

# 平面图像立体化关键技术研究\*

朱庆生<sup>1</sup> 支丽欧<sup>1</sup> 刘然<sup>1</sup> 许小艳<sup>1</sup> 谢洪涛<sup>1</sup> 谢洪波<sup>2</sup> 杨璐<sup>3</sup>

(重庆大学计算机学院 重庆 400044)<sup>1</sup> (天津大学精密仪器与光电子工程学院 天津 300072)<sup>2</sup>  
(同济大学电子与信息工程学院 上海 201804)<sup>3</sup>

**摘要** 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。本文以在平面图像中随机加入水平视差的立体化方法为基础,讨论当图像子块在水平方向上的随机位移分别服从不同分布时的成像情况,并采用交叉熵和均方根误差两个定量指标对转换效果进行评价。实验结果表明:(i)当随机变量都服从同一分布时,随机变量的取值对转换效果没有多大的影响;(ii)采用不同的分布时,立体效果化的效果会明显不同。一般的,采用正态分布立体效果较好。

**关键词** 立体视觉,立体化,转换,图像处理

## Research on Image Conversion from Planar into Stereo

ZHU Qing-Sheng<sup>1</sup> ZHI Li-Ou<sup>1</sup> LIU Ran<sup>1</sup> XU Xiao-Yan<sup>1</sup> XIE Hong-Tao<sup>1</sup> Xie Hong-Bo<sup>2</sup> YANG Jun<sup>3</sup>

(College of Computer Science, Chongqing University, Chongqing 400044)<sup>1</sup>

(College of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072)<sup>2</sup>

(School of Electronics and Information Engineering, Tongji University, Shanghai 201804)<sup>3</sup>

**Abstract** It's significant to convert planar images into stereo, and acquiring the stereo pairs of a scene is the key to binocular stereo imaging. Based on the method of adding horizontal parallax randomly to a planar image, this paper discusses the qualities of the conversion when applying different distributions to the horizontal offset of the sub-blocks images. Two quantitative criterions, cross-entropy and root-mean-square error, are introduced to evaluate the conversion. The results of the experiment show that, (i) when each random variable has the same distribution, different values of these random variables affect little the stereo effect; (ii) when applying different distributions to the random variables, the qualities of the conversion are different. Generally, we recommend normal distribution for better stereo effect.

**Keywords** Stereo vision, Planar to stereo, Conversion, Image processing

## 1 引言

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。如果能利用图像处理技术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。已有一些文献对相关的问题进行了探讨,这些文献中提出的方法大都是基于双目立体成像的基本原理。

双目立体成像是人类在三维世界中产生立体视觉的生理基础,由于两只眼睛视轴间距(瞳距)的存在,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成像的基本原理看到立体图像。因此,根据仅有的一幅平面图像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。

目前存在的立体图像大多是采用以下两种方式获得:(1)通过两部光学、机械特性完全一致且严格校准的摄像机,采用空间分离和同步旋转的方式拍摄<sup>[2]</sup>,这种方式成本高,难度大;(2)通过对坐标系中的物体进行双中心透视投影得到<sup>[3]</sup>,该方法对复杂的户外实景失去了作用。

我们知道平面图像本身蕴藏着丰富的三维信息,Shanks提出了一种马鞍曲面投影立体成像法<sup>[4]</sup>,侯春萍等人利用心理深度暗示(Psychological Depth Cue)和生理深度暗示(Physiological Depth Cue)原理,提出了一种在平面图像中随机加入双眼视差的立体化方法<sup>[1]</sup>。该方法将一幅图像划分成如图1所示的 $M \times N$ 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,得到的图像和原图像一起构成立体图像对(Stereo Pair)。处理的流程如图2所示,具体过程及参数的意义参见文<sup>[1]</sup>。

