

一种平面图像立体化的新方法

侯春萍, 俞斯乐

(天津大学电子信息工程学院, 天津 300072)

摘要: 基于对人类心理立体视觉和生理立体视觉的理论分析和实验研究, 发现在一幅平面图像的各个图像单元中引入随机交叉视差和随机非交叉视差之后, 其双眼深度暗示效果会明显增强的立体视觉现象, 并由此提出了一种新颖且实用的在仅含有心理立体视觉深度暗示的平面图像中随机加入双眼视差的平面图像立体化方法。

关键词: 体视学; 立体视觉; 平面图像立体化; 图像处理

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2002) 12-1861-04

A Novel Method of Picture Conversion from 2D to 3D

HOU Chun-ping, YU Si-le

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

Abstract: Based on the theoretical analysis and experimental research of the human binocular vision and the relationship between psychological stereoscopy and physiological stereoscopy, a novel practical method, to convert 2D pictures into 3D pictures by means of adding random binocular disparities to planar pictures only having psychological depth cue, is presented in this paper. Experiments show that the binocular depth cue can be obviously enhanced by this method.

Key words: Stereoscopy; Stereo vision; Binocular vision; 2D to 3D; Imaging processing.

1 引言

为了使屏幕上重现的自然景物更加生动、逼真, 人类一直在不懈的努力, 于是就出现了黑白电视、彩色电视和正在发展的高清晰度电视。由于人在自然界中所看到的是三维景物, 因此能够显示这种真实三维景物的立体电视必将成为电视技术下一个新的发展方向。与平面电视技术不同, 立体电视技术的发展更加依赖于人类对视觉系统的研究^[1]。目前世界上现存的大多数立体图像都是采用双摄像机拍摄, 也就是模拟人类视觉感知三维景物的方式, 用在水平方向上相隔一定距离的两部摄像机分别记录人左眼看到的平面图像(左图像)和右眼看到的平面图像(右图像)。在严格同步的情况下, 使人的左眼只观看左图像, 右眼只观看右图像, 最后在大脑视区合成立体图像。然而, 这种使用双摄像机获取立体图像的方法, 需要至少两部以上经过光学特性、机械特性和电子特性严格校准的摄像机, 它们的白平衡系统、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密的一致, 否则就无法获得理想的立体图像, 这样就增加了拍摄立体图像的难度。正是由于这个原因, 全世界的立体电视节目犹如凤毛麟角, 即使将人类发明立体电影以来的全部立体电影都包括在内, 也只是一百多部。实际上在平面图像中蕴含着大量的三维信息, 通过一定的方式可以将许多平面图像转换成立体图像。这样做不仅可以降低立体图像的制作成本、缩短制作周期, 而且还可以充分地利用人类社会多年来积累的大

量平面图像、图片。事实上, 平面图像立体化的方法很早就有人提出。1969年, L. Weiner 提出了倾斜投影屏立体成像法^[2], 后来 Shanks 对 Weiner 的方法进行了修改, 他用一个向后倾斜的马鞍型曲面代替了 Weiner 的倾斜投影屏^[3]。本文通过大量的实验研究和分析, 提出了一种实用的在仅含有心理立体视觉深度暗示的平面图像中随机加入双眼视差的平面图像立体化的新方法。

2 立体视觉

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。人类的立体视觉可以分为生理立体视觉和心理立体视觉。

2.1 生理立体视觉

生理立体视觉指由人眼的晶状体调节、双眼会聚和双眼视差等因素构成的立体视觉。其中, 双眼视差是人眼最强烈的生理立体视觉因素。Bela Julesz 利用随机点图已经证明^[4], 在排除掉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, 这说明双眼视差可以与任何视觉经验无关。由于双眼视差对立体视觉的贡献最大, 是最强烈的生理视觉暗示, 所以也是立体成像考虑的主要因素。另一个在立体成像中经常遇到的问题是在立体融像区问题, 又被称为 Penam 融像区^[5,7], 是 Penam 于 1858 年发现并提出的, 主要解决双象单视问题。

收稿日期: 2001-09-03; 修回日期: 2002-07-22

2.2 心理立体视觉

心理立体视觉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的双眼视差异等生理立体视觉识别的深度信息,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视觉记忆和立体视觉经验,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相对深度,人眼的这种立体视觉被称为心理立体视觉.

