

基于视觉重要区域的立体图像视觉舒适度客观评价方法

姜求平 邵枫* 蒋刚毅 郁梅 彭宗举
(宁波大学信息科学与工程学院 宁波 315211)

摘要: 先进的立体视频技术能给观众带来深度感和沉浸感,但也容易使人产生视觉疲劳,造成观看体验质量的下降,因此,如何对立体视频/图像的视觉舒适度进行有效评价是目前的研究难点。该文提出一种基于视觉重要区域的立体图像视觉舒适度客观评价模型,该模型主要包括3个部分:(1)利用图像显著图和视差图像得到立体图像视觉重要区域;(2)提取视觉重要区域的视差幅度特征、视差梯度边缘特征以及空间频率特征,作为反映立体图像视觉舒适度的感知特征信息;(3)通过支持向量回归建立特征信息与立体图像舒适度平均主观评分值的关系,预测得到立体图像视觉舒适度的客观评价。实验结果表明,与现有的方法相比较,在相同的立体图像测试库上,采用该文提出的客观评价模型可以获得更好的评价性能。

关键词: 立体图像;视觉舒适度评价;视觉重要区域;支持向量回归

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2014)04-0875-07

DOI: 10.3724/SP.J.1146.2013.00946

An Objective Stereoscopic Image Visual Comfort Assessment Metric Based on Visual Important Regions

Jiang Qiu-ping Shao Feng Jiang Gang-yi Yu Mei Peng Zong-ju
(Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China)

Abstract: Stereo video technology provides depth perception and immersion visual experience, but it also makes people feel visual fatigue and causes the decrease of the experience quality. Therefore, how to evaluate the visual comfort of stereoscopic image/video effectively is still a research focus. In this paper, an objective visual comfort assessment metric based on visual important regions is proposed. First, the Visual Important Regions (VIR) are obtained from image saliency and disparity map information. Then, disparity amplitude, disparity gradient, and spatial frequency features are extracted and fused into a feature vector. Finally, the values of objective assessment are predicted by Support Vector Regression (SVR). Experimental results show that compared with existing methods, the proposed metric achieves higher consistency with subjective visual comfort assessment of stereoscopic images.

Key words: Stereoscopic image; Visual comfort assessment; Visual Important Regions (VIR); Support Vector Regression (SVR)

1 引言

3维/立体视频技术能够提供多感知性,深度沉浸和临场体验,极大地提升观看者的视觉体验质量(Quality of Experience, QoE)。视觉舒适度(Visual Comfort, VC)是影响立体视频 QoE 的重要因素之一。目前对立体视频/图像的评价主要考虑内容失真对图像质量的影响,而较少考虑立体视频内容制作等因素导致的视觉不舒适度。因此,建立有效的立体视频/图像的视觉舒适度客观评价模型对3维内

容制作和后期处理具有十分重要的意义。

在观看立体视频时,过大的双目视差,双眼不对称,以及立体失真等因素都会导致视觉疲劳,引起视觉疲劳的根本原因是由于立体视频本身的制作原理导致观看者双眼焦点调节和辐辏的严重冲突^[1,2]。一般来说,立体图像的视差是决定人眼感知深度大小的重要因素,视差过大使得立体视频/图像的深度感增强,而当感知深度大小在大脑的 Panum 融合区之外时,就会导致双眼调节和辐辏的冲突,进而产生眼部甚至身体上的一系列疲劳症状。主观实验表明,随着立体视频/图像视差的增大,观看时的视觉舒适度会降低^[3-5]。同时,图像的空间频率也影响着人眼的深度感知和双目融合极限^[6]。Schor 等人^[7]利用高斯差分函数(Difference of Gauss, DoG)

2013-07-01 收到, 2013-11-15 改回

国家自然科学基金(61271021), 宁波市自然科学基金(2012A610039)

和宁波大学研究生科研创新基金资助课题

*通信作者: 邵枫 shaofeng@nbn.edu.cn

处理得到不同空间频率的立体图像,探究了图像的空间频率对双目融合极限的影响,图像的空间频率越低,双目的融合极限就越高。对于具有较大视差的立体图像,空间频率越低,观看的视觉舒适度越高^[8]。应当注意到,不同立体图像具有不同的空间频率特征,导致视觉舒适程度也不一致。

