平面图像立体化研究

朱庆生[①]　刘　然[①]　许小艳[①]　谢洪波[②]　杨　珺[③]
[①](重庆大学计算机学院　重庆　400044)
[②](天津大学精密仪器与光电子工程学院　天津　300072)
[③](同济大学电子与信息工程学院　上海　201804)

摘　要：将平面图像转换成立体图像很有意义。Hou 等人(2002)提出了一种平面图像立体化方法，称为 Hou 方法。在这个方法中 Hou 等人使用随机变量等参数控制转换过程，但没有讨论这些参数对立体效果的影响。该文利用 Hou 方法在标准的计算机监视器上对含有心理深度暗示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，并讨论了各参数对立体效果的影响。实验结果表明：当随机变量矩阵中的每一个随机变量都服从同一种均匀分布时，随机变量的取值对立体效果没有多大的影响；当监视器屏幕与观察者的距离为一个合适的值(例如 1m)时，图像子块的个数取不同的值对立体效果没有多大的影响。

关键词：图像处理；立体视觉；双目立体视觉；平面图像立体化；转换
中图分类号：TN911.73　　**文献标识码**：A　　**文章编号**：1009-5896(2007)12-2814-05

Research on Image Conversion from Planar to Stereo

Zhu Qing-sheng[①]　Liu Ran[①]　Xu Xiao-yan[①]　Xie Hong-bo[②]　Yang Jun[③]
[①](*College of Computer Science*, *Chongqing University*, *Chongqing* 400044, *China*)
[②](*College of Precision Instrument and Opto-Electronics Engineering*, *Tianjin University*, *Tianjin* 300072, *China*)
[③](*School of Electronics and Information Engineering*, *Tongji University*, *Shanghai* 201804, *China*)

Abstract: It is significant to convert planar images into stereo images. Hou *et al.* (2002) presented a method to realize the function and it is called Hou method. In this method, the parameters such as random variables are used to control the procedure, but the stereo effect that the parameters take on had not been discussed. Using Hou method, the planar image with psychological depth cue is converted into stereo on standard computer monitor. Quantitative criterions to evaluate the stereo effect after conversion are given. And further more, the stereo effect that the random variables take on is discussed. The results of the experiment show that, when every random variable in matrix has the same uniform distribution, different values of these random variables take little effect on stereo effect; when the distance between the monitor screen and the observer is a well chosen value (e.g. 1m), the number of the sub-block images takes little effect on stereo effect.

Key words: Image processing; Stereo vision; Binocular stereo vision; Planar to stereo; Conversion

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。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。已有一些文献对相关的问题进行了探讨[1−3]。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[2]。立体视觉技术有多种，目前使用较多的是双目立体视觉技术。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

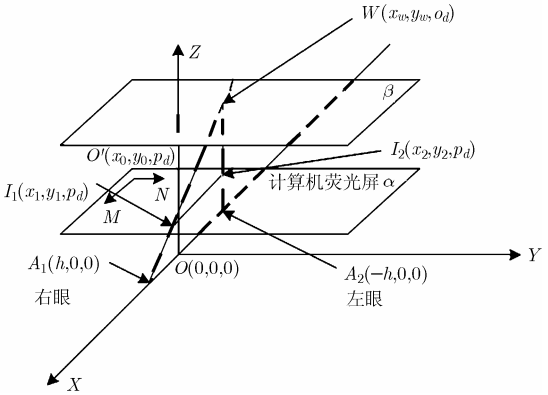

图 1 双目立体视觉模型

2006-04-26 收到，2006-10-16 改回
教育部高等学校博士学科点专项科研基金(20050611027)，重庆大学研究生科技创新基金(200701Y1A0080194)和重庆市自然科学基金(CSTC2006BB2229)资助课题

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，产生立体感。这便是双目立体视觉的基本原理。利用双目立体视觉的基本原理，可以通过标准的计算机监视器(又称为fishtank[4])让观察者看到立体图像。采用监视器作为显示设备的双目立体视觉模型如图 1 所示，所标参数的具体含义参见文献[2]。

