

文章编号: 0253-2239(2010)07-1989-05

基于 SIFT 匹配算法的多视点自由立体显示 视差图像的生成

张 超 王琼华 李大海 张映权

(¹ 四川大学电子信息学院, 四川 成都 610065
² 四川大学视觉合成图形图像技术国防重点学科实验室, 四川 成都 610065)

摘要 提出了一种利用同一场景的两幅视差图像生成用于多视点自由立体显示的多幅视差图像的方法。首先运用 SIFT 匹配算法从两幅视差图像中寻找匹配点, 并运用生长运算得到图像的稠密匹配; 然后根据匹配点坐标对视差进行了近似计算并得到左视差图像的深度图像; 最后提出一种基于投影原理的多幅视差图像生成方法, 由左视差图像和其深度图像生成了应用于多视点自由立体显示的多幅视差图像, 其立体显示效果良好。

关键词 图像处理; 自由立体显示; 视差图像; 尺寸不变特征变换(SIFT)匹配算法; 深度图像

中图分类号 TN141 **文献标识码** A **doi**: 10.3788/AOS20103007.1989

Acquisition of Parallax Images for Multi-View Auto-Stereoscopic Display Based on SIFT Matching Algorithm

Zhang Chao Wang Qionghua Li Dahai Zhang Yingquan

¹ School of Electronics and Information Engineering, Sichuan University, Chengdu, Sichuan 610065, China

² Key Laboratory of Fundamental Synthetic Vision Graphics and Image for National Defense, Sichuan University, Chengdu, Sichuan 610065, China

Abstract An efficient method to acquire multi-view parallax images for multi-view auto-stereoscopic display from two parallax images is presented. Firstly, the feature matching of the two given parallax images by using SIFT matching algorithm is done and the dense matching the images is accomplished after the propagation of the SIFT matching points. Then, the parallaxes are approximately calculated by the coordinates of the matching points to generate the depth map of the left image. At last, a method using an original image and its depth map to acquire multi-view parallax images is proposed. Good stereoscopic images are presented on an auto-stereoscopic display device after the parallax images were synthesized.

Key words image processing; auto-stereoscopic display; parallax image; SIFT matching algorithm; depth map

1 引 言

多视点自由立体显示器^[1~3]给人们带来身临其境的体验, 有着广泛的应用前景。多视点自由立体显示器的显示内容一般可由两种途径获得: 1) 立体相机阵列对真实场景进行拍摄, 由于采用了多个相

机, 其成本往往很高; 2) 建立计算机立体模型并采用虚拟相机阵列拍摄, 由于立体模型的建立比较复杂, 其成本也较高。这样, 立体片源制约了多视点自由立体显示器的推广使用。立体电影和眼镜/头盔式立体显示都只需要两幅视差图像, 若能利用两幅视

收稿日期: 2009-07-13; 收到修改稿日期: 2009-09-22

基金项目: 国家自然科学基金(60877004)和教育部新世纪优秀人才支持计划(NCET-07-0582)资助课题。

作者简介: 张 超(1983—), 男, 硕士研究生, 主要从事于三维立体显示技术方面的研究。

E-mail: chao831111@yahoo.com.cn

导师简介: 王琼华(1969—), 女, 博士, 教授, 博士生导师, 主要从事三维立体显示技术方面的研究。

E-mail: qhwang@scu.edu.cn (通信联系人)(中国光学学会会员号: S040421103M)

差图像生成多幅视差图像用于多视点自由立体显示,将大大节省拍摄成本,而且能充分利用人类已有的宝贵立体片源。因此,由两幅视差图像生成多幅视差图像技术成为立体显示领域的重要课题。国内已有该方向的研究工作,如基于视差场校正和区域分割的中间视图生成与内插方法^[4]和基于三角形网络的浓密视差图提取方法^[5]等,但这些方法的图像效果、生成速度等都有待提高。

本文通过对两幅视差图像进行特征匹配^[6~8]、区域生长算法^[9~11]、计算匹配点视差^[12]并生成反映场景深度信息的深度图像,最后利用投影关系生成多幅视差图像,用于多视点自由立体显示。

