$$

即

$$\tg(-\frac{0.8\pi}{180}) \leq \tg(\arctg \frac{E}{2p_d} - \arctg \frac{E}{2(p_d + d_p)}) \leq \tg(\frac{0.8\pi}{180}) \quad (6.10)$$

也就是

$$-1.2185 \leq \frac{Ed_p}{4p_d^2 + 4p_d d_p - E^2} \leq 1.2185 \quad (6.11)$$

根据公式(6.11)可以计算出 d_p 的取值范围。

在实际应用时,还应该考虑舒适 Panum 融合区(Panum's fusional area)的限制。

6.3 基于单目视频图像序列生成立体图像对

人类视觉具有很强的空间和时间感知能力。当追踪一个运动物体时，在每一时刻，双眼会看到视野中不同的部分，但却有可能在不同的时刻看到相同的场景。考虑一段通过摄像机运动拍摄到的静止场景的普通单目视频图像序列(monocular video frequency image sequence)，特别是当摄像机只有水平运动的时候，可以通过在序列中选择不同的帧组成立体图像对。基于这样一个事实，我们给出了一种基于单目视频图像序列生成立体图像对的方法：

对于给定的一段单目视频图像序列，我们将其作为左眼视频，通过以下两个步骤得到每帧对应的右眼视图：首先提取单目视频图像序列中的某一(些)前景对象，获得其在每帧中的位置信息；然后为每一帧在最小和最大帧间隔之间寻找右眼帧，使得对应两帧间的对象具有最适宜的视差，将该帧作为右眼视图和原来的帧一起构成立体图像对。

6.3.1 对象提取算法

基于单目视频图像序列生成立体图像对的方法首先要提取单目视频图像序列中的对象。对象提取是图像处理领域最难的课题之一，其性能直接影响到后续处理的准确性。现存在的一些全自动方法，大多是为特殊的图像处理服务的，并不具有通用性。本节采用一种结合了彩色边缘检测和种子区域增长的半自动算法。

为了减少图像噪声和初始种子的选择对提取质量的影响，首先对图像进行了中值滤波和量化处理；区域增长的每一步中，相邻像素是否合并到区域中是根据该像素点与已合并区域中心的距离以及局部边缘信息决定的。

为了包括进来更多潜在的边缘信息，我们采用改进的等方向性彩色边缘检测算子^[135]。

$(x-1, y-1)$	$(x-1, y)$	$(x-1, y+1)$
$(x, y-1)$	(x, y)	$(x, y+1)$
$(x+1, y-1)$	$(x+1, y)$	$(x+1, y+1)$

(a) 像素 8 邻域

1	2	1
0	0	0
-1	-2	-1

(b) 水平方向

1	0	-1
2	0	-2
1	0	-1

(c) 垂直方向

2	1	0
1	0	-1
0	-1	2

(d) 东北对角方向

0	1	2
-1	0	1
-2	-1	0

(e) 西北对角方向

图6.4 边缘检测算子

Fig 6.4 Edge detection operator

实验中采用 **RGB** 颜色空间，分别检测每个成分的边缘信息。一个边缘可能穿过水平、垂直、东北对角和西北对角四个方向，每个像素 (x, y) 的边缘强度用 $HE_i(x, y)$ 、 $VE_i(x, y)$ 、 $NE_i(x, y)$ 和 $SE_i(x, y)$ 表示，由 8 邻域内像素强度值加权计算得到：

$$\left\{ \begin{array}{l} HE_i(x, y) = |I_i(x-1, y-1) + 2I_i(x, y-1) + I_i(x+1, y-1) \\ \quad - I_i(x-1, y+1) - 2I_i(x, y+1) - I_i(x+1, y+1)| \\ VE_i(x, y) = |I_i(x-1, y-1) + 2I_i(x-1, y) + I_i(x-1, y+1) \\ \quad - I_i(x+1, y-1) - 2I_i(x+1, y) - I_i(x+1, y+1)| \\ NE_i(x, y) = |2I_i(x-1, y-1) + I_i(x, y-1) + I_i(x-1, y) \\ \quad - I_i(x+1, y) - I_i(x, y+1) - 2I_i(x+1, y+1)| \\ SE_i(x, y) = |I_i(x, y-1) + 2I_i(x+1, y-1) + I_i(x+1, y) \\ \quad - 2I_i(x-1, y+1) - I_i(x-1, y) - I_i(x, y+1)| \end{array} \right. \quad (6.12)$$

其中 $I_i(x, y) (i = 1, 2, 3)$ 分别表示像素 (x, y) 在 **R**、**G**、**B** 上的强度值。

像素 (x, y) 的局部最大边缘强度 $E_i(x, y)$ ，定义为上述四个边缘强度的最大值

$$E_i(x, y) = \max(HE_i(x, y), VE_i(x, y), NE_i(x, y), SE_i(x, y))$$

由于不同的视频对象，在颜色形状等各方面都具有不同的特性，因此为各种类型的对象寻找统一的规则，自动产生合适的种子，是非常困难的。这里采用交互方式，用户提示种子点的位置。区域增长的算法如下：

① 输入初始种子点 $S_k(x, y)$ ， $k = 1, \dots, n$ ；

② 依次判断每个与已合并区域相邻的像素 (x, y) ，计算其与区域中心的距离 $D_i(x, y) = |I_i^N - I_i(x, y)|$ 。 I_i^N 表示已合并区域在每个颜色分量上的强度均值。区域初始中心点由种子像素 $S_k(x, y)$ 决定： $I_i^1 = I_i(x, y)$ ， $i = 1, 2, 3$ 。

③ 将全局和局部边缘信息加权综合：

$$T_i(x, y) = \alpha D_i(x, y) + (1 - \alpha) E_i(x, y) \quad (6.13)$$

其中 $0 \leq \alpha \leq 1$ 为控制局部边缘梯度的权重, 可根据不同的图像进行调整。

④ 如果 $\max(T_1(x, y), T_2(x, y), T_3(x, y))$ 小于某个阈值, 则合并该像素, 并更新区域中心 $I_i^{N+1} = \frac{NI_i^N + I_i(x, y)}{N+1}$; 否则丢弃该像素。

⑤ 当相邻像素没有可以被合并的时候, 停止。

经过上述算法得到的对象区域边缘可能非常粗糙, 因此还有必要使用数学形态学中闭合运算来消除毛刺。最后得到对象的中心 (X_c, Y_c) 为:

$$\begin{cases} X_c = \frac{\sum_{(x,y) \in object} x}{\sum_{(x,y) \in object} \delta(x, y)} \\ Y_c = \frac{\sum_{(x,y) \in object} y}{\sum_{(x,y) \in object} \delta(x, y)} \end{cases} \quad (6.14)$$

其中 $\begin{cases} \delta(x, y) = 1, & \text{如果像素}(x, y) \in object \\ \delta(x, y) = 0, & \text{如果像素}(x, y) \notin object \end{cases}$ 。