Hou等人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(psychological depth cue)和生理深度暗示(physiological depth cue)的原理[1, 4]，将一幅图像划分成如图 2 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一起构成立体图像对。处理的流程图如图 3 所示，具体的过程参见文献[1]。

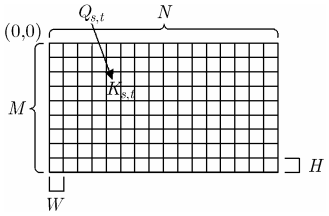

图 2 图像子块

图 3 平面图像立体化流程图

然而 Hou 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，也没有深入讨论随机变量、图像子块的个数对转换后的立体效果的影响。本文利用 Hou 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 节介绍了采用标准的监视器的双目分时立体成像系统，第 3 节给出了评价平面图像立体化的立体效果的两个定量指标——交叉熵和均方根误差，第 4 节讨论了随机变量对转换后的立体效果的影响，第 5 节讨论了图像子块的个数对转换后的立体效果的影响，最后是结束语。

2　双目分时立体成像系统

根据双目立体视觉的基本原理，考察图 1 所示的双目立体视觉模型可知：要让观察者通过监视器看到立体图像，必须使观察者双眼分别看到左右两幅稍有不同的图像。本文采用了双目分时立体成像系统(Time-sharing Binocular Stereo Imaging System)[5]。其中计算机的显卡为NVIDIA显卡，芯片类型为GeForce4 MX 440 with AGP8X。CRT监示器分辨率为 800×600 时刷新率最大能达到 100Hz。系统的组成及其连接方式类似于文献[5]中图 7 所示的方式。

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[5]。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[6]。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。

3　转换效果的评价指标

立体视觉可以看作是一门艺术。然而较少有文献对立体成像的立体视觉效果进行评价[2]。本文根据Hou方法的特点，以下面的定量指标来评价转换效果的好坏(设立体相机按人眼生理参数拍摄的左右眼视图分别为l Img和r Img。算法A对l Img进行立体化，得到右眼视图 $r\,\mathrm{Im}\,g'$，大小与r Img相同)：

(1)交叉熵(Cross-entropy)　设 X是具有有限 n个状态值的随机变量， $p_i=P\{X=x_i\}$; $i=0,1,\cdots,n-1$。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 $P=\{p_0,p_1,\cdots,p_{n-1}\}$ 和 $Q=\{q_0,q_1,\cdots,q_{n-1}\}$ 之间的信息量的差异：

$$I_{[P,Q]}=\sum_{i=0}^{n-1}p_i\mathrm{In}\frac{p_i}{q_i},\ q_i\neq0 \tag{1}$$

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。

从式(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 内容差别很大，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)　设图像的高和宽分别为 m 和 n(单位为像素)，r Im$g(i,j)$表示图像 r Img 在像素点(i,j)处的灰度值，则均方根误差：

$$E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)=\left\{\frac{1}{mn}\left[\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[r\,\mathrm{Im}\,g'(i,j)-r\,\mathrm{Im}\,g(i,j)\right]^2\right]\right\}^{1/2} \tag{2}$$

由式(1)和式(2)可知，如果 $I_{\{r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g\}}$ 和 $E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',\ r\,\mathrm{Im}\,g)$ 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 随机变量

Hou 方法在转换过程中，为了使图像子块在水平方向的移动具有随机性，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 $\boldsymbol{K}_{M\times N}$ [1]。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(每幅图像的大小为 800×600)所示，程序对左眼视图进行立体化。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参数相同。本实验过程中固定取 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 原始图像(右眼视图 $r\,\mathrm{Im}\,g$ ＋ 左眼视图 $l\,\mathrm{Im}\,g$)

