

## 结合立体视觉舒适度的立体图像显著性检测\*

周洋, 何永健, 刘晓琪, 唐向宏, 殷海兵



(杭州电子科技大学 通信工程学院, 浙江 杭州 310018)

通讯作者: 周洋, E-mail: zhouyang\_hz@126.com

**摘要:** 针对先前的立体图像显著性检测模型未充分考虑立体视觉舒适度和视差图分布特征对显著区域检测的影响, 提出了一种结合立体视觉舒适度因子的显著性计算模型. 该模型在彩色图像显著性提取中, 首先利用 SLIC 算法对输入图像进行超像素分割, 随后进行颜色相似区域合并后再进行二维图像显著性计算; 在深度显著性计算中, 首先对视差图进行预处理; 然后基于区域对比度进行显著性计算; 最后, 结合立体视觉舒适度因子对二维显著图和深度显著图进行融合, 得到立体图像显著图. 在不同类型立体图像上的实验结果表明, 该模型获得了 85% 的准确率和 78% 的召回率, 优于现有常用的显著性检测模型, 并与人眼立体视觉注意力机制保持良好的一致性.

**关键词:** 立体图像; 视觉显著性; 立体视觉舒适度; 深度显著性; 超像素分割

中文引用格式: 周洋, 何永健, 刘晓琪, 唐向宏, 殷海兵. 结合立体视觉舒适度的立体图像显著性检测. 软件学报, 2017, 28(Suppl. (2)): 1-10. <http://www.jos.org.cn/1000-9825/17013.htm>

英文引用格式: Zhou Y, He YJ, Liu XQ, Tang XH, Yin HB. Saliency detection for stereoscopic images by considering stereo visual comfort. Ruan Jian Xue Bao/Journal of Software, 2017, 28(Suppl. (2)): 1-10 (in Chinese). <http://www.jos.org.cn/1000-9825/17013.htm>

### Saliency Detection for Stereoscopic Images by Considering Stereo Visual Comfort

ZHOU Yang, HE Yong-Jian, LIU Xiao-Qi, TANG Xiang-Hong, YIN Hai-Bing

(School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** In view of the fact that the previous saliency detection models fail to fully consider the effect of stereo visual comfort and the distribution features of disparity values, a saliency computation model considering stereo visual comfort is proposed. In the extraction of color image's saliency, the model first segments an input image into super-pixel regions by using SLIC algorithm, and merges the regions according to color similarity among adjacent regions. After that, the computation of 2D image's saliency is conducted. In the computation of depth saliency, the model first preprocesses the disparity map, and then a regional disparity contrast-based saliency analysis is applied to compute the salient region of the depth map. Finally, the stereo visual comfort factor is embedded into the fusion of the 2D saliency map and depth map to obtain a final stereoscopic saliency image. We evaluated the proposed model for stereoscopic images with various scenarios. The experimental results indicate that the proposed model outperforms existing saliency detection models, yielding an 85% precision and 78% recall rate. Moreover, the saliency region distributions fit well with the human binocular visual attention.

**Key words:** stereoscopic image; visual saliency; stereoscopic visual comfort; depth saliency; super-pixel segmentation

视觉注意力是人类视觉系统处理视觉信息的核心机制, 它帮助人类快速地将人眼注视点集中在视觉显著区域, 进而从大量视觉信息中过滤出有价值的内容. 而图像/视频的视觉显著区域预测可以自动预测、定位和挖掘图像/视频信息中的显著区域, 是视频信号处理和计算机视觉领域的共同研究热点, 被广泛应用在物体探测、

\* 基金项目: 国家自然科学基金(61401132, 61572449); 浙江省自然科学基金(LY17F020027); “电子科学与技术”浙江省一流学科 A 类资助

Foundation item: National Natural Science Foundation of China (61401132, 61572449); Natural Science Foundation of Zhejiang Province of China (LY17F020027); Open Project of “Electronic Science and technology” Zhejiang first class discipline (Class A)

收稿时间: 2017-05-14; 采用时间: 2017-10-17

目标跟踪、视频图像压缩、视频图像修复和内容检索等领域<sup>[1]</sup>.而立体图像一般是由两台左右排列的摄像机(相机间基线距离常为人眼瞳距)对同一目标场景同时拍摄获取的图像,含有丰富的三维空间信息.人类视觉研究表明,对于立体图像,由于眼球的聚焦功能,双眼不能同时关注立体图像中的近处物体和远处物体,必须聚焦于某一区域,导致人类双眼的立体视觉选择性比平面视觉更强,表现在立体图像上其区域显著性更为突出.立体图像显著区域的获取对立体图像压缩、自然场景中物体的识别与跟踪、立体图像质量评价与提升、双目测距系统设计等研究都具有重要的指导意义<sup>[2]</sup>.

近十几年来,研究者们对自然场景的显著性检测展开了深入研究.根据显著性检测机理可将检测算法分成两类.

- (1) 一类是快速的、自底向上的、数据驱动的显著性提取;
- (2) 一类是慢速的、自顶向下的、任务驱动的显著性提取.

由于任务驱动的显著性检测算法通用性不强,大多数检测算法是基于自底向上的数据驱动模型,本文研究也面向自底向上的立体图像显著性提取.针对平面图像显著区域检测,学者们已提出了许多经典算法.Itti 等人<sup>[3]</sup>把各种显著性检测模型通过机器学习的方法进行融合来获取图像的显著区域.Hou 等人<sup>[4]</sup>通过计算傅里叶谱残差(SR)对图像背景剔除来提取显著区域.Achanta 等人<sup>[5]</sup>提出一种基于频率调谐(FT)的计算方法,通过计算图像亮度和颜色特征估计出中心四周差异值来获取图像显著目标.Zhang 等人<sup>[6]</sup>利用自然图像统计特性提出了基于贝叶斯框架的显著性提取模型.

