

基于双目神经元响应的无参考立体图像质量评价

叶蒙蒙, 胡晋滨, 王雪津, 邵枫*

宁波大学信息科学与工程学院, 浙江 宁波 315211

摘要 为了解决多失真图像的质量预测偏差问题,根据视觉生理心理学研究中人类大脑初级视皮层(V1)神经元细胞对视觉信息处理的过程,提出了一种无参考立体图像质量评价方法。首先对失真立体图像对进行 Gabor 滤波,构造了基于双目神经元响应的 V1 区模拟刺激模型;其次通过离散余弦变换(DCT)和去均值对比度归一化(MSCN),分别提取了失真立体图像对的 DCT 域和空间域的自然场景统计特征;最后采用支持向量回归(SVR)算法,建立了所提取特征和主观评价价值间的映射关系,进而构建了预测立体图像质量的客观评价模型。基于公开数据库对所提模型进行了验证和对比。结果表明,所提方法可统一预测单失真和多失真立体图像的感知质量,比现有评价方法的性能更优。

关键词 图像处理; 图像质量评价; 双目神经元; 离散余弦变换; 支持向量回归

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.2410007

No-Reference Stereoscopic Image Quality Assessment Based on Binocular Neuron Response

Ye Mengmeng, Hu Jinbin, Wang Xuejin, Shao Feng*

Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract In order to solve the problem of quality prediction deviation of multiply-distorted images, a method for no-reference stereoscopic image quality assessment is proposed according to the process of visual information processed by neurons in human primary visual cortex (V1) in the research of visual physiology and psychology. Firstly, Gabor filtering is performed on the distorted stereoscopic image pairs to construct a simulated stimulus model of the V1 layer based on the binocular neuron response. Second, with the discrete cosine transformation (DCT) and the mean subtracted contrast normalization (MSCN), the natural scene statistics features of those distorted stereoscopic image pairs in DCT domain and spatial domain are extracted, respectively. Finally, the support vector regression (SVR) is adopted to build the objective evaluation model for predicting stereoscopic image quality via establishing the mapping relationship between the extracted features and the subjective scores. The proposed model is verified and compared based on the public databases, and the results show that the proposed method can uniformly predict the perceptual quality of singly-distorted and multiply-distorted stereoscopic images with better performance than that of other existing evaluation methods.

Key words image processing; image quality assessment; binocular neuron; discrete cosine transformation; support vector regression

OCIS codes 100.2000; 100.6890

1 引言

随着立体图像生成技术的快速发展和显示设备

的更新换代,立体图像逐渐融入到我们的日常生活之中,3D 电影、3D 游戏、虚拟现实等越来越受欢迎^[1]。高质量的立体图像也成为人们关注的焦点,

收稿日期: 2021-01-04; 修回日期: 2021-01-28; 录用日期: 2021-03-05

基金项目: 国家自然科学基金(61622109)

通信作者: *shaofeng@126.com

而立体图像质量评价(SIQA)作为衡量立体图像失真程度的方法,是目前图像处理领域的一个研究热点。根据是否需要原始参考图像,图像质量评价方法通常分为全参考(FR)、半参考(RR)、无参考(NR)三类。其中,无参考图像质量评价方法无需原始图像的任何信息,就可预测图像质量,与人类视觉感知识别过程较为一致,具有较大的研究和应用价值^[2-5]。

目前,关于立体图像质量评价方法的研究主要是针对单失真图像^[6-7]。对于 FR-SIQA,Shao 等^[8]考虑了双眼感知和组合视觉特性,使用双眼差异模型反映双眼融合和抑制区域的视觉敏感性,提出了一种基于双目视觉特性的立体图像质量评价方法;Bensalma 等^[9]根据双目视觉特性提出了双目能量质量评价模型,通过计算原始图像和失真图像的双目能量图并比较其相似性,得到了立体图像的质量;Chen 等^[10]建立了双目抑制模型,通过模拟双目竞争行为,评价了立体图像的质量。对于 NR-SIQA,黄姝钰等^[11]利用小波变换分解重构了立体图像的左右视图,提出了一种基于图像融合的无参考立体图像质量评价算法;李一凡等^[12]通过四元数小波变换分解了左右视图,并结合人眼视觉特性生成了单目图,提出了一种无参考立体图像质量评价;Su 等^[13]提出了一种基于单变量和双变量自然场景统计(NSS)模型的无参考立体图像质量评价方法;Shao 等^[14]通过构建深度神经网络模型,提出了一种基于双目视觉融合的无参考立体图像质量评价方法,取得了较好的评价结果。

然而,立体图像在捕获和处理阶段容易产生多种类型的失真,因此研究多失真图像更为实用。以上 SIQA 方法可以用于评价多失真图像,但针对性不强且性能一般。目前,只有少数 SIQA 方法是专

门针对多失真立体图像的。文献[15]建立了一个对称多失真立体图像数据库,通过对不同失真类型下双目融合的过程进行建模分析,提出了一种无参考质量评价方法,评价结果与图像主观质量有较好的一致性,但其多模式稀疏表示框架需要预先了解失真类型,而且只考虑了对称多失真图像;文献[16]建立了一个非对称多失真立体图像数据库,并提出了一个多阶段池化模型,取得了较好的评价结果,但由于训练模型时未采用非对称多失真图像,因此预测感知和实际感知之间存在预测偏差。

上述方法缺乏对单失真和多失真立体图像的统一考虑。因此,本文提出了一种基于双目神经元响应的 NR-SIQA 方法,根据人类视觉系统(HVS)对立体图像信息处理的过程,建立了初级视皮层(V1)神经元响应的模型,提取了离散余弦变换(DCT)域和空间域的自然场景统计特征,最后采用支持向量回归(SVR)算法建立了立体图像特征与主观评价价值的映射关系,从而预测了立体图像质量的客观评价价值。实验结果表明,本文方法可以综合考虑单失真和多失真立体图像,能较好地预测图像的主观感知质量。