6.3.2 时序和帧间隔

基于单目视频图像序列生成立体图像对的关键问题是确定合适的时序和帧间隔, 以得到较适宜的视差。通过对由立体相机拍摄得到的原始立体图像对进行分析和实验, 我们发现, 采用前景对象对应点的视差等于-6 到-8(像素)的两帧构成立体图像对, 能得到较为稳定的位于头部和监视器屏幕之间的虚像。即对于由某一帧构成的左眼视图 F_l , 如果其中对象的中心点坐标为 (X_{cl}, Y_{cl}) , 则在其附近寻找该点横坐标在 $X_{cl}-6$ 到 $X_{cl}-8$ 之间的帧作为右眼视图 F_r , 这样构成的立体图像对立体效果较好。值得注意的是, 由视差的正负可以判断时序, 从而可以判断出摄像机移动的方向。

例如, 由立体相机拍摄得到的原始立体图像对如图 6.5 所示, 每幅图像大小为 800×600 , 颜色分辨率为 RGB $256 \times 256 \times 256$ 。我们利用 6.3.1 节所提到的对象提取算法, 将图中的棱形框分别提取出来, 如图 6.6 所示。表 6.1 给出了每个棱形框的中心点坐标及其对应的视差。

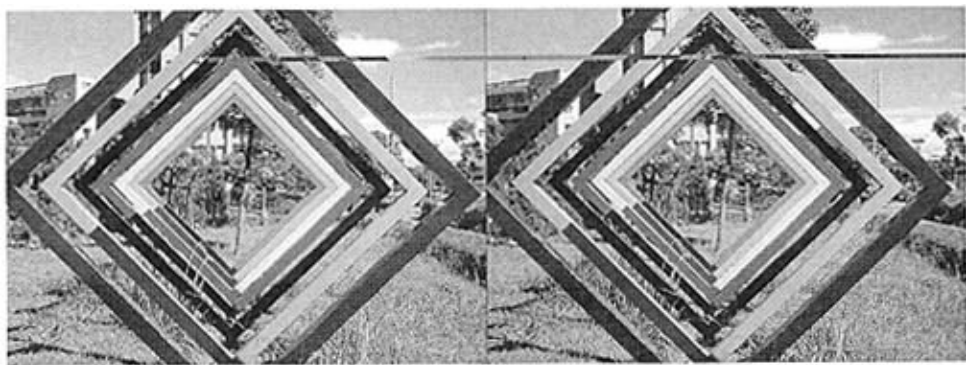


图 6.5 原始立体图像对

Fig 6.5 Original stereo pair

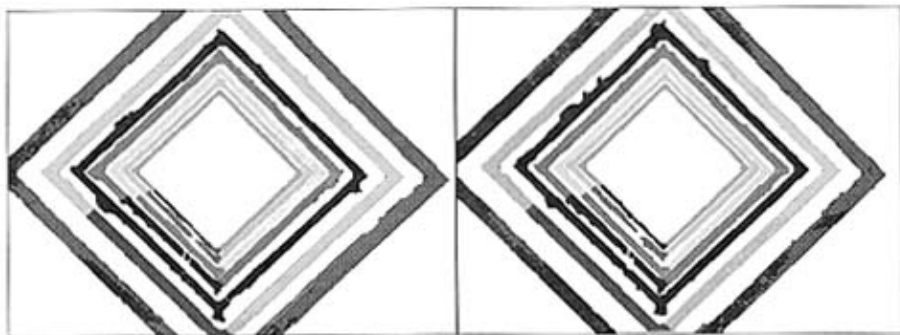


图 6.6 从立体图像对中提取出的对象

Fig 6.6 Objects extracted from the stereo pair

表 6.1 每个棱形框的中心点坐标及其对应的视差

Table 6.1 Coordinates of each square center point and their parallax

棱形框	左眼视图		右眼视图		视差	
	X_c	Y_c	X_c	Y_c	水平	垂直
1	383.3	300.4	379.0	302.3	-4.3	-1.9
2	381.6	300.9	377.2	299.6	-4.4	1.3
3	378.8	301.2	373.7	302.4	-5.1	-1.2
4	377.2	302.6	371.4	304.6	-5.8	-2.0
5	373.8	301.0	366.6	299.4	-7.2	1.6
6	378.6	297.0	367.2	296.2	-11.4	0.8

表 6.1 显示，该图像中引起人眼视觉映像的主要对象的水平视差大致在-11 到 -5 之间，从里到外对应棱形框的视差越来越大，所成的立体图像离屏幕也越来越

远。但是每一个棱形框都存在垂直视差，这是由于摄像机定标和对象提取算法的精确度所造成的。Sobel 通过实验得出：如果垂直视差在一定范围内，则对深度感知的影响非常小^[136]。

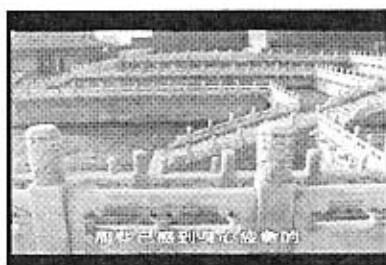
对其他立体图像对进行大量的实验和类似的统计，结果表明对应点的视差在-6到-8 之间的两帧构成立体图像对，能得到较为稳定的位于头部和监视器屏幕之间的虚像。事实的确如此。其他类型的立体图像分析后得到相似结果。

6.3.3 实验结果和结论

这里我们将一段有关中国传统园林的单目视频图像序列立体化。原始的单目视频图像序列作为左眼视图。

经过对象提取、中心点坐标计算、选择帧间隔、搜索最小间隔和最大间隔之间的帧等步骤，我们得到实验结果如图 6.7 所示。

图 6.7 给出了从第 31 帧到第 35 帧的左眼视图，以及程序为每一帧对应找到的右眼视图。将这些立体图像对分别放在分时双目立体成像系统中进行观察，发现可以产生一定程度稳定舒适的立体感。



第 31 帧



第 32 帧



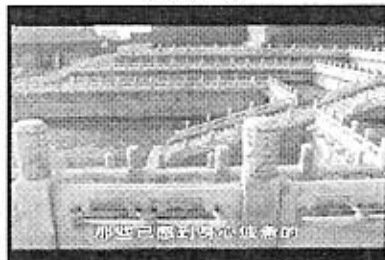
第 32 帧



第 33 帧



第 33 帧



第 35 帧

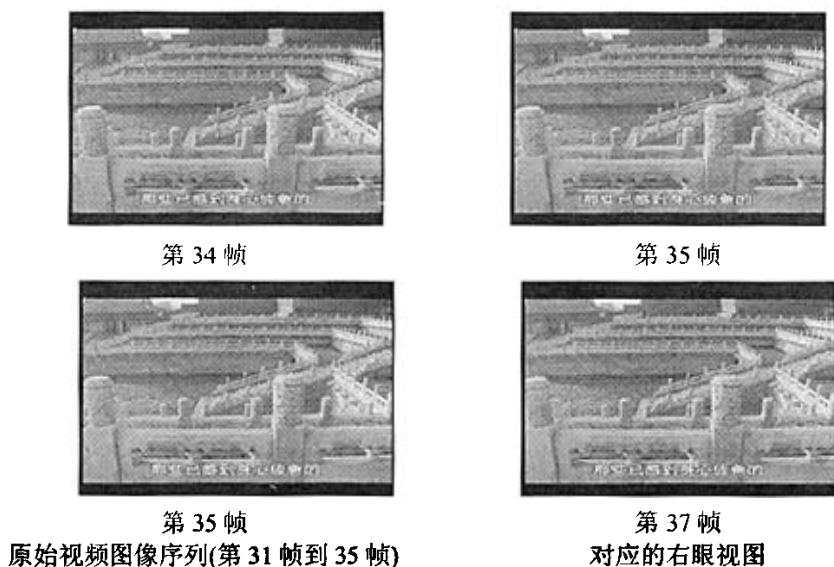


图 6.7 单目视频图像序列立体化

Fig 6.7 Stereo pairs created from the monocular video frequency image sequence

为了评价算法的稳定性,我们也对其他类型的视频进行了实验,结果显示当摄像机或场景中的对象进行近似水平方向上的运动时,立体图像对比较容易生成;当只存在垂直方向的运动时,或视频序列中有太多的关键帧时,立体图像对不容易找到,或者找到了但是立体效果差。

