

双眼立体感知几何模型的研究*

梁 栋 韦 穗

周敏彤

(安徽大学电子工程与信息科学系,合肥 230039) (苏州大学电子系,苏州 215000)

摘 要 利用普通针孔成象的双摄像模型对双眼立体感知几何模型进行了研究,详细讨论了体视象空间的视觉重构,最后通过实验分析了影响体视的因素。

关键词 立体视觉,双摄像模型,体视因素,虚拟现实

1 引 言

与在真实世界中相同,在虚拟环境中,视觉接口为观察者提供了使其身临其境的信息。一般来说,人的视觉可分为形觉、光觉、色觉和深度觉(也称为立体视觉)。形觉是人眼对外界物体各部分不同亮度的分辨能力,光觉是人眼对外界辐射光的感受能力,色觉是人眼分辨颜色的能力,立体视觉则是人眼分辨物体空间位置(包括远近、前后、高低等)的能力,这是人们进行正常操作的重要条件。目前在虚拟现实环境中,虚拟环境大多是由双摄像机实现的,一对摄像机从现场获取三维视景的体视信息,然后经过若干处理,最后显示给操作者。本文利用普通针孔成象的双摄像模型对双眼立体感知几何模型进行了研究,对整个体视象空间的视觉重构进行了系统、精确的描述,最后通过实验分析了影响体视的因素。

2 双眼立体视觉

人利用双眼观察空间物体的成象特征为:当双眼注视空间某点时,该点成象在视网膜中心凹上,在双眼与此点构成的三角区外的所有空间点中,只要这个点在双眼视网膜中心凹的同一侧,即可被融合,成为单一象,否则就成为双象。由于眼球的不停转动,改变注视点,所以在日常生活中我们不易察觉双象的那部分空间在眼中的影响。人的视觉的一个重要特性,是能分辨两物体的相对位置,即估计空间物

体的相对位置。双眼的立体感觉(深度感)来源于以下两个因素:

(1) 双眼辐辏

如果不是在左右眼视网膜的相应点上成象,即使同时从外界获取信息,看到的还是复象,为使这种复象成为单象,眼球进行旋转运动,使双眼视轴集中于眼的前方一点,称为辐辏,产生这种眼球运动的眼外肌张力感和双眼的融合情况提供立体视觉信息。两眼视轴的夹角称为辐辏角,由其大小可判断物体的远近,20m以内为其作用范围。辐辏与调节有着密切的关系,从理论上说,二者应具有 $3.6^\circ/D$ 的耦合关系,而实际上它们之间的关系为 $2.3^\circ/D$,即对于短观察距离而言,人眼的关注点在焦点之后。一般视觉条件下,对于一个被中心凹凝视的物体,眼睛的焦长与会聚角是一致的。

(2) 双眼视差

双眼观察物体时,物体在两眼视网膜上的象略有差异,视差经视觉皮层融合,产生立体视觉。双眼视差的大小由双眼视差角的大小决定,双眼视差可根据图1推导得出。令 α 为双眼的眼距, D 为双眼到 P_1 点的视距, ΔD 为 P_1 、 P_2 两点间的深度, α_1 、 α_2 分别为双眼到 P_1 、 P_2 点的辐辏角, θ_1 、 θ_2 分别为双眼到 P_1 、 P_2 点的视差角,则双眼视差为:

$$\Delta\theta = \theta_2 - \theta_1 = \alpha_2 - \alpha_1 \approx \frac{a}{D} - \frac{a}{D + \Delta D} \approx \frac{a\Delta D}{D^2} \quad (1)$$

* 本研究得到国家“863”计划资助(No. 863-512-27-03)

收稿日期:1998-02-26

由式(1)可知,当视差 $\Delta\theta$ 大时,人眼就感到两物体间纵向相对距离大,即深度大, $\Delta\theta$ 小时,则两物体间纵向相对距离就小。从式中还可以看出,视差随距离而降低,即 D 越大,对于同样深度的两物体间的视差就越小。 $\Delta\theta_{min}$ 被称为体视锐度或体视灵敏度,它指的是当 $\Delta\theta$ 小到一定程度,人眼刚能感到两物体间距离差异时的视差。体视锐度的阈值因个人、图象在视网膜上的成象位置、与参照物的远近以及光照强度的不同而不同。远离中心凹,体视锐度的阈值就迅速增大,至偏于中心凹 8° 时,体视锐度变成 $300''$ 。通常人眼的体视锐度约为 $30'' \sim 60''$,经过训练可达到 $5'' \sim 10''$,甚至 $3''$ 。这个值意味着在双眼注视情况下能分辨 $6m$ 左右 2 个相差为 $20mm$ 物体的前后关系,但体视(即立体视觉)在 $10m$ 处开始下降,到 $135m$ 处就无效了。

