

# 视差式立体信息采集、感知几何模型及 舒适观看区研究综述

侯春萍,袁万鑫,沈丽丽

(天津大学电信学院,天津 300072)

**摘 要:** 近年来,立体影视产业蓬勃发展,基于视差的立体成像技术成为市场主流.但是,基于视差的立体成像技术存在深度畸变、视觉不舒适等问题.针对这些问题,本文研究了立体视频系统几何模型与舒适观看区.首先,总结并比较了立体采集系统的各种配置模式;接着讨论了平移式立体系统中拍摄空间到感知空间的映射关系,即畸变模型,并用形状、深度畸变因子解释了木偶剧效应和纸板效应;然后分析了舒适观看区,对其三种定义做了比较,并基于国际上普遍采用的 $1^\circ$ 视差计算得出了舒适观看区域的几何范围;最后,本文将畸变模型和舒适观看区相关联,总结了立体图像拍摄应遵守的基本准则.

**关键词:** 立体系统;畸变模型;木偶剧效应;纸板效应;舒适观看区;立体拍摄规则

**中图分类号:** TP391.41 **文献标识码:** A **文章编号:** 0372-2112(2013)03-0561-09

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2013.03.024

## The Acquisition, Perception Geometric Model and the Comfortable Viewing Zone in Stereoscopic Video System: A Review

HOU Chun-ping, YUAN Wan-xin, SHEN Li-li

(School of Electronic and Information Engineering, Tianjin University, Tianjin 300072, China)

**Abstract:** With the prosperity of stereoscopic industry, the disparity-based stereoscopic technology becomes more and more popular. But this technology may bring viewing discomfort and various image distortions such as depth plane curvature, depth non-linearity and so on. To analyze the origins, characteristics and relations of these perceptual issues, we investigated the stereoscopic distortion model and the comfortable viewing zone. This paper first summarized and compared three different configurations of the disparity-based stereoscopic video capturing systems, and focused on the parallel configuration system. The geometry of stereoscopic camera and display systems was presented. The mapping relationship, known as the distortion model, between the camera space and the viewing space was discussed in the parallel-shifting stereoscopic video systems. This distortion model is the basis of the stereoscopic information processing. The shape distortion factor and the depth factor were employed to explain the puppet theater effect and the cardboard effect. The comfortable viewing zone was considered to reduce the problem of visual discomfort and visual fatigue. Its three different representations were compared both qualitatively and quantitatively. At last, the distortion model and the comfortable viewing zone were combined to draw a conclusion of the stereoscopic shooting rules.

**Key words:** stereoscopic system; distortion model; puppet theater effect; cardboard effect; comfortable viewing zone; stereoscopic shooting rule

## 1 引言

目前,市场上的立体电视终端几乎都采用视差式立体成像技术<sup>[1]</sup>.众所周知,人左、右眼的水平位置(观看视角或视点)略有不同,这导致自然场景在双眼视网膜

上形成的二维图像也略有差异,即存在双目视差(binocular disparity).大脑就是利用这种双目视差从而形成对景物的深度感知.视差式立体成像技术,就是模拟人眼视觉系统感知深度的方式,分别给左、右眼提供带有水平视差的图像,即左眼图像和右眼图像,使观看者获得

深度感知,或称之为立体感知,简称立体感.立体感是指视觉系统感知三维物体空间距离的能力<sup>[2,3]</sup>.显然,双目视差是形成立体视觉最重要的客观因素.

舒适的立体感知和完美的深度信息再现,不仅与立体信息系统的采集参数(如摄像机焦距长度、摄像机水平间距、姿态等)有关,还与观看参数(如观看距离、屏幕尺寸大小等)有关<sup>[4,5]</sup>.图1给出了视差式立体信息系统从自然场景的拍摄空间到在观看者大脑中形成的感知空间的映射过程,其中任何一个单元的信息畸变和失真都可能影响立体视觉感知的舒适性和深度信息再现的完美性.

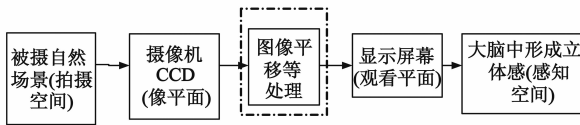


图1 视差式立体信息获取与立体视觉感知系统框图

## 2 视差式立体信息获取系统模型

### 2.1 立体信息采集系统配置模式分类

立体信息可以通过双摄像机拍摄<sup>[4]</sup>、多摄像机拍摄<sup>[6]</sup>和深度摄像机拍摄(time-of-flight 相机)<sup>[7]</sup>等多种方式获得,甚至可以从单幅图像中获取<sup>[8]</sup>.本文仅讨论最基本的视差式立体信息获取系统,即双摄像机系统,又被称为双视点摄像系统,双目摄像系统等.双视点摄像系统是模拟人眼从自然界真实三维景物中获取信息的方式,通过使用在水平方向上相隔一定距离的两部摄像机(与左、右眼相对应,分别称为左摄像机和右摄像机)来实现对三维场景,即拍摄空间信息的采集和记录.根据两部摄像机光轴的交点位置的不同,可将立体信息采集系统分为不同的配置模式.如果光轴夹角为零,即两部摄像机的光轴平行,称为平行(parallel)配置模式;如果光轴交于摄像机镜头前方,称为会聚(convergence 或内八字, toed-in)配置模式;如果光轴交于镜头后方,称为发散(divergence)配置模式<sup>[9]</sup>.在平行配置模式中,根据是否平移图像传感器或者是否在后期处理中平移图像,又分为平移模式和非平移模式.