传统的视觉舒适度客观评价方法主要基于全局的视差统计特性预测视觉舒适度,例如整幅视差图像的均值、方差、分布范围等特征^[9-11]。然而,根据人眼立体视觉注意力特性,人眼只对部分视觉重要区域的视觉舒适/不舒适比较敏感,如果以此全局的视差统计特征来预测视觉重要区域的视觉舒适程度,会导致无法精确预测得到客观评价价值。文献[12]在传统舒适度评价算法的基础上,提出一种基于相对视差和前景对象尺寸的立体图像视觉舒适度客观评价方法。该方法首先利用 mean-shift 分割法对视差图进行分割,有效地提取出对象的轮廓,然后计算到邻近对象间的相对视差和前景对象的尺寸,并将其作为舒适度感知特征,最后结合传统的全局视差统计特征,可以较大幅度地提高视觉舒适度预测的准确性。然而,该方法并未考虑空间频率特征对人眼深度感知的影响,此外,该方法也未考虑立体视觉显著/重要区域对视觉舒适度评价的影响。因此,如何在评价过程中有效地根据视觉重要区域来提取视觉舒适度特征信息,使得客观评价结果更加符合人类视觉系统,是立体图像视觉舒适度客观评价需要研究解决的问题。

2 基于视觉重要区域的立体图像视觉舒适度评价方法

根据人眼立体视觉注意力特性,人眼只对部分视觉重要区域的视觉舒适/不舒适比较敏感,为了更准确地预测立体图像视觉舒适度,如何在评价过程中有效地根据视觉重要区域来提取出视觉舒适度特征信息十分关键。本文提出了一种基于视觉重要区域(Visual Important Regions, VIR)的立体图像视觉舒适度客观评价模型,其框架图如图1所示。首先,利用平面图像显著图模型和视差信息得到视觉重要区域;然后,提取视觉重要区域的视差幅度特征、视差梯度边缘特征和空间频率特征,并以这些特征的组合作为反映立体图像舒适度的感知特征信息;最后,通过支持向量回归(Support Vector Regression, SVR)预测得到客观评价价值,避免了对人类视觉系统的相关特征与机理的复杂模拟过程。由于获得的立体图像的特征信息具有较强的稳定性,且能较好地反映图像的视觉舒适度变化情况,有效提高了客观评价情况与主观感知的相关性。

2.1 视觉重要区域提取

现有的研究表明,立体图像的视差特征和空间

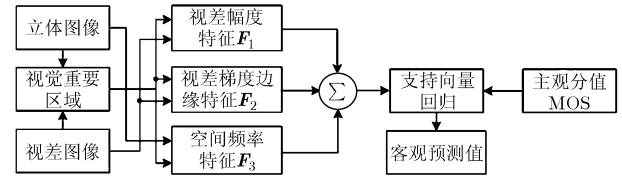


图1 立体图像视觉舒适度客观评价模型框图

频率都会影响视觉舒适度。立体图像视觉重要区域由两部分构成:2维图像的视觉显著图和深度显著图。图像显著图利用人眼视觉注意力模型计算图像的显著度,其代表性算法有 Itti 算法^[13], Hou 算法^[14], GBVS 算法^[15]等。本文采用 GBVS 算法计算得到右视点图像 $I_R(x,y)$ 的视觉显著图,记为 $SM_R(x,y)$,并采用线性方式将视觉显著图 $SM_R(x,y)$ 和右视差图像 $d_R(x,y)$ 进行结合,得到立体显著图,记为 $C(x,y)$ 。

$$C(x,y) = \omega_1 d_R(x,y) + \omega_2 SM_R(x,y) \quad (1)$$

其中 $\omega_1 + \omega_2 = 1$, 在本文中取 $\omega_1 + \omega_2 = 0.5$ ^[16]。

为了检测立体显著图的 VIR,采用阈值法提取出掩模图像,VIR 掩膜图像 $M(x,y)$ 定义为

$$M(x,y) = \begin{cases} 1, & C(x,y) > T \\ 0, & \text{其它} \end{cases} \quad (2)$$

其中 T 为通过最大类间方差法确立的阈值,对于立体显著图 $C(x,y)$ 中的任一像素点 (x,y) ,如果该像素点的显著值大于该阈值,则该像素点属于视觉重要区域,反之则属于非视觉重要区域。图2(a)~图2(e)分别给出了立体图像数据库中的3组立体图像的右视点图像、右视差图像、2维图像显著图、立体显著图和 VIR 掩膜图像,从图中可以看出,本方法提取的 VIR 掩膜图像能够很好地反映人类视觉的感知重要性情况。