相比于大量的平面图像显著性检测算法,立体图像的显著性检测算法相对较少.不同于平面图像,立体图像的显著性检测必须考虑深度(或视差)信息,它是影响人类立体感知的重要特性<sup>[7]</sup>.Niu 等人<sup>[8]</sup>提出了一种基于视差对比度和立体图像立体感知特征的立体图像显著性计算模型.Cong 等人<sup>[9]</sup>提出了基于多种显著特征混合的显著性检测模型.Wang 等人<sup>[10]</sup>在二维(two dimensions,简称 2D)图像显著图上融入了基于视觉跟踪实验的深度显著信息来获取立体显著区域.Fang 等人<sup>[11]</sup>提出了一种基于颜色、亮度、纹理、深度对比度的立体视频显著性检测框架.Maki 等人<sup>[12]</sup>将深度图作为权重因子加入到 2D 显著图中以获取立体图像显著图.Zhang 等人<sup>[13]</sup>提出了一种面向立体视频的视觉注意机制,它根据深度图的分布特征对二维显著性、运动显著性及深度图进行融合来获取立体视频显著图.Sheng 等人<sup>[14]</sup>基于显著物体本身范围较小以及其与背景处于不同的深度层次,根据显著物体与背景之间的深度对比度来分析立体显著性.Ferreira 等人<sup>[15]</sup>依据像素点被关注的可能性大小建立概率分布模型以求取深度显著图.

鉴于目前的立体图像显著性模型既缺乏考虑深度(视差)图质量高低及其视差值分布特征对深度显著性提取的影响,也很少考虑立体视觉舒适度对双目视觉注意力的调整影响,本文设计了一种结合立体视觉舒适度的立体图像显著性检测方法.在深度显著特征提取中,本文首先对视差图进行预处理,随后通过计算区域对比度进一步分离目标物体与背景来获取深度显著图,提高了深度图显著区域检测的准确性.在 2D 图像显著区域提取过程中,本文对彩色图像进行超像素分割,以块为单位计算显著区域,既提高了检测准确率,也降低了计算复杂度.最后在深度显著特征和 2D 图像显著特征融合部分,加入了立体视觉舒适度权重因子来优化立体图像的显著区域分布.

## 1 算法描述

本文算法总体框图如图 1 所示,它主要包括显著特征提取和显著特征融合两大部分,其中,显著特征提取又分两个模块.

- (1) 二维图像显著特征提取;
- (2) 深度显著特征提取.

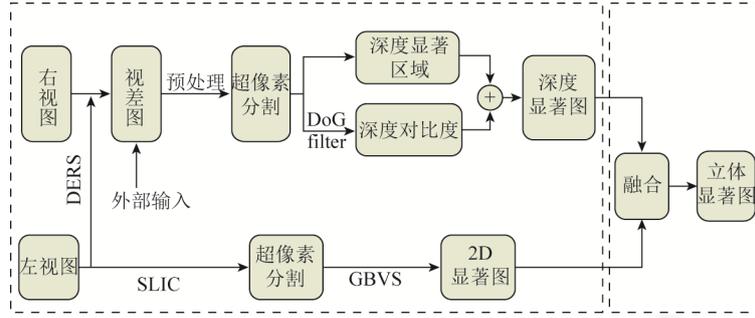


Fig.1 Framework of saliency detection for stereoscopic images

图1 立体图像显著性检测算法框图

### 1.1 二维显著区域提取

为降低显著区域提取的复杂度,本文首先对输入图像进行聚类分割.聚类分割是以区域为单位对图像进行划分,划分后的每个超像素区域都是由亮度、颜色等特征相近的像素点构成的图像块.本文采用 SLIC(simple linear iterative cluster)算法<sup>[16]</sup>对输入的立体图像左视图进行聚类分割来获取左视图的超像素区域.SLIC 算法能够将原图像分割成紧凑均匀的超像素图,并能良好地保留物体轮廓边缘特性.该算法的基本原理是先将原始彩色图像由 RGB 颜色空间转化到 CIELab 颜色空间,然后在  $x,y$  坐标下构造 5 维特征向量的距离计算公式.

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (1)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (2)$$

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2} \quad (3)$$

其中,  $d_c$  表示两点在 CIELab 颜色空间的距离,  $d_s$  表示两点在二维平面的空间距离,  $D$  为所求的距离度量,  $(l_i, a_i, b_i)$  和  $(l_j, a_j, b_j)$  分别表示二维空间上  $(x_i, y_i)$  和  $(x_j, y_j)$  两点像素在 CIELab 颜色空间对应的 3 个分量,  $m$  为迭代次数,默认值取 10;  $S$  为相邻种子点间距,  $S = \sqrt{N/K}$ ,  $N$  为图片像素点总个数,  $K$  为超像素块个数,文中  $K$  取 200.经过 SLIC 算法分割后的超像素块分布如下文中图 2(b)所示,图中的每个分割区域都拥有自己的种子点.接着,通过计算相邻种子点的颜色特征相似性对超像素图进行区域合并,并将初始超像素图中的各图像块  $R_k$  进一步合并为  $R_l$ ,其中,  $1 \leq k \leq K, K=200; 1 \leq l \leq L, L < K$ .超像素块合并后得到标签图像如图 2(c)所示.最后,采用基于图论的显著性计算算法(graph-based visual saliency,简称 GBVS)<sup>[17]</sup>对区域合并后的标签图进行二维空间显著特征计算,进而获取二维空间显著图  $S_{img}$ ,如图 2(d)所示,图中亮度越高代表该区域的显著性越强.

