

立体视觉中心凹 JND 模型及其图像压缩硬件实现

季渊^{1,2}, 郑志杰^{1*}, 吴浩¹, 张引¹, 陈文栋², 穆廷洲²¹上海大学微电子研究与开发中心, 上海 200072;²上海大学机电工程与自动化学院, 上海 200072

摘要 传统的二维 JND(Just Noticeable Difference)模型只能估计平面图像的最小可觉差,并不完全适用于虚拟立体视觉下的大视场角图像。首先根据人眼视觉特性,对亮度、对比度、中心凹和立体深度 4 种掩蔽特性设计相应的双目观测实验,通过实验数据建立 JND 数学模型并将其与当前其他 JND 模型进行对比,结果表明该模型在同等感知质量下可以去除更多的视觉冗余。然后将该视觉感知冗余模型应用于图像压缩,为此提出一种多重色阶压缩算法,该算法根据人眼色差阈值对图像的不同区域进行不同等级的色阶量化,量化过程结合 Floyd-Steinberg 误差分散算法可以去除视觉冗余数据。最后在 FPGA(Field-Programmable Gate Array)硬件平台上完成算法的验证,结果表明该算法的平均比特压缩率可以达到 61.65%,能够有效降低 VR(Virtual Reality)图像所需的传输数据量。

关键词 图像处理; 视觉光学; 中心凹 JND; 人眼视觉特性; 数据传输; 微显示器

中图分类号 TN27

文献标志码 A

doi: 10.3788/AOS202141.1210001

Foveated JND Model Based on Stereo Vision and Its Application in Image Compression with Hardware Implementation

Ji Yuan^{1,2}, Zheng Zhijie^{1*}, Wu hao¹, Zhang Yin¹, Chen Wendong², Mu Tingzhou²¹Microelectronics Research and Development Center, Shanghai University, Shanghai 200072, China;²School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China

Abstract The traditional two-dimensional JND (Just Noticeable Difference) model can only estimate the minimum noticeable difference of the planar image, and is not completely suitable for the large field of view image under virtual stereo vision. First, according to the visual characteristics of the human eye, corresponding binocular observation experiments are designed for the four masking characteristics of brightness, contrast, fovea and stereo depth, and the mathematical model of JND is established through experimental data, and compared with other current JND models. The results show that this model can remove more visual redundancy under the same perceptual quality. Then the visual perception redundancy model is applied to image compression. For this purpose, a multi-level compression algorithm is proposed. The algorithm performs different levels of color level quantization on different regions of the image according to the human eye color difference threshold. The quantization process which combines the Floyd-Steinberg error dithering algorithm can remove visual redundant data. Finally, the algorithm verification is completed on the FPGA (Field-Programmable Gate Array) hardware platform. The results show that the average bit compression rate of the algorithm can reach 61.65%, which can effectively reduce the amount of transmission data required for VR (Virtual Reality) images.

Key words image processing; vision optics; foveated JND; human visual characteristic; data transmission; microdisplay

OCIS codes 100.2000; 110.3000; 200.3050; 330.4595

收稿日期: 2020-12-09; 修回日期: 2021-01-04; 录用日期: 2021-01-22

基金项目: 国家自然科学基金(61774101)

* E-mail: 18800203705@163.com

1 引言

心理物理学中,最小可觉差(JND)可以表示人眼所能感受到的刺激的最小变化量,同时也能够体现人眼对图像失真的最大容忍度。目前,已有多个JND模型被提出。Chou等^[1]使用JND模型对亮度掩蔽特性和对比掩蔽特性的影响进行了研究,该模型是像素域2D-JND模型之一。Chen等^[2]提出了加入中心凹掩蔽特性的JND模型,结果表明JND值与偏心角有关。Wang等^[3]在Chen等^[2]的基础上加入了视觉注意特性。2D-JND模型对于平面图像是适用的,然而立体图像具有深度信息,所以人眼对其主观感知有所不同。Liu等^[4]提出了一种3D-JND模型,但该模型只研究了人眼舒适区的小范围深度变化,而且对比掩蔽特性中未考虑背景亮度的影响。Fan等^[5]提出了一种显著性加权立体JND模型,但其并未涉及中心凹掩蔽特性。

综上所述,提高JND模型与人眼特性的匹配程度十分重要,为此本文根据实验建立一种FD-JND(Foveation-Depth JND)模型。在相同的客观质量评价下,该模型的主观质量评价结果均高于其他JND模型,表明其更符合人眼视觉。JND体现了人眼对图像灰度变化的容忍度,其可以反映图像中的视觉冗余,因此应用于图像压缩。目前,虚拟现实(Virtual Reality, VR)设备普遍朝着高分辨率和广视角的方向发展^[6],然而当前显示技术的传输带宽无法承载过多的数据量^[7],为此季渊等^[8-9]根据人眼特性提出了一种基于多尺度金字塔的压缩方法,在分辨率上对VR图像进行了数据压缩。本文根据JND模型提出一种像素域的色阶压缩算法,使用三种单色的JND阈值加权得到色差阈值后对图像进行色阶压缩,采用误差分散算法减少压缩过程中的失真图像。针对该算法设计相应的微显示控制器,在现场可编程门阵列(Field-Programmable Gate Array, FPGA)平台上成功驱动硅基有机发光二极管(Organic Light-Emitting Diode, OLED)显示器,实现VR图像的实时压缩显示,有效减少传输的数据量,验证该算法在硬件方面实现的可行性。

2 双目观测实验

2.1 视觉掩蔽特性

JND会受到多种人眼视觉特性的影响,包括亮度掩蔽、对比掩蔽、中心凹掩蔽和深度掩蔽等特性。中心凹掩蔽特性描述人眼对中心凹区域的分辨力强

和外围区域弱的情况,深度掩蔽特性描述不同深度下人眼的感知情况。在立体视觉中,人眼有舒适区和非舒适区^[10],在舒适区范围内人眼视敏度随着视距的增加而减少^[11]。