6.4 小结

利用单目视频图像序列产生立体图像无论是在学术还是商业应用都具有很大的前景。本章提出了一种方法将单目视频图像序列立体化。该方法根据前景目标在帧中的相对位置,获取时序和帧间隔,为每帧在序列中寻找合适的右(左)眼视图。

在帧间选择立体图像对的重要选择依据是视差,它影响着人们的深度感知。对负视差而言,当视差的绝对值较小时,曲率很大,表明此时深度感知对视差很敏感;当视差的绝对值变大时,曲线变得平滑,此时视差的变化对深度感知的影响变小。对正视差也有类似结论。同时,视觉距离还应满足会聚角变化值的限制。

提取单目视频图像序列中的前景对象是基于单目视频图像序列生成立体图像对的方法重要的一步。我们采用一种结合了彩色边缘检测和种子区域增长的半自动算法来实现对对象的提取。

单目视频图像序列立体化的另一关键步骤是确定合适的时序和帧间隔,以得到较适宜的视差。实验和分析表明,多数情况下,采用前景对象对应点的视差等

于-6到-8(像素)的两帧构成立体图像对,能得到较为稳定的位于头部和监视器屏幕之间的虚像。由视差的正负可以判断时序。

本章最后给出了一个基于单目视频图像序列生成立体图像对的实验。实验结果显示,当摄像机或场景中的物体做近似水平方向上的运动时,具有较好的效果;当只存在垂直方向的运动时,或视频序列中有太多的关键帧时,立体图像对不容易找到,或者找到了但是立体效果差。

如何消除垂直方向运动限制是一件不太容易的事。一种可行的方法是采用视图变形(View Morphing)技术将某一帧图像作为原始图像 I_0 , 其它帧 I_i 通过前置变换方法变换到同一个平行平面上,得到图像 I'_0 和 I'_i , 且满足扫描线特性。即在不改变摄像机光学中心的前提下,将两幅图像进行对齐^[59, 137, 138]。此外,还可以把深度和视差结合到立体化方法中。这些是我们下一步研究工作的重点。

7 总结与展望

7.1 总结

本论文研究了计算机视觉领域中的热点问题之一——基于计算机立体视觉的双目立体成像，在对双目立体成像相关技术及理论进行了研究的基础上，提出了一个基于监视器的双目立体成像模型，详细研究了该模型所具有的成像性质。然后针对该模型提出了三种立体图像对获取方法：基于侯原理的立体图像对的生成方法、已建三维模型的情况下立体图像对的生成方法，以及基于单目视频图像序列的立体图像对的生成方法。这些工作是计算机视觉与计算机图形学等多个学科的交叉，它涉及到计算机视觉、模式识别、人工智能、认知心理学、计算机图形学等领域的研究。本文绝大部分篇幅都用于阐述自己所做的工作，论文的创新性工作主要包括以下几点：

① 针对遮挡问题，本文提出了利用空间关系的定性表示与推理来表示物体之间的位置关系以及让计算机理解这种位置关系的方法，从而大大节省计算量。在空间关系的定性表示方面，我们给出了一个 SR 表示模型，并提出了距离关系的定性表示模型和方向关系的定性表示模型，以方便进行定性三角推理。在空间关系的定性推理方面，本文定义了基于 SR 表示模型的组合运算，给出了相应的组合运算表；构造了定性三角推理所需要的 QTR 组合运算表、距离定性减法表和角度定性减法表。这些研究工作将为我们今后深入遮挡问题的研究提供基础。

② 在成像模型方面，本文提出了一个基于监视器的双目立体成像模型，运用数学方法探讨了在该模型下双目立体成像的性质，给出了立体图像对坐标的计算公式、视差性质、像点立体成像规律以及直线段立体成像规律，并设计实验验证了相关的结论。理论的分析 and 实验表明，当观察者保持双眼平行地远离或靠近监视器屏幕，他看到的各点深度都相同的图像虽然也会移动，但不会产生变形；但是，各点深度不同的图像则会产生形变，并且遵循一定的成像规律；所成的虚像在负视差的情况下是可测量的。这些规律可以用于虚拟植物的可视化立体展现，甚至假如我们利用虚拟植物软件系统在特定的参数下生成了一颗秧苗的立体图像，我们可以通过实测的方法测量这颗秧苗在垂直方向上能长多高。

③ 论文提出了基于侯原理的立体图像对的生成方法，给出了评价立体效果的定量指标（交叉熵和均方根误差），详细讨论了立体化参数对立体效果产生的影响，并指出采用这种立体化方法时，平滑与立体感是互相矛盾的，人们只能取一个折中值。该方法适合于版画等主要含有心理深度暗示的平面图像的立体化。

④ 论文提出了已建三维模型的情况下立体图像对的生成的方法，详细研究了

影响立体效果的重要因素——目标摄像机与三维模型的位置关系、基线长度、成像位置的控制、成像变形分析等等，并设计实验得出了相应的成像规律。根据得出的结论我们开发出了一个立体图像对生成插件，该插件功能是根据已有模型自动选择合适的参数，生成静态的立体图像对或者动态的立体视频。

⑤ 论文提出了一种基于单目视频图像序列生成立体图像对的方法。该方法首先采用改进的边缘检测算法提取单目视频图像序列中的某一(些)前景对象，获得其在每帧中的位置信息；然后为每一帧在最小和最大帧间隔之间寻找右眼帧，使得对应两帧间的对象具有最适宜的视差，将该帧作为右眼视图和原来的帧一起构成立体图像对。这种方法可用于摄像机或场景中的物体做近似水平方向上的运动的情形。

7.2 展望

本论文研究的基于计算机立体视觉的双目立体成像是计算机视觉与计算机图形学等多个领域交叉的学科之一，这些研究具有较为广泛的应用前景。本学位论文的研究工作是虚拟植物可视化展现研究工作的重要组成部分，论文涉及的内容只是其中的一些关键问题。基于本文的研究，我们认为还有必要针对以下问题进行深入研究：

① 由于自由立体显示器无需佩戴任何眼镜就能看到立体效果，因此，它代表了未来立体显示设备发展的方向。而本文的研究成果都是建立在以标准的 CRT 监视器(Fishtank)作为显示设备的基础之上的。因此，有必要研究基于自由立体显示器的双目立体成像，这将是我们的下一步工作的主要内容之一。