### 1.2 深度显著特征提取

不同于平面图像,立体图像中还包含深度信息(或视差信息),这是双目立体视觉感知的一个重要特征.因此,在立体图像显著性检测中必须考虑反映距离远近的深度信息对立体视觉注意力的影响.深度信息可由深度相机或光场相机拍摄获取,也可利用 DERS(depth estimation reference software)等视差计算软件根据输入的左视点图像和右视点图像计算出相应的视差图,图 3(c)所示为立体图像的深度信息.

通常来说,人眼更容易关注距离自己较近的物体,这意味着前景物体相对于背景物体更为显著.考虑到距离观察者较近的物体具有较大的视差值,而距离观察者较远的物体的视差值较小,因此,在先前的视差图显著性计算中,研究者通常先将视差值归一化,然后将视差图中视差值接近最大视差值  $d_{max}$  的像素点赋予高显著值;相反,将视差值中接近最小视差值  $d_{min}$  的像素点赋予低的显著值.然而值得指出的是,这种方法并不准确、有效.这是因为:(1) 距离近的物体不一定总是显著的,如图 3 所示的地面距离观察者较近而具有较大的视差值,但它却不是观察者所重点关注的区域;(2) 视差图的质量存在高低差异,有些视差图并不能有效地反映物体远近.因此,如何有效地对视差图进行预处理是深度显著性检测中关键的一步,它能有效地提高后续立体图像/视频显著性检

测的准确性.

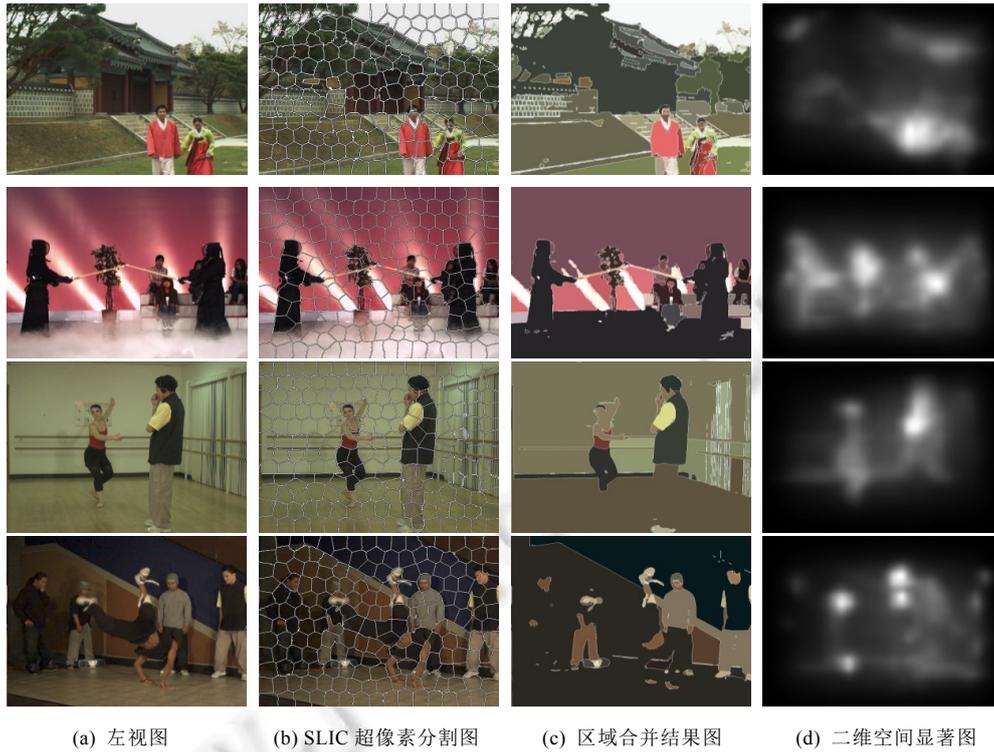


Fig.2 Examples of saliency detection for 2D image

图2 二维图像显著区域检测示意图

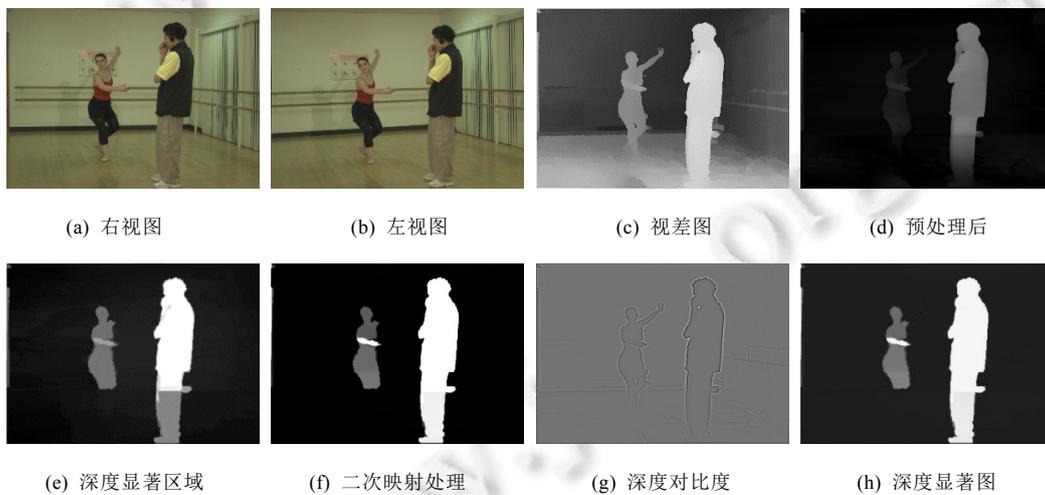


Fig.3 Detection of depth saliency map

图3 深度显著图检测

通过观察图 3(c)中的视差图可以发现:相较于地面背景,目标物体的视差变化相对较为唐突,而地面背景的视差改变较为平缓,即背景区域所在该行的视差值变化较小.因此,本文首先对原始视差图  $D$  进行预处理,即首先计算视差图中每行局部视差改变量,对背景与物体作初步分离,即