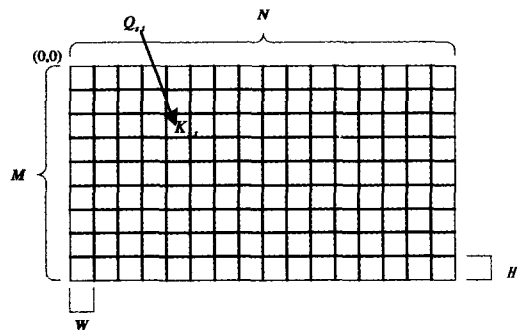


图1 图像分成 $M \times N$ 个子块

然而文<sup>[1]</sup>没有对平面图像立体化后的立体效果给出定

\*基金项目:重庆大学研究生科技创新基金(200701Y1A0080194),教育部高等学科点专项科研基金(20050611027)资助。朱庆生 教授,博士生导师,研究方向:图像工程、虚拟植物;支丽欧 硕士研究生,研究方向:数字图像处理。

量的评价指标,也没有深入讨论随机变量对立体效果的影响。本文利用文[1]的方法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。本文剩余部分组织如下:第2部分给出了评价立体效果的两个定量指标—交叉熵和均方根误差,第3部分讨论了随机变量对立体效果的影响,最后给出了结论以及需要进一步研究的工作。

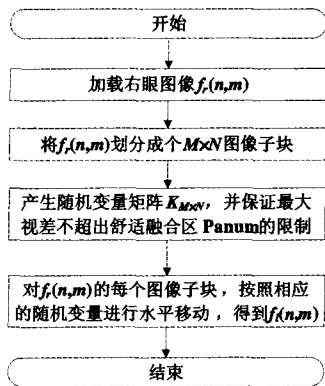


图2 平面图像立体化流程图

## 2 立体效果评价指标

由于每个观察者眼睛本身以及观察方式和习惯的不同,对同一幅图可能会有不同的评价。为了便于比较,我们采用交叉熵和均方根误差这两个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视觉的基本原理由软件生成的左右眼视图分别为  $lImg$  和  $rImg$ ,对  $rImg$  进行立体化,得到大小与  $lImg$  相同的左眼视图  $lImg'$ 。

### ① 交叉熵(Cross-entropy)

设  $X$  是具有  $n$  个有限状态值的随机变量,  $p_i = P\{X = x_i\}, i = 0, 1, \dots, n-1$ 。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:

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

这里  $p_i$  是对灰度图像  $P$  中灰度值为  $i$  的像素概率统计,  $n = 256$  表示灰度级数。交叉熵直接反映了两幅图像像素的信息量差异,是评价两幅图像差别的关键指标。一般情况下,交叉熵越小,两幅图越接近。

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  $Q$  内容差别很大,但是交叉熵  $I_{[P,Q]}$  却比较小,因此有必要参考另外一个指标——均方根误差。

### ② 均方根误差

设图像的高和宽分别为  $M$  和  $N$ (单位为像素),  $lImg(i, j)$  表示图像  $lImg$  在像素点  $(i, j)$  处的灰度值,则均方根误差:

$$E_{RMS}(lImg', lImg) = \left\{ \frac{1}{MN} \left[ \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [lImg'(i, j) - lImg(i, j)]^2 \right] \right\}^{1/2} \quad (2)$$

一般地,均方根误差越小,转换的效果越好。

## 3 随机变量

在图像的立体化过程中,为了使图像子块在水平方向上的移动具有随机性,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

$K_{M \times N}$ 。当随机变量的均值为 1 时,水平视差的期望为  $0^{[1]}$ ,图像子块在水平位置上随机地左右移动,由此产生水平方向的正视差或负视差。



图3 原始图像(左眼视图  $lImg$  + 右眼视图  $rImg$ )

下面讨论随机变量服从不同的分布时对立体效果的影响。实验使用的原始图像如图3(每幅图像的大小为  $800 \times 600$ )所示,程序对右眼视图进行立体化。为了便于比较,实验中固定取  $M=150, N=200$ 。

