# 基于 SIFT 匹配的多视点立体图像零视差调整

李实秋1,雷建军1,周志远2,张海龙3,范晓红1

(1. 天津大学 电子信息工程学院,天津 300072; 2. 天津津航技术物理研究所,天津 300308; 3. 中国电力科学研究院,北京 100192)

摘 要:针对多视点自由立体显示系统,提出一种基于 SIFT 匹配的平行多视点立体图像零视差调整方法。首先,通过 SIFT 变换提取图像特征关键点、对特征关键点进行匹配,获得精确匹配点。然后,运用频率调和显著性模型计算场景的视觉显著性、提取显著性掩膜,以筛选匹配点作为零视差调整的关键点。最后,依据 SIFT 匹配点计算视点间视差值,并基于视差调整原理对多视点立体图像进行零视差调整。实验结果表明,所提出的方法适合多视点自由立体显示系统,零视差调整后的多视点立体图像具有良好的自由立体显示效果。

关键词: 多视点立体图像; 零视差调整; SIFT 匹配; 自由立体显示

中图分类号: TN911.73 文献标志码: A 文章编号: 1007-2276(2015)02-0764-05

# Zero-disparity adjustment of multiview stereoscopic images based on SIFT matching

- Li Shiqiu<sup>1</sup>, Lei Jianjun<sup>1</sup>, Zhou Zhiyuan<sup>2</sup>, Zhang Hailong<sup>3</sup>, Fan Xiaohong<sup>1</sup>
- (1. School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China;
  - 2. Tianjin Jinhang Institute of Technical Physics, Tianjin 300308, China;
    - 3. China Electric Power Research Institute, Beijing 100192, China)

Abstract: A zero-disparity adjustment method based on SIFT matching was proposed for multiview stereoscopic images using in autostereoscopic display system. First, SIFT was introduced for pixel matching between adjacent views. Then, the result of SIFT matching was filtered by saliency mask which was extracted using frequency-tuned saliency model, and the key-point of disparity control was selected. Finally, the disparity between the neighboring views was computed based on SIFT matching points, and zero-disparity adjustment was conducted based on the principle of disparity control. The disparity of selected key-point was adjusted to zero. Experimental results demonstrate that the proposed method can effectively adjust the disparity of multiview stereoscopic images and generate vivid and comfortable 3D scenes for autostereoscopic display.

**Key words:** multiview stereoscopic images; zero-disparity adjustment; SIFT matching; autostereoscopic display

收稿日期:2013-09-04; 修订日期:2013-11-18

基金项目:国家自然科学基金(61271324,61202266,61101224,61202380);

天津市自然科学基金(12JCYBJC10400,12JCQNJC00300,12JCQNJC00500)

作者简介:李实秋(1989-),男,硕士生,主要从事立体图像处理、计算机视觉方面的研究。

通讯作者:雷建军(1980-),男,教授,硕士生导师,博士,主要从事立体显示、立体视频编码、计算机视觉方面的研究。

Email:jjlei@tju.edu.cn

# 0 引言

随着立体电影和电视的兴起,立体图像和视频越来越多地走进了人们的生活。立体图像提供三维立体的视觉效果,与人眼的视觉系统相契合,使观察者获得物体的远近和纵深信息,体验到逼真自然、身临其境的视觉感受[1-2]。

相应地,立体显示与图像处理技术的研究成为 了当前的研究热点。立体显示技术通常分为佩戴辅 助视具的显示方式和自由立体显示方式[3-4]。自由立 体显示技术又称为裸眼 3D 技术,用户无需佩戴眼镜 等辅助视具,即可观看到立体效果。同时,用于显示 的多视点立体图像能保证观众从多个角度舒适地观 看[5-6]。按照摄像机拍摄时的摆放布局,多视点立体 图像可分为会聚阵列多视点图像和平行阵列多视点 图像两大类。会聚多视点立体图像同时具有正负水 平视差,显示时可获得既有凹进屏幕内又有凸出屏 幕外的立体效果,但是存在垂直视差及楔形失真,会 增加观众的视疲劳鬥。平行摄像机拍摄的原始多视点 立体图像只有单向负水平视差,直接进行自由立体 显示只能观察到凸出屏幕的效果, 无法实现三维场 景的逼真再现,需要通过零视差调整,获得同时具有 正、负、零视差的平行多视点立体图像。目前,研究者 们对立体图像的零视差调整开展了一些研究工作。 Chamaret 等提出一种通过动态方式降低会聚-适应 不匹配来减轻视觉不舒适的方法,对两个视点图像 进行调节, 使图像中的感兴趣区域显示在零视差平 面上[8]。Kwon 等人提出了一种双视点零视差调整方 法,该方法对左右视点图像进行预处理后,标注提取 关键物体,计算其在左右视点中的位置差,通过调整 位置差实现零视差调整贸。以上方法主要针对双视点 立体图像对进行调整,缺少对多视点立体图像的研究。 在参考文献[10]中,分析了零视差控制原理的基础上, 通过均值漂移算法将图像分割为若干区域,选取图像 中心区域的几何中心作为图像零视差调整的关键点, 通过平移各视点图像将其视差值调整为零视差,实现 了多视点图像的零视差调整。该方法中的区域分割及 稠密视差图计算有较高的计算复杂度,同时基于图像 分割的方法未充分考虑人眼的视觉特性。