2.3 生理立体视觉与心理立体视觉之间关系的分析

通过大量的实验我们发现:当人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异时,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或心理立体视觉.当考虑在两只眼睛视网膜上形成的像差异时,它们构成了另一种深度暗示,即生理深度暗示或生理立体视觉.由于被观察的景物是同一个景物,其深度关系是同一个,因此心理立体视觉和生理立体视觉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立体视觉经验被记忆,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.本文所提出的平面图像立体化的方法正是基于这种分析和研究.

3 实验研究

前面分析指出人类的立体视觉是综合作用的结果,而生理立体视觉和心理立体视觉在不同的情况下有相互补充增强或相互抑制减弱的作用,图 1 是为说明两者之间的关系所设计的一个试验的描述,在试验中我们请受试者指出观看到图像的深度感觉.图中的横坐标 X 表示显示屏幕的水平方向,纵坐标 E 表示显示在电视屏幕上的图像给观察者造成的立体深度感觉程度,“•”表示观察者对屏幕上一个图像单元的心理深度感觉,“■”表示观察者对屏幕上一个图像单元的生理深度感觉,“△”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信息之后给观察者造成的综合立体深度感觉.如图 1 中的第一个图像单元给观察者造成的心理深度感觉是 E_{X1} ,随机加入的双眼视差异信息给观察者造成的生理深度感觉是 E_{S1} ,而在这个单元中加入的生理立体视觉信息与原有的心理深度暗示不矛盾,生理立体视觉和心理立体视觉共同形成的综合深度感觉为 E_{Z1} ,这个图像单元的深度感觉得到加强.再看图 1 中的第二个图像单元,其心理深度感觉是 E_{X2} ,随机加入的生理立体视觉给观察者造成的深度感觉为 E_{S2} ,由于心理立体视觉告诉观察者图像单元 2 的深度不可能单元 3 的更远,形成的综合深度感觉为 E_{Z2} ,这时心理立体视觉对生理立体视觉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觉,如图 1 所示.这样得到的综合立体深度感觉符合心理立体视觉深度的变化规律,由于心理立体视觉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眼视差异信息的深度感觉.

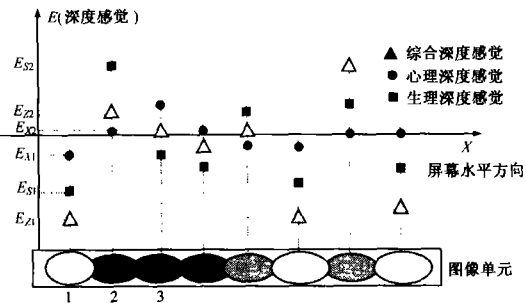


图 1 一个综合的立体视觉试验现象

由此得出结论:当生理深度感觉与心理深度感觉一致时,综合的深度感觉增强;而当生理深度与心理深度感觉矛盾时,生理立体视觉会被心理立体视觉抑制减弱,产生这一立体视觉现象的原因是,心理立体视觉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.

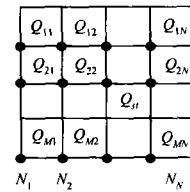


图 2 划分的图像单元

在每个图像单元中随机加入的双眼视差异信息,在心理立体视觉的作用下,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眼视差异图像时,心理立体视觉会自动地对加入的双眼视差异信息进行增强或抑制减弱,从而使观察者获得深度感觉逼真的立体图像.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非常有效.

4 在平面图像中加入双眼视差的方法

由于双眼视差异信息是由两帧平面图像提供的,所以首先应该得到两帧图像.假设要进行立体处理的平面图像为 $f(n, m)$,通过复制可以得到另外一帧完全相同的平面图像,记为 $f_i(n_i, m_i)$.为了加入双眼视差异深度信息,先对图像 $f(n, m)$ 进行处理,然后用处理得到的平面图像 $f_r(n_r, m_r)$ 与 $f_i(n_i, m_i)$ 一起来构成立体像对.