② 空间关系的定性表示与推理为我们解决遮挡问题提供了一种新的思路和新的手段。我们对空间关系的定性表示与推理作了一定基础性的研究，但还有不少地方需要改进。例如，如何把定性三角推理的推理规则推广到三维，还需要进一步研究。我们下一步的工作就是要将已有的成果完善，并具体应用到遮挡问题中去。

③ 基于单目视频图像序列的立体图像对的生成方法只适合于摄像机或场景中的物体做近似水平方向上的运动的情形。如何消除垂直方向运动限制是一件不太容易的事。一种可行的方法是采用视图变形(View Morphing)技术将某一帧图像作为原始图像 I_0 ，其它帧 I_i 通过前置变换方法变换到同一个平行平面上，得到图像 I'_0 和 I'_i ，且满足扫描线特性。即在不改变摄像机光学中心的前提下，将两幅图像进行对齐。能这样做是因为同一场景不同视点的两幅图像之间存在对极几何关系，它的代数描述形式是基本矩阵 $F(3 \times 3$ 阶、秩为 2)。这个矩阵描述了图像之间的对应点约束关系。对应于同一场景中某一点的两个图像点 $p_0 \in I_0$, $p_1 \in I_1$ ，基本矩阵

F 满足如下关系:

$$p_1^T F p_0 = 0 \quad (7.1)$$

如果已知8对以上匹配点对, 就可以利用改进八点算法^[137]计算出基本矩阵 F 。而我们知道, 当两幅图像平面平行时, 对应点满足扫描线特性的充分必要条件是基本矩阵具有如下形式:

$$F' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (7.2)$$

设 $p'_0 \in I'_0$, $p'_1 \in I'_1$, 则有

$$p'_0 = H_0 p_0 \Rightarrow p_0 = H_0^{-1} p'_0 \quad (7.3)$$

$$p'_1 = H_1 p_1 \Rightarrow p_1 = H_1^{-1} p'_1 \quad (7.4)$$

因为 p'_0 和 p'_1 也是同一场景中某一点的两个图像点, 所以有

$$p_1^T F' p'_0 = 0 = p_1^T F p_0 \quad (7.5)$$

将等式(7.3)和等式(7.4)代入等式(7.5)可得

$$H_1^{-1} F H_0^{-1} = F' \quad (7.6)$$

满足等式(7.6)的矩阵 H_0 和 H_1 都可作为投影变换矩阵。有了投影变换矩阵, 我们就可以将 I_0 和 I_1 变换到同一个平行平面上, 从而减小垂直视差。然而, 如何很好的选取匹配点、如何寻找最佳的 H_0 和 H_1 、各参数对减小垂直视差的影响……这些都有赖于我们进一步的工作。

此外, 我们认为在一定情况下, 从一段视频序列中得到每帧对应的深度图, 把深度和视差结合到基于单目视频图像序列的立体图像对的生成方法中, 可能会是一种比较有效的消除垂直方向运动限制的方案, 这也是我们接下来要开展的研究工作。

致 谢

在即将完成博士论文之际，回首二十多年的求学生涯，我要衷心感谢许多老师、同学、朋友和亲人的帮助与鼓励。

我觉得我非常幸运，在攻读博士学位期间遇到了好导师。首先，我要衷心感谢我的导师朱庆生教授。他那高瞻远瞩的眼光、开拓创新的气魄、严谨求是的治学态度、博大精深的学识，将使我在今后的生活与科研中终身受益。朱老师的每一次指点总能把我们的研究推向深入。是朱老师给了我机会加入他所指导的虚拟植物可视化课题组，把我带入计算机视觉研究领域。朱老师谦逊、随和的人格魅力，勤奋不倦的治学态度，对学术前沿敏锐的把握和业已存在的学术建树深深影响了我。他给予学生宽松愉快的研究环境、众多的机会以及生活上细心的照料，使我能够全心投入自己的研究。

感谢重庆大学—重庆市软件工程重点实验室所有的师兄弟姐妹！在这个具有优良传统和学术氛围的实验室内，大家团结合作、积极探讨、勇于探索科学真谛，在科研上给了我很多启发和帮助。非常感谢课题组的师兄弟姐妹们的帮助与支持，与他(她)们之间的良好合作给我留下了深刻的印象。很荣幸能够在学术上与他(她)们一起共同成长。感谢朱老师所指导的这个研究群体所焕发出来的活力。

难以忘记我的硕士导师郭平副教授对我的一贯的大力支持与鼓励！

我还要感谢王茜副教授、傅鹤岗副教授对我论文工作的支持！

感谢杨小帆教授、熊忠阳教授、何中市教授等多年来对我无私的帮助！

衷心感谢李传东副教授对我的鼓励与支持！

非常感谢与我奋战在同一子课题的张晓芸、支丽欧和胡章平同学！

感谢周树语、周庆、胡月、朱郑州、向涛、叶莲、刘君等博士生对我的帮助！

衷心感谢Information Sciences主编J.J. Albertsma、加拿大Guelph大学Simon X. Yang教授、天津大学精密仪器与光电子工程学院谢洪波副教授、同济大学电子与信息工程学院杨珺博士对我论文撰写的帮助！

感谢奥地利Habbware提供的技术支持！

感谢我的亲人对我的一贯支持与鼓励。二十年多的求学生涯，他们对我的理解与支持始终激励着我不断前进，不敢懈怠！

特别感谢妻子易琳在生活和学习上给我的莫大支持与帮助！

最后，感谢所有以上还未提及的曾经帮助过我的师长、同学和好友！

谨以此文献给所有帮助过我和关心我的人！

刻 然

二〇〇七年四月 于重庆

参考文献

- [1] Tim Thwaites. Down on the virtual farm, 1995.
- [2] Xilin He, Jun Wang. Introduction to virtual reality and its applications in agriculture. Chinese Agricultural Mechanization, 2004, (1): 22-24.
何喜玲, 王俊. 虚拟现实技术及其在农业中的应用. 中国农机化, 2004, (1): 22-24.
- [3] Bedrich Benes. Virtual climbing plants competing for space. Geneva, Switzerland 2002.
- [4] Renzhao Jin, Zhaoyi Wang. Introduction to Virtual Reality and Applications in Agriculture. Journal of Tianjin Agricultural College, 2001, 8(2): 27-32.
靳润昭, 王兆毅. 虚拟现实及其在农业上的应用. 天津农学院学报, 2001, 8(2): 27-32.
- [5] Xiaoqing Li, Xinyuan Huang. Realization of tree creating software based on 3DS MAX. Agriculture Network Information, 2005, (7): 14-17.
李晓青, 黄心渊. 基于 3DS MAX 脚本的树木生成软件的实现. 农业网络信息, 2005, (7): 14-17.
- [6] Yan Guo, Baoguo Li. New advances in virtual plant research. Chinese Science Bulletin, 2001, 46(4): 273-280.
郭焱, 李保国. 虚拟植物的研究进展. 科学通报, 2001, 46(4): 273-280.
- [7] Yunfeng Li. The Research on Extraction from Leaf Image and Visual Realization of Virtual Plant [Dissertation for Doctor Degree]. Chongqing: Chongqing University, 2005.
李云峰. 叶图像提取研究及虚拟植物可视化实现[博士学位论文]. 重庆: 重庆大学, 2005.
- [8] Yunfeng Li, Qingsheng Zhu, Yukun Cao, Xiping He. A fast visual modeling plant based on images. Application Research of Computers, 2005, 22(11): 253-254, 257.
李云峰, 朱庆生, 曹渝昆, 何希平. 一种基于图像的快速虚拟植物可视化重建. 计算机应用研究, 2005, 22(11): 253-254, 257.
- [9] Cyril S, Francois S, Frederic B, Philippe de. A physiological plant growth simulation engine based on accurate radiant energy transfer. Technical report. 2001. Report No.: 4116.
- [10] Weilong Ding. Research of the agricultural expert system based on artificial plant growth model. Journal of Zhejiang University of Technology, 2005, 33(5): 525-533.
丁维龙. 基于虚拟植物生长模型的农业专家系统研究. 浙江工业大学学报, 2005, 33(5): 525-533.
- [11] Yujin Zhang. Image understanding and computer vision. Beijing: Tsinghua University Press; 2000.