$$d'_p = d_p \bullet |d_p - \bar{d}_{row}| \quad (4)$$

其中,  $p$  表示视差图中的该点像素,  $d'_p$  和  $d_p$  分别表示预处理后  $p$  点像素的视差值和原始图中  $p$  点像素视差值,  $\bar{d}_{row}$  表示包含  $p$  点像素的某行像素的视差值均值. 经式(4)预处理后的视差图  $D'$  如图 3(d) 所示, 可以发现, 地面背景的视差值明显下降, 但保留了物体对象的视差.

接着, 本文采用区域对比度法<sup>[18]</sup>对预处理后的视差图  $D'$  进行显著性计算, 其实现过程主要分为两步: (1) 采用 SLIC 算法对输入图像  $D'$  进行区域分割, 种子点数为 200; (2) 计算  $D'$  中某区域  $R_i$  ( $1 \leq i \leq 200$ ) 与周围其他区域  $R_k$  ( $1 \leq k \leq 200$ ) 的视差对比度, 并进行显著性赋值:

$$S(R_i) = \sum_{R_i \neq R_k} n_k d'_R(R_i, R_k) \quad (5)$$

其中,  $S(R_i)$  代表区域  $R_i$  的显著值,  $n_k$  表示视差图中第  $k$  个区域  $R_k$  的像素点个数,  $d'_R(R_i, R_k)$  表示  $R_i$  和  $R_k$  两个区域之间的视差对比度大小. 通常认为较大的区域  $R_k$  对于显著性检测的计算贡献较大, 且视差对比度越大, 对比效果越突出, 相对显著性就更大. 其中, 视差对比度定义为

$$d'_R(R_i, R_k) = \frac{\sum_{p \in R_i, q \in R_j} \omega(p, q) d'(p, q)}{n_i n_k} \quad (6)$$

其中,  $d'(p, q)$  表示  $p$  和  $q$  两像素点视差值的差值, 定义为  $|d'_p - d'_q|$ ; 权重  $\omega(p, q)$  由  $p$  和  $q$  两点之间的空间距离决定, 即  $\omega(p, q) = \exp(-\|p - q\|_2^2 / \sigma^2)$ , 其中,  $\sigma^2$  为控制参数, 默认值取 0.4. 然后将计算得到的显著性图像归一化到  $[0, 1]$  区间, 处理结果如图 3(e) 所示, 可以看出, 本文方法很好地实现了目标背景分离.

最后, 本文基于人眼立体视觉感知特性中屏幕前凸起的物体更为显著的感知特征, 对求得的深度显著图进行二次映射赋值.

$$S'_R(x, y) = \frac{S_R(x, y) - S_{\min}}{S_{\max} - S_{\min}} \quad (7)$$

其中,  $S'_R(x, y)$  代表位于坐标  $(x, y)$  像素点映射后的显著值,  $S_{\max}$  和  $S_{\min}$  分别代表  $S_R(x, y)$  中的最大显著值和最小显著值, 相应的视差显著性计算结果如图 3(f) 所示.

然而, 仅将视差图中的距离信息作为深度显著特征会忽略掉其他重要的感知特征, 比如深度边缘特征和轮廓特征, 而忽略这些立体感知特征会导致检测精度有所降低<sup>[19]</sup>. 鉴于深度对比度能够有效反映物体的深度变化程度, 是深度感知特征的重要衡量指标, 本文采用 DoG (difference of Gaussians) 高斯差分滤波器计算视差图中各像素点的深度对比度值, 具体实现公式如下:

$$f(x, y) = \frac{1}{2\pi\sigma^2} \left[ \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) - \frac{1}{K^2} \exp\left(-\frac{x^2 + y^2}{2K^2\sigma^2}\right) \right] \quad (8)$$

其中,  $(x, y)$  表示该点像素的坐标位置,  $\sigma$  是滤波器尺度因子,  $K$  是高斯核半径比, 在本文的实验中,  $\sigma$  设置为 32,  $K$  值取 1.6 (近似于高斯拉普拉斯算子). 生成的深度对比度图  $S_c$  如图 3(g) 所示, 可以发现, 目标的深度边缘显著特性已得到良好的凸显.

最后, 将以上获取的视差显著图  $S'_R$  和深度对比度图  $S_c$  采用线性加权的融合的方式获取深度显著图  $S_{dep}$ .

$$S_{dep} = w_r S'_R + w_c S_c \quad (9)$$

其中,  $w_r$  和  $w_c$  分别为视差显著图  $S'_R$  和深度对比度图  $S_c$  的权重, 设  $w_r = w_c = 0.5$ . 图 3 描述了整个深度显著图的获取过程, 其中, 图 3(h) 为深度显著性计算的最终结果图.

