

基于视觉显著性的立体显示图像深度调整

夏振平, 程 成

苏州科技大学电子与信息工程学院, 江苏 苏州 215009

摘要 基于双目视差的立体显示系统中, 不恰当的图像景深分布是引起视疲劳的主要原因, 因此提出一种新颖的立体深度调整方法以提高立体显示观看视觉舒适度。该方法以原始立体图像作为输入, 应用图像谱残差提取的方法获得立体图像的视觉显著性, 结合特定的观看条件建立视疲劳指标模型, 获取特定图像在特定观看条件下的最佳零视差平面(ZDP)。根据最佳 ZDP 对原始输入左右视图进行视差调整, 从而获得视疲劳指标最小的立体图像。为了进一步验证该方法对立体显示视觉舒适度的提升效果, 设计并完成了相应的视觉感知实验。实验结果表明, 该方法显著提高了立体显示观看舒适度。

关键词 视觉光学; 立体显示; 深度调整; 视觉显著性; 视觉舒适度; 视觉感知

中图分类号 TN27 **文献标识码** A

doi: 10.3788/AOS201737.0133001

Stereoscopic Display Image Depth Adjustment Based on Visual Saliency

Xia Zhenping, Cheng Cheng

*College of Electronic and Information Engineering, Suzhou University of Science and Technology,
Suzhou, Jiangsu 215009, China*

Abstract In the stereoscopic display system based on binocular parallax, the improper depth distribution is the main cause of eye fatigue, therefore a novel stereoscopic depth adjustment method is proposed to improve the visual comfort. The proposed method takes the original stereoscopic image as the input. The visual saliency map is derived from original stereoscopic image by applying spectral residual extracting method. Combined with the specific viewing condition, a visual fatigue index model is established to calculate the location of optimum zero disparity plane (ZDP) for specific image in specific viewing condition. The original input stereoscopic image is adjusted according to the optimum ZDP to derive the output stereoscopic image with the minimum visual fatigue index. In order to verify the improvement on visual comfort of the proposed depth adjustment method, the visual perception experiment is designed and implemented. Experimental result indicates that the proposed method significantly improves the stereoscopic viewing comfort.

Key words visual optics; stereoscopic display; depth adjustment; visual saliency; visual comfort; visual perception

OCIS codes 330.7310; 100.6890; 110.2990; 110.3000; 120.2040

1 引 言

得益于电影等娱乐市场的强劲推动, 立体显示技术受到了持续的关注^[1-4]。目前, 大多数电影都采用立体模式, 观众也逐渐习惯这样的形式。由于现在应用的绝大多数立体显示技术都是基于双目视差, 而双目视差中的人眼辐辏和焦点调节的不一致性导致的视疲劳现象是不可避免的^[5]。为了尽可能减少立体显示的视觉疲劳, 立体显示内容拍摄参数的设置需要相当精确。即使能做到精确的拍摄设置, 所获得的立体显示内容也仅适用于特定的显示系统和观看条件, 并非所有条件都适用。除此之外, 显示内容的视觉显著性对视觉舒适度来说也是至关重要的因素^[6]。因此, 一个基于特定的立体显示内容和观看条件的视觉舒适度最优方案亟待产生。

收稿日期: 2016-07-19; **收到修改稿日期:** 2016-08-18

基金项目: 江苏省高校自然科学研究面上项目(15KJB510030)、苏州科技大学科研基金(XKQ201421)

作者简介: 夏振平(1985—), 男, 博士, 讲师, 主要从事立体显示图像质量测量、评价和优化方面的研究。

E-mail: xzp@usts.edu.cn

本文基于图像显著性获取并结合特定观看条件,提出了一种立体图像深度调整方法。该深度调节方法的核心是基于视觉显著性和特定观看条件,建立最佳零视差平面(ZDP)模型。获取该最佳零视差平面后再对输入立体图像进行深度调整,以期获得最佳观看舒适度。并设计了相应的视觉感知实验来验证所提出的方法,视觉感知实验的结果表明,该方法可以显著提升立体显示的观看舒适度。