- 章毓晋.图像理解与计算机视觉.北京:清华大学出版社, 2000.
- [12] Sun H. Rebuild 3D image of the contact surface of electrical apparatus using binocular 3D vision technique [Dissertation for Master Degree]. Hebei: Hebei University of Technology, 2002.
- 孙慧.用双目立体视觉技术重建电器触头立体图像[硕士学位论文].河北:河北工业大学, 2002.
- [13] Xiaobo Yang. Computer three-dimensional vision technique and its application in textile industry. Shanghai Textile Science & Technology, 2001, 29(2): 61-62.
- 杨小波.计算机立体视觉技术及其在纺织业中的应用.上海纺织科技, 2001, 29(2): 61-62.
- [14] Qi Li, Huajun Feng, Zhihai Xu, Yishi Han, Hongqiang Huang. Review of computer stereo vision technique. Optical Technique, 1999, (5): 71-73.
- 李奇, 冯华君, 徐之海, 韩一石, 黄红强.计算机立体视觉技术综述.光学技术, 1999, (5): 71-73.
- [15] Chunping Hou, Sile Yu. A novel method of picture conversion from 2D to 3D. Acta Electronica Sinica, 2002, 30(12): 1861-1864.
- 侯春萍, 俞斯乐.一种平面图像立体化的新方法.电子学报, 2002, 30 (12): 1861-1864.
- [16] Liping Zhou. Study on stereoscopic technique for virtual reality. Computer Applications, 1999, 19(4): 24-26.
- 周丽萍.虚拟现实中立体视觉的研究.计算机应用, 1999, 19(4): 24-26.
- [17] Qingsheng Zhu, Ran Liu, Xiaoyan Xu. Properties of a binocular stereo vision model. World Scientific Press; 2005 p. 831-834.
- [18] Mingyue Ding, Lixia Yang. 3D stereoscopic imaging and its application. Acta Electronica Sinica, 1995, 23(10): 124-128.
- [19] Somani R, Tennant M, Rudnisky C, Weis E, Ting A, Eppler J, et al. Comparison of stereoscopic digital imaging and slide film photography in the identification of macular degeneration. Can J Ophthalmol, 2005, Jun; 40(3): 293-302.
- [20] Sun H, Roberts DW, Farid H, Wu Z, Hartov A, Paulsen KD. Cortical surface tracking using a stereoscopic operating microscope. Neurosurgery, 2005, Jan, 56(1 Suppl): 86-97.
- [21] Jiang H, Chen WR, Wang G, Liu H. Localization error analysis for stereo X-ray image guidance with probability method. Med Eng Phys, 2001, Oct, 23(8): 573-581.
- [22] Ross MD. Medicine in long duration space exploration: the role of virtual reality and broad bandwidth telecommunications networks. Acta Astronaut, 2001, Aug, 49(3-10): 441-445.
- [23] Erlandsen SL, Greet BA, White J, Leith A, Marko M. High-resolution CryoFESEM of

- individual cell adhesion molecules (CAMs) in the glycocalyx of human platelets: detection of P-selectin (CD62P), GPI-IX complex (CD42A/CD42B alpha,B beta), and integrin GPIIbIIIa (CD41/CD61) by immunogold labeling and stereo imaging. *J Histochem Cytochem*, 2001, Jul, 49(7): 809-819.
- [24] Fitzke FW. Imaging the optic nerve and ganglion cell layer. *Eye*, 2000, Jun, 14 (Pt 3B): 450-453.
- [25] Richards RG, Wieland M, Textor M. Advantages of stereo imaging of metallic surfaces with low voltage backscattered electrons in a field emission scanning electron microscope. *J Microsc*, 2000, Aug, 199 (Pt 2): 115-123.
- [26] Galbraith W. Novel method for stereo imaging in light microscopy at high magnifications. *Neuroimage*, 1995, Mar, 2(1): 85-87.
- [27] Goh P, Tekant Y, Krishnan SM. Future developments in high-technology abdominal surgery: ultrasound, stereo imaging, robotics. *Baillieres Clin Gastroenterol*, 1993, Dec, 7(4): 961-987.
- [28] Mitchell TN, Robertson J, Nagy AG, Lomax A. Three-dimensional endoscopic imaging for minimal access surgery. *J R Coll Surg Edinb*, 1993, Oct, 38(5): 285-292.
- [29] Woodhouse JB, Eades JA. Comment on "Stereo-imaging made easy". *J Electron Microsc Tech*, 1990, Jan, 14(1): 85-86.
- [30] Zurek AK. Stereo-imaging made easy. *J Electron Microsc Tech*, 1989, Feb, 11(2): 174-177.
- [31] Turner JN. Introduction to stereo imaging. *Methods Cell Biol*, 1981, 22: 1-11.
- [32] William R.Sherman, AlanB.Craig. *Understanding Virtual Reality Interface, Application, and Design*. 1 edn. Beijing: Publishing House of Electronics Industry, 2004. (Translated)
- [33] William R.Sherman, Alan B.Craig. *Understanding virtual reality: interface, application, and design*. 1 edn. San Francisco, CA: Morgan Kaufmann Publishers, 2004.
- [34] David A.Forsyth, Jean Ponce. *Computer vision: a modern approach*. 1 edn. Pearson Education, Inc., 2004.
- [35] Chang Li, Zhengxing Zuo, Xinyuan Huang. Stereoscopy animation actualizing under practical virtual reality. *Computer Simulation*, 2002, 19(6): 84-86, 83.
- [36] Yang Liu, Xiaolin Zheng, Bing Li. An implementation of endoscope stereo vision. *Beijing Biomedical Engineering*, 2003, 22(3): 187-190.
刘洋, 郑小林, 李冰.内窥镜立体视觉的实现.北京生物医学工程, 2003, 22(3): 187-190.
- [37] Ran Liu, Qingsheng Zhu, Xiaoyan Xu, Jun Yang. Stereo effect of binocular stereo vision based on monitor. *Watam Press*, 2006 p. 1456-1460.
- [38] NVIDIA Corporation. *NVIDIA 3D Stereo User's Guide*. 2 ed, 2001.