### 3.1 均匀分布

当随机变量  $K_{s,t}$  服从  $[0.5, 1.5]$  上的均匀分布时,均值为 1。

同时无论服从何种分布的随机变量  $K_s$ ,还应该使每个图像子块  $Q_{s,t}$  在水平方向上的最大位移不超出舒适 Panum 融合区(Panum's fusional area)<sup>[1, 12]</sup>的限制,即要求转换后原图像的每一个像素对应的水平视差  $\Delta q_{i,k}$  满足不等式(3):

$$-\frac{2h\eta p_d}{EP_s} < \Delta q_{i,k} < \frac{2h\eta p_d}{EP_s} \quad (3)$$

其中  $2h$  是瞳距,一般为  $6.5 \text{ cm}$ ;  $\eta$  为视锐度,一般取  $1' \approx 2.907 \times 10^{-4} \text{ rad}$ ;监视器屏幕与观察者的距离为  $p_d (p_d > 0)$ ,这里取  $100 \text{ cm}$ ,  $E$  是瞳孔直径,一般为  $0.4 \text{ cm}$ ,  $P_s$  是监视器的像素间隔,这里取  $P_s = 3.53 \times 10^{-2} \text{ cm}^{[8]}$ ;因此可得,  $-13 \leq \Delta q_{i,k} \leq 13$ 。

产生一组符合要求的随机变量的算法如下:

算法 ProduceKmn:产生一个符合要求的随机变量矩阵  $K_{M \times N}$ 。

$K_{M \times N}$ :随机变量矩阵,共  $M$  行  $N$  列 ( $M \geq 1, N \geq 1$ )。矩阵中每一个元素  $K_{s,t}$  与  $Q_{s,t}$  对应。

输入:未填值的随机变量矩阵  $K_{M \times N}$ 。

输出:填好随机数的符合要求的随机变量矩阵  $K_{M \times N}$ 。

步骤:

```

for each row  $K_s$  in  $K_{M \times N}$  do
{
  repeat
  {
    for each element  $K_{[s, j]}$  in row  $K_s$  produce a random number  $K_{s,t} \in [0.5, 1.5], \alpha=1$ ;
    if  $K_{s,t}$  satisfies the constraint of Panum's Fusional Area then  $K_{[s, j]} := K_{s,t}$ ;
    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  $K_{[s, j]}$ 
  }
}
until all elements in row  $K_s$  have random numbers which satisfy the constraint;
};
    
```

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:每产生一个随机数,就立刻检查它是否符合要求;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。

表1中的“均匀分布”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(a)

所示。将其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 2 中“均匀分布”列所示。

表 1 四组符合要求的服从不同分布的随机变量取值

随机变量	取值			
	均匀分布	正态分布	三角分布	拟合灰度值的分布
$K_{1,1}$	0.6065	1.05552	0.9727	1.4412
$K_{1,2}$	1.2955	0.94092	1.4137	0.6862
$K_{1,3}$	1.3059	0.8338	1.1978	1.2761
...	...	...	...	...
$K_{150,200}$	1.1794	1.0885	0.7914	1.2405