2.2 视差幅度特征提取

一般来说,观看立体视频导致双眼的焦点调节和辐辏的冲突是产生视觉疲劳的主要原因。当人眼观看现实世界的场景,人眼的焦点调节和辐辏机制是一致的,而观看立体视频时,若左右视点的视差大小在融合区域内,尽管经过大脑融合仍可形成清晰的单视图像,但此时人眼的焦点调节处于屏幕上,而辐辏却处于因人脑融合而产生的“跃出”屏幕或“陷入”屏幕的虚拟立体图像上,造成焦点调节和辐辏的不一致,与人类长久形成的视觉生理特性相违背。为了减小这种因焦点调节和辐辏的冲突而引起的视觉疲劳,应该将立体图像感兴趣区域的视差大小控制在 Percival 舒适区域(Percival's Zone of Comfort, PZC)范围内^[17],因此本文提取视觉重要区域的视差均值 μ , 视差方差 δ , 最大交叉视差 v 和视

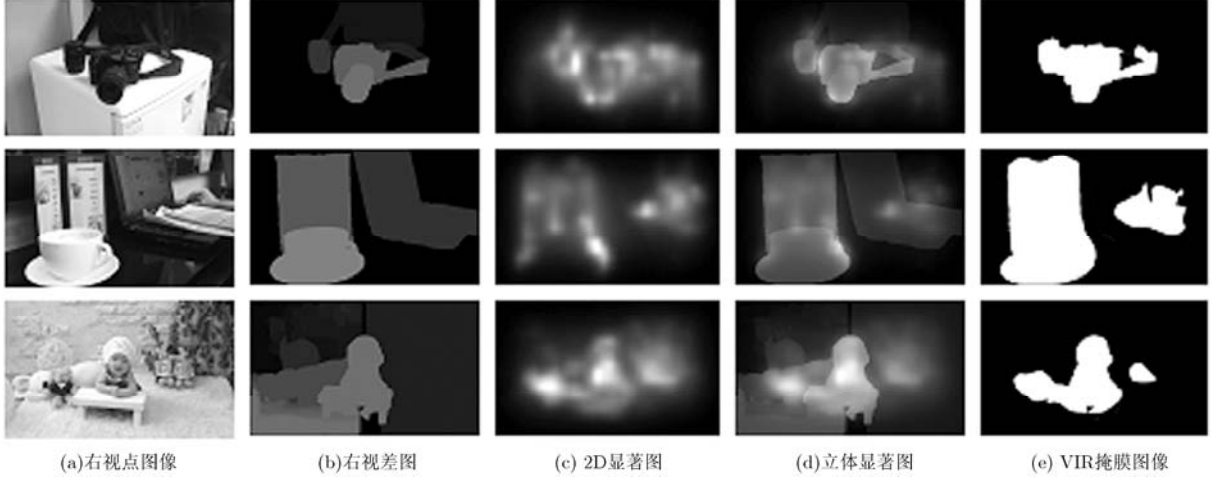


图2 立体图像视觉重要区域分割结果

差范围 τ 共 4 个特征，作为立体图像视差幅度特征矢量，记为 \mathbf{F}_1 , $\mathbf{F}_1 = [\mu, \delta, v, \tau]$ ，各特征的计算公式如下：

(1) VIR 视差均值 μ ：

$$\mu = \sum_{(x,y) \in \Omega} d_R(x,y) \times M(x,y) / \sum_{(x,y) \in \Omega} M(x,y) \quad (3)$$

(2) VIR 视差方差 δ ：

$$\delta = \sum_{(x,y) \in \Omega} (d_R(x,y) - \mu)^2 \times M(x,y) / \sum_{(x,y) \in \Omega} M(x,y) \quad (4)$$

(3) VIR 最大交叉视差 v : $v = d_{\min}$ ，其中， d_{\min} 表示视差图像 $d_R(x,y)$ 的视觉重要区域中视差值最小的 5% 像素点的视差均值。在此，交叉视差指负视差(即在观看时“跃出”屏幕的那部分内容)，较严重的负视差容易导致视觉疲劳。

(4) VIR 视差范围 τ : $\tau = d_{\max} - d_{\min}$ ，其中， d_{\max} 表示视差图像 $d_R(x,y)$ 的视觉重要区域中视差值最大的 5% 像素点的视差均值。

2.3 视差梯度边缘特征提取

基于视觉重要区域的视差边缘特征具有良好的稳定性，且能很好地反映立体图像视觉舒适度。本文通过计算右视差图像 $d_R(x,y)$ 的视差梯度幅值图像 $m(x,y)$ 和梯度方向图像 $\theta(x,y)$ ，利用空间域高斯函数和梯度方向距离的高斯函数对视差梯度幅值进行加权，得到视差梯度边缘图像 $E(x,y)$ ，最终利用 VIR 掩膜加权得到视觉重要区域的视差梯度边缘特征 ξ ，作为立体图像视差梯度边缘特征记为 \mathbf{F}_2 ，该特征的具体计算公式如式(5)~式(8)：