发散配置模式因不符合人眼的会聚特性并没有引起广泛关注, J. Son 在文献<sup>[9]</sup>中做了详细的分析.

会聚模式存在的主要缺点有<sup>[4,10,11]</sup>:

(1) 将直线映射为曲线,从而导致深度平面弯曲(depth plane curvature);

(2) 因存在梯形失真(keystone distortion),导致垂直视差,可能影响立体图像观看的舒适度,严重时甚至会引起视觉疲劳<sup>[12]</sup>. J. Son<sup>[13]</sup>用详细的理论推导,给出了消除梯形失真的条件,但实现起来很复杂.

平行模式的主要缺点是,拍摄获得的左、右图像间

公共部分比会聚模式小,致使左右图像无法配对的范围较大.但这个缺点可以通过图像平移来改善<sup>[14]</sup>.

综上所述,平行配置模式和会聚配置模式各有优劣.在实际应用中,平行配置模式因其易于实现、调整简单等优点而被广泛采用.

在两种平行配置模式中,非平移模式的主要缺点是只存在交叉视差(crossed disparity)或非交叉视差(uncrossed disparity)中的一种.为实现平移配置模式, K. Kwon 等人提出了一种采用平移图像传感器的方法<sup>[15]</sup>.但由于需要制作特定的摄像机,可能会带来精度、同步、噪声、反应滞后等问题,所以通常用与之等效的图像平移法来实现<sup>[16]</sup>.

基于上述分析,本文的后续章节只讨论分析平移配置模式下的立体信息采集和感知的几何模型建模问题.

### 2.2 平移配置模式下立体信息采集与感知系统模型

本文所采用的立体系统几何模型建立在 A. Woods<sup>[4]</sup>的立体系统模型的基础上,图2是平移配置模式下系统模型的俯视示意图.如果没有特殊说明,后文中的距离单位都为米(m).

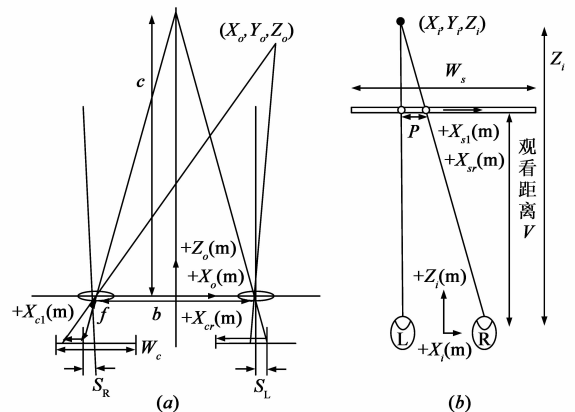


图2 平移配置模式的立体信息采集、感知系统

图2(a)是立体信息采集模型,其中两部摄像机在水平方向上间隔一定距离,平行放置,用于采集左、右眼图像或视频信息.设  $(X_o, Y_o, Z_o)$  是摄像机前被拍摄的自然场景的拍摄空间坐标,  $Y$  轴垂直纸面向外;  $(X_{cl}, Y_{cl})$ ,  $(X_{cr}, Y_{cr})$  分别表示左、右摄像机 CCD 像平面上的图像坐标,  $Y$  轴垂直纸面向里.该采集系统模型的功能是将自然场景中的拍摄空间点  $(X_o, Y_o, Z_o)$  转换成 CCD 像平面上的对应图像点  $(X_c, Y_c)$ .

采集模型的主要参数包括:

(1) 会聚距离  $C$  (convergence distance): 会聚点到两部摄像机透镜光学中心连线的距离,会聚功能是通过 CCD 平移或后期处理使图像出现相对位移实现

的,当图像左右平移距离相等的时候,从图 3(a)中可以得到会聚距离  $C = bf/(2S)$ ;

(2) 摄像机间距  $b$  (baseline distance): 两部摄像机透镜光学中心之间的距离;

(3) 摄像机的光学焦距  $f$  (focus length): 摄像机透镜中心到 CCD 成像平面的距离;

(4) 图像传感器 (CCD 或者 CMOS, 本文通用 CCD 指代) 像平面的宽度  $W_c$  (CCD width);

(5) 图像平移距离  $S_R$  (shifting right image),  $S_L$  (shifting left image): 为实现会聚功能, 左或右摄像机 CCD 中心向外 (远离右或左摄像机方向) 水平移动的距离; 或者等效的平移图像的距离. 当会聚点在  $Z_0$  轴上时,  $S_R = S_L = S$ .

图 2(b) 是立体信息感知系统模型的俯视视图, 其中,  $(X_i, Y_i, Z_i)$  是观看者的视觉系统通过立体显示器感知到的空间点坐标, 即感知空间的坐标,  $Y$  轴垂直纸面向外.  $(X_{sl}, Y_{sl})$ ,  $(X_{sr}, Y_{sr})$  分别表示呈现在显示器上的左、右眼图像坐标,  $Y$  轴垂直纸面向外.

感知系统的主要参数包括:

(1) 观看者双眼连线到显示器之间的距离  $V$  (viewing distance);

(2) 显示器屏幕宽度  $W_s$  (screen width);

(3) 观看者两眼之间的距离, 即瞳距  $I$  (inter-pupillary distance);

(4) 屏幕双目视差  $P$  (parallax),  $P = X_{sr} - X_{sl}$ , 即显示器屏幕上左、右眼图像对应点之间的水平距离;

(5) 图像放大因子  $M$  (magnitude),  $M = W_s/W_c$ .