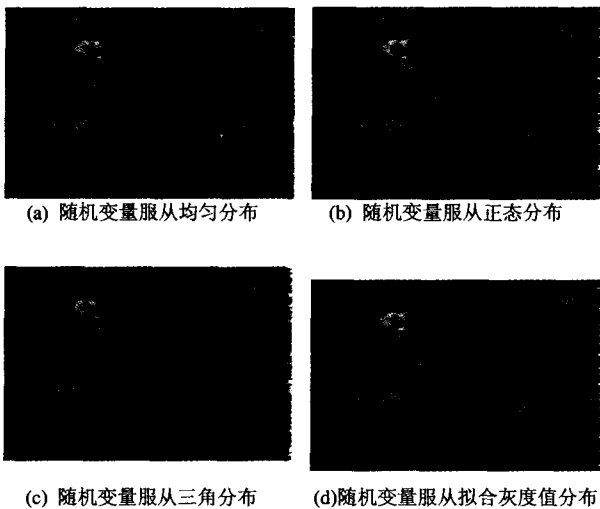


图 4 左眼视图  $lImg'$

表 2 随机变量服从不同分布时的交叉熵  $I_{[lImg', lImg]}$  和均方根误差  $E_{RMS}(lImg', lImg)$

实验次数	均匀分布		正态分布		三角分布		拟合灰度值的分布	
	$I_{[lImg', lImg]}$	$E_{RMS}(lImg', lImg)$	$I_{[lImg', lImg]}$	$E_{RMS}(lImg', lImg)$	$I_{[lImg', lImg]}$	$E_{RMS}(lImg', lImg)$	$I_{[lImg', lImg]}$	$E_{RMS}(lImg', lImg)$
1	0.005744	40.4866	0.000491	39.5727	0.001033	39.9397	0.003782	40.2754
2	0.005968	40.2931	0.000344	39.6942	0.000666	40.1731	0.002635	40.4307
3	0.006120	40.5028	0.000061	39.6601	0.000421	40.0783	0.002396	40.3512
4	0.006564	40.3817	0.000482	39.5812	0.000965	39.9811	0.002808	40.3192
5	0.006051	40.5021	-0.000098	39.9713	0.000887	39.9287	0.001578	40.3682
6	0.006559	40.2664	-0.000456	39.8983	0.000243	40.0196	0.003185	39.9659
7	0.006322	40.3727	0.000414	39.8635	0.001516	39.9939	0.003069	40.3905
8	0.005552	40.5113	-0.000287	39.8127	0.000860	39.9939	0.002649	40.4378
9	0.005656	40.3585	0.000537	39.6704	0.001437	39.7604	0.002567	40.5623
10	0.006666	40.2609	-0.000452	39.9525	0.000679	39.9413	0.002520	40.0976
平均值	0.006120	40.3936	0.000104	39.7677	0.000871	39.9810	0.002719	40.3199
标准差	0.000399	0.1010	0.000402	0.1502	0.000401	0.1069	0.0005733	0.1732

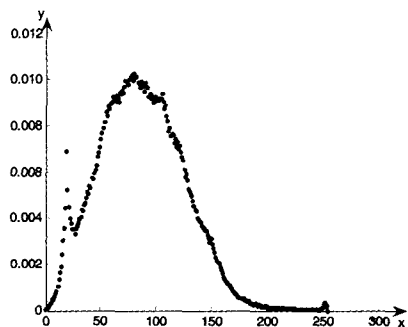


图 5  $rImg$  的灰度值分布

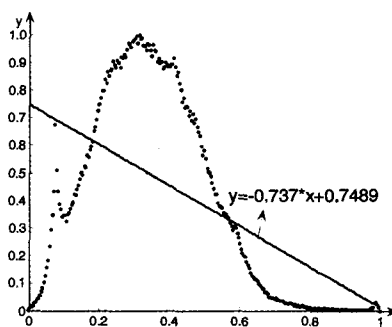


图 6 归一化灰度值分布及对应的拟合曲线

### 3.2 正态分布

当随机变量  $K_{s,t}$  服从均值  $\alpha = 1$  的正态分布时,为了便于比较,限制  $K_{s,t} \in [0.5, 1.5]$ 。根据“ $3\sigma$ ”原则,有  $P\{0.5 \leq K_{s,t} < 1.5\} = P\{\alpha - 3\sigma \leq K_{s,t} < \alpha + 3\sigma\} = 0.9974$ ,从中求得  $\sigma = 1/6, D(K_{s,t}) = \sigma^2 = 1/36 (\sigma > 0)$ 。

与 3.1 节相似,采用 ProduceKmn 算法,产生随机变量矩阵  $K_{M \times N}$ ,将实验得到的数据填入表 1 和表 2 相应的栏中,重复实验 10 次。和均匀分布相比,当随机变量服从正态分布时,得到的左眼视图  $lImg'$  与原始左眼视图  $lImg$  之间的交叉熵和均方根误差较小,立体化效果较好。

### 3.3 三角分布

为了进行对比,我们通过程序模拟了服从如下分布的随机变量  $K_{s,t}$ ,即

$$f(K_{s,t}) = \begin{cases} -\frac{\pi}{2} \cos(\pi k_{s,t}) & 0.5 \leq K_{s,t} \leq 1.5 \\ 0 & \text{else} \end{cases} \quad (4)$$