2 模型的基本原理

HVS 相关的生理心理学研究表明,双目神经元响应是人眼感知立体图像的主要行为^[17-18]。现有的研究表明,二维 Gabor 滤波器与哺乳动物视觉系统中的简单神经元细胞的感受野函数非常相似,在空间域和频率域可同时实现最优局部化,能够很好地抓取图像中对应于空间频率、空间位置及方向选择性的局部结构信息^[19]。在空间域中,二维 Gabor 滤波器是一个由正弦平面波调制的高斯核函数,即

$$\begin{cases} g(x, y; \theta) = \exp\left(-\frac{x_0^2 + \gamma^2 y_0^2}{2\sigma^2}\right) \exp\left[i\left(2\pi \frac{x_0}{\lambda} + \psi\right)\right] \\ x_0 = x \cos \theta + y \sin \theta \\ y_0 = -x \sin \theta + y \cos \theta \end{cases} \quad (1)$$

Gabor 函数分为实部和虚部,即偶函数部分(g_e)和奇函数部分(g_o),可分别表示为

$$g_e(x, y; \theta) = \exp\left(-\frac{x_0^2 + \gamma^2 y_0^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x_0}{\lambda} + \psi\right), \quad (2)$$

$$g_o(x, y; \theta) = \exp\left(-\frac{x_0^2 + \gamma^2 y_0^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x_0}{\lambda} + \psi\right), \quad (3)$$

式中: (x_0, y_0) 为 Gabor 滤波器在极坐标系下的空间位置; (x, y) 为空间域像素的位置; λ 为 Gabor 函

数的波长; θ 表征 Gabor 函数的滤波方向; ψ 为相位偏移; σ 为 Gabor 函数的高斯因子的标准差; γ 为

间纵横比,决定了 Gabor 函数的形状。

图像经 Gabor 滤波器滤波后,得到的复值输出响应分为偶响应和奇响应。偶响应是 Gabor 复值响应的实部,用以平滑图像;奇响应是 Gabor 复值响应的虚部,用以提取图像的边缘。对于输入的立体图像左视点 $f_L(x, y)$ 和右视点 $f_R(x, y)$, Gabor 滤波后的偶响应和奇响应可表示为

$$F_{Le}(x, y; \theta) = f_L(x, y) \otimes g_e(x, y; \theta), \quad (4)$$

$$F_{Lo}(x, y; \theta) = f_L(x, y) \otimes g_o(x, y; \theta), \quad (5)$$

$$F_{Re}(x, y; \theta) = f_R(x, y) \otimes g_e(x, y; \theta), \quad (6)$$

$$F_{Ro}(x, y; \theta) = f_R(x, y) \otimes g_o(x, y; \theta), \quad (7)$$

式中: \otimes 为卷积运算。

立体图像不同于平面图像,双目感知的深度信息是其最基本且最重要的信息之一。大多立体图像质量评价的研究^[20-23]仅利用 Gabor 幅值变换、小波变换、Bandelet 变换等对视皮层简单细胞的双目响应过程进行简单模拟,且大部分研究仅提取基于立体图像深度的感知特征,而忽略了贴近人类视觉感知机理模型的重要性。

人类具有独特且高度发达的视觉系统,而视皮层是其核心部分。立体感知涉及大量皮质区域以及它们之间的相互关系,V1 神经元具有水平视差调整功能,在立体图像深度信息处理中起着至关重要的作用。在视觉神经科学和相关心理学研究中,Ohzawa 等^[24]提出了一个双目能量模型,如图 1 所示,首先通过 Gabor 函数模拟 V1 区单目简单细胞

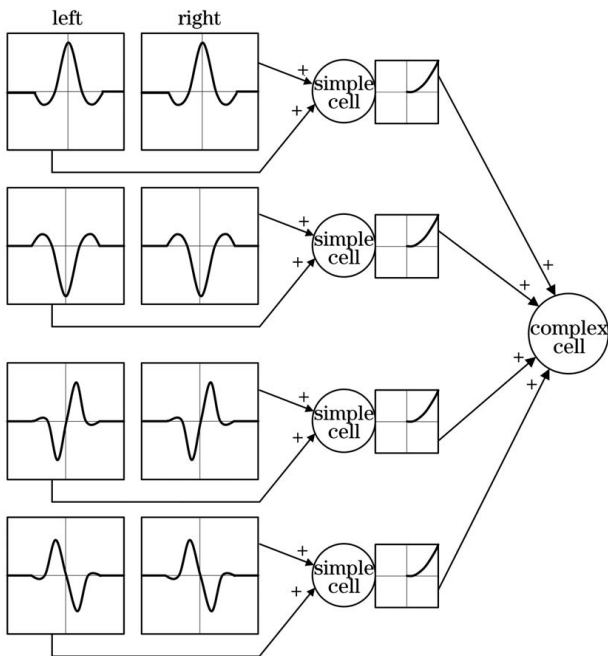


图 1 双目能量模型图

Fig. 1 Diagram of binocular energy model

对视觉刺激的响应,再将左视点和右视点信息组合,从而构建双目简单细胞模型,最后通过不同皮质区域简单细胞的共同作用,模拟双目复杂细胞的视差感知特性。Read 等^[25]对上述双目能量模型进行改进,其中单目简单细胞的响应在双目组合之前先经过半波整流,该模型可以解释复杂细胞对双目相关和反相关刺激的响应,更好地描述视皮层中视差调谐神经元细胞对立体图像的响应过程。

为了寻求一种贴近双目视觉感知的评价方法,本文仿照人类双目视觉系统对立体图像的识别处理过程,以及初级视皮层简单神经元在生理心理学上的数学模型,再结合上述两种视觉神经科学研究中简单细胞和复杂细胞的响应过程,考虑了以下五类 V1 区神经元模型^[26],分别为

$$\text{ODF TE: } R_1 = (F_{Le} + F_{Re})^2, \quad (8)$$

$$\text{ODF TI: } R_2 = (F_{Le} - F_{Re})^2, \quad (9)$$

$$\text{ODF ODD: } R_3 = (F_{Le} + F_{Ro})^2, \quad (10)$$

$$\text{RPC TE: } R_4 = ([F_{Le}] + [F_{Re}])^2, \quad (11)$$

$$\text{RPC ODD: } R_5 = ([F_{Lo}] - [F_{Ro}])^2, \quad (12)$$

式中:“ODF”表示 Ohzawa、DeAngelis 和 Freeman 引入的双目能量模型的原始形式;“RPC”表示 Read、Parker 和 Cumming 提出的修改模型形式;“TE”表示调谐兴奋,指单眼感受野具有相同相位并且视差调谐曲线关于中心峰对称的细胞^[27];“TI”表示调谐抑制,指单眼感受野具有相反相位并且视差调谐曲线关于中心谷对称的细胞;“ODD”指单眼感受野异相 $\pi/2$ 的细胞,在 ODF 模型中,此类细胞具有奇对称调整的特性; $[\]$ 为半波整流,即

$$[x] = \begin{cases} x, & x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