2 立体显示技术

2.1 立体图像获取

立体拍摄技术大体可以分为平行配置模式和会聚配置模式^[7],如图 1 所示。前者,两摄像机所拍摄的公共区域较小,同时空间中的同一点在左图像中的成像位置总是在右图像成像位置的右侧,如图 1(a)所示。如果左右图像不经过平移而直接通过立体显示系统呈现,则立体场景将全部呈现在立体显示屏前。后者,会带来显著的图像畸变,如图 1(b)所示。如果不是经验丰富的拍摄者,将很难理解图 1(b)中两相机光轴的交叉点以及两相机间的距离对最终成像的作用,因此不能拍摄出舒适安全的立体影像。

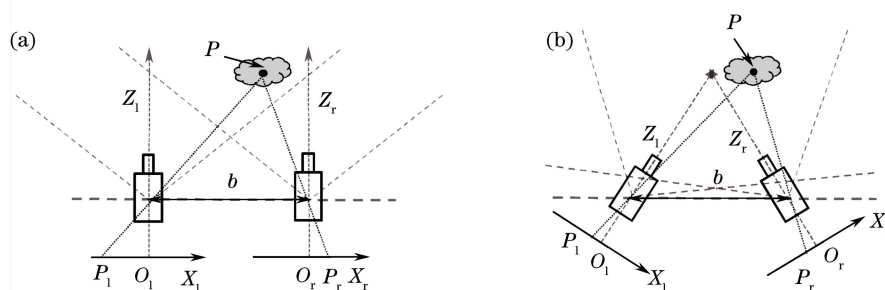


图 1 立体拍摄相机配置模式。(a)平行相机配置;(b)会聚相机配置

Fig. 1 Stereo filming camera configuration. (a) Parallel camera configuration; (b) toed-in camera configuration

2.2 立体图像显示和观看

立体显示技术所呈现的双目视差效果是图像感知中重要的立体线索,如图 2(a)所示。当今的立体显示技术可以简单地分为助视型立体显示和自由(裸眼)立体显示技术。助视型立体显示技术包括最常见的两种形式,即偏振式和主动快门式,观看者需要配戴相应的眼镜来帮助区分左右视图。自由(裸眼)立体显示技术克服了立体显示对眼镜等助视设备的依赖,观看者直接像观察真实世界一样利用裸眼即可观看到立体显示效果。最常用的自由(裸眼)立体显示技术主要有光栅式和透镜式两种类型,两者以不同的形式将具有视差的立体视图投射到不同的空间位置^[8]。

与传统平面显示技术相比,立体技术带来了更强烈的沉浸式临场体验。然而,最常用的基于双目视差的立体显示技术都存在着由视觉上辐辏和焦点调节不匹配造成的视觉不舒适问题。其中,焦点调节是一种人眼想在视网膜上呈现更加锐利的聚焦图像的视觉原理。立体显示的左右视图都呈现在显示屏平面上,因此焦点调节距离是显示屏到人眼的一个固定的距离。双眼辐辏原理是通过眼球围绕垂直轴旋转的方式使观察

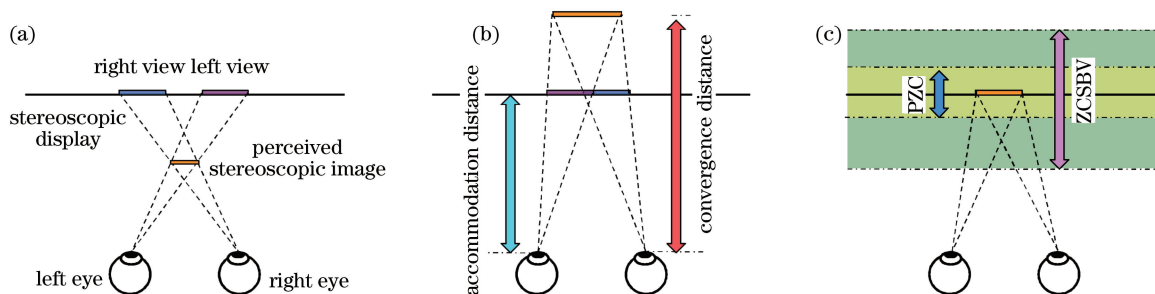


图 2 立体显示以及相应观看视觉机理。(a)立体显示和感知;(b)辐辏和焦点调节的视觉机理;(c) ZCSBV 和 PZC 的定义

Fig. 2 Stereoscopic display and relative observation vision mechanism. (a) Stereoscopic display and perception; (b) vision mechanism with convergence and accommodation; (c) definition for ZCSBV and PZC