- [39] Bo Yu, Liding Wang. Development of a binocular field sequential stereo microscopic imaging system. *Optics and Precision Engineering*, 1997, 5(4): 49-52.
- [40] Toni Emerson, Jerry Prothero, Suzanne Weghorst. *Medicine and Virtual Reality: A Guide to the Literature (MedVR)*, 2006.
- [41] Frank L.Kooi, Alexander Toet. Visual comfort of binocular and 3D displays. *Displays*, 2004, 25: 99-108.
- [42] Yulian Gu, Xuanping Cai. The implementing techniques of stereoscopic computer graphics. *Reference Periodical of National Defense Technology*, 1998, 19(1): 63-70.
顾郁莲, 蔡宣平. 计算机立体视图绘制技术. *国防科技参考*, 1998, 19(1): 63-70.
- [43] Xiaoge Yang. Microcomputer time-separating stereo animation implementation and stereo error analysis, 1996, p. 31-39.
- [44] Eaglesham BS, Lion LW, Ghiorse WC. An aufwuchs chamber slide for high-resolution confocal laser scanning microscopy and stereo imaging of microbial communities in natural biofilms. *Microb Ecol*, 2004, Apr, 47(3): 266-270.
- [45] Erlandsen SL, Greet BA, White J, Leith A, Marko M. High-resolution CryoFESEM of individual cell adhesion molecules (CAMs) in the glycocalyx of human platelets: detection of P-selectin (CD62P), GPI-IX complex (CD42A/CD42B alpha,B beta), and integrin GPIIbIIIa (CD41/CD61) by immunogold labeling and stereo imaging. *J Histochem Cytochem*, 2001, Jul, 49(7): 809-819.
- [46] Habbware. XIdMary Stereo Camera. <http://www.habware.at/duck6.htm>, 2003.
- [47] Xinyuan Huang, Shihong Chen. Implementation of stereodisplay animation generating module. *Computer Engineering*, 2002, 28(4): 178-179, 257.
- [48] Debevec P E, Taylor C J, Malik J. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image-based Approach, 1996.
- [49] Yu Y, Malik J. Recovering Photometric Properties of Architecture Scenes from Photograph, 1996.
- [50] Shum H Y, Szeliski R. Stereo Reconstruction from Multiperspective Panaramas, 1999.
- [51] Chris B, Michael B, Leonard Mc. Non-Metric Image-Based Rendering for Video Stabilization, 2001.
- [52] Chen S, Williams L. View Interpolation for Image Synthesis, 1993.
- [53] Chen S E. Quicktime VR: An Image-based Ap-proach to Virtual Environment Navigation, 1995.
- [54] Jianjun Wang, Qing Li. Image-based rendering technology. *Fire Control & Command Control*,

- 2003, 28(5): 328-331.
- [55] Ruiyuan Yang, Jianxiong Qiu. A survey of image-based modeling and rendering. *Journal of Computer-Aided Design & Computer Graphics*, 2002, 14(2): 185-188.
- [56] Bjaalie JG, Leergaard TB, Pettersen C. Micro3D: computer program for three-dimensional reconstruction visualization, and analysis of neuronal populations and brain regions. *Int J Neurosci*, 2006, Apr, 116(4): 515-540.
- [57] Daul C, Graebler P, Tiedeu A, Wolf D. 3-D reconstruction of microcalcification clusters using stereo imaging: algorithm and mammographic unit calibration. *IEEE Trans Biomed Eng*, 2005, Dec, 52(12): 2058-2073.
- [58] Xin Zheng, Wencheng Wang, Enhua Wu. Image-based rendering through layered texture mapping. *Journal of Software*, 2001, 12(11): 1647-1653.
- [59] Seitz S M, Dyer C R. *View Morphing*. NY, USA, 1996, p. 21-42.
- [60] Yingkai Guo. *Image generation and object recognition on geometry and illumination changing*, Shanghai Jiao Tong University, 2001.
- [61] Zengbo Qian, Zhengeng Qiu, Yongqiang Zhang. Reviewing the difficult problems facing the researchers of stereoscopic vision. *Journal of Institute of Surveying and Mapping*, 2001, 18(4): 267-272.
- [62] Zhu YM. A Java program for stereo retinal image visualization. *Comput Methods Programs Biomed*, 2007, Mar, 85(3): 214-219.
- [63] Huaizhou Yang, Yunbo Hua, Lusheng Xu, Zhiqian Yie. A new methodology research about realizing a stereoscopic system with RDS and liquid-crystal switches. *CHINESE J MED PHYS*, 1997, 14(1): 49-54.
- [64] Yuanqing Wang. Auto-stereoscopic Display Based on LCD. *Chinese Journal of Liquid Crystals and Displays*, 2003, 18(2): 116-120.
- [65] Guoying Zhao, Shihong Chen. Introduction of Autostereoscopic Display Technology. *J NORTH CHINA UNIV OF TECH*, 2002, 14(1): 51-56.
- [66] Fayun Liang, Shanxi Deng, Yongyue Yang. Optical Structure and Designing of Naked-eye Stereoscopic LCD Panel. *Chinese Journal of Liquid Crystals and Displays*, 2005, 20(6): 544-548.
- [67] Nick Holliman. Stereo3D Displays. <http://www.dur.ac.uk/n.s.holliman/Presentations.htm>, 2006, Available from: URL: <http://www.dur.ac.uk/n.s.holliman/Presentations.htm>
- [68] Peimin Hu, Huibiao Chen. The double colour method for the stereoscopic display of three-dimensional curves. *Computer Engineering and Applications*, 1999, 30-32.

- [69] Bing Li. Development of The Dual-eye Imaging Techniques for The Stereo Display of Virtual Endoscopes, Chongqing University, 2002.
- [70] Song CG, Park SK. Performance evaluation of stereo endoscopic imaging system incorporating TFT-LCD. J Med Eng Technol, 2005, Sep, 29(5): 244-250.
- [71] Kebin Li, Shiqi Li. The Newest Research of 3D Display. Computer Engineering, 2003, 29(12): 3-4.
- [72] Jungong Han, Zhaoyang LU. Methods for stereo video image compression. Journal of China Institute of Communications, 2003, 24(6): 113-123.
- [73] Yongdong Zhang, Guiling Li. Stereo Video Coding Based on Disparity Compensated Prediction. Signal Processing, 2001, 17(4): 335-339, 370.
- [74] Dimitrios Tzovaras, Nikos Grammalidis, Michael G.Strintzis. Joint three-dimensional motion/disparity segmentation for object-based stereo image. Optical Engineering Bellingham, 1996, 35(1): 137-144.
- [75] John E W Mayhew, John P Frisby. 3D Model Recognition from Stereoscopic Cues. London, England: The MIT Press Cambridge, Massachusetts, 1991.
- [76] S. B. Pollard, J. E. W. Mayhew, and J. P. Frisby. PMF: A Stereo Correspondence Algorithm using a Disparity Gradient Limit. Perception, 1985, 14: 449-470.
- [77] Victor S Grinberg, Gregg Podnar, Siegel M W. Geometry of binocular imaging, 1994, p. 56-65.
- [78] Alex Paul Pentland. A new sense for depth of field. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(4): 523-531.
- [79] Skerjanc Robert. Combined motion and depth estimation based on multiocular image sequences for 3D TV, 1994, p. 35-44.
- [80] Chunping Hou, Azimov Nurlan, Sile Yu. Mathematical Models of Stereoscopic Imagery System and Methods of Controlling Stereo Parallax. Journal of Tianjin University, 2005, 38(5): 455-460.
侯春萍, 阿陆南, 俞斯乐. 立体成像系统数学模型和视差控制方法. 天津大学学报, 2005, 38(5): 455-460.
- [81] Chunping Hou. A Research on Vision Stereoscopy of Planar Image School of Electronic Information Engineering, Tianjin University, 1998.
- [82] Marr D, Poggio T. Cooperative Computation of Stereo Disparity. Science, 1976.
- [83] Marr D, Poggio T. A Computational Theory of Human Stereo Vision. Proc Royal Soc London B, 1979.
- [84] Fua P. A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image