视觉感知质量与人的视觉神经系统和心理感受有关,本文通过模拟视皮层简单神经元的响应并结合多种 V1 神经元,对神经元之间的激励作用或抑制作用进行建模,以表征复杂神经元对立体图像的双目响应特性。

3 基于双目神经元响应的无参考立体图像质量评价

本文通过模拟人类视觉系统刺激,设计双目神经元响应并提取特征,提出了一种 NR-SIQA 方法,其总体框架如图 2 所示。首先,对失真的立体图像对进行 Gabor 滤波,通过模拟 V1 区五类神经元响

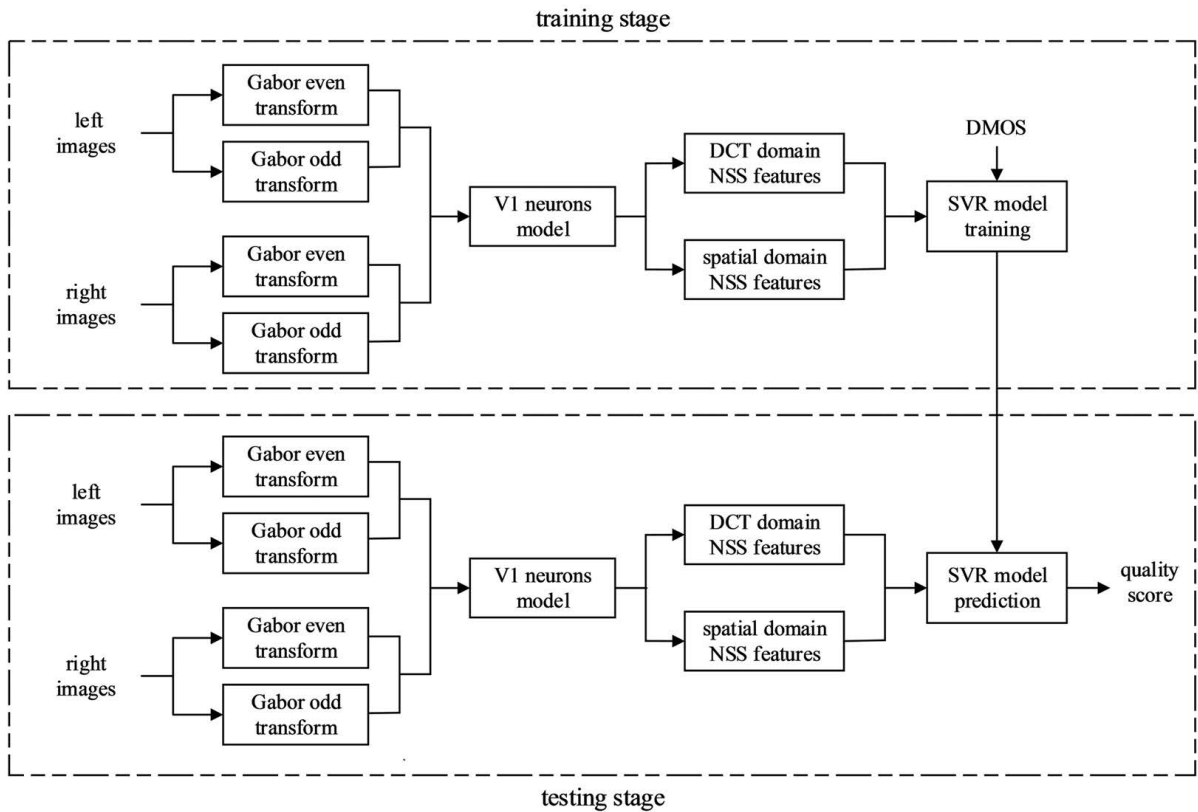


图 2 立体图像质量评价模型框图

Fig. 2 Diagram of stereoscopic image quality assessment model

应,提取出不同滤波方向、不同神经元响应图的特征,最后通过 SVR 算法建立立体图像特征与主观评价值的映射关系,从而预测得出立体图像质量的客观评价价值。

3.1 模拟双目神经元响应图像

对于立体图像对 $f_L(x, y)$ 和 $f_R(x, y)$, 其双目神经元响应图像表示为 $R_i(x, y; \theta_j)$ ($i=1, \dots, m$;

$j=1, \dots, n$)。其中, θ_j 表征第 j 个滤波方向, i 为 V1 区神经元模型的索引, j 为 Gabor 滤波器滤波方向的索引, m 为 V1 区神经元模型的个数, n 为 Gabor 滤波器滤波方向的个数, 本文考虑五类神经元模型, 四个滤波方向 $0, \pi/4, \pi/2, 3\pi/4$, 即 $m=5$, $n=4$ 。不同滤波方向下的五类 V1 神经元响应图像如图 3 所示。