## 2 立体图像显著区域融合

人眼立体视觉感知特性表明, 过大的双目视差会造成视觉不舒适, 进而影响人类的立体视觉注意力. 所以,

即使物体距离观察者很近并具有较大的显著性,但如果当人们注视时感觉到眼睛不舒服,就会自主地调整视觉注意区域而不再去注视它.因此,在获取立体图像显著区域分布时,必须考虑物体的视差大小是否位于人眼的视差舒适范围之内.因此,本文在融合深度显著图和二维空间显著图时,结合立体视觉感知舒适度因子 $\lambda$ ,并采用线性加权融合的方式获取立体图像显著图.

$$S_{SM}(x,y) = \lambda(\gamma S_{img}(x,y) + (1-\gamma)S_{dep}(x,y)) \quad (10)$$

其中, $S_{SM}(x,y)$ 、 $S_{img}(x,y)$ 和  $S_{dep}(x,y)$ 分别表示立体图像显著图、二维空间显著图和深度显著图, $\gamma$ 表示权重, $\gamma=0.5$ , $\lambda$ 表示 $(x,y)$ 像素点的视觉舒适度.

$$\lambda = \begin{cases} 1 - \frac{|d_p(x,y) - T_{neg}|}{T_{pos} - T_{neg}}, & d_p(x,y) < T_{neg} \\ 1, & T_{neg} \leq d_p(x,y) \leq T_{pos} \\ 1 - \frac{d_p(x,y) - T_{pos}}{T_{pos} - T_{neg}}, & d_p(x,y) > T_{pos} \end{cases} \quad (11)$$

其中, $T_{pos}$ 和  $T_{neg}$ 分别表示视差舒适区域的正、负门限值,其大小通过实验统计手动设置,图 4(e)所示为通过式(11)融合获取的立体显著图.

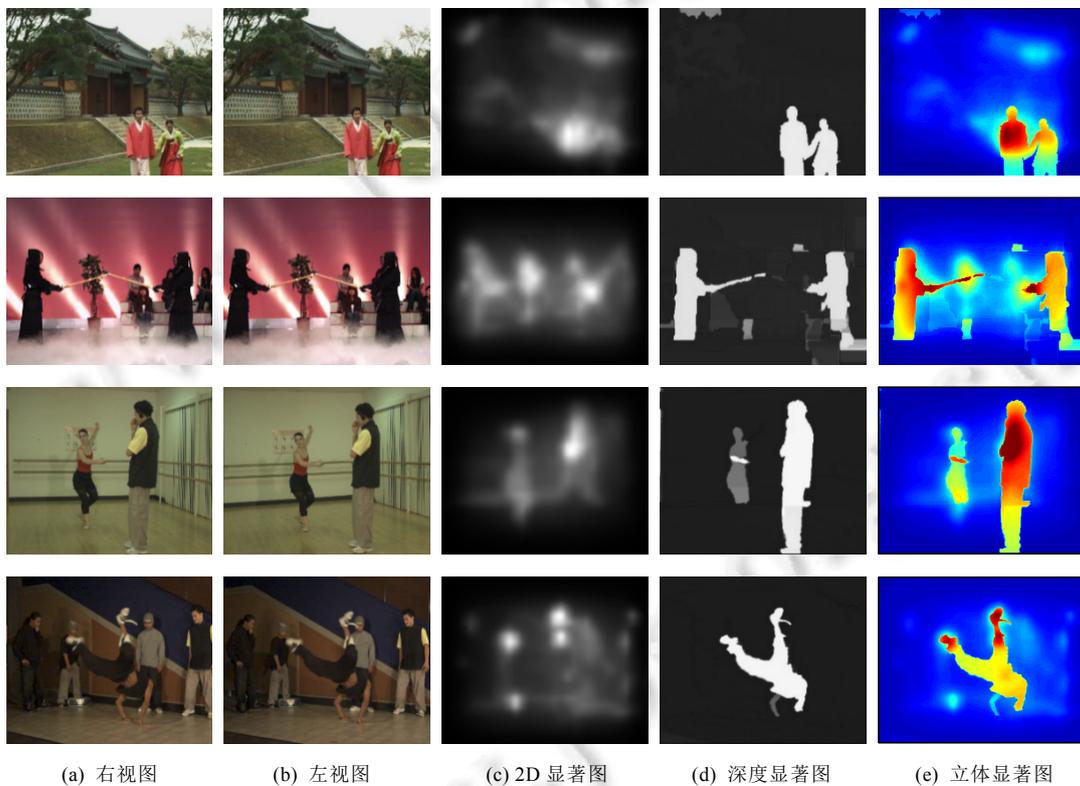


Fig.4 Saliency detection of stereoscopic image

图4 立体图像的显著性检测结果

### 3 实验结果与分析

#### 3.1 立体视频图像测试集

为了验证本文计算模型的有效性,我们采用了 Middlebury 2005/2006 立体图像数据集中的 10 组图像,其中,

这 10 组图像均为室内场景且图像内容特征各不相同.另外,还采用 Lovebird、Café、Pantomime、Balloons、Kendo、Champagne\_tower、Ballet 和 Breakdancers 这 8 组不同场景下的立体序列图像进行实验,除 Lovebird 立体序列图像为室外场景外,其他序列图像均为室内场景,其中,Kendo、Café、Breakdancers 和 Balloons 序列图像的场景较为复杂.

为了得到立体图像显著性的评价基准(ground truth)图,本文首先选用 4 位同学采用文献[20]的方法用矩形框来标注立体图像的显著区域分布,其中,矩形框标签交集度高区域为显著性强的区域;然后召集 20 位同学进行立体视觉追踪实验,获得立体视觉注意力强度分布图;接着根据注意力强度分布图对人工标签出的显著区域进行修正,剔除不准确的标注区域;最后,使用 Photoshop 软件,先根据之前获得的显著区域位置标出显著区,再分割得到显著性评价基准图.图 5 所示为部分立体序列图像的显著性基准图.