在感知空间中, 视觉系统感知到的景物点可能位于显示屏的前后方, 因此双目视差可分为正视差 (positive disparity) 和负视差 (negative disparity), 也分别称为非交叉视差和交叉视差.

### 3 拍摄空间到感知空间的几何映射模型及畸变分析

#### 3.1 映射模型

由图 1 可知, 拍摄空间点到感知空间点的几何映射可表示为 3 个映射过程:

$$(X_o, Y_o, Z_o) \rightarrow (X_{cl}, Y_{cl}), (X_{cr}, Y_{cr}) \rightarrow (X_{sl}, Y_{sl}), (X_{sr}, Y_{sr}) \rightarrow (X_i, Y_i, Z_i)$$

首先来求物空间点  $(X_o, Y_o, Z_o)$  到 CCD 平面像点  $(X_{cl}, Y_{cl})$ ,  $(X_{cr}, Y_{cr})$  的映射关系. 假设摄像机镜头符合理想的小孔成像模型 (pin-hole model), 则由图 2(a) 可以得到 CCD 像平面上的水平坐标:

$$X_{cl} = \frac{f(b + 2X_o)}{2Z_o} - S_L \quad (1)$$

$$X_{cr} = \frac{f(-b + 2X_o)}{2Z_o} + S_R \quad (2)$$

对应像点的垂直坐标是相同的, 为:

$$Y_{cl} = Y_{cr} = \frac{Y_o f}{Z_o} \quad (3)$$

接着求 CCD 平面像点  $(X_{cl}, Y_{cl})$ ,  $(X_{cr}, Y_{cr})$  到显示器平面像点  $(X_{sl}, Y_{sl})$ ,  $(X_{sr}, Y_{sr})$  的映射关系. 假设 CCD 平面和显示器平面满足相似关系, 则 CCD 图像到显示器图像可以被认为是一个简单的放大过程, 所有坐标均线性增大  $M$  倍:

$$X_{sl} = MX_{cl}, X_{sr} = MX_{cr} \quad (4)$$

$$Y_{sl} = MY_{cl}, Y_{sr} = MY_{cr} \quad (5)$$

最后求显示器像点  $(X_{sl}, Y_{sl})$ ,  $(X_{sr}, Y_{sr})$  到感知空间点  $(X_i, Y_i, Z_i)$  的映射关系. 由图 2(b) 的几何关系可得显示器像点与感知空间的立体像点间的关系为:

$$X_i = \frac{I}{I - P} \cdot \frac{X_{sl} + X_{sr}}{2} \quad (6)$$

$$Y_i = \frac{I}{I - P} \cdot \frac{Y_{sl} + Y_{sr}}{2} \quad (7)$$

$$Z_i = \frac{I}{I - P} \cdot V \quad (8)$$

综上所述可得:

$$X_i = \frac{MI \left[ f \frac{X_o}{Z_o} + \frac{1}{2} (S_R - S_L) \right]}{I - M(S_R + S_L) + \frac{Mfb}{Z_o}} \quad (9)$$

$$Y_i = \frac{MfI \frac{Y_o}{Z_o}}{I - M(S_R + S_L) + \frac{Mfb}{Z_o}} \quad (10)$$

$$Z_i = \frac{VI}{I - M(S_R + S_L) + \frac{Mfb}{Z_o}} \quad (11)$$

拍摄空间中的物点  $(X_o, Y_o, Z_o)$  到感知空间的立体像点  $(X_i, Y_i, Z_i)$  间的几何映射关系比较复杂, 并不是简单的线性关系, 它还与拍摄场景的参数和显示器参数, 甚至观看者所处的位置都有关. 此映射模型是研究、设计和制作立体成像系统、处理立体信息、实现立体舒适显示的理论基础.

#### 3.2 从映射几何模型得出的几个结论

##### 3.2.1 拍摄空间中的会聚点被映射在显示器屏幕上

假设左右图像平移距离一致 (方向相反), 即  $S_R = S_L = S$ , 并设拍摄空间中的会聚点为

$$Z_o = C = \frac{bf}{2S} \quad (12)$$

代入式(11), 有

$$Z_i = V \quad (13)$$

式(13)说明,拍摄空间中的会聚点被映射在显示器屏幕上,形成零视差点(zero disparity point).从式(9)~(11)可以看出,映射函数是单调递增的,因此,拍摄空间中会聚点前的物体被映射在显示屏幕的前方,会聚点后的物体被映射在屏幕的后方.那么,适当的控制拍摄空间中的会聚点的位置可以很好的掌握立体显示时物体的凹凸感,这种控制被称为会聚控制<sup>[16]</sup>或者是零视差控制.

### 3.2.2 非平移配置模式仅有交叉视差

对于非平移配置模式,  $S = 0$ , 将其带入式(11), 有

$$Z_i = \frac{V}{1 + \frac{Mfb}{Z_o I}} < V \quad (14)$$

这说明拍摄空间中的点都被映射在显示屏幕的前方,即只存在交叉视差.在实际应用中,常出现感知到的所有物体都出现在显示屏幕后方的情况,这是由于采集到的左、右眼图像与呈现在显示器上的左、右眼图像发生了翻转的缘故.

### 3.2.3 深度非线性失真

式(11)说明深度存在非线性失真,并且拍摄空间的无限远点被映射成感知空间中距离值有限的一个点:

$$Z_{\text{infinite}} = \frac{VI}{I - 2MS} \quad (15)$$

式(15)说明,显示器尺寸越大,图像位移量越大,在感知空间中的最远距离也越大,而且这个距离与观看距离成正比.