为了能够利用平面图像中的心理深度暗示信息,先把平面图像分解成图像单元,将 $f(n, m)$ 划分成如图 2 所示的 $M \times N$ 个图像子块,即图像单元.每个子块用矩阵 $Q_{11}, Q_{12}, \dots, Q_{M1}, \dots, Q_{MN}$ 表示.假设每个图像子块的水平宽度为 W ,子块中的水平像素数为 K ,垂直宽度为 H .由于 $f(n, m)$ 与左图像 $f_i(n_i, m_i)$ 是同一帧图像,因此只要图像 $f(n, m)$ 中的每个子块 $Q_{11}, Q_{12}, \dots, Q_{M1}, \dots, Q_{MN}$ 在水平方向上随机地偏离了它们原来所处的位置,就等于在立体图像中的各图像子块对应的区域内加入了随机的双眼视差异深度信息,用这样处理得到的图像 $f_r(n_r, m_r)$ 与 $f_i(n_i, m_i)$ 就可以构成立体像对.各图像子块的位移应该满足如下条件:

- (1) 为了使构成的立体图像不存在垂直图像视差,因此每

个图像子块 Q_{st} 只能在水平方向上产生位移,不能在垂直方向上产生位移。

(2)每个图像子块 Q_{st} 在水平方向上的位移应该是随机变化的。

(3)每个图像子块 Q_{st} 在水平方向上最大的位移量不能超出舒适 Penam 融像区的限制。

这三个条件称为图像子块的移动条件.如果各图像子块的位移满足了移动条件,在形成的立体图像中,不存在垂直图像视差,水平图像视差不超出舒适 Penam 融像区的限制。

5 图像子块随机移动的方法及水平视差的控制

在前面的讨论中并没有规定在平面图像中加入双眼视差信息的方法和图像子块的移动方式,因此只要各图像子块的移动满足上一节所述的三个条件,可以采用任何方法使图像子块在水平方向上产生随机移动.本文采用了使各图像子块在水平方向上随机放大或缩小的方法来实现子块的水平位移.这样形成的立体图像,在每个图像点都存在随机变化的双眼视差信息.采用的方法可以用图 3 说明.图 3(a)是原始的图像子块;图 3(b)是各图像子块在水平方向上被随机放大或缩小不同倍数的情况;图 3(c)是在电视监视器上合成左、右图像之后显示的水平视差异.图 3 表明用随机放大或缩小图像子块水平宽度的方法可以达到随机加入双眼视差深度信息的目的.由于图 3(b)中的各图像子块在垂直方向上没有位移,所以在构成的立体图像中也不存在垂直视差异,能够满足图像子块移动条件的第一个条件。

为了使图像子块在水平方向的移动具有随机性,在算法中对每一个图像子块 Q_{st} 都设置一个随机变量 K_{st} ,其中 $s = 1, 2, \dots, M, t = 1, 2, \dots, N$. 为了方便起见,假设随机变量 $K_{11}, K_{12}, \dots, K_{st}, \dots, K_{MN}$ 的概率分布、均值 a 和均方差 σ 相同,且互相独立,在离散集合 $|C| = \{C_1, C_2, \dots, C_i, C_L\}$ 内随机地取值,其中 $C_1, C_2, \dots, C_i, C_L$ 是放大或缩小倍数;当 $C_i < 1$ 时,图像子块的水平宽度缩小,当 $C_i > 1$ 时,图像子块的水平宽度增加.图 3 中第 s 行、第 t 列的图像子块 Q_{st} 经变换后得到其像素点的水平坐标为:

$$q_{tk}^* = W \times (K_{s1} + K_{s2} + \dots + K_{st-1}) + k \times WK_{st} / K \quad (1)$$