对于立体图像中的某一像素点 $P(i, j)$,深度值的计算公式为

$$d(i, j) = \frac{b \times f}{D(i, j)}, \quad (1)$$

式中: b 为相邻两台相机的基线距离; f 为相机焦距; D 为像素点 P 对应左右图像之间的视差值。

2.2 实验建立

为了研究立体视觉场景下4种掩蔽特性对JND的影响,对亮度和对比度进行两种类型的心理学实验,每种实验又对偏心角和深度进行改变。实验采用的设备为Oculus Rift CV1,该VR设备的单眼屏幕分辨率为1080 pixel×1200 pixel,水平视角为110°,角分辨率(PPD)约为10。实验测试人员共14名,其中男性10人,女性4人,年龄均在21~25岁之间。通过弗莱堡视力测试,所有人员的视力表现均正常。为了避免心理影响,测试人员仅知晓基本的信息。

2.3 中心凹-深度-亮度掩蔽特性实验

实验中,采用左右图片的格式在VR中进行显示,左右图的分辨率均为1080 pixel×1200 pixel,测试的左眼图片如图1所示,右眼图类似,图片的中心点用于帮助测试者进行凝视。图片的背景亮度为 l ,用于测取JND的噪点亮度 $l+k$,两者的灰度亮度值均在0~255之间,其中 k 为噪点幅度。本次主观实验仅对深度掩蔽特性进行探究,因此注入的测试噪声是相同的,即左右图片中的噪点幅度相等。噪点区域为10 pixel×10 pixel大小的方块,其分布于偏心角为 e 的圆形边界,对应的视差值为

$$D = D_L - D_R + 2a, a = 60, \quad (2)$$

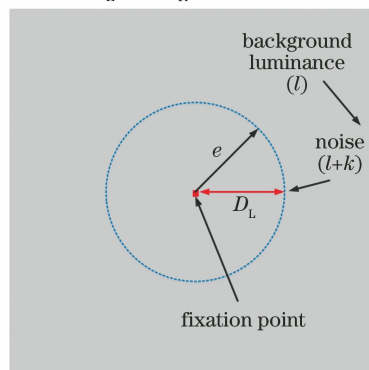


图1 亮度掩蔽特性实验的测试图像

Fig. 1 Test image of luminance masking characteristic experiment

式中: D_L 为左图中噪点到凝视点的像素差值; D_R 为右图中噪点到凝视点的像素差值; a 为凝视点到人眼中心的像素差值。令凝视点所对应的成像处的深度值为 1.0, 则 $D = 120$, 此时噪点相对凝视点的偏心角按照两者的像素距离与 PPD 之比进行计算, PPD 值设为 10。

图 2 为 3D 成像原理。A 点和 B 点均为屏幕上的噪点, A 点在深度为 1.0 和偏心角为 e 处的成像点为 A' , B 点在深度为 d 和偏心角为 e 处的成像点为 B' , A 点在深度为 d_0 和偏心角为 0 处的成像点为 E。设 A 点和 B 点分别到凝视点的像素差值为 D_{L0} 和 D_{L1} , 则由(1)式和(2)式可得

$$d_0 = \frac{120}{2D_{L0} + 120} \quad (3)$$

根据三角形相似原理可以得到相同偏心角下不同深度的像素差值, 表达式为

$$D_{L1} = (60 + D_{L0})/d - 60. \quad (4)$$

实验中设置的亮度组为 11 组, 偏心角组为 13

表 1 亮度掩蔽特性和对比掩蔽特性实验的具体数值

Table 1 Specific values of luminance masking characteristic and contrast masking characteristic experiments

Attribute	Luminance masking	Contrast masking
Background luminance	0, 10, 20, 60, 100, 140, 180, 200, 220, 230, 240	100, 140, 180, 220
Contrast luminance	0	20, 40, 60, 80
Eccentricity / ($^\circ$)	1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50	1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50
Depth	0.2, 0.4, 0.6, 0.8, 1.0	0.2, 0.4, 0.6, 0.8, 1.0
Disparity / pixel	600, 300, 200, 150, 120	600, 300, 200, 150, 120

2.4 中心凹-深度-对比掩蔽特性实验

对比掩蔽特性实验相比于亮度特性实验, 实验中进一步增加了对比区域, 测试过程中对比区域为方型, 与凝视点同样处于图像的中心位置, 如图 3 所示。对比区域的亮度为 $l-c$, 其中 c 为亮度对比值, 即与背景亮度的差值。当 c 为 0 时, 则对应亮度

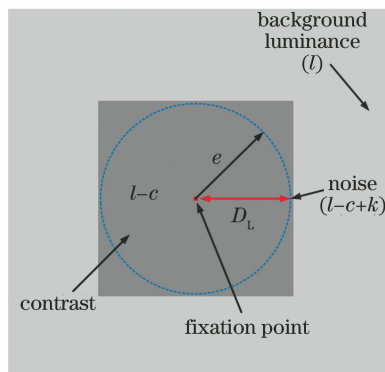


图 3 对比掩蔽特性实验的测试图像

Fig. 3 Test image of contrast masking characteristic experiment

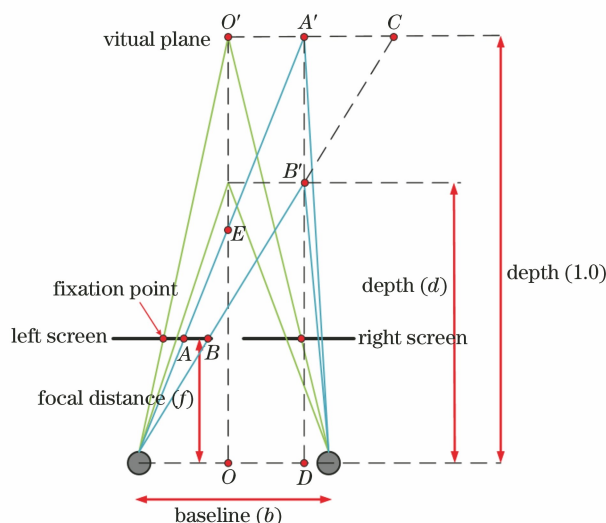


图 2 3D 成像原理

Fig. 2 Principles of 3D imaging

组, 深度组为 5 组, 具体数值如表 1 所示。测定过程中采用的方法为心理学测定阈值的阶梯法, 即多轮增减噪点灰度法, 最后取均值得到 JND 阈值。

掩蔽特性实验, 噪点位于对比区域与背景区域的边界处, 测试噪点幅度基于对比区域的亮度值。实验中设置的亮度组为 4 组, 对比组为 4 组, 偏心角组为 13 组, 深度组为 5 组, 具体数值如表 1 所示。