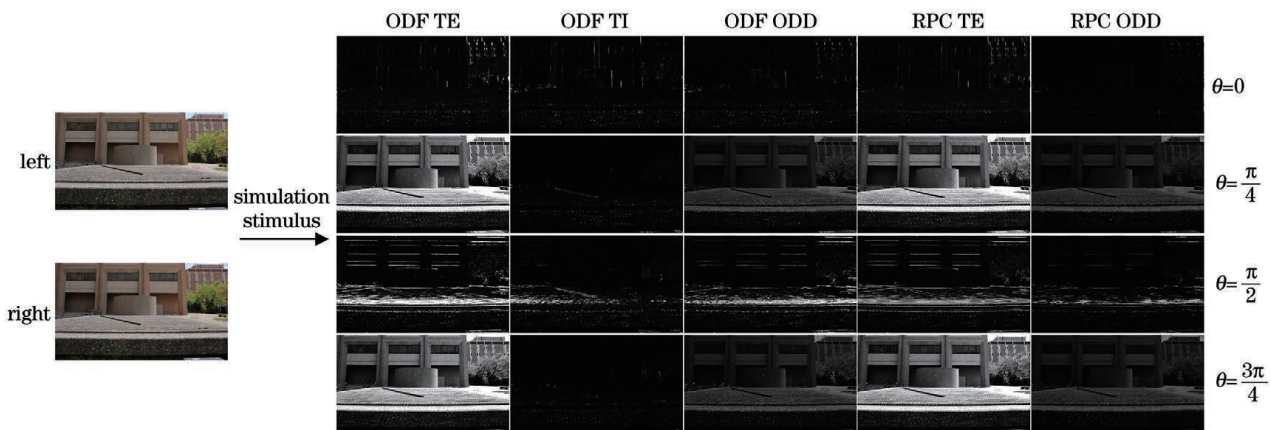


图 3 五类 V1 神经元响应图像

Fig. 3 Five types of V1 neuron response images

3.2 神经元响应图像特征提取

NSS 模型通常采用广义高斯分布 (GGD) 和非

对称广义高斯分布 (AGGD) 模型对自然图像的统计特征进行建模, 可有效地捕获失真图像的统计数据。

GGD 模型可以表示为

$$f(z; \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left[-\left(\frac{|z|}{\beta}\right)^\alpha\right], \quad (14)$$

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}, \quad (15)$$

$$\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt, z > 0, \quad (16)$$

式中: z 为输入图像的像素强度; β 为控制分布的尺度参数; t 为伽马函数的积分变量; α 为分布的形状参数; σ^2 为分布的方差; $\Gamma(\cdot)$ 为伽马函数。AGGD 模型可表示为

$$f(z; \nu, \sigma_l^2, \sigma_r^2) = \begin{cases} \frac{\nu}{(\beta_l + \beta_r)\Gamma(1/\nu)} \exp\left[-\left(\frac{-z}{\beta_l}\right)^\nu\right], & z < 0 \\ \frac{\nu}{(\beta_l + \beta_r)\Gamma(1/\nu)} \exp\left[-\left(\frac{z}{\beta_r}\right)^\nu\right], & z \geq 0 \end{cases}, \quad (17)$$

$$\begin{cases} \beta_l = \sigma_l \sqrt{\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}} \\ \beta_r = \sigma_r \sqrt{\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}} \end{cases}, \quad (18)$$

$$\eta = (\beta_r - \beta_l) \frac{\Gamma(2/\nu)}{\Gamma(1/\nu)}, \quad (19)$$

式中: β_l 和 β_r 分别为控制模型两侧分布的尺度参数; η 为经验所得的特征元素; ν 为分布的形状参数; σ_l^2, σ_r^2 分别为模型两侧分布的方差。

3.2.1 DCT 域特征

由于 DCT 具有良好的去相关和能量集中特性, 自然图像的能量都集中在离散余弦变换后的低频部分, 因此采用二维 DCT 对双目神经元响应图像 $R_i(x, y; \theta_j)$ ($i=1, \dots, m; j=1, \dots, n$) 进行变换, 并对提取的局部 DCT 系数进行建模, 再将模型的参数用于提取 DCT 域的 NSS 特征。

首先对双目神经元响应图像进行分块(块大小为 5 pixel×5 pixel, 块间重合为 2 pixel), 对每个图像块进行 DCT 处理, 再用 GGD 模型对 DCT 系数进行拟合, 获得每个图像块的形状参数 τ , 分别用最小的 10% τ 值的均值和所有 τ 值的均值表示第 1、2 个特征, 记为 f_1, f_2 。第二个参数为频率变化系数 ζ , 计算公式为

$$\zeta = \frac{\sigma |X|}{\mu |X|} = \sqrt{\frac{\Gamma(1/\tau)\Gamma(3/\tau)}{\Gamma^2(2/\tau)}} - 1, \quad (20)$$

式中: $\mu |X|$ 和 $\sigma |X|$ 分别为 DCT 系数幅值 $|X|$ 的均值和标准差。计算每个图像块的 ζ 值, 分别用最大的 10% ζ 值的均值和所有 ζ 值的均值表示第 3、4 个特征, 记为 f_3, f_4 。

图像失真通常会改变图像的局部光谱特征, 使其与原始图像的光谱特征不同。为了衡量图像光谱特征的变化, 采用局部 DCT 子带能量比进行度量。

将每个图像块的所有 DCT 系数按子带频率分为三类, 对每个子带的 DCT 系数分别求方差 σ_k^2 , 将子带 k 的平均能量定义为其方差 ($k=1, 2, 3$ 分别表示低频、中频、高频), 即 $E_k = \sigma_k^2$ 。第三个参数为子带能量比 Q_k , 计算公式为

$$Q_k = \frac{\left|E_k - \frac{1}{k-1} \sum_{l < k} E_l\right|}{E_k + \frac{1}{k-1} \sum_{l < k} E_l}, k = 2, 3, \quad (21)$$

式中: E_l 为子带 l 的平均能量; l 为比当前子带频率更低的子带。计算所有图像块的 Q_2 和 Q_3 的平均值 \bar{Q} , 分别用最大的 10% \bar{Q} 值的均值和所有 \bar{Q} 值的均值表示第 5、6 个特征, 记为 f_5, f_6 。

图像失真常以非自然的方式改变局部各个方向的能量, 为了捕获图像中与人类主观感知相关的方向信息, 将图像块的 DCT 系数按照频谱分为三个方向信息子带。对每个子带用 GGD 模型进行拟合, 得到参数 ζ_k ($k=1, 2, 3$), 计算其方差 $D(\zeta)$, 分别用最大的 10% $D(\zeta)$ 值的均值和所有 $D(\zeta)$ 值的均值表示第 7、8 个特征, 记为 f_7, f_8 。

3.2.2 空间域特征

空域失真是导致图像失真的直接原因, 其特征提取算法的复杂度较低, 执行速度较快, 与变换域特征提取算法相比, 更加适用于实时系统。首先对双目神经元响应图像 $R_i(x, y; \theta_j)$ ($i=1, \dots, m; j=1, \dots, n$) 进行对比度归一化处理, 得到去均值对比度归一化(MSCN)系数, 计算公式为

$$\hat{R}_i(x, y; \theta_j) = \frac{R_i(x, y; \theta_j) - \mu_i(x, y; \theta_j)}{\sigma_i(x, y; \theta_j) + C}, \quad (22)$$

$$\mu_i(x, y; \theta_j) = \sum_{a=-3}^3 \sum_{b=-3}^3 w_{(a,b)} R_i^{(a,b)}(x, y; \theta_j), \quad (23)$$

$$\sigma_i(x, y; \theta_j) = \sqrt{\sum_{a=-3}^3 \sum_{b=-3}^3 w_{(a,b)} [R_i^{(a,b)}(x, y; \theta_j) - \mu_i(x, y; \theta_j)]^2}, \quad (24)$$