图像在双眼中都能呈现在视网膜的中心位置。因此,辐辏距离是人眼到双眼视线交叉点的距离,如图 2(b)所示。人眼观察真实世界的辐辏和焦点调节位置是一致的,而立体显示系统中的这种不一致是造成立体显示观看视疲劳等一系列视觉问题的主要原因^[5]。这种不一致越大,视疲劳越严重。根据已有视觉感知研究结果,研究人员分别定义了双眼单视清晰区(ZCSBV)和珀西瓦尔舒适区(PZC)^[9],如图 2(c)所示。

2.3 视觉显著性

近年来涌现出大量视觉显著性的计算方法和模型^[10-13]。其中,绝大多数方法都基于自下而上的视觉注意力原理,它的特点是速度快、前意识和数据驱动^[14]。这些模型中的图像显著性用于估计人的视觉注意力,同时基于一种人脑只集中于分析观看到图像的一小部分内容的假设,图像中这样的一小部分内容通常由低层次图像特征的变化而获得,这些低层次图像特征包括亮度、颜色、方向、边缘等。图像显著性被广泛应用于图像识别、图像分割、图像尺寸调整、图像压缩等图像处理技术中^[13,15-16]。立体图像显著性包含了立体深度信息,其是由平面图像显著性和从深度信息提取的深度显著性融合后而获得。然而,深度显著性对整个立体图像显著性的贡献如何表示尚未清晰,目前大多采用简单平均的方式,即各占 50% 的贡献度^[17]。

3 立体深度调节方法

为了降低由于人眼在观看立体显示图像时辐辏和焦点调节不一致所产生的视觉不舒适度,基于图像显著性的检测,同时融合特定观看条件提出了一种立体图像深度调整方法。

3.1 方法架构

图 3 以两视图立体图像为例给出立体深度调整方法的实施流程。原始左右视图作为整个系统的输入部分。首先,通过对左视图应用图像谱残差提取的方法获取图像显著性图。然后,结合特定的观看条件建立最佳零视差平面模型,根据该模型获取特定立体图像在特定观看条件下的最佳零视差平面。最后,原始输入立体图像基于获取的最佳零视差平面经过视差调整后获得调整后的立体图像左右视图,即为系统的输出部分。

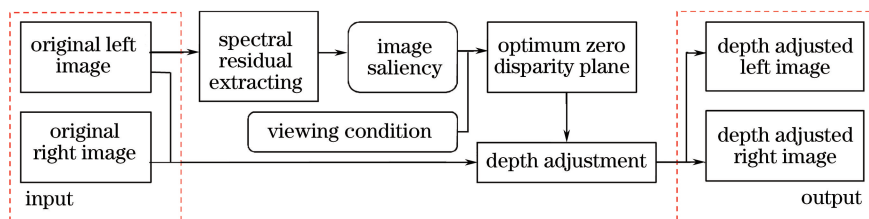


图 3 立体图像深度调整方法实施流程图

Fig. 3 Flow chart of the stereoscopic depth adjustment method

3.2 图像显著性检测

在众多的图像显著性检测方法中,本文系统采用了一种简单可行的方法,即谱残差提取方法^[11]。这种图像显著性检测方法不依赖于目标的特征、类别或者其他先验知识,仅分析输入图像的对数谱以获取谱残差。图像显著性图再基于谱残差在空间域进行重构获得。谱残差 $R(f)$ 表示为