### 3.2.4 X Y Z 轴方向放大倍数不同会导致感知空间物体发生畸变

令  $S_R = S_L = S$ , 再由式(9)~(11), 可得 XYZ 方向上的放大倍数之比:

$$\frac{X_i}{X_o} : \frac{Y_i}{Y_o} : \frac{Z_i}{Z_o} = Mf : Mf : V \quad (16)$$

由式(16)可知,只有  $Mf = V$  是个固定值时,图像在 XYZ 方向上的放大倍数才会一致;但是,即使 XYZ 三个方向的放大倍数一致,立体图像还会产生畸变,这是因为在每一个方向上还存在着非线性畸变.当  $Mf < V$  时,感知到的立体图像被拉伸;当  $Mf > V$  时,感知到的立体图像被压扁.

上面分析了感知图像的整体放大倍数比.通常情况下,人们更关心物体局部的失真情况.在局部范围内,可以用微分来分析物体的畸变情况. X 和 Y 方向上的局部变化率是一样的,即

$$D_x = \frac{DX_i}{DX_o} = \frac{\frac{MfI}{Z_o}}{I - 2MS + \frac{Mfb}{Z_o}} = \frac{DY_i}{DY_o} = D_y \quad (17)$$

深度畸变因子  $D_z$  被定义为 Z 深度方向的变化率:

$$D_z = \frac{DZ_i}{DZ_o} = \frac{VMfb}{(IZ_o - 2MSZ_o + Mfb)^2} \quad (18)$$

形状畸变因子  $D_s$  被定义为深度方向的变化率和 X(Y)方向上的变化率之比,即

$$D_s = \frac{D_z}{D_x} = \frac{Vb}{IZ_o - 2MSZ_o + Mfb} \quad (19)$$

用  $D_z$  和  $D_s$  可以较清楚的描述立体图像的畸变特性.

对于会聚系统,深度畸变因子  $D_z$  和形状畸变因子  $D_s$  的表达式非常复杂,朱效明<sup>[17]</sup>给出了近似的表达式:

$$D_z \approx \frac{DZ_i}{DZ_o} \approx \frac{VMfb}{IZ_o^2} \quad (20)$$

$$D_s \approx \frac{D_z}{D_x} \approx \frac{Vb}{IZ_o} \quad (21)$$

立体视觉系统存在很多失真,如梯形失真<sup>[18~20]</sup>、深度平面弯曲<sup>[4]</sup>、剪切失真<sup>[4,21,22]</sup>(shear distortion),木偶剧效应(puppet theater effect)和纸板效应(cardboard effect)等等.这些失真都可以用上述畸变模型来解释.在会聚配置系统中,可使用式(20)、(21)来解释木偶效应和纸板效应.在接下来的 3.3 和 3.4 节,将在平行配置系统中使用  $D_z$  和  $D_s$  来解释木偶剧效应和纸板效应产生的原因.

## 3.3 木偶剧效应

木偶剧效应很早就被注意到并有许多研究<sup>[21,23~26]</sup>.在观看立体电视、电影时,会出现前景物体相对于后景物体显得非常大的情况,就如同在木偶剧表演中,为了突出木偶,使木偶比背景大很多一样.这种立体图像的失真现象称为木偶剧效应.

木偶剧效应是由于前、后景物体的尺寸放大率不同造成的,用深度畸变因子  $D_z$  可以很方便的解释其成因.由式(10)、(12)得到  $D_z$  的另外一种形式:

$$D_z = \frac{DZ_i}{DZ_o} = \frac{VMfb}{\left(IZ_o + Mfb\left(1 - \frac{Z_o}{C}\right)\right)^2} \quad (22)$$

在实际应用中,瞳距( $I$ )一般取值为 65mm,放大倍数( $M$ )在 30 左右,焦距( $f$ )一般在 50mm 左右,摄像机间距( $b$ )通常接近  $I$ ,典型值为 70mm,会聚距离( $C$ )与具体拍摄场景有关,一般为 1~3m,而被拍摄场景离摄像机的距离( $Z_o$ )通常接近会聚距离.将上述参数应用到式(22),得:

$$D_z \approx \frac{VMfb}{IZ_o^2} \propto \frac{1}{Z_o^2} \quad (23)$$

由此可知,感知到的物体的尺寸放大率反比于被摄物体距离的平方,当前景较近,而后景较远,前景的尺寸放大率将远大于后景尺寸放大率,这样就可能出现木偶剧效应了.如果要减少甚至消除木偶剧效应,可

以根据式(22)适当的调整拍摄参数和显示参数,使得前后景物体的尺寸放大率不要相差太大。

### 3.4 纸板效应

纸板效应很早就受到了关注<sup>[23,24,26~29]</sup>。在立体图像显示时,某些物体在形状上可能会被感知的比正常形状扁很多,如同扁平的纸板被堆砌在不同深度平面上一样,这种立体图像的失真现象被称为纸板效应。纸板效应这种主观感受可用客观的形状畸变因子  $D_s$  来解释。由式(11)、(12),可以得到:

$$D_s = \frac{D_z}{D_x} = \frac{Vb}{IZ_o + Mfb \left(1 - \frac{Z_o}{C}\right)} \quad (24)$$