3 FD-JND 模型的建立及分析

3.1 亮度掩蔽特性模型

实验结果如图 4 所示。当深度为 1.0 和偏心角为 20° 时, JND 与背景亮度之间的关系曲线如图 4(a) 所示。从图 4(a) 可以看到, 在背景亮度的低区域和高区域处 JND 值较大, 中间区域较小, 这与 2D 图像的观测结果^[1] 基本一致。当深度为 1.0 和背景亮度为 140 时, JND 与偏心角之间的关系曲线如图 4(b) 所示。从图 4(b) 可以看到, JND 在偏角为 5° 以内基本保持不变, 之后开始增加, 在 $15^\circ \sim 25^\circ$ 之间变化平缓, 40° 之后开始急剧增大。当背景亮度为 140 和偏心角为 20° 时, JND 与深度之间的关系曲线如图 4(c) 所示。从图 4(c) 可以看到, JND 在深

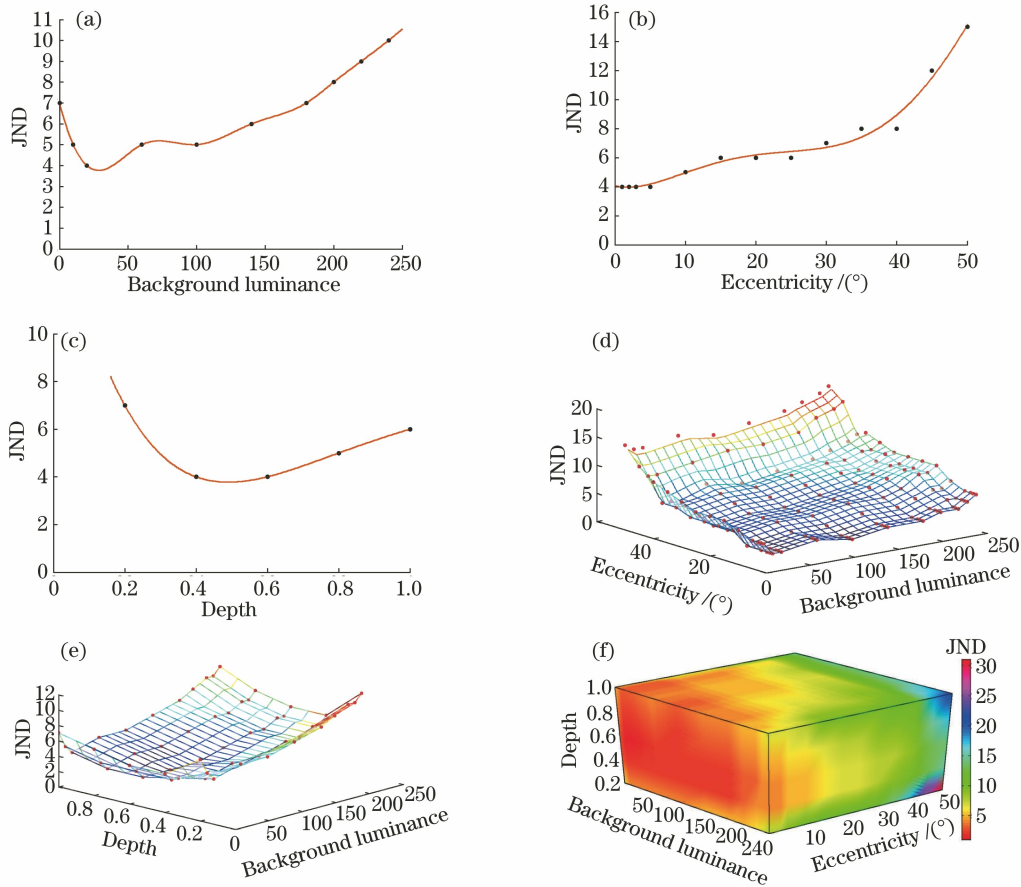


图 4 亮度掩蔽特性的实验结果。(a) JND 与亮度的关系;(b) JND 与偏心角的关系;(c) JND 与深度的关系;(d) 在 $d=1.0$ 的情况下,JND 与亮度和偏心角的关系;(e) 在 $e=20$ 的情况下,JND 与亮度和深度的关系;(f) 亮度掩蔽特性实验的 4D 图像

Fig. 4 Experimental results of luminance masking characteristic experiment. (a) Relationship between JND and luminance; (b) relationship between JND and eccentricity angle; (c) relationship between JND and depth; (d) relationship among JND, luminance, and eccentricity angle at $d=1.0$; (e) relationship among JND, luminance, and depth at $e=20$; (f) 4D image of luminance masking characteristic experiment

度为 0.5~1.0 之间逐渐增加,这与视敏度随视距的增大而减小的情况相符^[11];在小于 0.5 以内,深度值越小,JND 反而越大,这是由于成像过近,人眼开始进入非舒适区^[10],所以视敏度出现减小。图 4(d)和图 4(e)为 JND 与不同参数的变化关系,图 4(f)为整个亮度掩蔽特性的 4D 图像,其加入了三项掩蔽特性的影响参数。

根据实验数据建立相应的亮度掩蔽特性模型,模型的方均根误差(Root Mean Squared Error, RMSE)值为 1.0171,JND 与背景亮度 l 、偏心角 e 和深度 d 的关系分别为

$$l_1 = a_1 \times l^4 + a_2 \times l^3 + a_3 \times l^2 + a_4 \times l + a_5, \quad (5)$$

$$e_1 = b_1 \times e^4 + b_2 \times e^3 + b_3 \times e^2 + b_4 \times e + b_5, \quad (6)$$

$$d_1 = c_1 \times d^4 + c_2 \times d^3 + c_3 \times d^2 + c_4 \times d + c_5, \quad (7)$$

$$J_{FD-JND,l}(l, e, d) = l_1 \times \left(\frac{e_1}{e_1(0)} + \mu_1 \right) \times \frac{d_1}{d_1(1.0)} + \mu_2, \quad (8)$$