文中针对多视点自由立体显示系统,提出了一

种基于尺度不变特征变换 (Scale Invariant Feature Transform, SIFT)匹配的平行多视点立体图像零视差 调整方法。首先对相邻视点进行 SIFT 匹配,获得特 征匹配点;然后采用频率调和(Frequency-Tuned,FT) 模型从图像中提取显著性掩膜,筛选 SIFT 匹配点, 选择出合适的零视差调整关键点;最后根据关键点 的视差值,依据视差调整原理对各视点进行平移,实 现零视差调整。文中提出的方法具有如下优点:(1) 通过 SIFT 匹配能提取图像的关键特征点,从而可以 获得良好的特征匹配点作为零视差调整关键点;(2) SIFT 匹配获得关键点的同时实现了视差计算,简化 了原有算法的视差计算过程;(3) 利用 FT 显著性掩 膜筛选 SIFT 匹配点,考虑了人眼的视觉特性,保证 了零视差调整后的立体效果。多视点自由立体显示 实验结果表明,采用文中提出方法获得的多视点立 体图像具有良好的立体显示效果。

## 1 基于 SIFT 匹配的零视差调整

## 1.1 视点间 SIFT 匹配

David Lowe 于 1999 年提出了尺度不变特征变换算子,并在 2004 年加以完善<sup>[11]</sup>。SIFT 算法是一种检测图像局部特征的算法,该算法通过求取图像中的特征点及其有关尺度和方向的描述子得到特征并进行图像特征点匹配<sup>[12]</sup>。SIFT 特征具有尺度、旋转、仿射及亮度不变性,对拍摄视角变化、图像噪声具有较高的稳定性,因此适合图像的高精度匹配。SIFT 算法的主要步骤如下:

(1) 尺度空间极值检测。首先对图像建立高斯差分(Difference of Gaussian, DoG)金字塔尺度空间,可由下式表示:

 $D(x,y,\sigma)=L(x,y,k_i\sigma)-L(x,y,k_j\sigma)$  (1) 式中: $L(x,y,k\sigma)$ 由原图像 I(x,y)经尺度为  $k\sigma$  的高斯 滤波器而得,即:

$$L(x,y,k\sigma) = G(x,y,k\sigma) * I(x,y)$$
 (2)

然后在金字塔相邻层中寻找极值点,作为潜在 的对于尺度和旋转不变的兴趣点。

- (2) 特征点定位。筛选出对比度、稳定度较高的 非边缘兴趣点,在兴趣点位置上,确定特征点的位置 和尺度。
- (3) 特征点方向确定。基于图像局部的梯度方向,给每个特征点分配方向,使其保持旋转不变性。

特征点的方向  $\theta(x,y)$ 按下式计算:

$$\theta(x,y) = \text{atan}(L(x,y+1) - L(x,y-1), L(x+1,y) - L(x-1,y))$$
 (3)

式中: $atan(\cdot, \cdot)$ 为反正切函数;L(x,y)则是特征点所在的高斯滤波图像。

(4) 特征点描述。在每个特征点的 16×6 邻域内 计算图像局部梯度的幅值及方向。最终构建一个 128 维的特征向量。

在对相邻视点分别提取 SIFT 特征向量后,采用特征向量的欧式距离作为两个视点中特征点的相似性判定度量。取一个视点的某个特征点,找出其与另一视点特征向量欧式距离最近的前两个特征点,在这两个特征点中,如果最近的距离除以次近的距离少于设定的比例阈值,则接受这一对匹配点。重复以上操作,遍历所有特征点,完成 SIFT 匹配。

#### 1.2 基于 FT 显著性的零视差调整关键点选择

文中采用图像显著性[18-14]对提取的 SIFT 匹配点进行筛选,以选取具有显著性的零视差调整关键点。图像显著性的计算中,采用 FT 模型获得图像的显著性图[15]。像素的显著性定义如下式:

$$S(x,y) = ||I_{\mu} - I_{\omega hc}(x,y)|| \tag{4}$$

式中: I. 为图像平均特征向量, 按下式计算:

$$I_{\mu} = |I|^{-1} \sum_{(x,y) \in I} I_F(x,y) \tag{5}$$

式中:II表示图像 I 中的像素个数。公式(4)中的  $I_{obc}(x, y)$ 是经过高斯滤波后的特征向量,即:

$$I_{\omega hc}(x,y) = \left(\frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)\right) *I_F(x,y)$$
 (6)