将 3.3 节提供的参数应用到式(24)中,可得:

$$D_s \approx \frac{Vb}{IZ_o} \propto \frac{1}{Z_o} \quad (25)$$



图3 平滑深度图以减小纸板效应

## 4 立体舒适观看区

立体图像的视觉舒适程度会直接影响立体产业的发展。梯形失真、木偶剧效应、纸板效应、会聚与调节之间的冲突 (vergence-accommodation conflict)<sup>[31,32]</sup>、串扰 (crosstalk)<sup>[33~36]</sup>、观看距离<sup>[37]</sup>、个人生理差异<sup>[12]</sup>等各种因素都会影响立体视觉舒适度。与立体舒适度直接相关的一个概念是立体舒适观看区 (comfortable viewing zone)<sup>[12,37~39]</sup>, 立体舒适观看区被认为是这样一种区域:

在观看距离一定的情况下,在这个区域内,观看者可以容易地将左、右眼图像融合成单一的立体图像,不会产生融像困难;融合后的立体图像在深度方向上没有模糊的感觉;不会产生强烈的视觉不舒适感和疲劳感。

立体舒适观看区域的大小及其定量描述是一个长期没有得到完美解决的问题。目前,有三种方式来定量分析舒适观看区域的大小:感知空间中的距离区间、相对屈光度范围、用角度定义的视差范围。下面将分别介绍这三种表示方法,并给出它们之间的转换关系及定量的比较结果。图 4 给出了舒适观看区示意图,图中观看者在立体显示器正前方观看立体视频,接下来的 4.1

式(25)说明,形状畸变因子反比于被摄物体距离,当  $Z_o$  的取值使  $D_s < 1$  时,随着  $Z_o$  值的增大,  $D_s$  迅速减小,导致物体显得越来越扁平,这样就可能导致纸板效应。

纸板效应很容易出现在“左(右)视图 + 深度图 (depth map)”格式的立体视频中。在图 3(a)是一段立体视频中某个视点的图像,(b)图是相应深度图。在图 3(b)中,人物的深度值单一,而其背景较远,使得纸板效应比较明显。为了减少纸板效应,可采用的方法之一是对深度图做模糊处理,得到如图 3(c)所示的深度图,经过处理后纸板效应就会得到明显的改善<sup>[30]</sup>。如何在深度图质量和纸板效应之间做出权衡是值得深入研究的课题。

~ 4.3 节会给出具体的解释。

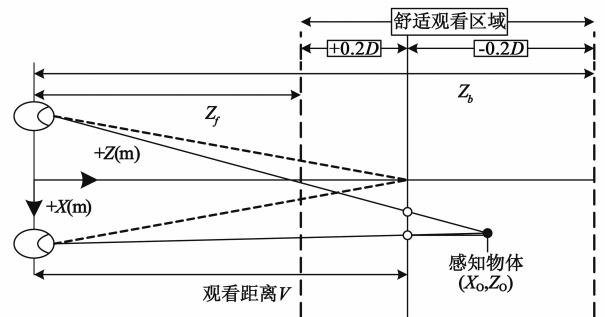


图4 舒适观看区示意图

### 4.1 舒适观看区的距离区间表示

在感知空间中,可用图 4 所示的距离区间  $[Z_f, Z_b]$  来表示立体舒适观看区域<sup>[40]</sup>,其中  $Z_f, Z_b$  分别表示舒适观看区离人眼的最近距离和最远距离。舒适观看区可以用相对于屏幕的距离来表示:  $[Z_f - V, Z_b - V]$ ,这时很容易将它与屏幕视差( $P$ )相关联,由图 4 可以计算得到  $P$  的范围区间:

$$P \in \left[ \frac{I(Z_f - V)}{Z_f}, \frac{I(Z_b - V)}{Z_b} \right] \quad (26)$$

### 4.2 舒适观看区域的屈光度范围表示

外界物体的光线在眼睛光学系统中各种界面处

发生折射的现象被称为屈光,光线在各界面上的折射程度,通常用屈光力来表达,在眼球光学系统中,用屈光度(Diopter,简称为D)作为屈光力的单位.在几何光学中,屈光度被定义为焦距(m)的倒数,即单位屈光度( $D$ ) =  $1/f^{[41]}$ .可以用相对屈光度区间[ + DIO, - DIO]来表示舒适观看区域,如图4所示,从图中可以看出在相同的屈光度绝对值的情况下,屏幕后方的相对距离要大于屏幕前方的相对距离.屈光度区间与距离区间的转换关系如式(27)、(28)所示.

$$Z_f = \frac{1}{\frac{1}{V} + \text{DIO}} \quad (27)$$

$$Z_b = \begin{cases} \frac{1}{\frac{1}{V} + \text{DIO}}, & \text{如果 } V < \text{DIO}^{-1} \\ \infty, & \text{如果 } V \geq \text{DIO}^{-1} \end{cases} \quad (28)$$

#### 4.3 舒适观看区域的视差范围表示

用视差(disparity)也可以表示立体舒适观看区间,而视差通常是用观看角度差来定义.