式中: $a_1 = 2.046 \times 10^{-8}$; $a_2 = -9.44 \times 10^{-6}$; $a_3 = 0.001437$; $a_4 = -0.07248$; $a_5 = 5.273$; $b_1 = 3.099 \times 10^{-6}$; $b_2 = -0.0001298$; $b_3 = 8.477 \times 10^{-6}$; $b_4 = 0.1435$; $b_5 = 4.663$; $c_1 = 26.04$; $c_2 = -93.75$; $c_3 = 124$; $c_4 = -66.25$; $c_5 = 15$; μ_1 和 μ_2 均为校正参数,值分别为 -1.637 和 0.224 ; $J_{FD-JND,l}$ 为中心凹深度(Foveated Depth,FD)影响下由 l 得到的 JND 值。由于亮度与灰度值呈正相关,因此 l 可以使用当前邻域内像素灰度的加权均值来代替,通常取 $5 \text{ pixel} \times 5 \text{ pixel}$ 大小的加权模板^[1]。

3.2 对比掩蔽特性模型

对比掩蔽特性的实验结果如图 5 所示。当深度为 1.0、偏心角为 20° 和背景亮度为 140 时, JND 与对比度之间的关系曲线如图 5(a) 所示。从图 5(a) 可以看到, 对比掩蔽特性可以使 JND 值增大, 且对比度越大 JND 值越大。对比掩蔽特性实

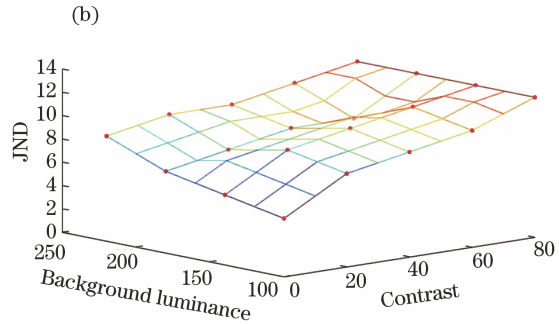
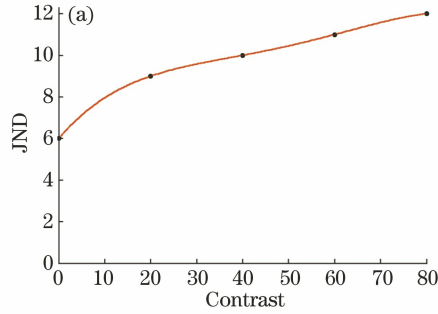


图 5 对比掩蔽特性的实验结果。(a) JND 与对比度的关系; (b) JND 与亮度和对比度的关系

Fig. 5 Experimental results of contrast masking characteristic experiment. (a) Relationship between JND and contrast; (b) relationship among JND, luminance, and contrast

根据测得的实验数据来建立相应的对比掩蔽特性模型, JND 与背景亮度 l 、对比度 c 、偏心角 e 和深度 d 的关系分别为

$$M = \begin{cases} (u_1 \times l^2 + u_2 \times l + u_3) \times \sqrt{c} + u_4, & (0 \leq e < 10) \\ (v_1 \times l^2 + v_2 \times l + v_3) \times \sqrt{c} + v_4, & (10 \leq e < 30) \\ (w_1 \times l^2 + w_2 \times l + w_3) \times c + w_4, & (30 \leq e) \end{cases}, \quad (9)$$

$$J_{\text{FD-JND},c}(l, c, e, d) = J_{\text{FD-JND},l} + M, \quad (10)$$

式中: M 为相较于亮度掩蔽, 由对比掩蔽所引起的 JND 变化值; $J_{\text{FD-JND},c}$ 为中心凹深度影响下由 c 得到的 JND 值; $u_1 = -2.206 \times 10^{-5}$; $u_2 = 0.006143$; $u_3 = -0.07232$; $u_4 = -1.06$; $v_1 = -1.731 \times 10^{-5}$; $v_2 = 0.004102$; $v_3 = 0.3572$; $v_4 = -0.65$; $w_1 = 3.983 \times 10^{-6}$; $w_2 = -0.001496$; $w_3 = 0.2457$; $w_4 = -0.46$ 。对比度 c 可以使用边沿梯度来表示, 本文使用 $5 \text{ pixel} \times 5 \text{ pixel}$ 区域内的 Sobel 算子进行计算。

最终的 FD-JND 模型可表示为

$$J_{\text{FD-JND}}(l, c, e, d) = \begin{cases} J_{\text{FD-JND},l}, & \text{if } c = 0 \\ J_{\text{FD-JND},c}, & \text{if } c > 0 \end{cases} \quad (11)$$

由(11)式可知, 模型的 RMSE 值为 1.4042, 表明模型的拟合程度较好。

验中, 当深度为 1.0 和偏心角为 20° 时, JND 与背景亮度和对比度之间的关系曲线如图 5(b) 所示。从图 5(b) 可以看到, 当对比度大于 0 和背景亮度大于 100 时, JND 随着背景亮度呈现先增加后减小再增加的变化趋势, 当对比度过大时 JND 变化不明显。

4 JND 模型的比较

为了验证所建立的 FD-JND 模型的准确性, 将其与 Liu^[4] 和 Wang^[3] 提出的模型进行对比。测试图片均来自 Waterloo IVC 3D 数据集^[12], 图片的分辨率为 $1920 \text{ pixel} \times 1080 \text{ pixel}$ 。与文献[13]类似, 本文对原始图片注入相同的噪声量进行污染, 用来比较不同 JND 模型下的主观视觉质量。为了便于测试和比较, 不同 JND 模型的深度值均设为 1.0, 凝视点固定在图像的中心, 注入过程采用大小为 $5 \text{ pixel} \times 5 \text{ pixel}$ 的滑块。污染后的图片在像素 (i, j) 处的灰度值为

$$I^*(o, i, j) = I(o, i, j) + N_{\text{rand}} \times J_{\text{JND}}(o, i, j), \quad (12)$$

$$J_{\text{JND}}(k) =$$

$$\begin{bmatrix} J_{\text{JND},l} & J_{\text{JND},l} & J_{\text{JND},l} & J_{\text{JND},c} & J_{\text{JND},c} \\ J_{\text{JND},l} & J_{\text{JND},l} & J_{\text{JND},l} & J_{\text{JND},c} & J_{\text{JND},c} \\ J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},l} & J_{\text{JND},l} \\ J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},l} & J_{\text{JND},l} \\ J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},c} & J_{\text{JND},l} & J_{\text{JND},l} \end{bmatrix}, \quad (13)$$