$$R(f) = L(f) - A(f), \quad (1)$$

式中 $L(f)$ 是对数谱, $A(f)$ 是对数谱的一般形状。

然后,图像显著性图 S 可以表示为

$$S(x) = g(x) \mathcal{F}^{-1} \{ \exp [R(f) + P(f)] \}^2, \quad (2)$$

式中 $g(x)$ 是用来平滑显著性图的高斯滤波器, $P(f)$ 是图像的相位谱, \mathcal{F}^{-1} 是傅里叶逆变换符号。

3.3 最佳零视差平面模型

建立最佳零视差平面模型是为了获取针对特定立体图像和特定观看条件下的最佳立体深度调整参照标准,从而使调整后的立体图像呈现最佳的观看舒适度。图 4 是立体显示及其观看系统的几何结构。立体图像在左右视图存在视差的情况下,人眼感知到的立体图像(以 P 点表示)就会在距离显示屏平面一段距离处融合,观看到的立体图像与显示屏平面之间的距离可以表示为

$$Z_d = \frac{D_d \times V_d}{D_d - D_{IP}}, \quad (3)$$

式中 D_d 是立体图像左右视图的视差在距离上的表达, V_d 是从人眼到显示屏的观看距离, D_{IP} 是双眼瞳孔间的距离。

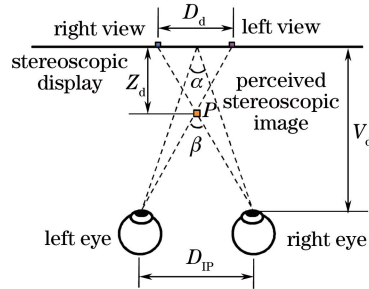


图 4 立体显示观看系统几何结构

Fig. 4 Geometry of stereoscopic display viewing system

考虑到立体图像的观看条件可能随着显示器尺寸、分辨率、观看距离等的变化而变化,以实际距离 Z_d 的方式表达人眼辐辏和焦点调节机理之间的冲突程度并不恰当。因此,提出以角度差的方式来表达双眼视差,表达式为

$$D_a = \alpha - \beta = 2 \times \left[\arctan\left(\frac{D_{IP}/2}{V_d}\right) - \arctan\left(\frac{D_{IP}/2}{V_d - Z_d}\right) \right], \quad (4)$$

式中 α 是双眼视线在显示屏面相交时的夹角, β 是立体图像呈现在显示屏面以外空间平面时的双眼视线夹角。当观看的立体图像恰好位于显示屏平面时,角度视差 D_a 为零,如图 2(c)所示;当立体图像在屏前融合时,角度视差 D_a 为负,如图 2(a)所示;若角度视差 D_a 为正,则意味着立体图像呈现在显示屏面后方,如图 2(b)所示。

根据角度视差的定义,角度视差图可以根据立体视差图(基于像素差)和设定的观看条件计算而得。本文方法中,立体视差图(基于像素差)利用基于分割的立体图像匹配方法获得^[18]。由于立体显示视觉不舒适度与角度视差绝对值大小的正相关性,本文定义的视疲劳指标图是图像显著性图与角度视差图绝对值的乘积,表示为

$$I_{vf} = S \times |D_a|. \quad (5)$$

系统通过调整零视差平面的位置,使整个立体场景的角度视差范围从 $-R_{da}$ 变化到 R_{da} , 每调整一次零视差平面都需要计算相应视疲劳指标图,最终整个变化范围内最小的视疲劳指标图对应的零视差平面即为最佳零视差平面。其中, R_{da} 是特定立体图像在特定观看条件下的角度视差范围,表示为

$$R_{da} = D_{a\max} - D_{a\min}, \quad (6)$$

式中 $D_{a\max}$ 和 $D_{a\min}$ 分别是特定立体图像场景中的最大角度视差和最小角度视差。

3.4 模型及方法应用

根据所提出的立体图像深度调整方法,实施结果如图 5 所示。基于图像显著性检测[图 5(b)],可以确定视觉注意区域。结合图像角度视差[图 5(c)]和图像显著性图,可以获得视疲劳指标,视疲劳指标的最小值即对应最佳零视差平面。获得的最佳零视差平面用于对原始输入立体图像[图 5(a),立体图像为了直观表达均采用红蓝补色形式]进行视差调整,从而获得深度调整后的立体图像[图 5(d)],对应最佳视觉舒适度。