式中： μ_i 和 σ_i 分别为第 i 个神经元响应图像的均值和标准差； $R_i^{(a,b)}$ 为高斯加权的第 i 个神经元响应图像； $x \in \{1, 2, \dots, M\}$ ， $y \in \{1, 2, \dots, N\}$ ，其中 M 和 N 分别为图像的高度和宽度； C 为常数 1； $w = \{w_{(a,b)} | a = -3, \dots, 3; b = -3, \dots, 3\}$ 为一个二维圆形对称的高斯加权函数，其中 $w_{(a,b)}$ 为加权因子， a 和 b 为高斯核尺寸。

自然图像的 MSCN 系数统计直方图呈高斯分布，图像失真会破坏其统计特性。本文对上述 MSCN 系数采用 GGD 模型进行拟合，将模型的拟

合参数 (α, σ^2) 作为两个分布特征，记为 f_9, f_{10} 。另外，相邻的 MSCN 系数之间也显示出规则的统计特性，图像失真也会导致其受到破坏，因此采用 AGGD 模型对四个方向（水平、垂直、主对角线、副对角线）的相邻 MSCN 系数的乘积进行拟合，每个方向得四个拟合参数 $(\eta, \nu, \sigma_1^2, \sigma_r^2)$ ，记为 $f_{11} \sim f_{26}$ 。

本文对立体图像对进行四个方向的 Gabor 滤波后，又建立了五类 V1 区神经元模型，一共得到 20 幅双目神经元响应模拟刺激图像，每幅图像提取 8 个 DCT 域特征和 18 个空间域特征，如表 1 所示。

表 1 所提取特征的总结

Table 1 Summary of extracted features

Feature	Feature description	Symbolic representation
DCT domain feature	10% and 100% pooled shape parameter	f_1 and f_2
	10% and 100% pooled coefficient of frequency variation	f_3 and f_4
	10% and 100% pooled energy sub-band ratio	f_5 and f_6
	10% and 100% pooled orientation feature	f_7 and f_8
Spatial feature	(α, σ^2) from GGD fit of MSCN coefficients	f_9 and f_{10}
	$(\eta, \nu, \sigma_1^2, \sigma_r^2)$ from AGGD fit of four adjacent MSCN coefficients	$f_{11} \sim f_{26}$

3.3 特征回归及质量分数预测

SVR 算法是基于统计学习理论的新型机器学习方法，具有出色的学习性能。基于 SVR 的立体图像质量预测过程主要包含训练阶段和测试阶段。

3.3.1 训练阶段

输入 SVR 的训练数据集可以表示为

$$\begin{bmatrix} \mathbf{F} \\ q^{\text{DMOS}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \dots & \mathbf{F}_S \\ q_1^{\text{DMOS}} & q_2^{\text{DMOS}} & \dots & q_S^{\text{DMOS}} \end{bmatrix}, \quad (25)$$

式中： \mathbf{F} 为训练图像的特征矢量； q^{DMOS} 为对应主观标签值； S 为训练图像对的个数； \mathbf{F}_s 为第 s ($s = 1, 2, \dots, S$) 个训练图像的特征矢量， s 为训练图像的索引； q_s^{DMOS} 为对应的主观标签值，其中 DMOS 为平均主观得分差异。SVR 运用非线性映射把数据集从原始特征空间映射到高维特征空间中，再进行线性

回归以构造最优的线性函数，其表达式为

$$f(\mathbf{F}) = \sum_{s=1}^S \omega_s \cdot \kappa(\mathbf{F}, \mathbf{F}_s) + \delta, \quad (26)$$

式中： $\kappa(\mathbf{F}, \mathbf{F}_s)$ 为特征矢量 \mathbf{F} 的核函数，本文采用高斯径向基函数 (RBF) 作为核函数； $\omega = (\omega_1, \omega_2, \dots, \omega_S)$ 为核函数的权重矢量，其中 ω_s 为第 s 个训练图像的权重； δ 为常数。

3.3.2 测试阶段

提供 C' 个测试图像的特征矢量 \mathbf{F}_c ($c = 1, 2, \dots, C'$) (c 为测试图像的索引)，通过测试集的特征回归得到测试图像的客观评价值 q 。主观质量评价过程会引入一系列非线性因素，为了避免其对客观质量评价方法性能的影响，本文采用 5 参数 Logistic 函数对客观评价值 q 进行非线性拟合，得出客观评价预测值 q' 为

$$q' = \beta_1 \cdot \left\{ \frac{1}{2} - \frac{1}{1 + \exp[\beta_2 \cdot (q - \beta_3)]} \right\} + \beta_4 \cdot q + \beta_5, \quad (27)$$

式中： $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ 为训练时设置的常量。

4 实验设计及结果分析

4.1 SIQA 数据库介绍

本文实验选取 LIVE 实验室 SIQA 数据库 LIVE 3D Phase-I、II 和宁波大学多失真 SIQA 数据库 NBU-MDSID Phase-I、II。

LIVE 3D Phase-I^[28] (LIVE-I) 包含 20 组原始立体图像对和 365 组失真立体图像对,有 JPEG2000 压缩(JP2K)80 组、JPEG 压缩(JPEG)80 组、高斯白噪声(WN)80 组、高斯模糊(GB)45 组以及快衰落信道失真(FF)80 组五种失真类型,均是左视点和右视点失真程度相同的对称失真。

LIVE 3D Phase-II^[29] (LIVE-II) 包含 8 组原始立体图像对和 360 组失真立体图像对(对称失真和非对称失真均有),失真类型和 LIVE-I 相同,每种失真类型各 72 组。