此时

$$E(K_{s,t}) = \int_{-\infty}^{+\infty} K_{s,t} f(K_{s,t}) dK_{s,t} = \int_{0.5}^{1.5} K_{s,t} \cdot \left(-\frac{\pi}{2} \cos(\pi K_{s,t})\right) dK_{s,t} = 1$$

我们称之为三角分布。

同样,我们采用该分布重复实验 10 次,将实验结果分别填入表 1 和表 2 相应的栏中。结果表明,当随机变量服从三角分布时,立体化效果介于均匀分布和正态分布之间。

### 3.4 拟合灰度值的分布

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  $lImg'$ 。

$rImg$  的灰度直方图如图 5 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 6 所示, 采用最小二乘法对图 6 中的离散点进行线性拟合。设  $p(k) = p_1 \cdot k + p_2$ , 根据  $rImg$  的灰度值的分布可求得:  $p(k) \approx -0.737 \cdot k + 0.7489$ 。它所对应的曲线如图 6 所示。

我们希望转换后的图像大小不变, 通过将该曲线向右平移  $k_0$ , 向纵轴上移  $c_0$ , 使得随机变量  $k$  的数学期望为 1。设调整后的曲线为  $f(k) = p(k - k_0) + c_0$ , 那么  $f(k)$  必须同时满足

$$\int_{k_0}^{k_0+1} k f(k) = \int_{k_0}^{k_0+1} k [p(k - k_0) + c_0] dk = 1 \quad (5)$$

和

$$\int_{k_0}^{k_0+1} f(k) = \int_{k_0}^{k_0+1} [p(k - k_0) + c_0] dk = 1 \quad (6)$$

根据这两个方程可求得  $k_0 = 0.5614, c_0 = 0.6196$ , 因此

$$f(k) = \begin{cases} -0.737k + 1.782 & 0.5614 \leq k \leq 1.5614 \\ 0 & \text{else} \end{cases} \quad (7)$$

$f(k)$  即为随机变量  $K_{s,t}$  的概率密度函数。

然后可用如下方法构造服从概率密度函数为  $f(k)$  的随机数:

① 求  $f(k)$  的分布函数  $F(k)$ :

$$F(k) = \int_{-\infty}^k f(k) dk = \int_{0.5614}^k (-0.737k + 1.782) dk = -0.3685k^2 + 1.782k - 0.8844 \quad (8)$$

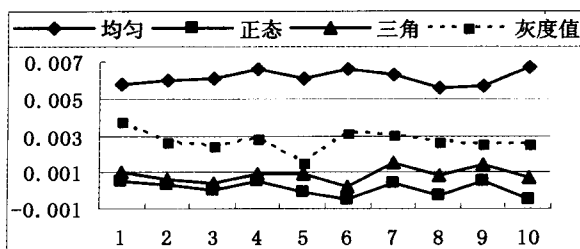
② 因为  $0 < F(k) < 1$ , 令  $r = F(k)$ , 求  $F(k)$  的反函数  $F^{-1}(r)$ :

$$k + F^{-1}(r) = 2.4183 - \sqrt{3.4481 - 2.7137r} \quad (9)$$

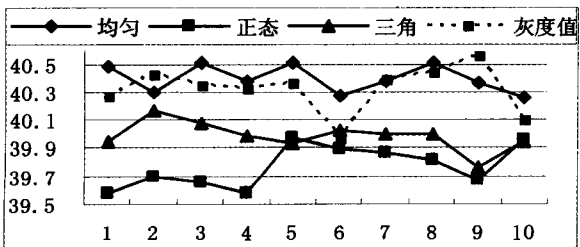
先产生  $[0, 1)$  区间上服从均匀分布的随机数, 再代入 (7) 就可计算得到服从概率密度函数为  $f(k)$  的随机数。

模拟产生的随机数同样也必须不超出舒适 Panum 融合区的限制, 重复实验 10 次, 将实验得到的数据分别填入表 1 和表 2 相应的栏中。根据第一组数据所得到的图像见图 4 (d)。