2 基于 SIFT 匹配算法的深度图像计算

2.1 SIFT 特征匹配算法

尺寸不变特征变换(SIFT)算法是一种提取局部特征的算法,在尺度空间寻找极值点,提取位置、尺度、旋转不变量。对图像旋转、尺度缩放和亮度变化保持不变性,对视角变化、仿射变换、噪声也保持一定程度的稳定性。该算法的主要步骤为:

1) 检测尺度空间极值点,初步确定特征点位置和所在尺度。利用不同尺度的高斯差分核与图像 $I(x, y)$ 卷积,生成高斯差分尺度空间,并在相邻的

3 层尺度空间中寻找极值点作为特征点。其中高斯差分尺度空间表示为

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y), \quad (1)$$

尺度不同的高斯卷积核表示为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2], \quad (2)$$

式中 σ 为尺度。

2) 对特征点进行筛选,去除边缘点和低对比度点,使其具有稳定性;

3) 确定特征点方向,利用特征点邻域梯度分布特性为其指定方向参数,使其具有旋转不变性;

4) 生成特征描述子,首先将特征点旋转至主方向,对特征点邻域 16×16 的窗口,划分 16 个 4×4 子块,计算子块内各点的梯度幅值和方向,统计得到 $0 \sim 360^\circ$ 内间隔为 45° 的 8 个方向的特征向量,为每个特征点分配 128 维特征向量。

以图 1 所示的海豚图像为例,首先利用 SIFT 算法对两幅视差图像进行特征点提取,其中左视差图像找出特征点 157 个,右视差图像找出特征点 156 个,然后通过最近邻域匹配得到匹配点 57 个,其中匹配点在图 1 中用点标出。

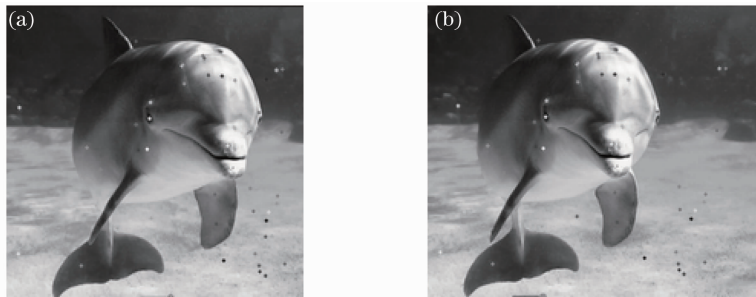


图 1 两幅视差图像的 SIFT 算法匹配结果

Fig. 1 SIFT matching result of the two parallax images

2.2 区域生长算法

SIFT 算法只是完成了离散的特征点匹配,要得到连续的匹配点就需要对离散的匹配点进行区域生长运算。区域生长算法是在已有匹配点的基础上,

判断其邻域内是否有满足相似度条件的点,并将满足条件的点记为新的匹配点。采用零均值归一化互相关(ZNCC) f_{ZNCC} 来衡量两个点的相似度,定义为

$$f_{ZNCC} = \frac{\sum_i [I(X+i) - \bar{I}(X)][I'(X_1+i) - \bar{I}'(X_1)]}{\sqrt{\sum_i [I(X+i) - \bar{I}(X)]^2 \sum_i [I'(X_1+i) - \bar{I}'(X_1)]^2}}, \quad (3)$$

式中 $\bar{I}(X)$ 和 $\bar{I}'(X_1)$ 分别为两点邻域的灰度均值, i 值决定了匹配窗口的大小,本文的匹配窗口为以这

两点为中心的 5×5 区域, f_{ZNCC} 值越大,表示两点相似程度越高。

对两幅视差图像 SIFT 匹配点的 3×3 邻域 8 点依次进行计算, 将非纹理平滑区 ZNCC 值最大且大于 0.5 的两点作为新的匹配点。通过区域生长算法找到两幅视差图像的匹配点如图 2 所示。由图 2