4 视觉感知实验验证

4.1 实验设计

为了验证所提出的立体图像深度调整方法对观看舒适度的提升效果,设计了相应的视觉感知实验。为了验证所提方法在静态立体图像和立体视频中的表现,实验分为两组实施。实验一挑选了 8 幅立体图像,包含室内图像、户外图像、自然图像和计算机生成图像,包含原始立体图像和深度调整后图像,实验中被试观看

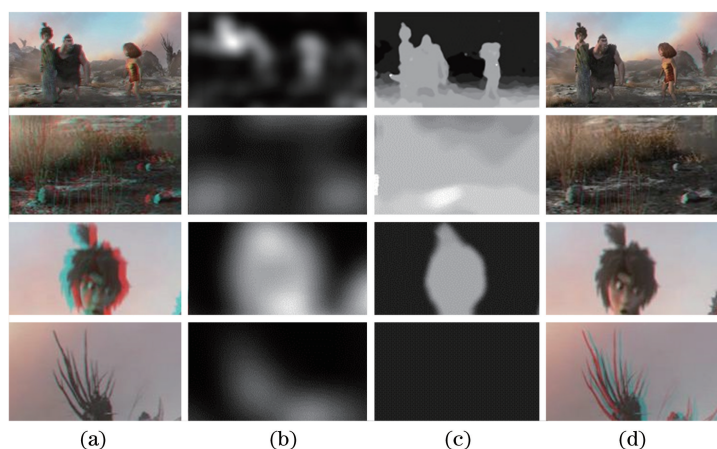


图 5 立体深度调整方法应用。(a)原始立体图像;(b)图像显著性图;(c)角度视差图;(d)深度调整后的立体图像

Fig. 5 Implementation results of proposed method. (a) Original stereoscopic image; (b) image saliency map; (c) angular disparity map; (d) depth adjusted stereoscopic image

的图像共有 16 幅。实验采用了偏振式立体显示系统,测试图像随机出现,被试观察每幅图像 8 s 时间后对视觉舒适度进行评分。实验设置如图 6 所示,被试者端坐在显示屏 3.2 倍屏高的距离处^[19],正对显示屏中央位置观看立体图像。评分标准采用 5 分制^[20],5 分为毫无影响,4 分为轻微影响,3 分为中度影响,2 分为明显影响,1 分为难以忍受,具体分值对应视觉舒适度指标如表 1 所示,总共 16 名被试者参与实验,包含 8 名女性,平均年龄 20 岁。

表 1 视觉感知实验评分标准

Table 1 Grading standard of visual perception experiment

Grade	Visual comfort
5	Comfortable
4	Hardly uncomfortable
3	Slightly uncomfortable
2	Uncomfortable
1	Extremely uncomfortable

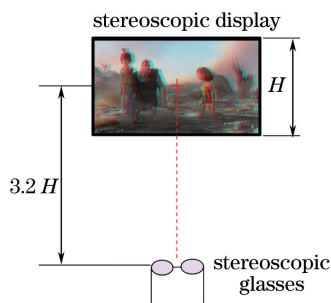


图 6 实验设置

Fig. 6 Experimental setting

实验二选择了 4 段时长都为 3 min 的立体视频作为测试内容,包含原始立体视频和深度调整后的立体视频,被试观看的立体视频共 8 段。为了更全面地验证视疲劳的改善效果,实验结合眼科学范畴内的生理因素,采用 6 项症状作为主观评价指标,这 6 项涉及视疲劳的症状分别是眼疲劳、视力模糊、集中注意力困难、头晕、头疼、恶心^[21-23]。每一症状的描述同样采用 5 分制。被试者观看随机播放的每段视频后进行问卷评分,随后有足够的休息时间以恢复状态。

4.2 实验结果

实验结果采用 SPSS 软件(IBM SPSS Statistics 22.0)进行分析。实验一中,方差分析(ANOVA)(表 2)结果表明,所采用测试图像原始和调整后的立体深度对视觉舒适度评分有高度显著的影响($p < 0.01$),而测