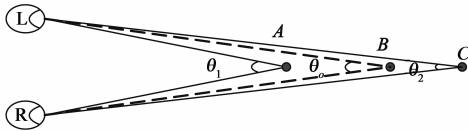


图5 用观看角度定义的视差

在图5中,如果人眼注视B点,则A点的视差(负视差)被定义为:

$$\text{deg}_- = \theta_0 - \theta_1 \quad (29)$$

同样,C点的视差(正视差)被定义为:

$$\text{deg}_+ = \theta_0 - \theta_2 \quad (30)$$

如果用视差(|deg| (degree,度)作为舒适观看的范围,可以计算出它与立体舒适观看区距离区间之间的关系:

$$Z_f = \frac{1}{2} \cdot \cot\left(\arctan \frac{I}{2V} + \frac{|\text{deg}|}{2} \cdot \frac{\pi}{180}\right) \quad (31)$$

$$Z_b = \frac{1}{2} \cdot \cot\left(\arctan \frac{I}{2V} - \frac{|\text{deg}|}{2} \cdot \frac{\pi}{180}\right) \quad (32)$$

注意到在这里只分析了屏幕中央的位置的情况,但可以在下面的表1中看到这是对真实情况的一个很好的假设,而且人眼确实更容易注视在眼前中央的区域.

长期以来,人们对立体舒适度进行了许多研究.早在1992年,S. Pastoor<sup>[42]</sup>就讨论了影响视差范围的人的生理因素,总结出立体影像的舒适度是它与普通二维影像竞争的关键.N. Holliman<sup>[40]</sup>用特定的显示器试验得出舒适观看区域为显示器屏幕前后的一个小区域[-50mm, +60mm]的结论.M. Wöpping<sup>[43]</sup>通过主观评

价,提出了立体舒适度的视差上限为70arcmin(1arcmin = 1°/60).S. Pastoor<sup>[44]</sup>, F. Speranza<sup>[45]</sup>和 M. Lambooi<sup>[12]</sup>建议使用60arcmin作为双眼视差的上限.在ITU-R BT.1438<sup>[46]</sup>中,建议在±0.3D范围内显示立体图像,S. Yano<sup>[47]</sup>和 W. Chen<sup>[5]</sup>认为±0.2D更合适.国际上,对立体舒适观看范围的争论从未停止过.

表1以1°视差为绝对上界,给出了距离区间和屈光度之间的对应关系.

表1 1°视差为上界的舒适观看区间

观看距离 V(mm)	屈光度 (D)	Z <sub>f</sub> (mm)	屈光度 (D)	Z <sub>b</sub> (mm)	屈光度 (D)
500	2	441	2.27	578	1.73
1000	1	788	1.27	1368	0.73
2000	0.5	1301	0.77	4321	0.23
3000	0.33	1661	0.6	15430	0.065

从表1可以看出,1°视差与+ - 0.27D表示的范围基本一致.

## 5 考虑畸变模型和舒适观看区域等因素的综合拍摄规则

立体成像技术应用有两个重要指标:立体感的真实性和观看的舒适性.

针对真实性和舒适性,有两个问题须考虑:(1)是否应该将所有场景都映射在舒适观看区间中,以保证舒适度?(2)如何保证感兴趣区域(region of interest)的真实性?

G. Jones<sup>[40]</sup>介绍了如何通过调整立体摄像机的参数来控制感知的深度范围,以避免出现过大的深度视差,导致不舒适情况.N. Holliman<sup>[48]</sup>提出了一个非对称畸变算法,以控制立体图像中不同位置的畸变程度:增大非感兴趣区的畸变程度,以控制视差范围,保证舒适性,同时降低感兴趣区域的畸变程度,使得感兴趣区域更加逼近真实场景.图6(a)表示在未做处理前,由于系统畸变,球状物体被感知为椭球状物体.图6(b)是用了非对称畸变算法后的结果,感兴趣区域更加逼近真实场景,而整体的视差范围却缩小了.

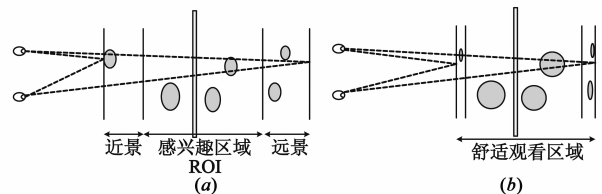


图6 非对称畸变算法示意图

上述的非对称畸变算法对于立体动画、立体电影制作是有指导价值的,尤其是对于计算机生成的场景,此时很容易控制物体的畸变.

W. Chen<sup>[5]</sup>进一步完善了这个工作,总结了立体视

频拍摄的两个规则和它们的优先级:

(1)调整可变的参数,使在特定的观看条件下感兴趣区域的畸变值最小.

(2)调整可变的参数,使场景中的物体都尽可能地映射在舒适观看区域内.

如果两个规则不能同时被满足,则优先考虑规则 2. 规则 1 是为了保证感兴趣区域的真实性,而规则 2 是为了保证立体影像的舒适性,舒适性应优于真实性.

## 6 结论

本文总结了视差式立体信息采集系统的配置模式,比较了各种配置模式的优缺点;同时讨论了平移式平行拍摄系统下拍摄空间到感知空间之间的映射关系,即畸变模型,并用它解释了木偶剧效应和纸板效应的成因.然后,本文讨论了舒适观看区的范围,分析比较了多数文献中普遍使用的 3 种舒适观看区的定义,给出了这 3 种定义之间的转换式.最后,将畸变模型和舒适观看区域联系起来,给出了立体图像/视频的拍摄规则:拍摄时尽可能地兼顾舒适性和真实性,在两者发生冲突时应优先考虑立体图像的舒适性.

值得一提的是, $1^\circ$ 视差舒适观看区范围和 0.2 屈光度舒适观看区范围目前在国际上尚有争论,还需要更多的立体视觉舒适融像理论来支持.本文提到的拍摄空间到感知空间的映射模型是几何模型.遗憾的是,人类感知立体信息方式是否完全服从几何映射的规律,学术界尚无明确的定论.