在 Lab 色彩空间中,每个像素的 L、a、b 值构成了该像素的特征向量  $I_F = [L, a, b]^T$ ,二范数 $\|\cdot\|$ 则表示特征向量之间的欧氏距离。

获得显著性图后,首先选择局部邻域显著性点。 文中采用像素八邻域显著性极大值点作为局部邻域 显著性点,进而获取显著性掩膜。然后,用显著性掩 膜对上文获取的 SIFT 匹配点进行筛选,缩小零视差 调整关键点的候选集 {P}, 限定关键点为视觉注意 点。最后,对于候选集中的候选点,选择距离图像中 心最近的点 Pkey 作为零视差调整关键点。

#### 1.3 关键点视差计算及零视差调整

由于  $P_{key}$ 是 SIFT 匹配点,其在两个视点中的位置已经确定,因此关键点的水平视差  $d_{key}$  可以按照下

式计算而得:

$$d_{\text{key}} = |x(1, P_{\text{key}}) - x(2, P_{\text{key}})|$$
 (7)

式中: $x(i, P_{key})$ 表示  $P_{key}$  在视点 i 中的水平坐标。

依据平行多视点立体图像零视差调整原理[10],以平行八视点立体图像为例,设视点 1 至视点 8 的水平坐标分别为  $V_1,V_2,\cdots,V_8$ ,进行视差调整后的坐标依次为  $V_1',V_2',\cdots,V_8'$ ,依次将视点 2 至视点 8 平移  $\Delta x, 2 \times \Delta x, \cdots, 7 \times \Delta x$ ,则调整后的视差值为:

$$d_{i}' = V_{i}' - V_{i-1}' = (V_{i} - i \times \Delta x) - (V_{i-1} - (i-1) \times \Delta x) = d - \Delta x, \quad i = 1, 2, \dots, 7$$
(8)

将零视差调整的关键点调整为零视差点,能够获得同时存在正、负、零视差的场景,同时使关键点显示在屏幕平面上,给观众以良好的立体视觉效果。因此,这里令  $\Delta x=d_{key}$ ,将视点 1 至视点 7 的图像按照等差数列分别平移  $d_{key}$ ,  $2\times d_{key}$ , …,  $7\times d_{key}$  的距离,实现多视点立体图像的零视差调整。

# 2 实验结果及分析

为了验证文中提出的零视差调整方法的有效性,采用 ETRI 和 MPEG Korea 提供的 Lovebird1 多视点序列<sup>[16]</sup>和 Middlebury 立体数据库中的 Middl 场景<sup>[17]</sup>进行实验。Lovebird1 序列包含 12 个分辨率为 1 024×768 的平行视点,选择相邻的 8 个视点的图像帧作为待调整的多视点立体图像。Middl 数据库包含 7 个分辨率为 1 366×1 110 的平行视点,采用第 1 视点和第 5 视点及对应的原始视差图,利用虚拟视点生成算法在视点 1 和 5 之间生成 6 个新的视点,组成八视点立体图像作为待调整的多视点立体图像。在对相邻两个视点图像进行 SIFT 匹配时,SIFT 匹配阈值 Th 设为 0.4。

SIFT 匹配结果如图 1 所示,由于相邻视点图像 画面相似度较高,初始的 SIFT 匹配点较多。在 SIFT





图 1 初始 SIFT 匹配点 Fig.1 Original SIFT matching points

匹配点中,一些难以引起人眼视觉注意的不显著点不适合作为零视差点,因此利用 FT 显著性掩膜对 SIFT 匹配点进行筛选,筛选结果如图 2 所示,可见筛选后的候选点大幅减少。对于筛选后的 SIFT 匹配点中,选择距离图像中心最近的 SIFT 匹配点作为零视差调整的关键点,其位置如图 3 所示。





图 2 基于 FT 显著性掩模筛选后的 SIFT 匹配点 Fig.2 SIFT matching points filtered by FT saliency mask