NBU-MDSID Phase-I^[15] (NBU-MDSID I) 包含 10 组原始立体图像对、270 组多失真立体图像对和 90 组单失真立体图像对,其中多失真立体图像对是通过按顺序在每组原始立体图像对中添加 GB、JPEG 和 WN 失真而生成的,单失真立体图像对是通过在相同的原始立体图像对中仅添加一种失真而产生的,均为对称失真。

NBU-MDSID Phase-II^[16] (NBU-MDSID II) 包含 10 组原始立体图像对和 300 组非对称多失真立体图像对,多失真立体图像对是通过按不同顺序在每组原始左右图像中添加 GB、JPEG 和 WN 失真而

生成的。

4.2 各图像数据库性能比较与分析

本文实验采取相同的训练测试策略,随机选取 80%失真立体图像对作为训练集,构造 SVR 训练模型,剩余 20%作为测试集用以评价算法性能,实验进行 1000 次,选取平均值作为最终实验结果。在单失真立体图像数据库和多失真立体图像数据库上均进行了实验,为了客观评价本文方法的性能,在两类数据库上对本文方法与现有的相关算法进行了对比。

实验对比采用三个常用的衡量客观图像质量评价算法性能的评价指标,分别为 Pearson 线性相关系数(PLCC)、Spearman 等级系数(SROCC)、均方根误差(RMSE)。其中,PLCC 描述算法预测值与人眼主观评价价值之间的相关性,其绝对值越接近 1,表明算法性能越好;SROCC 衡量算法预测的单调性,其绝对值越接近 1,表明算法预测值与人眼主观评价价值的单调性越好;RMSE 衡量算法预测的准确性,其值越接近 0,表明算法预测值与人眼主观评价价值的准确性越好。

为了验证本文提出的客观评价模型对单失真立体图像质量预测的有效性,表 2 给出了本文方法与五种 FR 质量评价方法、四种 NR 质量评价方法在 LIVE-I 和 LIVE-II 上的各项性能指标,其中最优结果由粗体显示。从表 2 可以看出,本文方法在 LIVE-I 上的 PLCC 超过 0.93, SROCC 超过 0.92, RMSE 接近 5.5;对于 LIVE-II, PLCC、SROCC 均超过 0.93, RMSE 接近 3.9,均优于其他方法。表 3

表 2 各评价方法在 LIVE-I 和 LIVE-II 上的性能比较

Table 2 Performance comparison among different assessment metrics on LIVE-I and LIVE-II

Method	Type	LIVE-I			LIVE-II		
		PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
FI-PSNR	FR	0.865	0.856	8.242	0.658	0.638	8.496
FI-SSIM	FR	0.870	0.861	8.087	0.684	0.680	8.230
Method in Ref. [8]	FR	0.925	0.922	-	0.759	0.745	-
Method in Ref. [9]	FR	0.889	0.877	7.519	0.770	0.751	7.204
Method in Ref. [10]	FR	0.918	0.909	6.501	0.907	0.901	4.766
Method in Ref. [3]	NR	0.917	0.911	5.864	0.737	0.701	7.665
Method in Ref. [4]	NR	0.910	0.901	6.794	0.750	0.701	7.198
Method in Ref. [13]	NR	0.922	0.903	6.258	0.913	0.905	4.657
Method in Ref. [15]	NR	0.891	0.885	-	0.784	0.805	-
Method in Ref. [22]	NR	0.938	0.868	-	0.851	0.831	-
Proposed method	NR	0.938	0.927	5.583	0.937	0.931	3.901

表 3 各评价方法在 LIVE-I 和 LIVE-II 库的不同失真类型下的 SROCC 值

Table 3 SROCC results of different assessment metrics on LIVE-I and LIVE-II for different distortion types

Method	LIVE-I					LIVE-II				
	JPEG	JP2K	WN	GB	FF	JPEG	JP2K	WN	GB	FF
FI-PSNR	0.207	0.839	0.928	0.935	0.658	0.613	0.719	0.907	0.711	0.701
FI-SSIM	0.241	0.822	0.928	0.879	0.687	0.564	0.700	0.909	0.739	0.735
Method in Ref. [8]	0.615	0.875	0.943	0.938	0.781	0.720	0.848	0.846	0.801	0.851
Method in Ref. [9]	0.347	0.819	0.908	0.918	0.653	0.846	0.804	0.939	0.884	0.874
Method in Ref. [10]	0.440	0.865	0.937	0.924	0.758	0.840	0.833	0.955	0.910	0.889
Method in Ref. [3]	0.699	0.890	0.899	0.922	0.649	0.566	0.704	0.459	0.896	0.711
Method in Ref. [4]	0.570	0.812	0.940	0.878	0.784	0.605	0.695	0.440	0.860	0.683
Method in Ref. [13]	0.603	0.838	0.906	0.791	0.679	0.818	0.845	0.946	0.903	0.899
Method in Ref. [15]	0.693	-	0.899	0.853	-	0.622	-	0.803	0.713	-
Method in Ref. [22]	0.633	-	0.920	0.903	-	0.788	-	0.929	0.909	-
Proposed method	0.742	0.891	0.920	0.867	0.737	0.847	0.850	0.942	0.914	0.903

所示为各评价方法在 LIVE-I 和 LIVE-II 库的不同失真类型下的 SROCC 值,其中最优结果由粗体显示。从表 3 可以看出,本文方法在五种失真类型下的大部分 SROCC 值在 0.89 以上,且多次成为评价指标最优的评价方法。尤其是在包含对称失真和非对称失真的 LIVE-II 上,本文方法对 JPEG、JP2K、GB、FF 失真图像的评价指标均最优,对 WN 失真图像的评价指标也接近最优。尽管对 LIVE-I 库的 GB 失真图像的评价性能稍差,但本文方法在大部分失真类型上的评价性能和总体性能都是更优的。

为了验证本文方法对单失真和多失真立体图像

质量预测的通用性,表 4 给出了本文方法与三种 FR 质量评价方法、六种 NR 质量评价方法在多失真立体图像数据库 NBU-MDSID I 和 NBU-MDSID II 上的各项性能指标,其中最优结果由粗体显示。从表 4 可以看出,本文方法在 NBU-MDSID I 上的 PLCC 超过 0.96, SROCC 超过 0.95, RMSE 接近 3.8;对于 NBU-MDSID II, PLCC 值和 SROCC 值均超过 0.85, RMSE 接近 6.1, 大部分评价指标均优于其他方法,其余评价指标也接近最优。可见,本文方法对多失真立体图像的质量预测也较为准确,证明了本文方法的通用性和有效性。