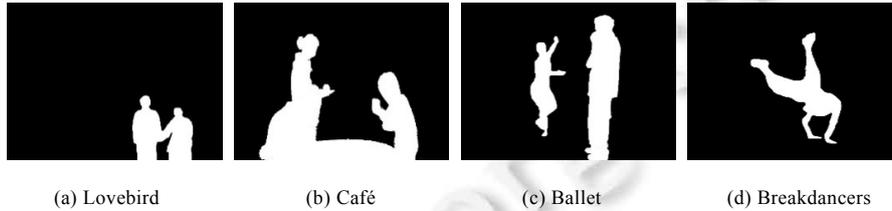


Fig.5 Some saliency benchmarks of the stereo images

图 5 部分立体序列图像的显著性基准图

### 3.2 主观性能比较

为了验证本文提出的显著性检测算法的性能,我们选择了 5 种典型的自底向上显著性预测方法进行性能比较,它们分别是 Zhang 等人<sup>[6]</sup>提出的 SUN 模型、Guo 等人<sup>[21]</sup>提出的 PQFT 模型、Achanta 等人<sup>[5]</sup>提出的 FT 法、Harel 等人<sup>[17]</sup>提出的 GBVS 模型以及 Wang 等人<sup>[10]</sup>提出的 Jun 立体检测模型.

下文的图 6 为上述显著性预测模型和本文方法的显著性预测结果对比图,比较各子图后可以发现:图 6(b)所示的 SUN 模型虽能检测出显著物体以及物体的边缘轮廓特性,但没有很好地将显著目标凸显出来;图 6(c)所示的 PQFT 预测模型只能大致预测出目标物体,显著区域比较模糊;图 6(d)所示的 FT 模型算法运行速度较快,且能消除图像中的一些噪声干扰从而使得显著区域变得清晰,但存在显著区域散乱、显著物体不完整的缺点;图 6(e)所示的 GBVS 预测模型获取的显著区域比较符合人类视觉感知特性,颜色及亮度对比度较强的地方获得较高的显著性,所以本文采用该模型进行二维空间显著性检测,但该模型的显著物体边缘轮廓特征没有得到很好的凸显;图 6(f)所示为 Jun 立体显著性检测模型,它能够完成立体图像的显著性检测,但是由于在深度显著性检测方面,仅仅通过简单的视差映射获取,无法有效地划分目标与背景的显著性,降低了深度显著性检测的准确性,进而误将视差大的地面区域也赋予了较高的显著性.此外,在立体显著性计算中也没有考虑立体视觉舒适度对立体视觉注意力的影响.本文提出的立体图像显著性预测结果如图 6(g)所示,相比于上述方法,由于本文算法在检测过程中同时考虑了视差特征分布、深度对比度以及立体视觉舒适度等多个立体视觉感知因素的影响,所以能够更准确地计算出立体图像的显著性区域,实验结果能与人眼的立体视觉注意区域保持较好的一致性.

### 3.3 客观性能比较

为了进一步客观地对比本文方法与其他模型的立体图像显著性预测性能,本文采用准确率(precision)、召回率(recall)及综合评价指标( $F$ -measure)这 3 种常用的评估准则进行评估,各自相应的计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$$F = \frac{(1 + \beta^2) \times Precision \times Recall}{\beta^2 \times Precision + Recall} \quad (14)$$

其中, $TP$ 表示某像素点位于显著区且也准确判别在显著区域的像素个数, $FP$ 表示位于非显著区域而被误判为显著区域的像素个数, $FN$ 表示位于显著区域而被误判为非显著区域的像素个数;式中参数 $\beta^2$ 用来控制准确性,默认值大小取 0.3.