## 参考文献

[1] Andrew Woods. The Illustated 3D HDTV List [EB/OL]. <http://www.3dmovielist.com/3dhdtvs.html>, 2012-12-10/2012-12-14.

[2] Ian P Howard, Brian J Rogers. Seeing in Depth: Basic Mechanics [M]. New York: Oxford University Press, 2008. 1-10.

[3] Farley Norman, Hideko F Norman, Amy E Craft, et al. Stereopsis and aging [J]. Vision Research, 2008, 48(23): 2456-2465.

[4] Andrew Woods, Tom Docherty, Rolf Koch. Image distortions in stereoscopic video systems [A]. Stereoscopic Displays and Applications [C]. San Jose, USA: SPIE, 1993. 36-48.

[5] Wei Chen, Jérôme Fournier, Marcus Barkowsky, et al. New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone [A]. Stereoscopic Displays and Applications XXII [C]. San Francisco, USA: SPIE, 2011. Vol. 7863; Paper No. 78631O.

[6] LIU Yebin, DAI Qionghai, XU Wenli. A wide base line multiple camera system for high performance 3D video and free viewpoint video [J]. Chinese Journal of Electronics, 2009, 18(2): 275-280.

[7] Park J, Kim H, Tai Y, et al. High quality depth map upsampling for 3D-TOF cameras [A]. IEEE International Conference on Computer Vision [C]. Barcelona, Spain: IEEE, 2011. 1623-1630.

[8] GONG Xun, WANG Guoyin, LI Xinxin, et al. A statistical two-step method for 3D face reconstruction from a single image [J]. Chinese Journal of Electronics, 2011, 20(4): 671-675.

[9] Jung-Young Son, Seokwon Yeom, Dong-Su Lee, et al. A stereoscopic camera model of focal plane detector array [J]. Display Technology, 2011, 7(5): 281-288.

[10] Seung W Lee, Nam Kim. A method for precise depth detection in stereoscopic display [J]. Journal of the Optical Society of Korea, 2006, 10(1): 37-41.

[11] Robert S Allison. The camera convergence problem revisited [A]. Stereoscopic Displays and Virtual Reality Systems XI [C]. San Jose, USA: SPIE, 2004. 167-178.

[12] Marc Lambooi, Wijnand IJsselsteijn, Marten Fortuin, et al. Visual discomfort and visual fatigue of stereoscopic displays: A review [J]. Imaging Science and Technology, 2009, 53(3): 1-14.

[13] Jung-Young Son, Yuri N Gruts, Kae-Dal Kwack, et al. Stereoscopic image distortion in radial camera and projector configurations [J]. Journal of Optical Society of America, 2007, 24(3): 643-650.

[14] Zahir Y. Alpaslan, Alexander A. Sawchuk. Multiple camera image acquisition models for multi-view 3D display interaction [A]. IEEE 6th Workshop on Multimedia Signal Processing [C]. Siena, Italy: IEEE, 2004. 256-262.

[15] Ki-Chul Kwon, Jae-Kwang Choi, Nam Kim. Automatic control of horizontal-moving stereoscopic camera by disparity compensation [J]. Journal of the Optical Society of Korea, 2002, 6(4): 150-155.

[16] Ki-Chul Kwon, Young-Tae Lim, Nam Kim, et al. Vergence control of binocular stereoscopic camera using disparity information [J]. Journal of the Optical Society of Korea, 2009, 13(3): 379-385.

[17] 朱效明, 高雅允. 双 CCD 立体视觉系统的理论研究 [J]. 光学技术, 2003, 29(3): 298-300.

Zhu Xiaoming, Gao Zhiyun. Theoretic research on double-CCD stereoscopic system [J]. Journal of Optical Technique, 2003, 29(3): 298-300. (in Chinese)

[18] Wijnand A IJsselsteijn, Huib de Ridder, Joyce Vliegen. Subjective evaluation of stereoscopic images: effects of camera parameters and display duration [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(2): 225-233.

[19] Lew B Stelmach, Wa J Tam, Filippo Speranza, et al. Improving the visual comfort of stereoscopic images [A]. Stereoscopic Displays and Virtual Reality Systems X [C]. Santa Clara,