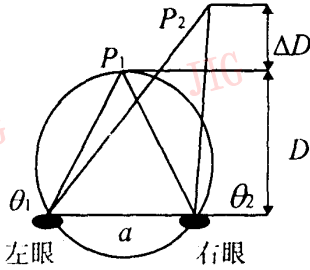


图 1 双眼视觉示意图

3 双摄像模型

在虚拟现实,由于体视是由双摄像机实现的,因此双摄像模型是体视模型的基础。为了得到物体 P 在两摄像机成象平面上的投影值,我们以摄像机的成象平面作为观察坐标面,按右手法则形成一个新的坐标系,物体 P 以此坐标系被投影到成象平面,如图 2 所示。

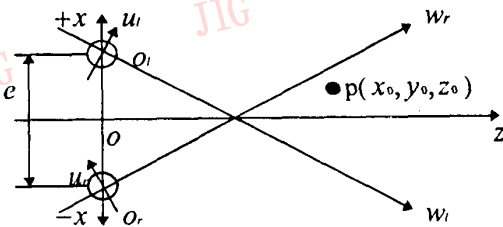


图 2 双摄像模型示意图

其中, $O_l - u_l v_l w_l$ 为左摄像机的观察坐标系; $O_r - u_r v_r w_r$ 为右摄像机的观察坐标系; $O - xyx$ 为用户坐标系; $p(x_0, y_0, z_0)$ 为物体在用户空间的坐标; e 为两摄像机的间距; θ 为两摄像机的夹角; v_l, v_r 与 y

轴平行。具体推导过程如下:

(1) 将用户坐标系的原点平移到左视点,即将坐标 $(0,0,0)$ 平移至 $(e/2,0,0)$,则平移变换矩阵为:

$$T_{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{e}{2} & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

(2) 令平移后的新坐标系绕 v_l 轴旋转一角度 $\theta/2$,则物体上的点 P 也旋转了 $\theta/2$,旋转变换矩阵为:

$$T_{12} = \begin{bmatrix} \cos(-\frac{\theta}{2}) & 0 & -\sin(-\frac{\theta}{2}) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(-\frac{\theta}{2}) & 0 & \cos(-\frac{\theta}{2}) & 0 \\ -\frac{\theta}{2} & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \cos(\frac{\theta}{2}) & 0 & \sin \frac{\theta}{2} & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\frac{\theta}{2}) & 0 & \cos \frac{\theta}{2} & 0 \\ -\frac{\theta}{2} & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

所以,以 $(e/2,0,0)$ 为原点,在 $O_l - u_l v_l w_l$ 坐标系下:

$$\begin{aligned} [u_l \ v_l \ w_l \ 1] &= [x_0 \ y_0 \ z_0 \ 1] T_{11} T_{12} = \\ &[(x_0 - \frac{e}{2}) \cos \frac{\theta}{2} - z_0 \sin \frac{\theta}{2}, y_0, \\ &(x_0 - \frac{e}{2}) \sin \frac{\theta}{2} + z_0 \cos \frac{\theta}{2}, 1] \end{aligned}$$

其中 (u_l, v_l, w_l) 为点 P 在左摄像机成象平面上的投影点坐标。由此可以得到点 P 在左摄像机显示平面上的坐标为:

$$x_{sl} = \frac{1}{S_x} \frac{(x_0 - \frac{e}{2}) \cos \frac{\theta}{2} - z_0 \sin \frac{\theta}{2}}{(x_0 - \frac{e}{2}) \sin \frac{\theta}{2} + z_0 \cos \frac{\theta}{2}} \quad (4)$$

$$y_{sl} = \frac{1}{S_y} \frac{y_0}{(x_0 - \frac{e}{2}) \sin \frac{\theta}{2} + z_0 \cos \frac{\theta}{2}} \quad (5)$$