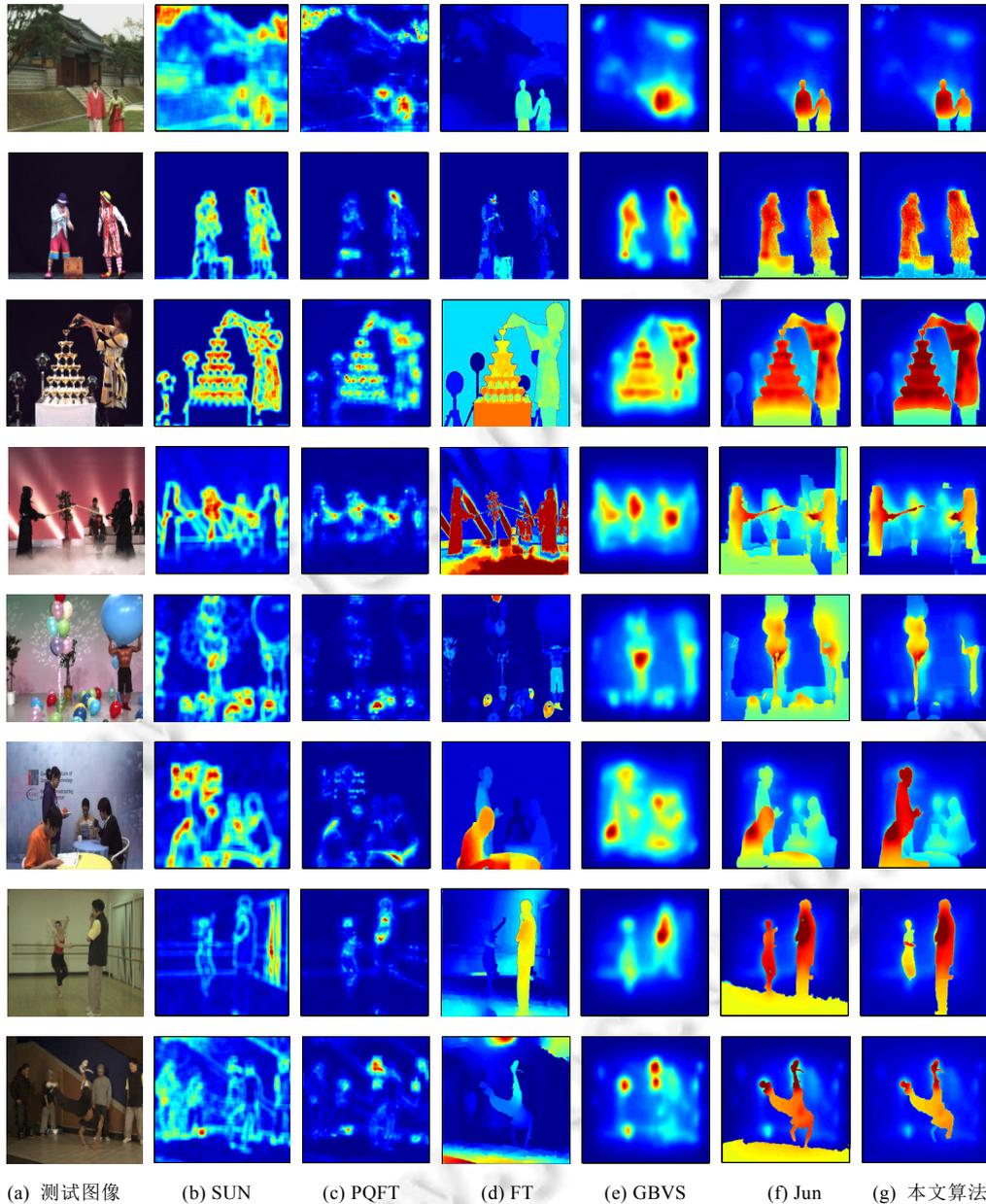


Fig.6 Comparison of saliency detection results of different methods

图 6 不同显著性检测结果对比图

图 7(a)和图 7(b)分别表示本文模型与其他模型计算上述几组立体图像的 Precision-Recall 曲线和  $F$ -Measure 曲线分布图.从图中可以看出,通过对上述多组不同场景下的立体图像进行显著性计算,本文的预测模型由于结合了多种立体视觉注意力感知因素,获得了 85%的准确率和 78%的召回率,相比于其他显著性检测算法,本文算

法获得了更高的准确率和召回率.此外,本文算法获取的  $F$  度量值也优于其他算法,这都从客观评价角度验证了本文方法能够很好地完成立体图像的显著性预测.

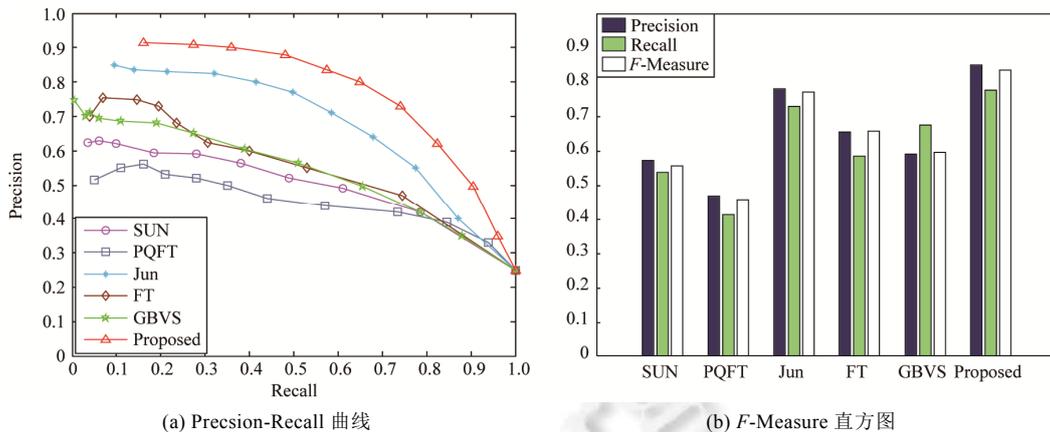


Fig.7 Objective performance evaluation results

图7 客观性能评价结果示意图

## 4 结论

根据人眼立体视觉注意力机制以及立体图像的特点,提出了一种结合立体视觉舒适度的立体图像显著性计算模型,该模型采用了二维显著性与深度显著性相融合的计算框架来求取立体图像显著性.在深度显著性计算中,首先考虑到离观察者近的大视差值区域不一定显著,如舞台地面,此外还存在一些深度图质量较差的问题,不宜直接用来计算深度显著性.因此,本文首先对深度图进行预处理后再计算深度显著性.同时,在显著性计算过程中,本文利用 SLIC 算法对输入的二维彩色图像和预处理后的深度图进行超像素块分割,然后基于图像块进行显著性检测,降低了算法复杂度;同时,在二维显著性与深度显著性相融合的过程中,本文考虑了立体视觉舒适度对立体显著性分布的影响.实验结果表明:本文提出的显著性预测模型能在不同的场景下有效地计算出立体图像的显著性区域,预测结果具有显著对象突出、更符合人眼立体注意力机制的特点,优于当前常用的图像视觉显著性检测模型.