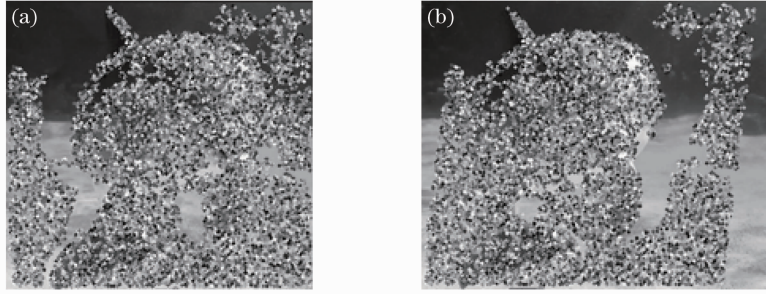


图 2 两幅视差图像生长算法匹配结果

Fig. 2 Matching result of two parallax images by using propagation algorithm

2.3 深度计算

匹配点的坐标差异可以近似为空间场景该点的视差, 视差描述了该点距离观看者的远近^[9], 可分为负视差、正视差和零视差。在多视点自由立体显示时, 负视差、正视差和零视差分别表示再现物体凸出显示屏、凹进显示屏和位于显示屏上。

视差的模值由匹配点坐标的距离表示(单位为 pixel), 视差值正负则由匹配点横向坐标确定, 当左视差图像匹配点横向坐标小于右视差图像对应匹配点横向坐标时, 定义为正视差, 反之为负视差, 若相等则为零视差。视差值反映了该点距离观看者的远近信息, 将最大正视差到最大负视差区间映射到 0~255 的灰度区间内, 从而得到其深度图像。图 3 为左视差图像与其对应的深度图像。深度图像中最亮点表示最近点, 最暗点表示最远点。由于未匹配点无法进行深度计算, 将其灰度值以零表示, 造成了深度的不连续, 因此需要对深度图像进行改进。为了减少人工干预, 采用类似生长算法的灰度生长, 对未匹配点的灰度进行邻域赋值, 最后对整幅图像进行高斯平滑滤波。改进后的左视差图像的深度图像如图 4 所示, 该深度图像可以正确表示场景中各个部

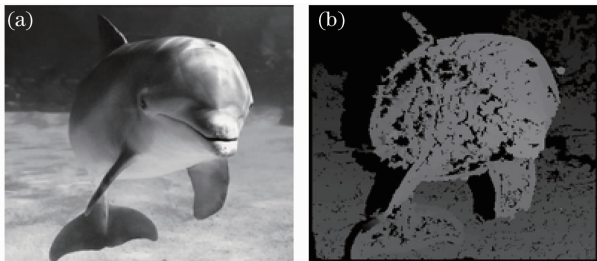


图 3 左视差图像与其对应的深度图像

Fig. 3 Left parallax image and its depth map

中可看出, 匹配点连续并覆盖了图像的大部分区域, 但仍存在未匹配的区域, 这主要是由遮挡和低纹理区域造成的, 它们在图像匹配上属于硬性问题, 将在计算深度的步骤中作近似计算。



图 4 改进后的左视差图像的深度图像

Fig. 4 Improved depth map of the left parallax image 分的相对位置关系。

3 多幅视差图像的生成

利用左视差图像和其深度图像可以生成适用于多视点自由立体显示的多幅视差图像, 该方法主要依据如图 5 所示的投影原理。图 5 中显示屏上黑白间隔的四个矩形表示屏上某一行的四个相邻的水平像素, 显示屏后方的折线为待显示立体场景的四个像素, 它们与显示屏上的像素一一对应。当观看者从观看点 A 观看这四个像素的立体场景时, 其观看效果等同于在点 A 观看它们在显示屏上的投影。其中第一个像素投影到显示屏的 CD 段, 第二个像素投影到显示屏的 DE 段, 依次类推。设折线上第一个像素的灰度值为 X, 第二个像素灰度值为 Y, 观看点 A 到显示屏的垂直距离为 L, Z 为第一个像素相对于显示屏的距离, 设 CD 段长度与显示屏上单个像素长度的比为 a, 则有