#### 4 实验结果分析和讨论



(a) 交叉熵  $I[Ilmg', Ilmg]$



(b) 均方根误差  $ERMS(Ilmg', Ilmg)$

图 7 根据表 2 中的数据得到的图

理论上,  $I_{[Ilmg', Ilmg]}$  和  $E_{RMS}(Ilmg', Ilmg)$  的最佳值应该都

是 0, 这时表明  $Ilmg'$  和  $Ilmg$  几乎没有差别。但从图 7 可以看出, 在每次实验中  $I_{[Ilmg', Ilmg]}$  和  $E_{RMS}(Ilmg', Ilmg)$  基本上都稳定在平均值附近, 没有衰减为 0 的趋势。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, 实际观察也证实了这一点。

平面图像立体化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后的立体效果是不准确的。我们首先可以考察交叉熵, 在交叉熵比较小的情况下, 再考察均方根误差, 一般地, 均方根误差的值越小效果越好。可以用这个指标控制平面图像立体化的过程。我们分析比较了随机变量服从四种不同分布的情况, 从实验结果可知, 在大多数情况下, 当图像子块在水平方向上的偏移量服从均值为 1 的正态分布时, 具有最优的转换效果, 其余从优到差依次为三角分布, 拟合灰度值分布和均匀分布。

**结论与展望** 通过在平面图像中随机加入水平视差是一种有效的立体化方法, 我们讨论了当随机变量服从不同分布时转换效果。一般情况下, 采用正态分布立体效果较好。

我们注意到算法 ProduceKmn 需要对产生的随机数进行检查, 因此这个算法是不稳定的。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  $\Delta q_{i,s}$  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加快算法的收敛速度。

#### 参考文献

- 侯春萍, 俞斯乐. 一种平面图像立体化的新方法[J]. 电子学报, 2002, 30(12): 399~402
- Peleg S, Pritch Y, Ben-Ezra M. Cameras for stereo panoramic imaging [J]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000 (1): 208~214
- Hodges L F. Tutorial: Time-Multiplexed Stereoscopic Computer Graphs [J]. IEEE Computer Graphics & Application, 1993, 3:10~20
- Shanks, Ian A. Method and apparatus for producing three dimensional displays; United States Patent, 4414565 [P], 1981-11-8
- Zhu Qingsheng, Liu Ran, Xu Xiaoyan. Properties of a Binocular Stereo Vision Model [A]. In: Proceedings of the 11th Joint International Computer Conference [C], Chongqing, China, World Scientific Press, 2005. 831~834
- Peleg S, Ben-Ezra M. Stereo Panorama with a Single Camera [J]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, 1:395~401
- Sherman W R, Craig A B, 魏迎梅, 等译. Understanding Virtual Reality Interface, Application, and Design[M]. 北京: 电子工业出版社, 93~94
- Atmosukarto I, Cavender A, Jayant C. Reconstructing Antique Stereo Pairs. <http://www.cs.washington.edu/homes/cjayant/finalproject/0-paper.html>, 2006, 3
- Brooks, Frederich P Jr. What's real about virtual reality? [J]. IEEE Computer Graphics and Applications, 1999, 19(6): 16~27
- Yamaguchi K, Takemura H, Yamazawa K, Yokoya N. Real-time Generation and Presentation of View-dependent Binocular Stereo Images Using a Sequence of Omnidirectional Images [J]. In: Proceedings. 15th International Conference on Pattern Recognition, 2000, 4:589~593
- Liu Ran, et al. Stereo Effect of Binocular Stereo Vision Based on Monitor [A]. International Conference on Sensing, Computing and Automation [C]. Watam Press, 2006. 1456~60
- 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[J]. 天津大学学报, 2005, 38(5): 455~465
- 周印, 谢叻. 虚拟现实场景中立体视差生成算法的研究[J]. 计算机仿真, 2005, 22(12): 185~190
- 李畅. 体视图生成算法的研究及应用[J]. 计算机工程, 2005, 31(22): 197~201
- 李斌, 庄天戈. 一种新的立体视觉系统的分析与设计方法[J]. 数据采集与处理, 2000, 15(4): 417~421