- Features. *Machine Vision and Applications*, 1993, (6): 35-49.
- [85] Kanade T, Okutomi M. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment. *Proc IEEE Trans Pattern Analysis and Machine Intelligence*, 1994, 61(9).
- [86] Ohta Y, Kanade T. Stereo by Intra- and Inter-Scanline Search using Dynamic Programming. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1985, (7).
- [87] Okutomi M, Kanade T. A Multiple Baseline Stereo. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1993, 15(4): 353-363.
- [88] Scharstein D, Szeliski R. Stereo Matching with Nonlinear Diffusion. *Int'l J Computer Vision*, 1998, 28(2).
- [89] Belhumeur P N, Mumford D A. A Bayesian Treatment of the Stereo Correspondence Problem using Half-Occluded Regions. *Proc IEEE Conf Computer Vision and Pattern Recognition*, 1992.
- [90] Intille S, Bobick A. Disparity-Space Images and Large Occlusion Stereo. *Proc European Conf Computer Vision*, 1994.
- [91] Geiger D, Ladendorf B, Yuille A. Occlusions and Binocular Stereo. *Int'l J Computer Vision*, 1995, 14.
- [92] Beveridge J R, Riseman E. Optimal geometric model matching under full 3D perspective. *CVGIP: Image Understanding*, 1995, 61(3).
- [93] Fisher A, Kolbe T, Lang F. Integration of 2D and 3D reasoning for building reconstruction. Bonn, Germany: 1997.
- [94] Collins Robert T. The Ascender System: Automated Site Modeling from Multiple Aerial Image. *Computer Vision And Image Understanding*, 1998, 72.
- [95] Jun Chen, Renliang Zhao. Spatial relations in GIS: a survey on its key issues and research progress. *ACTA GEODAETICA et CARTOGRAPHICA SINICA*, 1999, 28(2): 95-102.
- [96] Yabin Liu, Dayou Liu. A Review on Spatial Reasoning and Geographic Information System. *Journal of Software*, 2000, 11(12): 1598-1606.
- [97] Anthony G.Cohn, Brandon Bennett, John Gooday, Nicholas Mark Gotts. Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. *Geoinformatica*, 1997, (1): 1-44.
- [98] Wei Guo, Jun Chen. the Formal Description of Topological Spatial Relationship in 3d Based on Point Set Topology. *ACTA GEODAETICA et CARTOGRAPHICA SINICA*, 1997, 26(2): 122-127.
- [99] Han Cao, Jun Chen. Qualitative Description and Reasoning of Spatial Distance and Direction.

- Journal of Xi'an Petroleum Institute(Natural Science Edition), 2001, 16(1):68-72.
- [100] F.Wolter, M.Zakharyashev. Spatial reasoning in RCC-8 with Boolean region terms. Berlin: IOS Press, 2000, p. 244-248.
- [101] Longin Latecki, Ralf Rohrig. Orientation and qualitative angle for spatial reasoning, 1993, p. 1544-1549.
- [102] J.Liu. A Method of Spatial Reasoning based on Qualitative Trigonometry. Artificial Intelligence, 1998, 98: 137-168.
- [103] Ping Guo, Li Fan, Lian Ye. Obtain topological Relations from GIS spatial Database. Soft Computing as Transdisciplinary Science and Technology, 2005, 1109-1118.
- [104] Frank A.U. Qualitative spatial reasoning with cardinal directions. Journal of Visual Languages and Computing, 1992, (3): 343-371.
- [105] Nerode A. Some lectures on intuitionistic logic. Logic and Computer Science, 1990, 1429: 12-59.
- [106] H.Wang, S.McCloskey. A relation - algebraic approach to the region connection calculus. Theor Comput Sci, 2001, 255(1-2): 63-83.
- [107] S.Skiadopoulos, M.Koubarakis. Consistency Checking for Qualitative Spatial Reasoning with Cardinal Directions. London, UK: Springer-Verlag; 2002, p. 341-355.
- [108] Ghosh SK. Quantitative methods applied to stereo imaging. Application to single specimens. Methods Cell Biol, 1981, 22: 193-198.
- [109] Lehmann S, Bradley AP, Clarkson IV, Williams J, Kootsookos PJ. Correspondence-free determination of the affine fundamental matrix. IEEE Trans Pattern Anal Mach Intell, 2007, Jan, 29(1): 82-97.
- [110] Jiang H, Chen WR, Wang G, Liu H. Localization error analysis for stereo X-ray image guidance with probability method. Med Eng Phys, 2001, Oct, 23(8): 573-581.
- [111] Jiang H, Liu H, Wang G, Chen W, Fajardo LL. A localization algorithm and error analysis for stereo x-ray image guidance. Med Phys, 2000, May, 27(5): 885-893.
- [112] Kearfott KJ, Juang RJ, Marzke MW. Implementation of digital stereo imaging for analysis of metaphyses and joints in skeletal collections. Med Biol Eng Comput, 1993, Mar, 31(2): 149-156.
- [113] Ghosh SK. Quantitative methods applied to stereo imaging. Hardware and methods. Methods Cell Biol, 1981, 22: 177-192.
- [114] Ghosh SK. Quantitative methods applied to stereo imaging. Theory. Methods Cell Biol, 1981, 22: 155-176.