实验中，程序产生的 $K_{s,t}$ 服从均匀分布，且满足 $0.5\le K_{s,t}\le 1.5$。之所以要求 $0.5\le K_{s,t}\le 1.5$，是因为要让程序产生的随机变量的均值 α=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(Panum's fusionalarea)[1, 7]的限制。Hou等人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。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[7]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(3)：

$$-\frac{2h\eta p_d}{EP_\sigma}<\Delta q_{t,k}<\frac{2h\eta p_d}{EP_\sigma} \tag{3}$$

其中$2h$是瞳距，一般为6.5cm；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100cm，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[3]；因此可得，$-13<\Delta q_{t,k}<13$。由于立体化的转换过程是一张原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[8]得到的，可以不考虑文献[7]中提到的理论限制。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 $\boldsymbol{K}_{M\times N}$。

$\boldsymbol{K}_{M\times N}$：随机变量矩阵，共 M 行 N 列($M\ge 1$，$N\ge 1$)。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 $\boldsymbol{K}_{M\times N}$。

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

步骤：

for each row K_s in $\boldsymbol{K}_{M\times N}$ do

{

repeat

{

for each element $K_{[s,t]}$ in row K_s produce a random number $K_{s,t}\in[0.5,1.5)$, which has uniform distributions;

if $K_{s,t}$ satisfies the constraint of Panum's Fsional 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;

};

从文献[1]的平面图像立体化方法可以看出，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3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。

表 1 是一组符合要求的随机变量取值。根据这组取值得到的右眼视图如图 5 所示。

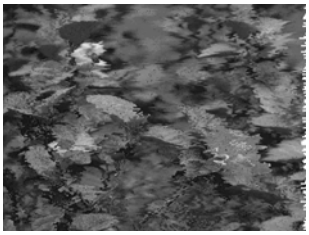

图 5 根据表 1 立体化得到的右眼视图 $r\,\mathrm{Im}\,g'$

表 1　一组符合要求的随机变量取值

随机变量	取值
$K_{1,1}$	0.985380
$K_{1,2}$	1.496020
⋮	⋮
$K_{200,150}$	1.339008

将图 5 和原始图像中的右眼视图进行对比，可得到相应的交叉熵和均方根误差。重复进行上述实验，可得到每次实验的数据如表 2 所示。

表 2　交叉熵和均方根误差

实验次数	交叉熵 $I_{[r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g]}$	均方根误差 $E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)$
1	0.005744	40.48661
2	0.005968	40.29314
3	0.006120	40.50277
4	0.006564	40.38166
5	0.006051	40.50212
6	0.006559	40.26641
7	0.006322	40.37273
8	0.005552	40.51129
9	0.005656	40.35854
10	0.006666	40.26092
平均值	0.006120	40.39362
标准差	0.000326	0.085663

理论上，$I_{[r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g]}$ 和 $E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)$ 的最佳值应该都是 0，这时表明 $r\,\mathrm{Im}\,g'$ 和 $r\,\mathrm{Im}\,g$ 几乎没有差别。但从表 2 可以看出，随机变量取不同的值时，$I_{[r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g]}$ 和 $E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)$ 基本上都稳定在平均值附近，没有衰减为 0 的趋势。这表明同一种分布下，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。

5　图像子块个数

Hou 方法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，随机变量 $K_{s,t}$ 服从均匀分布，且满足 $0.5\le K_{s,t}\le 1.5$。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 节，实验结果如表 3 所示。

表 3　M 和 N 取不同值时的交叉熵和均方根误差

实验次数	M	N	交叉熵 $I_{[r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g]}$	均方根误差 $E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)$
1	24	32	0.018224	40.63781
2	60	80	0.016221	40.67361
3	150	200	0.005744	40.48661
4	300	400	−0.000419	40.12642
5	600	800	0.002914	38.92406
6	1	32	0.026581	40.50308
7	1	80	0.020986	39.33110
8	1	200	0.009365	38.82928
9	1	400	0.001465	39.05174
10	1	800	0.001650	39.41228
…	…	…	…	…