式中, S_x, S_y 为摄像机在 x, y 方向的数字采样间隔。

同理,点 P 在右摄像机显示平面上的坐标为:

$$x_{sr} = \frac{1}{S_x} \frac{(x_0 + \frac{e}{2}) \cos \frac{\theta}{2} + z_0 \sin \frac{\theta}{2}}{z_0 \cos \frac{\theta}{2} - (x_0 + \frac{e}{2}) \sin \frac{\theta}{2}} \quad (6)$$

$$y_{sr} = \frac{1}{S_y} \frac{y_0}{z_0 \cos \frac{\theta}{2} - (x_0 + \frac{e}{2}) \sin \frac{\theta}{2}} \quad (7)$$

$$\frac{1}{S_y} \frac{-e \sin \frac{\theta}{2}}{(z_0 \cos \frac{\theta}{2} - \frac{e}{2} \sin \frac{\theta}{2})^2 - x_0^2 \sin^2 \frac{\theta}{2}} \quad (9)$$

4 体视象空间的视觉重构

双眼视差是人产生立体感(深度感)的最重要因素,立体显示正是利用了人的双眼视差因素。立体显示的主要特征是它产生了图象对应点之间的位差(即左、右 2 幅图象对应点之间的距离)。位差和双眼视差的概念类似,只不过位差是指显示屏上的差异,而双眼视差是指两眼视网膜之间的差异。位差可分为水平位差和垂直位差,垂直位差与深度恢复无关。立体显示的原理是通过体视设备使左、右眼看到两幅图象,由于 2 幅图象之间存在的水平位差在人眼中形成视差,使得原本在 2D 平面的图象被重构成一虚拟的立体空间。

假定空间上的任意一点 $P(x, y, z)$ 影射到显示平面上的体视点为 $L(x_{sl}, y_{sl})$ 和 $R(x_{sr}, y_{sr})$, 则操作者通过体视设备观察时,将重构出视觉空间中的象点 A , 由于体视点对的相对位置不同,则象点 A 的位置也不同,如图 3 所示。

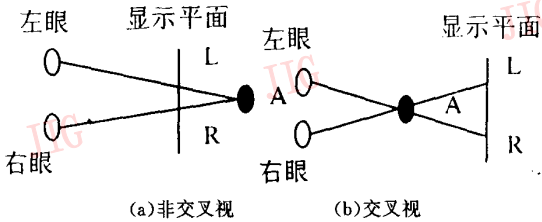


图 3 体视图象重构示意图

在图 3(a)中,立体成象点位于显示平面之后,此时视差被称为正视差,人感觉象点陷进了显示平面;在图 3(b)中,立体成象点位于显示平面之前,此时视差被称为负视差,感觉象点浮现出显示平面。不难看出,体视点对 L 和 R 的距离越远,即水平位差越大,所成立体象点距显示平面也越远,对于整幅图象而言,则纵深层次越强。

由式(4)、(5)、(6)、(7),我们可以得到:

$$\text{水平位差 } H = x_{sl} - x_{sr} =$$

$$\frac{1}{S_x} \left[\frac{-(x_0^2 - \frac{e^2}{4}) \sin \theta - z_0^2 \sin \theta - z_0 e \cos \theta}{(z_0 \cos \frac{\theta}{2} - \frac{e}{2} \sin \frac{\theta}{2})^2 - x_0^2 \sin^2 \frac{\theta}{2}} \right] \quad (8)$$

$$\text{垂直位差 } V = y_{sl} - y_{sr} =$$

由式(8),我们可以推导出当水平位差为零时的物体平面,这对分析物体的空间位置与重构深度的关系是很有意义的。令 $H=0$ 可得:

$$-(x_0^2 - \frac{e^2}{4}) \sin \theta - z_0^2 \sin \theta - z_0 e \cos \theta = 0$$

两边同除 $\sin \theta$ 并整理得:

$$(z_0 + \frac{e}{2} \tan \theta)^2 + x_0^2 = \frac{e^2}{4} \sec^2 \theta \quad (10)$$

由式(10)可知,在显示平面上为零视差而使象落在显示平面上的点,在物体空间中都落在 $(z_0 + \frac{e}{2} \tan \theta)^2 + x_0^2 = \frac{e^2}{4} \sec^2 \theta$ 这个柱面上,且这个面是 e 和 θ 的函数,即只与摄像机的位置有关。由于零视差点是落在一柱面上,而显示平面为一平面,则必然会引起体视重构上的失真现象。当 θ 越小时,柱面越接近于平面,由式(9)可知,垂直位差也越小,图象越容易融合。