其中, $k = 0, 1, \dots, K - 1$, 是子块中的水平像素序号. 这样可以得到立体图像的水平视差为:

$$\Delta q_{tk} = q_{tk} - q_{tk}^* = W \times [(t-1) - (K_{s1} + K_{s2} + \dots + K_{st-1})] + k \times W(1 - K_{st}) / K$$

$$k = 0, 1, \dots, K - 1; \quad t = 1, 2, \dots, N; \quad (2)$$

用这种方法构成的立体图像,水平视差是随机变化的,因而满足了图像子块移动的第二条件.由式(2)可以得到水平图像视差的数学期望值:

$$E[\Delta q_{tk}] = W \times (t-1) \times (1-a) + k \times W(1-a) / K \quad (3)$$

当 $a = 1$ 时, $E[\Delta q_{tk}] = 0$, 说明图像子块 Q_{st} 在原始的水平位置上随机地左、右移动,产生水平交叉视差或非交叉视差,这正是我们所期望的.根据图像子块移动的第三个条件,每个图像子块随机移动形成的水平图像视差不能超出舒适立体融像区的限制,可以用均方差实现对水平图像视差的控制.由式(3)可以得到各图像块水平视差 Δq_{tk} 的最大方差为:

$$D[\Delta q_{tk}]_{\max} = W^2 \times T \times \sigma^2 \quad (4)$$

只要最大视差不超出立体融像区的限制,即有 $\sqrt{D[\Delta q_{tk}]_{\max}} < \Delta q_{\max}$, 即可以实现对各图像子块的水平视差控制,满足图像子块移动的第二条件.当 W 和 Δq_{\max} 确定之后,由式(4)可以得到 σ 的值。

6 验结果和结论

图 4、5、6 是用本文提出的方法和算法对平面图像进行计算机处理之后得到的立体图像对.由于人眼对人物面部特征识别较对自然景物特征的识别更为敏感,因此对含有人物面部特征的平面图像进行立体化处理是一个难点.由于篇幅有限,这里特意挑选了 3 幅含有人物面部特征内容的画面,对其它类型平面图像的处理请参见文献[8].

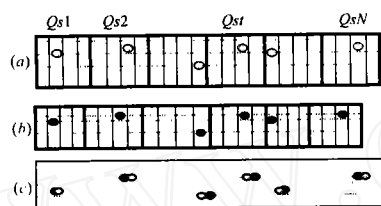


图 3 图像子块随机移动的方法及形成的水平图像视差



图 4 菲律宾小女孩(用交叉视差观看)



图 5 孩子和他们的母亲们(用交叉视差观看)



图 6 调皮的中学生(用交叉视差观看)

试验结果表明,本文提出的平面图像立体化方法是可行的,其立体效果较好,与双摄像机拍摄的立体图像相类似,是一种具有实用前景的平面图像立体化新方法。

参考文献:

- [1] T Troscianko, R Montagnon, J Leclerc, E Malbert, P - L Chanteau. The role of colour as a monocular depth cue [J]. Vision Research, 1991, 31 (11): 1923 - 1930.
- [2] Melvin L Weiner. The Development of a new 3D photographic process Through Utilization of a Projection Screen Method without the Use of Stereoscopic Vision[J]. Applied Optics. 1969, 6: 120 - 125.
- [3] Iana, Shanks. Malvern England [P]. US Patent: 4414565.
- [4] Julesz Bela. A New Sense for Depth of Field[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, PAMI-9(4): 523 - 530.
- [5] John E W Mayhew, John P Frisby. 3D Model Stereoscopic Cues[M]. London, England: The MIT Press Cambridge, Massachusetts, 1991.
- [6] Dimitrios Tzovaras, Nikos Grammalidis, Michael G. Strintzis. Joint three-dimensional motion/disparity segmentation for object-based stereo image [J]. Optical Engineering Bellingham, 1996, 35(1): 137 - 144.
- [7] 荆其诚, 等. 人类的视觉 [M]. 北京: 科学出版社, 1987.
- [8] 侯春萍. 平面图像视觉立体化技术的研究[D]. 天津: 天津大学研究生院, 1998.

作者简介:



侯春萍 女, 1957 年生于吉林, 博士, 副教授. 主要研究领域为三维图像处理、移动通信技术和系统.



俞斯乐 男, 1930 年生于苏州, 教授, 博士生导师. 主要研究领域为数字视频技术, 电视与图像信息系统.

www.cnki.net