从表 3 可以看出：

(1) 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，从交叉熵和均方根误差这两个指标来看，M 取 1 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 $K_{s,t}$ 服从均匀分布，并且都是对同一副图像 $l\ \mathrm{Im}g$ 立体化。可以计算出 $I_{[l\,\mathrm{Im}\,g,r\,\mathrm{Im}\,g]}$= 0.005149，对应的 $E_{\mathrm{RMS}}(l\,\mathrm{Im}\,g,r\,\mathrm{Im}\,g)$=38.92347。根据上述立体化方法可知，$E_{\mathrm{RMS}}(r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g)$ 不应偏离 $E_{\mathrm{RMS}}(l\,\mathrm{Im}\,g,r\,\mathrm{Im}\,g)$ 太远，$I_{[r\,\mathrm{Im}\,g',r\,\mathrm{Im}\,g]}$ 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，在 M=1 的情况下，当 N 趋于 $r\mathrm{Im}g$ 的宽度时，看到的立体图像虽然要显得平滑，但立体感要减弱。从理论上分析这种现象是可以理解的：当 N 趋于 $r\ \mathrm{Im}g$ 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 $r\,\mathrm{Im}\,g'$ 与 $l\ \mathrm{Im}g$ 相差不大，立体感就会减弱。平滑与立体感是矛盾的，人们只能取一个折中值。

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效果没有多大的影响。

6　结束语

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。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。本文只讨论了当随机变量服从同一种均匀分布时，随机变量的取值对立体效果的影响；当随机变量服从其他分布时对立体效果有什么影响，还需要进一步研究。

参 考 文 献

[1] 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.
Hou Chun-ping and Yu Si-le. A Novel Method of Picture Conversion from 2D to 3D. *Acta Electronica Sinica*, 2002, 30(12): 1861–1864.

[2] Zhu Qingsheng, Liu Ran, and Xu Xiaoyan. Properties of a Binocular Stereo Vision Model. In: Proceedings of the 11th Joint International Computer Conference, Chongqing, China, 2005, Nov., 831–834.

[3] Atmosukarto Indri, Cavender Anna, and Jayant Chandrika. Reconstructing Antique Stereo Pairs. http://www.cs.washington.edu/homes/cjayant/finalproject/0-paper.html, 2006, March.

[4] Sherman William R and Craig Alan 著. 魏迎梅，杨冰等译. 虚拟现实系统——接口、应用与设计. 第一版, 北京: 电子工业出版社，2004: 93–94.

[5] 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.
Liu Yang, Zheng Xiao-lin, and Li Bing. An Implementation of Endoscope Stereo Vision. *Beijing Biomedical Engineering*, 2003, 22(3): 187–190.

[6] NVIDIA Corporation. NVIDIA 3D Stereo User's Guide. Revision 2.0, 2001.

[7] 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): 455–465.
Hou Chun-ping, Azimov Nan, and Yu Si-le. Mathematical Models of Stereoscopic Imagery System and Methods of Controlling Stereo Parallax. *Journal of Tianjin University*, 2005, 38(5): 455–460.

[8] 李自力，朱光喜，朱耀庭. 一种基于会聚双目立体视觉的用户化身模型. 电子与信息学报, 2003, 25(6): 763–770.
Li Zili, Zhu Guang-xi, and Zhu Yao-ting. A User Avatar Model Based-on Convergent Binocular Stereo Vision. *Journal of Electronics & Information Technology*, 2003, 25(6): 763–770.

朱庆生： 男，1956 年生，教授，博士生导师，中国计算机学会和中国电子学会高级会员，中国计算机学会常务理事兼教育委员会副主任和多媒体专委会常委，全国高等学校计算机教育研究会学术委员会主任，《计算机科学》杂志编委。研究方向为图像工程、虚拟植物.

刘 然： 男，1978 年生，博士生，研究方向为虚拟植物可视化.