该模型可应用于立体图像编码、自动驾驶导航系统、立体图像质量评价等领域,在下一步的研究工作中,我们将对视差图质量进行评估,并对质量差的视差图进行预处理,进而更有效地挖掘出深度信息与立体图像/视频显著性的关联度,提升立体图像显著性预测的精度和速度.

## References:

- [1] Zhang GM, Yin JP, Zhu E, Mao L. Fast visual attention model algorithm based on approximate Gaussian pyramids. Ruan Jian Xue Bao/Journal of Software, 2009,20(12):3240–3253 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3649.htm> [doi: 10.3724/SP.J.1001.2009.03649]
- [2] Song HK, Liu Z, Du H, Sun GL, Meur OL, Ren TW. Depth-Aware salient object detection and segmentation via multiscale discriminative saliency fusion and bootstrap learning. IEEE Trans. on Image Processing, 2017,26(9):4204–4216.
- [3] Wang J, Borji A, Kuo C-CJ, Itti L. Learning a combined model of visual saliency for fixation prediction. IEEE Trans. on Image Processing, 2016,25(4):1566–1579.
- [4] Hou XD, Zhang LQ. Saliency detection: A spectral residual approach. In: Proc. of the 2007 IEEE Conf. on Computer Vision and Pattern Recognition. 2007. 1–8.
- [5] Achanta R, Hemami S, Estrada F, Susstrunk S. Frequency-Tuned salient region detection. In: Proc. of the 2009 IEEE Conf. on Computer Vision and Pattern Recognition. 2009. 1597–1604.
- [6] Zhang LY, Tong MH, Marks TK, Shan HH, Cottrell GW. SUN: A Bayesian framework for saliency using natural statistics. Journal of Vision, 2008,8(7):1–20.

- [7] Lang C, Nguyen TV, Katti H, Yadati K, Kankanhalli M, Yan S. Depth matters: Influence of depth cues on visual saliency. In: Proc. of the European Conf. on Computer Vision. Springer-Verlag, 2012. 101–115.
- [8] Niu Y, Geng Y, Li X, Liu F. Leveraging stereopsis for saliency analysis. In: Proc. of the 2012 IEEE Conf. on Computer Vision and Pattern Recognition. 2012. 454–461.
- [9] Cong RM, Lei JJ, Zhang CQ, Huang QM, Cao XC, Hou CH. Saliency detection for stereoscopic images based on depth confidence analysis and multiple cues fusion. IEEE Signal Processing Letters, 2016,23(6):819–823.
- [10] Wang JL, Silva DMP, Callet PL, Ricordel V. Computational model of stereoscopic 3D visual saliency. IEEE Trans. on Image Processing, 2013,22(6):2151–2165.
- [11] Fang YM, Wang JL, Narwaria M, Callet PL, Lin WS. Saliency detection for stereoscopic images. IEEE Trans. on Image Processing, 2014,23(6):2625–2636.
- [12] Maki A, Nordlund P, Eklundh J. A computational model of depth based attention. In: Proc. of the IEEE Int'l Conf. on Pattern Recognition. 1996. 734–739.
- [13] Zhang Y, Jiang GY, Yu M, Chen K. Stereoscopic visual attention model for 3D video. In: Proc. of the 16th Int'l Multimedia Modeling Conf. on Advances in Multimedia Modeling. 2010. 314–324.
- [14] Sheng H, Liu XY, Zhang S. Saliency analysis based on depth contrast increased. In: Proc. of the 2016 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. 2016. 1347–1351.
- [15] Ferreira L, da Silva Cruz L A, Assuncao P. A method to compute saliency regions in 3D video based on fusion of feature maps. In: Proc. of the 2015 IEEE Int'l Conf. on Multimedia and Expo. 2015. 1–6.
- [16] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S. SLIC superpixels. In: Ecole Polytechnique Fedrale de Lausanne (EPFL). 2010. 1–15.
- [17] Harel J, Koch C, Perona P. Graph-Based visual saliency. In: Proc. of the NIPS. 2006. 545–552.
- [18] Cheng MM, Mitra NJ, Huang X, Torr PH, Hu SM. Global contrast based salient region detection. IEEE Trans. on Pattern Analysis & Machine Intelligence, 2015,37(3):569–582.
- [19] Didyk P, Ritschel T, Eisemann E, Myszkowski K, Seidel HP. A perceptual model for disparity. ACM Trans. on Graphics, 2011, 30(4):1–10.
- [20] Liu T, Sun J, Zheng NN, Tang X, Shum HY. Learning to detect a salient object. In: Proc. of the IEEE CVPR. 2007. 1–8.
- [21] Guo CL, Zhang LM. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. IEEE Trans. on Image Processing, 2010,19(1):185–198.

#### 附中文参考文献:

- [1] 张国敏,殷建平,祝恩,毛玲. 基于近似高斯金字塔的视觉注意模型快速算法. 软件学报, 2009,20(12):3240–3253. <http://www.jos.org.cn/1000-9825/3649.htm> [doi: 10.3724/SP.J.1001.2009.03649]



周洋(1979—),男,浙江杭州人,博士,副教授,CCF 专业会员,主要研究领域为图像/视频特征分析,三维视频编码与传输.



唐向宏(1962—),男,博士,教授,主要研究领域为信号检测,数字图像修复,数字水印技术,扩频通信技术.



何永健(1992—),男,硕士,主要研究领域为图像/视频处理,视频大数据分析.



殷海兵(1974—),男,博士,教授,CCF 高级会员,主要研究领域为视频编解码算法,芯片结构设计,多媒体通信.



刘晓琪(1994—),女,硕士生,主要研究领域为图像/视频显著性检测,视频大数据分析.