表 4 各评价方法在 NBU-MDSID I 和 NBU-MDSID II 上的性能比较

Table 4 Performance comparison among different assessment metrics on NBU-MDSID I and NBU-MDSID II

Method	Type	NBU-MDSID I			NBU-MDSID II		
		PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
Method in Ref. [8]	FR	0.919	0.905	3.687	0.802	0.862	7.212
Method in Ref. [9]	FR	0.856	0.834	4.943	0.820	0.780	7.110
Method in Ref. [10]	FR	0.885	0.877	4.385	0.763	0.749	7.560
Method in Ref. [3]	NR	0.937	0.930	3.244	0.835	0.800	6.589
Method in Ref. [4]	NR	0.920	0.900	3.701	0.824	0.799	6.827
Method in Ref. [14]	NR	0.934	0.920	3.348	0.791	0.763	7.180
Method in Ref. [15]	NR	0.878	0.882	4.570	0.606	0.627	9.586
Method in Ref. [16]	NR	0.916	0.922	3.836	0.785	0.765	7.442
Method in Ref. [30]	NR	0.938	0.926	3.062	-	-	-
Method in Ref. [22]	NR	0.940	0.936	3.804	0.845	0.819	7.020
Proposed method	NR	0.963	0.952	3.887	0.859	0.851	6.110

为了进一步验证本文神经元模型的有效性,表 5 给出了五类 V1 神经元单独响应和本文模型的性能比较,其中最优结果由粗体显示。可以看出,对于多失真立体图像库 NBU-MDSID I、NBU-MDSID II 和非对称单失真立体图像库 LIVE-II,本文构造的基于 V1 区复杂神经元响应的模拟刺激模型的性能均优于五类 V1 神经元单独响应模型的性能;而对于对称单失真立体图像库 LIVE-I,本文模型的性能略差于五类 V1 神经元单独响应模型的性能,因为 LIVE-I 库的左、右视点均为对称单失真类型,单个简单 V1 神经元即可对图像进行

感知处理。表 6 给出了五类神经元响应下 LIVE-I、LIVE-II 库的不同失真类型的 SROCC 值,其中最优和次优结果由粗体显示。从表 6 可以看出,每类神经元响应对不同失真类型图像的评价性能有不同程度的贡献,对于 LIVE-I, RPC TE、RPC ODD 神经元响应表现出较好的评价性能;对于 LIVE-II, ODF TI、RPC ODD 神经元响应模型的预测较为准确。总体而言,本文构造的 V1 区模拟刺激模型在单失真和多失真图像库上都取得了较好的评价结果,该方法对立体图像的客观质量评价更为有效。

表 5 五类神经元单独响应和综合响应性能比较

Table 5 Performance comparison among individual response and comprehensive response of five types of neurons

Database	Index	Neuron response					
		ODF TE	ODF TI	ODF ODD	RPC TE	RPC ODD	All
LIVE-I	PLCC	0.944	0.942	0.944	0.943	0.946	0.938
	SROCC	0.934	0.925	0.933	0.936	0.938	0.927
LIVE-II	PLCC	0.899	0.909	0.912	0.905	0.936	0.937
	SROCC	0.864	0.883	0.880	0.872	0.926	0.931
NBU-MDSID I	PLCC	0.960	0.950	0.952	0.959	0.955	0.963
	SROCC	0.939	0.922	0.922	0.936	0.926	0.952
NBU-MDSID II	PLCC	0.807	0.702	0.729	0.810	0.748	0.859
	SROCC	0.788	0.673	0.712	0.793	0.736	0.851

表 6 五类神经元响应下不同失真类型的 SROCC 值

Table 6 SROCC results of different distortion types under five types of neuron responses

Database	Type	Neuron response				
		ODF TE	ODF TI	ODF ODD	RPC TE	RPC ODD
LIVE-I	JPEG	0.730	0.671	0.702	0.765	0.751
	JP2K	0.886	0.892	0.874	0.895	0.881
	WN	0.901	0.885	0.908	0.902	0.915
	GB	0.880	0.864	0.874	0.885	0.884
	FF	0.762	0.787	0.758	0.767	0.713
LIVE-II	JPEG	0.719	0.781	0.722	0.716	0.798
	JP2K	0.762	0.819	0.744	0.768	0.885
	WN	0.948	0.917	0.947	0.945	0.943
	GB	0.927	0.932	0.915	0.931	0.920
	FF	0.861	0.916	0.880	0.857	0.916

为了衡量算法的计算效率,从 LIVE-I 图像库选取一幅分辨率为 640 pixel × 360 pixel 的图像计算预测其质量分数的时间,实验在 Win10 下的

Matlab R2016a 中完成,使用的设备为 Intel (R) i5-8400 CPU @ 2.8 GHz,内存为 8 GB。各算法的运行时间如表 7 所示,可见本文算法不仅能达到较好

表 7 各算法的运行时间

Table 7 Running time of each algorithm

Algorithm	Method in Ref. [10]	Method in Ref. [3]	Method in Ref. [4]	Method in Ref. [13]	Method in Ref. [22]	Proposed method
Time /s	20.799	31.295	0.542	153.356	7.760	8.703

的评价效果,计算效率也较高。

5 结 论

针对多失真立体图像的质量预测偏差问题,提出了一种基于双目神经元响应的无参考立体图像质量评价方法。基于视觉生理心理学研究中 V1 区神经元细胞对视觉信息处理的过程,构建了 V1 区模拟视觉刺激模型,分别提取了失真立体图像对的 DCT 域和空间域的自然场景统计特征,并采用 SVR 算法建立了所提取特征与主观评价价值间的映射关系,从而预测了立体图像的客观质量评价价值。实验结果表明,所提方法能够准确地预测立体图像的感知质量。人类视觉系统的神经活动非常复杂,所构建的模型是一种较为简化的模型。未来的研究重点是针对视觉感知机理建立更加深入而准确的模型,进一步提高评价方法的性能。