5 实验结果及结论

双眼视差是立体显示中最重要的因素,但影响观察者立体感的还有许多其它因素,如观看距离、融合极限等等。为此,我们做了一系列实验来说明这些因素的影响。

实验 1 观看距离与深度感的关系

在实验过程中,让观察者(8人)在显示屏前每隔 1 米测一次图象重构后所在的空间位置,重构位置正确定位时,要保证图象完全被融合且位置不随时间改变,此时记录重构位置。实验所用图象对的水平位差为 1.4m,实验结果见表 1。

表 1

观差距离(m)	1	2	3	4	5	6	7	8
实测均值(cm)	16.4	30.5	50.0	62.3	62.4	61.3	65.0	65.8
理论值(cm)	17.7	35.4	53.2	70.9	88.6	106.3	124.1	141.8
误差(cm)	1.3	4.9	3.2	8.6	26.2	45.0	59.1	76.0

由上表可知:

- (1) 随着观看距离增加,可以对体视重构做准确定位的人越少;
- (2) 体视重构的深度在一定范围内随距离增大而显著递增,但在一定距离之后(在实验中为 4~5m)深度变化就不再明显,即观察距离不超过 5m 时,观察者有较强的深度感,能更准确地判断物体的

远近。

在实验中我们还发现,水平位差越大,体视重构深度显著递增的距离范围越大。

实验 2 水平位差融合极限的测试

我们在实验中发现对于小的水平位差不管站在什么位置都可以融合,此时融合极限的概念似乎没有了意义,但水平位差的绝对值(即 2 点之间的距离)却很直观地表现出了融合的情况,因此,用水平位差的绝对值来定义融合极限。结果见表 2。

表 2

观察距离(m)	1	2	3	4	5	6	7	8
水 2.8cm	√	√	√	√	√	√	√	√
3.7cm	√	√	√	√	√	√	√	√
平 4.6cm	4.0	5.5	3.8	√	√	2.0	1.3	4.4
5.5cm	5.0	6.0	5.5	1.4	√	2.4	1.6	5.0
位 6.2cm	5.8	×	7.0	2.4	×	3.8	3.6	6.0
6.8cm	×	×	×	6.2	×	6.5	3.6	6.5
差 7.4cm	×	×	×	×	×	×	×	×

√不管在任何位置均能融合 ×不管在任何位置均不能融合

由上表可知:

(1) 在水平位差小于 4cm 时,水平位差的融合是轻而易举的;

(2) 随着水平位差的增大,融合所需要的观察距离也要增大,每个人的融合极限与观察距离之间的关系不大,但不同的人有不同的值。

我们在实验中还发现:位差越大,则融合所需时间越长;图象一旦融合后,融合极限可以增大,但此时眼睛很容易疲劳。

实验 3 垂直位差的融合极限及与深度感的关系

表 3

观察距离(m)	1	2	3	4	5	6	7	8
垂 0.8cm	√	√	√	√	√	√	√	√
	无	无	无	无	无	无	无	无
直 1.1cm	3.4	2.3	√	3.0	2.0	×	√	4.1
	无	无	无	无	无	无	无	无
位 1.5cm	4.1	3.8	3.0	4.6	2.8	×	2.0	5.0
	无	无	无	无	无	无	无	无
差 1.7cm	5.1	5.2	3.4	4.5	3.8	×	3.6	5.0
	无	无	无	无	无	无	无	无

√不管在任何位置均能融合 ×不管在任何位置均不能融合

在实验过程中,让观察者看着一系列垂直位差不同的图象对,就每一幅图象对测出一可融合极限位置,并记录下来。所用图象对是一组水平位差为零且垂直位差变化的图象对。实验结果如表 3。

由表 3 可知:

- (1) 垂直位差不能产生深度感;
- (2) 垂直位差大于 1cm 时,融合起来较为困难;
- (3) 垂直位差不会引起不适的感觉。

6 结束语

本文对双眼立体感知几何模型及影响体视的因素进行了一些初步研究,目的是更好地利用视觉特征进行图象编码,并通过改善显示格式以更好地达到视觉匹配,为进一步开展深入研究打下坚实的基础。

参考文献

- 1 Kotulak J C. The effect of perceived distance on accommodation under binocular conditions. *Vision Research* 1995,35,791~795.
- 2 Andersen G J. Focused attention in three-dimensional space. *Perception and Psychophysics*,1980,9,651~662.
- 3 Kruger P B. Eye accommodation to head-up virtual image. *Human Factors*, 1988,30,689~702.
- 4 Krol J D. The double nail illusion; experiments on binocular vision with nail, needles and pins. *Perception* 1980,9,651~662.
- 5 Inoue T. Accommodation and convergence when looking at binocular 3D images. *Human Factors in organizational design and management* ■,1990,249~252.