$$a = \frac{L}{L + Z} \tag{4}$$

从观看点 A 观看折线, 根据投影关系, 显示屏

上第一个像素的灰度值由折线上的前两个像素共同确定,为

$$G_1 = aX + (1 - a)Y. \quad (5)$$

按照(4)式,(5)式可以依次求得显示屏上各像素的投影灰度。

以上是对立体场景中某一行像素的投影,对于整个立体场景,即每个像素都有一定的深度,根据这种投影方式,在不考虑垂直视差的情况下,可以得到整幅图像的投影图像,即视差图像。在观看区域内设定多个平行放置、间距为人眼瞳距的观看点后,按照上述方法,就可以得到多幅视差图像。

利用两幅视差图像生成适用于分辨率为 $1680 \text{ pixel} \times 1050 \text{ pixel}$ 的 22 inch (55.88 cm) 的 9 视点自由立体显示器的 9 幅视差图像。在视差图像生成之前,确定待显示场景的空间位置,该位置由深度图像和显示器参数共同确定。由于多视点自由立体显示器最

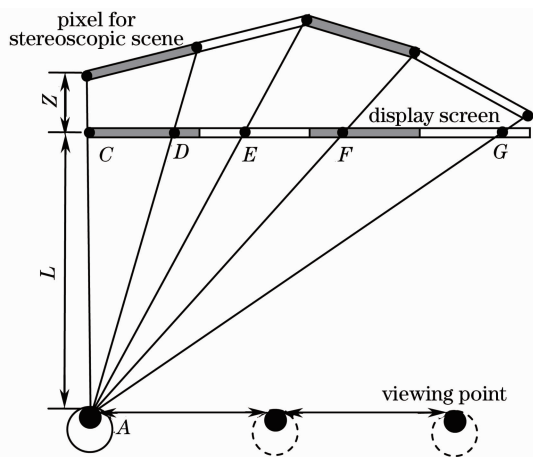


图 5 多幅视差图像生成的投影原理

Fig. 5 Projection principle of generating multi-view parallax images

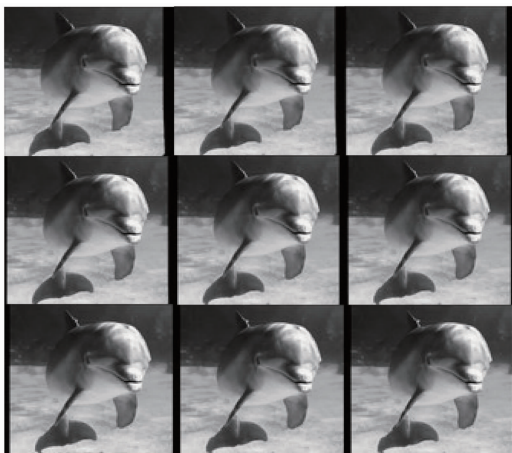


图 6 生成的 9 幅视差图像

Fig. 6 Generated 9 parallax images

大视差的限制,立体场景纵深设定为 500 pixel 的实际宽度,显示屏与场景位置关系设定为:场景纵深的 40% 凸出显示屏,剩余 60% 凹进显示屏。根据投影关系生成的 9 幅视差图像如图 6 所示,将这 9 幅视差图像合成^[13]后在该立体显示器上显示,观看者在立体观看区域内可以明显看到场景中海豚的立体轮廓,海豚头部凸出显示屏,尾部及背景则凹进显示屏,整个场景立体效果明显,图 7 为实拍显示效果图。这种方法的处理速度很快,从匹配开始到视差图像生成只花费了约 4 s 时间。



图 7 自由立体显示效果图

Fig. 7 Stereoscopic effect of the synthesized image on an auto-stereoscopic display device