式中: I 为原始图片; o 为第 o 个滑块; N_{rand} 为随机数, 其从 0、0.382、0.574、0.700 和 1.000 中随机选取。测试图片以 110° 的视角角在 VR 设备中进行全屏显示, 测试方法选用主观测试建议书 ITU-R BT.500 中所建议的双刺激连续质量尺度法, 评分分为 5 个等级, 如 5(两张图的差异完全不可见), 4(差异可

见但噪声不明显),3(可察觉少量噪声),2(可以察觉的较多噪声),1(可察觉的噪声很多)。注入的噪声量以峰值信噪比(PSNR)作为衡量标准,不同 JND 模型下的平均主观得分(MOS)如表 2 所示。

从表 2 可以看到,建立的模型与 Liu^[4] 和 Wang^[3] 提出的模型相比,在注入相同的噪声量下,每幅图像的平均主观得分更高,同时也反映 FD-JND 模型在主观得分较高的情况下,能够注入的噪声量较多,PSNR 的平均值可以低至 31.41 dB。

不同 JND 模型污染后的噪声分布如图 6 所示,三个模型基于图 6(a)注入噪声。从图 6 可以看到,建立的 JND 模型在中心区域的噪声注入较少,在远离中心区域的噪声注入较多。图 7 为不同 JND 模型污染后图片中心区域的放大图。从图 7 可以看到,建立的模型在中心区域的图像清晰度高于其他两个 JND 模型,与原图差别很小,而人眼观察区域主要集中在视敏度高的中心区域,因此视觉感知质

表 2 不同模型的平均主观得分对比

Table 2 Comparison of average subjective scores of different models

Image	PSNR /dB	MOS		
		Ref. [4]	Ref. [3]	Proposed model
Gate	32.01	4.20	3.50	4.80
Woman	32.33	4.30	3.50	4.90
Column	30.89	4.20	3.40	4.70
Hall	30.72	4.30	3.50	4.90
Lab	30.71	4.00	3.30	4.60
Repair	32.48	4.30	3.40	4.80
Report	30.24	4.20	3.30	4.80
Soccer	31.22	4.40	3.60	5.00
Branch	31.42	4.30	3.40	4.90
Umbrella	32.03	4.20	3.40	4.80
Average	31.41	4.24	3.43	4.82

量更高,这与主观评测结果相符。以上结果表明,相比于其他两个 JND 模型,FD-JND 模型更符合人眼视觉。

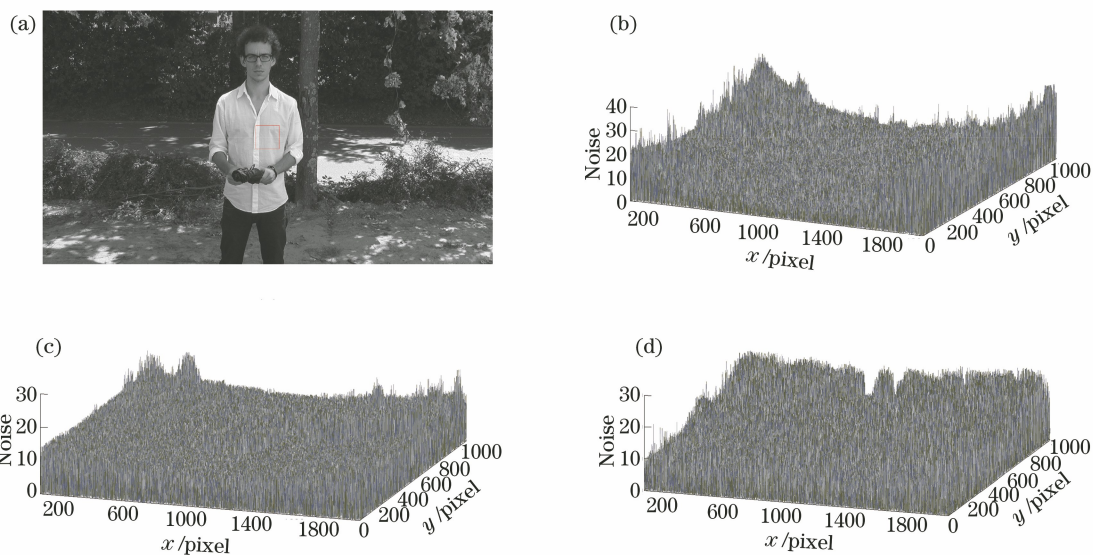


图 6 不同 JND 模型的噪声注入分布。(a)原始图像;(b)所提模型;(c)文献[4];(d)文献[3]

Fig. 6 Noise injection distributions of different JND models. (a) Original image; (b) proposed model; (c) Ref. [4]; (d) Ref. [3]

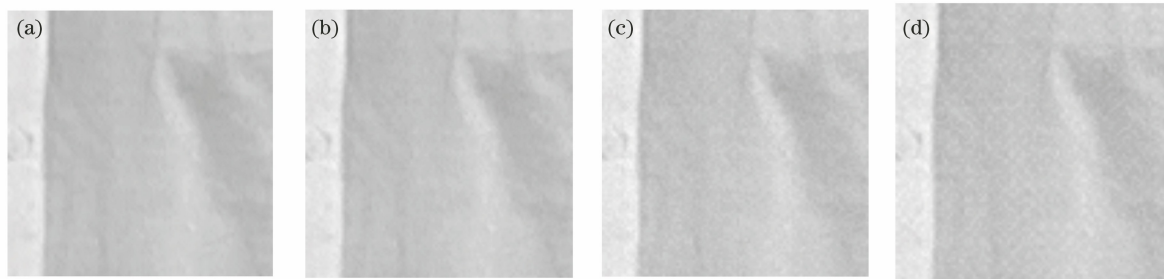


图 7 不同 JND 模型污染后图片的中心区域放大图。(a)原始图像;(b)所提模型;(c)文献[4];(d)文献[3]

Fig. 7 Enlarged view of central area of picture after pollution by different JND models. (a) Original image; (b) proposed model; (c) Ref. [4]; (d) Ref. [3]