参 考 文 献

- [1] Zhang W, Qu C F, Ma L, et al. Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network [J]. *Pattern Recognition*, 2016, 59: 176-187.
- [2] Moorthy A K, Bovik A C. Blind image quality assessment: from natural scene statistics to perceptual quality [J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3350-3364.
- [3] Saad M A, Bovik A C, Charrier C. Blind image quality assessment: a natural scene statistics approach in the DCT domain [J]. *IEEE Transactions on Image Processing*, 2012, 21(8): 3339-3352.
- [4] Mittal A, Moorthy A K, Bovik A C. No-reference image quality assessment in the spatial domain [J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.
- [5] He L H, Tao D C, Li X L, et al. Sparse representation for blind image quality assessment [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 1146-1153.
- [6] Shao F, Li K M, Lin W S, et al. Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties [J]. *IEEE Transactions on Image Processing*, 2015, 24(10): 2971-2983.
- [7] Fu Z Q, Fei Y J, Yang Y, et al. No-reference stereoscopic image quality assessment based on deep feature learning [J]. *Journal of Optoelectronics & Laser*, 2018, 29(5): 545-552.
富振奇, 费延佳, 杨艳, 等. 基于深层特征学习的无参考立体图像质量评价 [J]. *光电子·激光*, 2018, 29(5): 545-552.
- [8] Shao F, Lin W S, Gu S B, et al. Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics [J]. *IEEE Transactions on Image Processing*, 2013, 22(5): 1940-1953.
- [9] Bensalma R, Larabi M C. A perceptual metric for stereoscopic image quality assessment based on the binocular energy [J]. *Multidimensional Systems and Signal Processing*, 2013, 24(2): 281-316.
- [10] Chen M J, Su C C, Kwon D K, et al. Full-reference quality assessment of stereopairs accounting for rivalry [J]. *Signal Processing: Image Communication*, 2013, 28(9): 1143-1155.
- [11] Huang S Y, Sang Q B. No-reference stereo image quality assessment based on image fusion [J]. *Laser & Optoelectronics Progress*, 2019, 56(7): 071004.
黄姝钰, 桑庆兵. 基于图像融合的无参考立体图像质量评价 [J]. *激光与光电子学进展*, 2019, 56(7): 071004.
- [12] Li Y F, Li C F, Sang Q B. No-reference stereo image quality assessment of cyclopean images optimized using quaternion wavelet transform [J]. *Laser & Optoelectronics Progress*, 2019, 56(18): 181006.
李一凡, 李朝锋, 桑庆兵. 四元数小波变换优化单目图的无参考立体图像质量评价 [J]. *激光与光电子学进展*, 2019, 56(18): 181006.
- [13] Su C C, Cormack L K, Bovik A C. Oriented correlation models of distorted natural images with application to natural stereopair quality evaluation [J]. *IEEE Transactions on Image Processing*, 2015, 24(5): 1685-1699.
- [14] Shao F, Tian W J, Lin W S, et al. Toward a blind deep quality evaluator for stereoscopic images based on monocular and binocular interactions [J]. *IEEE*

- Transactions on Image Processing, 2016, 25(5): 2059-2074.
- [15] Shao F, Tian W J, Lin W S, et al. Learning sparse representation for No-reference quality assessment of multiply distorted stereoscopic images[J]. IEEE Transactions on Multimedia, 2017, 19(8): 1821-1836.
- [16] Shao F, Gao Y, Jiang Q P, et al. Multistage pooling for blind quality prediction of asymmetric multiply-distorted stereoscopic images[J]. IEEE Transactions on Multimedia, 2018, 20(10): 2605-2619.
- [17] Ding Y, Zhao Y, Chen X D, et al. Stereoscopic image quality assessment by analysing visual hierarchical structures and binocular effects[J]. IET Image Processing, 2019, 13(10): 1608-1615.
- [18] Yang J C, Sim K, Gao X B, et al. A blind stereoscopic image quality evaluator with segmented stacked autoencoders considering the whole visual perception route [J]. IEEE Transactions on Image Processing, 2019, 28(3): 1314-1328.
- [19] Kong R, Zhang B. Design of Gabor filters' parameter[J]. Control and Decision, 2012, 27(8): 1277-1280.
孔锐, 张冰. Gabor 滤波器参数设置[J]. 控制与决策, 2012, 27(8): 1277-1280.
- [20] Shao F, Lin W S, Wang S S, et al. Learning receptive fields and quality lookups for blind quality assessment of stereoscopic images[J]. IEEE Transactions on Cybernetics, 2016, 46(3): 730-743.
- [21] Liu Y, Kong F H, Zhen Z Z. Toward a quality predictor for stereoscopic images via analysis of human binocular visual perception[J]. IEEE Access, 2019, 7: 69283-69291.
- [22] Wang X J, Qi M L, Shao F, et al. Blind quality assessment for multiply distorted stereoscopic images towards IoT-based 3D capture systems[J]. Journal of Visual Communication and Image Representation, 2020, 71: 102868.
- [23] Fang Y M, Yan J B, Wang J H, et al. Learning a no-reference quality predictor of stereoscopic images by visual binocular properties[J]. IEEE Access, 2019, 7: 132649-132661.
- [24] Ohzawa I, DeAngelis G C, Freeman R D. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors[J]. Science, 1990, 249(4972): 1037-1041.
- [25] Read J C, Parker A J, Cumming B G. A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images[J]. Visual Neuroscience, 2002, 19(6): 735-753.
- [26] Read J C A, Cumming B G. The psychophysics of stereopsis can be explained without invoking independent ON and OFF channels [J]. Journal of Vision, 2019, 19(6): 7.
- [27] Read J C, Cumming B G. Ocular dominance predicts neither strength nor class of disparity selectivity with random-dot stimuli in primate V1 [J]. Journal of Neurophysiology, 2004, 91(3): 1271-1281.
- [28] Moorthy A K, Su C C, Mittal A, et al. Subjective evaluation of stereoscopic image quality [J]. Signal Processing: Image Communication, 2013, 28(8): 870-883.
- [29] Chen M J, Cormack L K, Bovik A C. No-reference quality assessment of natural stereopairs [J]. IEEE Transactions on Image Processing, 2013, 22(9): 3379-3391.
- [30] Jiang Q P, Shao F, Gao W, et al. Unified no-reference quality assessment of singly and multiply distorted stereoscopic images[J]. IEEE Transactions on Image Processing, 2019, 28(4): 1866-1881.