4 结 论

研究了由两幅视差图像生成用于多视点自由立体显示的多幅视差图像的方法。首先利用图像匹配算法对两幅视差图像进行了特征点匹配,然后根据立体视觉原理对匹配点进行了深度计算,并由其中一幅视差图像和其深度图像获得了多视点自由立体显示所用的多幅视差图像。通过本文方法,为一台分辨率为 $1680 \text{ pixel} \times 1050 \text{ pixel}$ 的 56 cm 九视点自由立体显示器生成了 9 幅视差图像,得到良好的立体显示效果,并且该方法处理速度很快,从匹配开始到视差图像生成只花费了约 4 s 时间。

参 考 文 献

- 1 Wang Qionghua, Tao Yuhong, Li Dahai *et al.*. 3D auto-stereoscopic liquid crystal display based on lenticular lens [C]. Proc. of Asia Display, Shanghai, 2007; 445~455
- 2 N. A. Dodgson. Auto-stereoscopic 3D displays [J]. *Computer*, 2005, **38**(8): 31~36
- 3 Zhao Renliang, Zhao Wuxiang, Wang Qionghua *et al.*. Research on stereo viewing zone in auto-stereoscopic display based on parallax barrier [J]. *Acta Photonica Sinica*, 2008, **37**(5): 960~963
- 4 赵仁亮, 赵悟翔, 王琼华 等. 狭缝光栅自由立体显示器立体可视区域的研究[J]. *光子学报*, 2008, **37**(5): 960~963
- 5 Luo Yan, An Ping, Zhang Zhaoyang. An algorithm for intermediate view image creation and interpolation based on disparity field calibration and region segmentation [J]. *J. China*

- Institute of Communications*, 2004, **25**(10): 127~133
- 骆 艳, 安 平, 张兆扬. 基于视差场校正和区域分割的中间视图图像生成与内插方法[J]. 通信学报, 2004, **25**(10): 127~133
- 5 Yin Yong, Zhang Jianhua. Dense disparity map extraction based on triangulation mesh [J]. *Computer Technology and Development*, 2007, **17**(11): 52~56
- 印 勇, 张建华. 基于三角形网络的浓密视差图提取[J]. 计算机技术与发展, 2007, **17**(11): 52~56
- 6 Li Haibin, Zhang Qiang. A new stereo matching algorithm based on sub-line segment [J]. *Acta Optica Sinica*, 2007, **27**(5): 907~912
- 李海滨, 张 强. 一种新的基于子线段的立体匹配算法[J]. 光学学报, 2007, **27**(5): 907~912
- 7 Wang Xiangjun, Wang Yan, Li Zhi. Fast target recognition and tracking method based on characteristic corner [J]. *Acta Optica Sinica*, 2007, **27**(2): 360~364
- 王向军, 王 研, 李 智. 基于特征角点的目标跟踪和快速识别算法研究[J]. 光学学报, 2007, **27**(2): 360~364
- 8 D. G. Lowe. Distinctive image features from scale-invariant key points [J]. *Int. J. Comput. Vision*, 2004, **60**(2): 91~110
- 9 M. Lhuillier, L. Quan. Robust dense matching using local and global geometric constraints [C]. 15th International Conference on Pattern Recognition, Barcelona, Spain, 2000, **1**: 968~972
- 10 Tang Li, Wu Chengke, Liu Daigang *et al.*. Image dense stereo matching by technique of region growing [J]. *Chinese J. Computers*, 2004, **27**(7): 936~942
- 唐 丽, 吴成柯, 刘侍刚 等. 基于区域增长的立体像对稠密匹配算法[J]. 计算机学报, 2004, **27**(7): 936~942
- 11 G. P. Otto, T. K. Chau. A region-growing algorithm for matching of terrain images [J]. *Image Vision Comput.*, 1989, **7**(2): 83~94
- 12 Ma Songde, Zhang Zhengyou. *Computer Vision Calculation Theory and Algorithm Foundation* [M]. Beijing: Science Press, 1998
- 马颂德, 张正友. 计算机视觉-计算理论与算法基础[M]. 北京: 科学出版社, 1998
- 13 J. Konrad, P. Agniel. Subsampling models and anti-alias filters for 3-D automultiscopic displays [J]. *IEEE T. Image Process.*, 2006, **15**(1): 128~140