试图像及其与深度的交叉对评分没有显著性影响($p > 0.05$)。表 2 中 X_{DOF} 表示自由度, F 表示组方差值, p 表示显著性水平。不同立体深度下视觉舒适度评分的平均值及其 95% 置信区间, 如图 7 所示, 表明所提出的立体图像深度调整方法对立体观看舒适度有显著的提高。实验二中, ANOVA 分析结果表明, 立体深度对眼疲劳、视力模糊、集中注意力困难和头晕 4 项症状有显著的影响($p < 0.01$), 而对头疼和恶心症状没有显著影响($p > 0.05$), 这是由于测试内容基本没有产生明显的头疼和恶心现象。6 项视疲劳症状指标的评分平均值及其 95% 置信区间, 如图 8 所示, 所有症状评分中深度调整后的内容都大于原始内容, 表明所提出的深度调整方法对视频内容的观看舒适度也有显著的提高。

表 2 各变量因素对视觉舒适度评分的 ANOVA 分析结果

Table 2 Results of ANOVA analysis for different factors on visual comfort

Factor	Visual comfort		
	df	F	p
Image	7	1.168	0.322
Depth	1	7.526	0.007
Image \times Depth	7	0.056	1.000

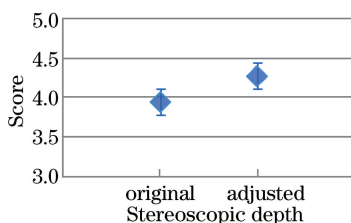


图 7 不同立体深度下视觉舒适度评分平均值

Fig. 7 Mean scores of visual comfort for different stereoscopic depths

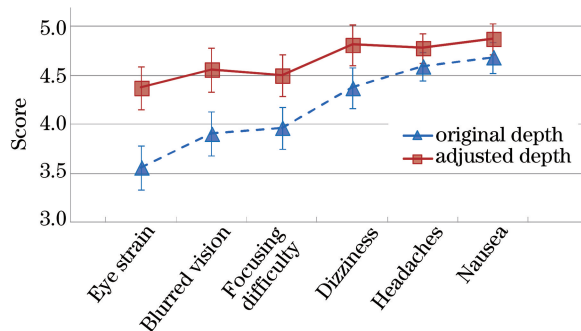


图 8 不同立体深度下 6 种视疲劳症状评分平均值

Fig. 8 Mean scores of six visual fatigue symptoms for different stereoscopic depths

5 结 论

双目视差的立体显示技术都存在人眼辐辏和焦点调节不一致的问题, 这种视觉机理的冲突是立体显示中视疲劳的主要原因。为了尽可能地避免辐辏和焦点调节冲突带来的视觉不舒适现象, 提出了一种基于图像显著性检测并融合观看条件的立体图像深度调整方法。该方法的核心是基于视觉显著性和视差程度与视疲劳的相关性建立视疲劳指标模型, 从而获取特定图像内容在特定观看条件下的最佳零视差平面, 并根据该最佳零视差平面对立体图像深度进行调整。视觉感知实验结果表明, 所提出的立体图像深度调整方法对立体观看舒适度有显著性的提高。

参 考 文 献

[1] Xia Zhenping, Li Xiaohua, Cui Yuan, *et al.* Crosstalk analysis and cancellation method study in stereoscopic displays [J]. Acta Optica Sinica, 2012, 32(10): 1012002.