梁 栋 副教授,1963 年 6 月生,1985 年毕业于安徽大学电子工程与信息科学系,1990 年在校获硕士学位。主要从事计算机视觉、图象处理和模式识别等方面的研究教学工作。先后参加或主持过国家 863、安徽省自然科学基金及其它科研开发等项目 7 项,其中一项获安徽省教委科技进步三等奖,已发表学术论文 10 多篇。





韦穗 博士生导师, 1946 年生, 现任安徽大学副校长、国家 863 计划智能机器人主题专家。长期从事计算机视觉、图象处理、模式识别和人工智能等方面的研究和教学工作。在美国进修期间因“邻域拟合算子”的研究获 R. B. Harxli 等三位著名教授的联合签名奖, 1993 年被中科院定为“最肯定、最急需支持的学科类和学术带头人”。已发表学术论文 20 多篇。

周敏彤 硕士, 1972 年 9 月生, 1993 年毕业于安徽大学自动化系, 1996 年在中科院合肥智能所获硕士学位, 现任苏州大学工学院助教。主要从事计算机视觉、模式识别等方面的研究和教学工作。在虚拟现实研究方面发表论文 3 篇。

Research of Geometric Model for Binocular Perception

Liang Dong, Wei Sui

(Dept. of Electronic Engineering and Information Science, Anhui university, Hefei 230039)

Zhou Mintong

(Dept. of Electron, Suzhou University, Suzhou 215000)

Abstract The geometry model of binocular perception are researched in this paper by utilizing double video camera model of normal needle aperture. It discusses visual reconstruction of stereographics vision image space on detail and analyzes the factors influencing stereographics vision by experiments finally.

Keywords Binocular perception, Double video camera model, Stereo vision factor, Virtual reality

虚拟现实(VIRTUAL REALITY, VR)

虚拟现实(VR)技术是一种高级的人与计算机交流的技术,它通过计算机生成一种模拟人工环境——即虚拟环境。虚拟现实技术的提出已有 10 多年,技术先进国家正式列入研究计划已有 8 年以上,90 年代开始有了 VR 商品系统,到 1993 年,其应用范围已涉及到军事、航天、医疗、通信、教育等诸多领域。在我国应用才刚刚起步。

VR 技术有 3 大特征,即临场感(Immersion),交互性(Interaction),想象性(Imagination),简称“3I”。

VR 技术与当今几项热门的技术诸如仿真技术、计算机图形、多媒体技术有很大的差异。仿真技术(Simulation)是利用计算机模拟实际环境进行科学实验的技术,从这点讲,二者有类似之处。但仿真技术一般仅以视觉和听觉为主要感知,很少用到触觉等其它感知。另外,仿真的视景不随用户的视点移动而实时变化,因此缺乏临场感。计算机图形技术

(Computer Graphics)是一个图形生成与显示的技术,它具有良好的实时交互性和一定的自主性,但不能把人的感觉与三维图形世界有机地联系起来。所以在感知方式与临场感方面与 VR 有很大的差距。多媒体技术(Multimedia)是利用计算机综合组织、处理和操作多媒体如视频、音频、图象、文字等信息的技术。虽然具有多种媒体,但感觉类别仍没有 VR 广泛,且缺乏临场感和交互性。总之,VR 技术与上述 3 种技术既不同又有联系。VR 是在众多相关技术的基础上发展起来的,是高技术发展的必须结果。VR 技术不同于以往其它任何技术,而是新技术的集成。它汇集了计算机图形图象、三维视觉、立体音响、人工智能、测控通信等方面技术以及生理学、心理学、认识科学等理论。VR 技术充分发挥各项领域新技术的优势,将它们融为一体,为人类开创了一个崭新的认识世界、改造世界的领域。