- USA: SPIE, 2003. 269 – 282.
- [20] Hoonjong Kang, Namho Hur, Seunghyun Lee, et al. Horizontal parallax distortion in toed-in camera with wide-angle lens for mobile device [J]. *Optical Communications*, 2008, 281(6): 1430 – 1437.
- [21] Lydia M Meesters, Wijnand A IJsselsteijn, Pieter J Seuntjens. A survey of perceptual evaluations and requirements of three dimensional TV [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(3): 381 – 391.
- [22] J Daly, Robert T Held, David M Hoffman. Perceptual issues in stereoscopic signal processing [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 347 – 361.
- [23] Kenichiro Masaoka, Atsuo Hanazato, Masaki Emoto, et al. Spatial distortion prediction system for stereoscopic images [J]. *Electronic Imaging*, 2006, 15(1): 1 – 11.
- [24] Hirokazu Yamanoue. The relation between size distortion and shooting conditions for stereoscopic images [J]. *Journal of the Society of Motion Picture and Television Engineers*, 1997, 106(4): 225 – 232.
- [25] David L MacAdam. Stereoscopic perceptions of size shape distance and direction [J]. *Journal of the Society of Motion Picture and Television Engineers*, 1954, 62(4): 271 – 289.
- [26] Hirokazu Yamanoue, Masaki Emoto, Yuji Nojiri. *Stereoscopic HDTV* [M]. Japan: Springer, 2012. 11 – 44.
- [27] Hirokazu Yamanoue, Makoto Okui, Fumio Okano. Geometrical analysis of puppet-theater and cardboard effects in stereoscopic HDTV images [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, 16(6): 744 – 752.
- [28] Mikko Kytö, Jussi Hakala. Geometric and subjective analysis of stereoscopic I3A cluster images [A]. *Stereoscopic Displays and Applications XXII* [C]. San Francisco, USA: SPIE, 2011. Vol. 7863: Paper No. 786319.
- [29] Hirokazu Yamanoue, Makoto Okui, Ichiro Yuyama. A study on the relationship between shooting conditions and cardboard effect of stereoscopic images [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2000, 10(3): 411 – 416.
- [30] Shimono Koichi, Wa T James, Vazquez Carlos, et al. Removing the cardboard effect in stereoscopic images using smoothed depth maps [A]. *Stereoscopic Displays and Applications XXI* [C]. San Jose, USA: SPIE, 2010.
- [31] Takashi Shibata, Joohwan Kim, David M Hoffman. The zone of comfort: Predicting visual discomfort with stereo displays [J]. *Journal of Vision*, 2011, 11(8): 1 – 29.
- [32] David M Hoffman, Ahna R Girshick, Kurt Akeley, et al. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue [J]. *Journal of Vision*, 2008, 8(3): 1 – 30.
- [33] Liyuan Xing, Junyong You, Ebrahimi T, et al. Assessment of stereoscopic crosstalk perception [J]. *IEEE Transactions on Multimedias*, 2012, 14(2): 326 – 337.
- [34] Janusz Konrad, Bertrand Lacotte, Eric Dubois. Cancellation of image crosstalk in time-sequential displays of stereoscopic video [J]. *IEEE Transactions on Image Processing*, 2000, 9(5): 897 – 908.
- [35] Lili Wang, Teunissen K, Yan Tu, et al. Crosstalk evaluation in stereoscopic displays [J]. *Display Technology*, 2011, 7(4): 208 – 214.
- [36] Tsirlin I, Wilcox L M, Allison R. The effect of crosstalk on the perceived depth from disparity and monocular occlusions [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 445 – 453.
- [37] Wei Chen, Jérôme Fournier, Marcus Barkowsky, et al. New requirements of subjective video quality assessment methodologies for 3DTV [A]. *Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics* [C]. Scottsdale, USA: VPQM, 2010. 1 – 6.
- [38] Stephan Reichelt, Ralf Häussler, Gerald Fütterer, et al. Depth cues in human visual perception and their realization in 3D displays [A]. *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV* [C]. Orlando, USA: SPIE, 2010. Vol. 7690: Paper No. 76900B.
- [39] Junle Wang, Marcus Barkowsky, Vincent Ricordel, et al. Quantifying how the combination of blur and disparity affects the perceived depth [A]. *Human Vision and Electronic Imaging XVI* [C]. San Francisco, USA: SPIE, 2011. Vol. 7865: No. 78650K.
- [40] Graham Jones, Delman Lee, Nicolas Holliman, et al. Controlling perceived depth in stereoscopic images [A]. *Stereoscopic Displays and Virtual Reality Systems VIII* [C]. Bellingham, WA: SPIE, 2001. 42 – 53.
- [41] 赵堪兴, 杨培增. *眼科学 (第七版)* [M]. 北京: 人民卫生出版社, 2008. 226 – 229.
- [42] Siegmund Pastoor. Human factors of 3DTV: An overview of current research at Heinrich-Hertz-Institute Berlin [A]. *IEE Colloquium on Stereoscopic Television* [C]. London: IEEE, 1992. 1 – 4.
- [43] Matthias Wöpking. Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus [J]. *Journal of the Society for Information Display*, 1995, 3(3): 101 – 103.
- [44] Siegmund Pastoor. Human factors of 3D imaging: Results of recent research at Heinrich-Hertz-Institute Berlin [A]. *2nd International Display Workshops* [C]. Hamamatsu, Japan: IDW'95 Committee, 1995. 69 – 72.
- [45] Filippo Speranza, Wa J Tam, Ron Renaud, et al. Effect of disparity and motion on visual comfort of stereoscopic images [A]. *Stereoscopic Displays and Virtual Reality Systems XIII*

[C]. San Jose, USA: SPIE, 2006. Vol. 6055: Paper No. 60550B.

[46] Recommendation ITU-R BT. 1438 – 2000, Subjective Assessment of Stereoscope Television Pictures[S].

[47] Sumio Yanoa, Masaki Emotoa, Tetsuo Mitsuhashi. Two factors in visual fatigue caused by stereoscopic HDTV images[J].

Journal of Displays, 2004, 25(4): 141 – 150.

[48] Nick Holliman. Mapping perceived depth to regions of interest in stereoscopic images[A]. Stereoscopic Displays and Virtual Reality Systems XI[C]. San Jose, USA: SPIE, 2004. 117 – 128.

### 作者简介



侯春萍(通信作者) 女, 1957年6月生于吉林省长春市. 现为天津大学电子信息工程学院教授、博士研究生导师. 主要研究领域为立体图像传输与处理以及宽带无线通信.

E-mail: hep@tju.edu.cn



袁万鑫 男, 1987年6月生于四川成都. 现为天津大学电子信息工程学院在读研究生. 研究领域为立体图像处理.

E-mail: yuanwanxin@tju.edu.cn