5 基于 FD-JND 模型的多重色阶压缩算法

5.1 图像色觉冗余

JND 是人眼视觉特性形成掩蔽效应的临界值,其反映了人眼对图像不同区域中信号失真的容忍程度,利用该人眼生理特性对图像所包含的信息进行压缩,可以大幅度优化图像信息的表现方式,达到去除图像中的视觉信息冗余而且人眼无法察觉的目的。由于 JND 反映的是灰度图像的视觉冗余,因此不能直接应用于彩色图像,需要对其进行一定的变换处理。对于彩色图像,人眼对于红绿蓝(RGB)三色的敏感程度不一,若使用 JNCD(Just Noticeable Color Differences)来表示人眼的色差阈值,当察觉到一个单位的色差时,则对应三个颜色分量的刺激增量不同,为此 JNCD 在 RGB 空间中会呈现椭球形状分布。对于彩色图像中总的颜色亮度,通常按照 $Y=0.30R+0.59G+0.11B$, Y 体现了 RGB 三基色的综合能量强度,三个比例系数(R 、 G 和 B)对应人眼对红绿蓝三色的不同敏感程度。实验结果表明,当 $R:G:B=2:1:4$ 时,模型能够较好地保证图像质量^[14],因此可以定义 JNCD 为三色分量中每种颜色下的加权 JND 值,计算公式为

$$J_{\text{JNCD}} = 2J_{\text{JND,R}}/7 + J_{\text{JND,G}}/7 + 4J_{\text{JND,B}}/7。 \quad (14)$$

5.2 图像低色阶化

若将 8 bit 的灰度图像降低灰阶,最简单的方法是直接舍弃像素的低数据位,灰度图像的低灰阶量化公式为

$$g(i, j) = \sum_{i=1}^M \sum_{j=1}^N \text{round}[f(i, j)/2^n] \times 2^n, \quad (15)$$

式中: $g(i, j)$ 为低灰阶量化后的图像; $f(i, j)$ 为原始 8 bit 的图像; $\text{round}(\cdot)$ 为四舍五入的取整函数; M 和 N 分别为图像的长和宽; n 为低灰阶量化等级, $n=0, 1, \dots, 7$ 。

对于彩色图像,可以在三个通道中同时进行量化以降低色阶,舍弃低位的数据越多,则图像失真的情况越多,因此可以采用误差分散算法,即 Floyd-Steinberg 抖动算法,其基本思想是将像素点的量化误差分散到邻近像素点,并对舍弃的数据进行补偿。 2×2 大小的误差传递模板可表示为

$$\mathbf{T} = \begin{bmatrix} 0 & 3/8 \\ 3/8 & 1/4 \end{bmatrix}, \quad (16)$$

式中: $3/8$ 、 $3/8$ 和 $1/4$ 为权重系数。

算法的处理过程:首先将当前像素的量化误差分别乘以 \mathbf{T} 中的三个权重系数,然后加到右、下和右下的像素中。根据实际实验结果发现,对于彩色图像,Floyd-Steinberg 抖动算法可以将色阶量化至 16,此时人眼能够观察到较小的颗粒噪点,当色阶量化至 8 时,噪点颗粒较明显。

5.3 多重色阶压缩算法

彩色图像的多重色阶压缩过程:采用 $5 \text{ pixel} \times 5 \text{ pixel}$ 大小的滑块从图像的左上角开始进行移动,分别计算出 RGB 三个通道下滑块中心所对应的背景亮度 L_R 、 L_G 和 L_B 以及偏心角 e 。由于低色阶量化是对整个滑块进行处理,为了加快运算速度,所有滑块均使用固定深度的亮度掩蔽 JND 函数。由 l 和 e 可以得到当前滑块中心在三个通道下的 JND 值,加权后得到最终的 JNCD 值。根据 JNCD 值使用 2×2 模板的 Floyd-Steinberg 抖动算法进行不同程度的色阶量化,在量化过程中三个通道的色阶均保持一致。根据 JND 值的分布以及抖动算法在低色阶量化过程中的失真情况,多次实验后按照表 3 进行压缩,图像质量和压缩率均能够较好平衡。

表 3 不同 JNCD 取值的色阶

JNCD value	Color scale	Bit width
(0,4]	256	8
(4,6]	128	7
(6,8]	64	6
(8,10]	32	5
(10,11]	16	4
(11,∞]	8	3

6 FPGA 系统验证

6.1 硬件整体框架

为了验证算法在硬件方面实现的可行性,采用以 FPGA 为核心的硅基 OLED 微显示器测试平台,结构框架如图 8 所示。硬件系统主要包括三个部分,分别为视频输入源、微显示控制器以及 OLED 微显示器。视频的输入源主要由 PC(Personal Computer)产生,其将视频流通过高清多媒体接口(High Definition Multimedia Interface, HDMI)传输至视频解码电路中,将其转换为标准的 VSEA(Video Electronics Standards Association)视频信号后送入 FPGA 内部的微显示控制器中。微显示控制器包括图像压缩模块和扫描控制模块。当传入

每帧图像时,将图像数据保存在两个缓存块中。第一个数据缓存块采用 4 个 FIFO(First Input First Output Storage)和 25 个移位寄存器,用来获取 $5 \text{ pixel} \times 5 \text{ pixel}$ 的像素值,然后通过加权模板计算出背景亮度和偏心角,由此得到当前图像块的 JNCD 阈值。第二个数据缓存块采用 SDRAM(Synchronous Dynamic Random Access Memory)存储器,传入 JNCD 值后控制存储管理器以读取 SDRAM 的每个像素值,接着根据查找表(Look Up Table, LUT)进行色阶量化,量化完成后再进行误差分散以更新 SDRAM 中的相邻像素值,同时将量化后的数据保存在输出缓存 SDRAM 中。由于使

用两个数据缓存块,因此 JNCD 阈值产生和色阶量化可以并行执行,同时每帧图像也可以进行流水线处理,这减少了输出延时。整帧图像完成量化压缩后,输出使能信号至扫描控制器中。扫描控制模块采用脉冲宽度调制(Pulse Width Modulation, PWM)数字驱动 OLED 微显示器^[15],收到使能信号后对输出的缓存数据进行接收和预处理,之后扫描 LUT 查找表对 SDRAM 进行寻址,最终完成图像数据的读取,最后以低压差分信号(Low Voltage Differential Signaling, LVDS)的方式将同步信号和数据信号传输到硅基 OLED 微显示器的接收端中。