- [115] Jiang H, Chen WR, Wang G, Liu H. Localization error analysis for stereo X-ray image guidance with probability method. *Med Eng Phys*, 2001, Oct, 23(8): 573-581.
- [116] Huanming Wu, Yi Fang. A Measuring Technology Based on Image Processing and Computer Stereo Vision. *Journal of Engineering Graphics*, 2002, (4): 60-67.
吴焕明, 方漪. 基于计算机立体视觉的图像测量技术. *工程图学学报*, 2002, (4): 60-67.
- [117] Ruihe Zhang. Study on acquiring coordinates of objects in 3D world based on Computer Vision [Dissertation for Master Degree]. Nanjing: Nanjing Agricultural University, 2000.
张瑞合. 基于计算机视觉的空间目标定位系统研究[硕士学位论文]. 南京: 南京农业大学, 2000.
- [118] Lehmann S, Bradley AP, Clarkson IV, Williams J, Kootsookos PJ. Correspondence-free determination of the affine fundamental matrix. *IEEE Trans Pattern Anal Mach Intell*, 2007, Jan, 29(1): 82-97.
- [119] Bo Wang, Shensheng Zhang, Yue Huang. A Stereo Display Platform Based on VRML. *Computer Engineering and Applications*, 2003, (12): 136-227.
王波, 张申生, 黄越. 基于 VRML 文件的立体显示平台及相关技术研究. *计算机工程与应用*, 2003, (12): 136-227.
- [120] Indri A, Anna C, Anna C. Reconstructing Antique Stereo Pairs. <http://www.cs.washington.edu/homes/cjayant/finalproject/0-paper.html> 2006 Available from: URL: [http://www.cs. Washington.edu/homes/cjayant/finalproject/0-paper.html](http://www.cs.washington.edu/homes/cjayant/finalproject/0-paper.html)
- [121] Xinyuan Huang. *Virtual Reality: Technology of and Applications*. 1 edn. Beijing: Science Press, 1999.
- [122] Xinyuan Huang, Liangrui Tang. The generating method of parallax animation. *Journal of Computer-Aided Design & Computer Graphics*, 1999, 11(5): 409-411.
- [123] Xinyuan Huang, Chang Li. Image generation and image composition in parallax dynamic simulation. *Computer Simulation*, 1999, 16(3): 26-28.
- [124] Shejiao Xu. The coordinate transformation between the two coordinate systems of 3D graphic systems. *Journal of Xidian University*, 1996, Sep, 23(3): 429-432.
- [125] Shejiao Xu. Coordinate Transformation Based on Direction Cosine Parameters between the Object Coordinate System and the Word Coordinate System. *Journal of Engineering Graphics*, 2004, (1): 123-127.
- [126] Ping An, Zhaoyang Zhang. A fast algorithm for disparity correspondence based on geometric properties of stereo camera system. *Journal of Shanghai University (Natural Science)*, 2000, 6(5): 399-402.

- [127] Yicheng Jin, Yong Yin, Yi Wei, Zhenning Dong, Changwen Chen. Study on developing 3D view in marine simulator by using OpenGL. Journal of Dalian Maritime University, 1999, 25(2): 22-27.
- [128] Chios PS, Tan AC. Implementation of a dedicated optical fibre network for real time stereo imaging. Stud Health Technol Inform, 2000, 77: 1151-1155.
- [129] Cassandra T. Swain. Integration of Monocular Cues to Create Depth Effect. IEEE, 1997, 2745-2748.
- [130] K Yamada, K Suehiro, H Nakamura. Pseudo 3D Image Generation with Simple Depth Models, 2005, p. 277-278.
- [131] S. Curti, D. Sirtori, F. Vella. 3D Effect Generation from Monocular View, 2002, p. 550-553.
- [132] Yue Feng, Jianmin Jiang, Ipson SS. A Shape-Match Based Algorithm for Pseudo-3D Conversion of 2D Videos, 2005, p. 808-811.
- [133] T. Okino. New Television with 2D/3D Image Conversion Technologies. SPIE Photonic West, 1995, 2653: 96-103.
- [134] Jianping Fan. Automatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region Growing. IEEE Transactions on Image Processing, 2001, 10(10): 1454-1466.
- [135] Erik C. Sobel, Thomas S. Collett. Does Vertical Disparity Scale the Perception of Stereoscopic Depth? Proceeding: Biological Sciences, 1991, 244(1310): 87-90.
- [136] Hartley R. In Defense of the Eight-point Algorithm. IEEE Transaction on and Machine Intelligence, 1997, 19(6): 580-593.
- [137] Seitz SM, Dyer C R. Physically-valid View Synthesis by Image Interpolation. Cambridge MA, USA, 1995, p. 18-25.

附 录

A. 作者在攻读博士学位期间发表的论文目录

- [1] 朱庆生, 刘 然等.平面图像立体化研究.电子与信息学报.(已录用)
- [2] 刘 然, 朱庆生等.基于监视器的双目立体视觉的立体效果.同济大学学报. (已录用, EI 核心刊源)
- [3] R Liu, QS Zhu, XY Xu *et al.* Stereo Effect of Image Converted from Planar. INFORMATION SCIENCES. (已录用, SCI 核心刊源)
- [4] Qingsheng Zhu, Ran Liu, and Xiaoyan Xu. Properties of a binocular stereo vision model. Proceedings of the 11th Joint International Computer Conference , 831-834. 2005 (ISTP 检索号: BDF35)
- [5] Qingsheng Zhu, Ran Liu, Xiaoyan Xu, *et al.* Image conversion from planar into stereo: random variables. Proceedings of the 2006 International Conference on computational intelligence and security , 1793-1796. 2006. (ISTP 检索号: BFP93)
- [6] Ran Liu, Qingsheng Zhu, Xiaoyan Xu, and Jun Yang. Stereo effect of binocular stereo vision based on monitor. Proceedings of the International Conference on Sensing, Computing and Automation Added volume to DCDIS series B: Applications and Algorithms, 1456-1460. 2006, ISBN: 1492-8760
- [7] Qingsheng Zhu, Ran Liu, Xiaoyan Xu *et al.* Stereo effect of the stereo image created from a planar. Proceedings of the International Conference on Mechanical Transmissions, 1665-1669. 2006
- [8] 郭平, 刘然, 林勇, 董晓华.方向关系的定性表示与推理.计算机工程与科学, Vol.27(8), 2005: 81-84
- [9] 郭平, 刘然, 董晓华.二维空间对象空间关系的定性表示.计算机工程, Vol.31(2), 2005: 39-41. (EI 检索号: 05098867002)
- [10] 周树语, 许小艳, 刘然.基于 GIS 的分布式实时协同制图系统的研究. 计算机工程与设计, 2005, 26(1): 55-60
- [11] Ping Guo, Xiaoyan Xu, and Ran Liu. Binocular stereo vision model based on computer monitor. WAVELET ACTIVE MEDIA TECHNOLOGY AND INFORMATION PROCESSING, VOL 1, 405-412, 2006. (ISTP 检索号: BEZ27)
- [12] 朱庆生, 支丽欧, 刘 然等.平面图像立体化关键技术研究.计算机科学.(已录用)
- [13] Ping Guo, Xiaoyan Xu, Ran Liu. Spatial Reasoning Based on Qualitative Trigonometry. Proceedings of International Symposium on Robotics and Automation, ISRA-2006, 405-412

B. 作者在攻读博士学位期间参加的科研项目

- [1] SRFDP(高等学校博士学科点专项科研基金, 项目编号: 20050611027): 虚拟作物生长可视化关键技术研究;
- [2] Natural Science Foundation Project of CQ CSTC (重庆市自然科学基金, 项目编号: CSTC2006BB2229): 基于神经网络和小波的图像分割
- [3] Chongqing University Postgraduates' Science and Innovation Fund(重庆大学研究生科技创新基金, 项目编号: 200701Y1A0080194): 基于计算机立体视觉的双目立体成像研究