- 夏振平, 李晓华, 崔 渊, 等. 眼镜式立体显示中的串扰分析及消除方法研究[J]. 光学学报, 2012, 32(10): 1012002.
- [2] Xia Zhenping, Li Xiaohua, Chen Lei, *et al.* Objective evaluation of flicker in stereoscopic display with active shutter glasses[J]. *Acta Optica Sinica*, 2013, 33(12): 1233001.
夏振平, 李晓华, 陈 磊, 等. 快门眼镜式立体显示中闪烁的客观评价[J]. 光学学报, 2013, 33(12): 1233001.
- [3] Xia Zhenping, Li Xiaohua, Chen Lei, *et al.* Study on the effect of ambient light flicker in stereoscopic display with active shutter glasses[J]. *Acta Optica Sinica*, 2014, 34(7): 0733002.
夏振平, 李晓华, 陈 磊, 等. 对快门式立体显示中环境光闪烁的影响研究[J]. 光学学报, 2014, 34(7): 0733002.
- [4] Xia Zhenping, Li Xiaohua, Chen Lei, *et al.* Study on evaluation of motion blur in binocular parallax based stereoscopic displays[J]. *Acta Optica Sinica*, 2015, 35(1): 0111001.
夏振平, 李晓华, 陈 磊, 等. 基于双目视差的立体显示运动模糊评价方法研究[J]. 光学学报, 2015, 35(1): 0111001.
- [5] Hoffman D M, Girshick A R, Akeley K, *et al.* Vergence-accommodation conflicts hinder visual performance and cause visual fatigue[J]. *Journal of Vision*, 2008, 8(3): 33.1-30.
- [6] Jung Y, Sohn H, Lee S, *et al.* Predicting visual discomfort of stereoscopic images using human attention model[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013, 23(12): 2077-2082.
- [7] Kishi S, Abe N, Shibata T, *et al.* Stereoscopic camera system with creator-friendly functions[C]. *SPIE Electronic Imaging, International Society for Optics and Photonics*, 2009, 7237: 72371M.
- [8] Urey H, Chellappan K V, Erden E, *et al.* State of the art in stereoscopic and autostereoscopic displays [C]. *Proceedings of the IEEE*, 2011, 99(4): 540-555.
- [9] Jiang Q, Shao F, Jiang G, *et al.* A depth perception and visual comfort guided computational model for stereoscopic 3D visual saliency[J]. *Signal Processing: Image Communication*, 2015, 38: 57-69.
- [10] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254-1259.
- [11] Hou X, Zhang L. Saliency detection: A spectral residual approach [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [12] Bruce N D, Tsotsos J K. Saliency, attention, and visual search: An information theoretic approach[J]. *Journal of Vision*, 2009, 9(3): 5.1-24.
- [13] Cheng M, Zhang G, Mitra N J, *et al.* Global contrast based salient region detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2011, 37(3): 409-416.
- [14] Yarbus A L. *Eye movements and vision*[M]. New York: Plenum Press, 1967.
- [15] Wang Shiyi, Wang Shuang, Zhang Libao. A road extraction algorithm with saliency analysis in high-resolution remote sensing images[J]. *Acta Optica Sinica*, 2015, 35(s2): s210001.
王士一, 王 双, 张立保. 一种结合显著性分析的高分辨率遥感图像道路提取算法[J]. 光学学报, 2015, 35(s2): s210001.
- [16] Mo Shaowen, Deng Xinpu, Wang Shuai, *et al.* Moving object detection algorithm based on improved visual background extractor[J]. *Acta Optica Sinica*, 2016, 36(6): 0615001.
莫邵文, 邓新蒲, 王 帅, 等. 基于改进视觉背景提取的运动目标检测算法[J]. 光学学报, 2016, 36(6): 0615001.
- [17] Wang J, Da Silva M P, Le Callet P, *et al.* Computational model of stereoscopic 3D visual saliency [J]. *IEEE Transactions on Image Processing*, 2013, 22(6): 2151-2165.
- [18] Klaus A, Sormann M, Karner K. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure[C]. *18th International Conference on Pattern Recognition*, 2006, 3: 15-18.
- [19] BT Series. Subjective methods for the assessment of stereoscopic 3D TV systems: ITU-R BT.2021-1 [S]. Geneva: International Telecommunication Union, 2015.
- [20] BT Series. Methodology for the subjective assessment of the quality of television picture: ITU-R BT.500 [S]. Geneva: International Telecommunication Union, 2012.
- [21] Kishi S, Yamazoe T, Shibata T, *et al.* An ergonomic evaluation system for stereoscopic 3D images[J]. *Journal of the Institute of Image Information and Television Engineers*, 2006, 60(6): 934-942.
- [22] Atallah P, Pelah A, Wilkins A J. Visual stress symptoms from stereoscopic television[C]. *International Conference on 3D Imaging*, 2012: 1-7.
- [23] Tan Zefu, Ding Yanzhi, Lei Guoping, *et al.* Attention model-based visual discomfort evaluation model of stereoscopic display[J]. *Video Engineering*, 2015, 39(7): 108-110.
谭泽富, 丁妍芝, 雷国平, 等. 基于视觉显著性图的立体图像视疲劳评价[J]. 电视技术, 2015, 39(7): 108-110.