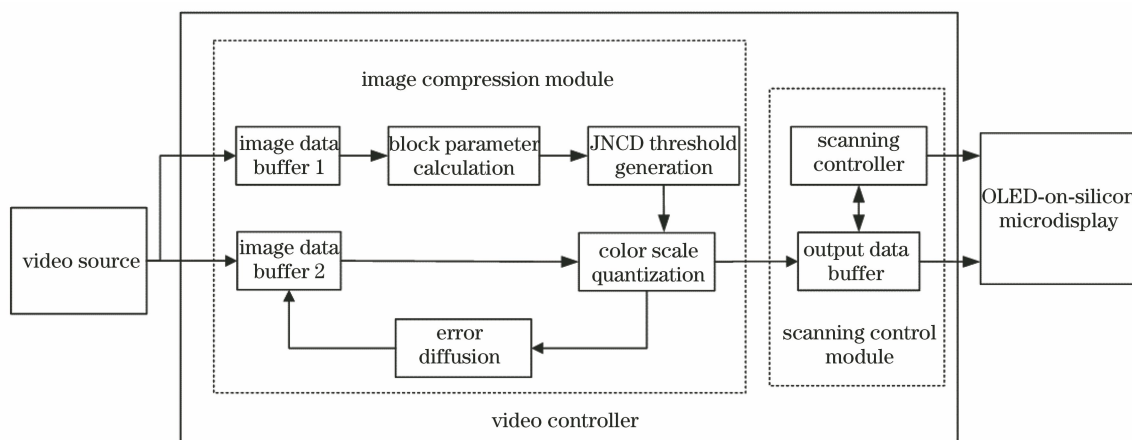


图 8 基于 FPGA 的硬件系统结构框架

Fig. 8 Hardware system architecture framework based on FPGA

6.2 压缩结果与分析

压缩测试所用的 9 张测试图片均来自 Middlebury Stereo 2005 数据集^[12],图片的分辨率约为 $695 \text{ pixel} \times 555 \text{ pixel}$,当深度值为 1 时,按照 $w/110$ (w 为图片宽度)的角分辨率进行数据压缩,图 9 为在 FPGA 平台上显示图像压缩后的效果,验证了所提算法在硬件方面实现的可行性。

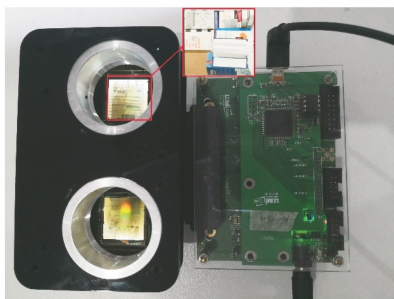


图 9 硬件测试平台

Fig. 9 Hardware test platform

为了对图像质量进行分析,采用 PSNR、中心凹的峰值信噪比(FPSNR)^[16]和 SSIM(Structural

Similarity)三个客观质量评价参数进行评价。由于图像是在大视场角的 VR 中进行显示,因此对不同偏心角的图像计算 PSNR。FPSNR 将 PSNR 与视敏度函数 f 结合,用来评估图像的整体质量。SSIM 对两幅图的亮度、对比度和结构进行相似性评估,值越接近 1 表示相似度越高。

图 10 为算法的处理结果。从图 10 可以看到,相比于原图,算法压缩后的图片在人眼视敏度高的中心区域可以较好地保留清晰度,在边缘区域降低清晰度。表 4 为图像的各项评价指标。从表 4 可以看到在偏心角为 30° 以内,处理后的彩色图像的 PSNR 基本在 50 dB 以上,能够保证近周边视觉^[17]良好的主观感受,FPSNR 和 SSIM 的平均值分别为 40.65 dB 和 0.87,表明压缩后的 VR 图像整体质量较好。

在压缩率方面,以处理后图像的总比特与原始图像的总比特之比作为压缩指标,彩色图像的平均比特压缩率可以达到 61.65%,较大程度地减少了 VR 图像的数据量。

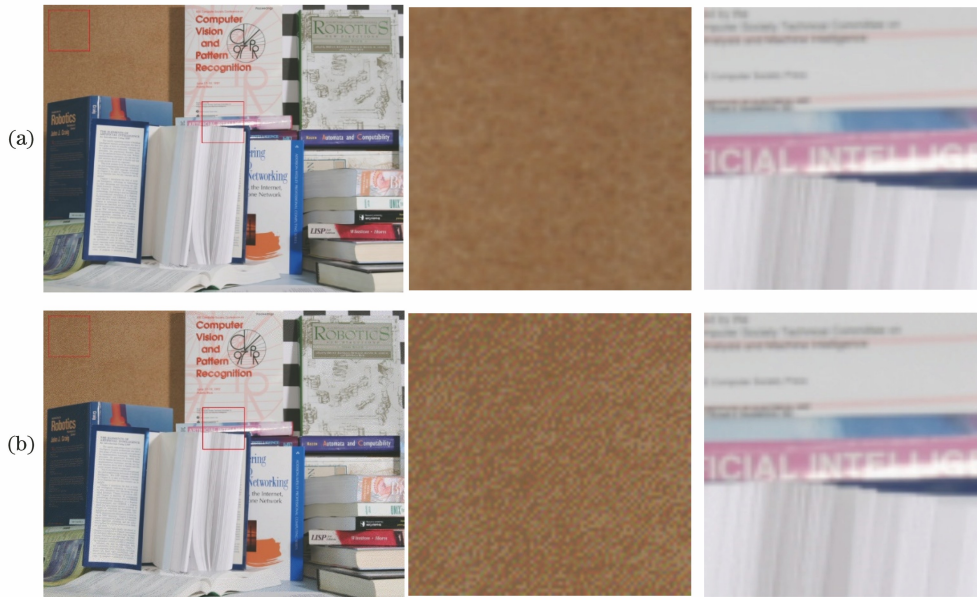


图 10 图像压缩效果的对比及其局部放大图。(a)原始图像;(b)压缩后的图像

Fig. 10 Comparison of image compression effect and local enlarged view. (a) Original images; (b) compressed images