图 3 零视差调整关键点 Fig.3 Key-points of zero-disparity adjustment

文中采用多视点自由立体显示实验来验证零视 差调整算法的有效性。未进行零视差调整的八视点 图像合成图如图 4 所示,采用文中方法调整后的八 视点图像合成图如图 5 所示。由图 4 可见,未进行零 视差调整的八视点图像的合成图中所有区域都是模 糊的,表明立体图像中不存在零视差点。由图5可 见,零视差调整关键点位置附近(Lovebird1中的屋檐 和 Middl 中的灰色帽子)的区域比较清晰,证明其视 差为零或者接近于零。将 Lovebird1 的合成图输入八 视点自由立体显示器显示时, 屋檐部分能够显示在 屏幕平面上; 图像右侧区域的树等显示在屏幕平面 后,具有凹进屏幕的视觉效果;而场景中的情侣则显 示在屏幕平面前,具有凸出屏幕的视觉效果。相似 地,将 Middl 的合成图输入八视点自由立体显示器 显示时, 玩偶小狗、MIDDLEBURY 抱枕都具有凹进 屏幕的效果,灰色帽子的关键点区域则能够显示在 屏幕平面上。经过零视差调整的多视点立体图像观 看感受更加舒适,显示效果更加逼真。





图 4 未经过零视差调整的八视点合成图 Fig.4 Composite images of original eight views





图 5 零视差调整后的八视点合成图 Fig.5 Composite images of adjusted eight views

# 3 结 论

文中提出一种基于 SIFT 匹配的多视点立体图像零视差调整方法。通过 SIFT 匹配实现了相邻视点间特征点匹配及视差计算。同时,结合人眼的视觉注意机制,采用 FT 显著性模型,从 SIFT 匹配点中筛选出合适的零视差调整关键点。最后,根据零视差调整原理,依据视差平移各视点图像,实现了多视点立体图像的零视差调整。实验表明,采用本文提出方法获得的多视点图像适合多视点自由立体显示系统,立体显示效果良好。

#### 参考文献:

- [1] Bose E, Pepion R, Le Callet P, et al. Towards a new quality metric for 3 -D synthesized view assessment [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5 (7): 1332-1343.
- [2] Hu S, Kwong S, Zhang Y, et al. Rate-distortion optimized rate control for depth map-based 3 –D video coding [J]. IEEE Transactions on Image Processing, 2013, 22 (2): 585–594.
- [3] Grau O, Borel T, Kauff P, et al. 3D-TV R&D activities in europe [J]. *IEEE Transactions on Broadcasting*, 2011, 57 (2): 408-420.
- [4] Holliman N S, Dodgson N A, Favalora G E, et al. Threedimensional displays: a review and applications analysis [J].

- *IEEE Transactions on Broadcasting*, 2011, 57 (2): 362 371.
- [5] Dodgson N A. Autostereoscopic 3D displays [J]. *Computer*, 2005, 38(8): 31–36.
- [6] Cellatoglu A, Balasubramanian K. Autostereoscopic imaging techniques for 3D TV: proposals for improvements [J]. *Journal of Display Technology*, 2013, 9(8): 666–672.
- [7] Woods A, Docherty T, Koch R. Image distortions in stereoscopic video systems [C]//SPIE, Stereoscopic Displays and Applications, 1993: 36–48.
- [8] Chamaret C, Godeffroy S, Lopez P, et al. Adaptive 3D rendering based on region-of-interest [C]//SPIE, Stereoscopic Displays and Applications XXI, 2010, 7524: 1–12.
- [9] Kwon K, Lim Y, Kim N. Vergence control of binocular stereoscopic camera using disparity information [J]. *Journal* of the Optical Society of Korea, 2009, 13(3): 379–385.
- [10] Lei J, Zhang H, Hou C, et al. Segmentation-based adaptive vergence control for parallel multiview stereoscopic images [J]. Optik-International Journal for Light and Electron Optics, 2013, 124(15): 2097–2100.
- [11] Lowe D G. Object recognition from local scale-invariant features [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1999: 1150–1157.
- [12] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision,

2004, 60(2): 91-110.

1635-1641.

- [13] Ji Chao, Liu Huiying, Shao Gang, et al. Research on biologically-inspired computional model for image saliency detection [J]. *Infrared and Laser Engineering*, 2013, 42(3): 823-828. (in Chinese)
  - 纪超, 刘慧英, 邵刚, 等. 基于生物激励计算模型在图像显著性提取中的研究[J]. 红外与激光工程, 2013, 42(3): 823-828.
- [14] Xu Dejiang, Shi Zelin, Luo Haibo. Image structural difference clutter metric based on human visual properties [J]. Infrared and Laser Engineering, 2013, 42(6): 1635–1641. (in Chinese) 徐德江, 史泽林, 罗海波. 利用人眼视觉特性的图像结构差异性杂波度量 [J]. 红外与激光工程, 2013, 42(6):
- [15] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009: 1597–1604.
- [16] Um G M, Bang G, Hur N, et al. Contribution for 3D video test material of outdoor scene [S]. ISO/IEC JTC1/SC29/ WG11, M15371, 2008.
- [17] Hirschmuller H, Scharstein D. Evaluation of cost functions for stereo matching [C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2007: 1–8.