表 4 彩色图像压缩后的各项评价指标和比特率

Table 4 Evaluation index and bit rate of color image compression

Image	PSNR /dB					FPSNR / dB	SSIM	Bit rate /%
	e is $0^{\circ}-5^{\circ}$	e is $5^{\circ}-10^{\circ}$	e is $10^{\circ}-15^{\circ}$	e is $15^{\circ}-30^{\circ}$	$e > 30^{\circ}$			
Art	70.5936	58.9329	57.1360	52.1277	33.2020	40.4823	0.8675	61.59
Book	55.2470	51.6498	52.2067	49.4993	32.8974	39.8006	0.8686	57.15
Computer	65.8224	54.3767	53.7429	52.6272	33.1890	40.4265	0.8783	60.84
Doll	66.6142	54.6717	55.0177	51.9896	33.3387	40.6038	0.8799	61.53
Drumstick	62.9970	57.3913	56.8413	53.1031	33.7317	41.0152	0.8980	62.02
Dwarves	77.8636	62.7846	62.1044	53.2000	33.5468	40.9747	0.8669	63.92
Laundry	65.5591	53.5276	54.2809	52.7823	33.7454	41.0754	0.8861	62.26
Moebius	73.0924	66.3706	60.1905	53.2447	33.3951	40.6560	0.8637	61.78
Reindeer	64.3097	56.8130	58.1696	52.8886	33.3697	40.8316	0.8606	63.74
Average	66.8999	57.3909	56.6322	52.3847	33.3795	40.6518	0.8744	61.65

7 结 论

根据人眼视觉特性,对立体图像下的 JND 进行研究并设计 4 种掩蔽特性实验,用来分析不同掩蔽特性对 JND 的影响,建立 FD-JND 模型并与另外两种中心凹 JND 模型进行对比,结果表明该模型具有更好的准确性。将感知冗余模型应用于图像压缩,用来去除图像中的冗余数据,为此提出一种 FD-JND 结合误差分散的多重色阶压缩算法,并在 FPGA 平台上完成该算法的设计与验证。实验结果表明,在保证图像质量较高的同时,该算法能够较大程度地降低 VR 图像的数据量,为超高清近眼显示设备面临的传输数据量过大问题提供一种良好的解决方案。

参 考 文 献

- [1] Chou C H, Li Y C. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, 5(6): 467-476.
- [2] Chen Z Z, Guillemot C. Perceptually-friendly H.264/AVC video coding based on foveated just-noticeable-distortion model [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(6): 806-819.
- [3] Wang H K, Yu L, Wang S W, et al. A novel foveated-JND profile based on an adaptive foveated weighting model [C] // 2018 IEEE Visual Communications and Image Processing (VCIP),

- December 9-12, 2018, Taichung, Taiwan, China. New York: IEEE Press, 2018: 18620452.
- [4] Liu D, Wang Y B, Chen Z Z. Joint foveation-depth just-noticeable-difference model for virtual reality environment [J]. *Journal of Visual Communication and Image Representation*, 2018, 56: 73-82.
- [5] Fan Y, Larabi M C, Cheikh F A, et al. Just noticeable difference model for asymmetrically distorted stereoscopic images [C] // ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 12-17, 2019, Brighton, UK. New York: IEEE Press, 2019: 2277-2281.
- [6] Shi X W, Yuan H, Lü M X, et al. Current status and progress of virtual reality technology in medical field [J]. *Laser & Optoelectronics Progress*, 2020, 57(1): 010006.
石晓卫, 苑慧, 吕茗萱, 等. 虚拟现实技术在医学领域的研究现状与进展 [J]. *激光与光电子学进展*, 2020, 57(1): 010006.
- [7] Fan L Y, Ma J Y, Zhang K F, et al. The development status and prospect of virtual reality hardware industry [J]. *Science & Technology Review*, 2019, 37(5): 81-88.
范丽亚, 马介渊, 张克发, 等. 虚拟现实硬件产业的发展 [J]. *科技导报*, 2019, 37(5): 81-88.
- [8] Ji Y, Yu Y S, Gao Q, et al. System of OLED-on-silicon micro display based on visual perception of human eye [J]. *Acta Photonica Sinica*, 2019, 48(4): 0411001.
季渊, 余云森, 高钦, 等. 基于人眼视觉特性的硅基 OLED 微显示器系统 [J]. *光子学报*, 2019, 48(4): 0411001.
- [9] Ji Y, Gao Q, Yu Y S, et al. OLED-on-silicon microdisplay controller based on multi-scale pyramid [J]. *Acta Optica Sinica*, 2019, 39(12): 1223001.
季渊, 高钦, 余云森, 等. 基于多尺度金字塔的硅基 OLED 微显示控制器 [J]. *光学学报*, 2019, 39(12): 1223001.
- [10] Tam W J, Speranza F, Yano S, et al. Stereoscopic 3D-TV: visual comfort [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 335-346.
- [11] Alfonso J F, Fernández-Vega L, Señaris A, et al. Prospective study of the Acri. LISA bifocal intraocular lens [J]. *Journal of Cataract & Refractive Surgery*, 2007, 33(11): 1930-1935.
- [12] Wang J L, 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.
- [13] Wu J J, Li L D, Dong W S, et al. Enhanced just noticeable difference model for images with pattern complexity [J]. *IEEE Transactions on Image Processing*, 2017, 26(6): 2682-2693.
- [14] Wang X Y, Yang H Y, Chen L K. New color image watermarking based on adaptive quantization [J]. *Mini-Micro Systems*, 2005, 26(9): 1525-1529.
王向阳, 杨红颖, 陈利科. 基于人眼视觉系统的自适应量化数字水印算法研究 [J]. *小型微型计算机系统*, 2005, 26(9): 1525-1529.
- [15] Mu T Z, Ji Y, Chen W D, et al. Organic-light-emitting-diode-on-silicon microdisplay based on double-frame digital-analog-hybrid scanning strategy [J]. *Laser & Optoelectronics Progress*, 2019, 56(9): 092302.
穆廷洲, 季渊, 陈文栋, 等. 基于双帧数模融合扫描的硅基有机发光二极管微显示器 [J]. *激光与光电子学进展*, 2019, 56(9): 092302.
- [16] Lee S, Pattichis M S, Bovik A C. Foveated video quality assessment [J]. *IEEE Transactions on Multimedia*, 2002, 4(1): 129-132.
- [17] Bringmann A, Syrbe S, Görner K, et al. The primate fovea: structure, function and development [J]. *Progress in Retinal and Eye Research*, 2018, 66: 49-84.