(1) 右视差图像 $d_R(x,y)$ 的视差梯度幅值图像 $m(x,y)$ ：

$$m(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)} \quad (5)$$

其中 $G_x(x,y)$ 表示水平梯度值， $G_y(x,y)$ 表示垂直梯度值。

(2) 右视差图像 $d_R(x,y)$ 的视差梯度方向图像 $\theta(x,y)$ ：

$$\theta(x,y) = \arctan(G_x(x,y), G_y(x,y)) \quad (6)$$

(3) 右视差图像 $d_R(x,y)$ 的梯度边缘图像 $E(x,y)$ ：

$$E(\mathbf{p}) = \sum_{\mathbf{q} \in N(\mathbf{p})} G_s(\|\mathbf{p} - \mathbf{q}\|) G_o(\|\theta(\mathbf{p}) - \theta(\mathbf{q})\|) m(\mathbf{q}) \quad (7)$$

其中 \mathbf{p} 代表当前像素点坐标位置 (x,y) , $N(\mathbf{p})$ 表示以坐标位置为 \mathbf{p} 的像素点为中心的 3×3 邻域窗口，

$$G_s(\|\mathbf{p} - \mathbf{q}\|) = \exp\left[-\frac{\|\mathbf{p} - \mathbf{q}\|^2}{2\sigma_s^2}\right], \quad G_o(\|\theta(\mathbf{p}) - \theta(\mathbf{q})\|) =$$

$$\exp\left[-\frac{\|\theta(\mathbf{p}) - \theta(\mathbf{q})\|^2}{2\sigma_o^2}\right], \quad \theta(\mathbf{p}) = [\sin(\theta(\mathbf{p})), \cos(\theta(\mathbf{p}))],$$

$$\theta(\mathbf{q}) = [\sin(\theta(\mathbf{q})), \cos(\theta(\mathbf{q}))], \quad \text{本文取 } \sigma_s^2 = 0.4, \sigma_o^2 = 0.4.$$

(4) 视差梯度边缘图像 $E(x,y)$ 的视觉重要区域的均值，作为视差梯度边缘特征 ξ ：

$$\xi = \sum_{(x,y) \in \Omega} E(x,y) \times M(x,y) / \sum_{(x,y) \in \Omega} M(x,y) \quad (8)$$

2.4 空间频率特征提取

立体图像的空间频率与人眼对视差的融合极成反比，立体图像的低空间频率所致的模糊会增大人眼对焦点调节和辐辏冲突的融合范围。本文通过计算右视点图像 $I_R(x,y)$ 的水平方向频率值 $\text{HF}(x,y)$ ，垂直方向频率值 $\text{VF}(x,y)$ ，对角方向频率值 $\text{DF}(x,y)$ ，得到 $I_R(x,y)$ 的空间频率图像 $\text{SF}(x,y)$ ，然后提取 $\text{SF}(x,y)$ 视觉重要区域的空间频率均值 η ，空间频率方差 ρ ，空间频率范围 ζ 和空间频率敏感因子 λ 共 4 个特征，作为立体图像空间频率特征矢量记为 \mathbf{F}_3 ， $\mathbf{F}_3 = [\eta, \rho, \zeta, \lambda]$ ，该特征矢量具体计算过程如下：

(1) $I_R(x,y)$ 的空间频率图像 $\text{SF}(x,y)$ ：

$$\text{SF}(x,y) = \sqrt{\text{HF}(x,y)^2 + \text{VF}(x,y)^2 + \text{DF}(x,y)^2} \quad (9)$$

其中 $HF(x,y)$, $VF(x,y)$ 和 $DF(x,y)$ 分别为右图像 $I_R(x,y)$ 的水平、垂直和对角方向的频率值^[18]。

(2) VIR 空间频率均值 η :

$$\eta = \frac{\sum_{(x,y) \in \Omega} SF(x,y) \times M(x,y)}{\sum_{(x,y) \in \Omega} M(x,y)} \quad (10)$$

(3) VIR 空间频率方差 ρ :

$$\rho = \frac{\sum_{(x,y) \in \Omega} (SF(x,y) - \eta)^2 \times M(x,y)}{\sum_{(x,y) \in \Omega} M(x,y)} \quad (11)$$

(4) VIR 空间频率范围 ζ : $\zeta = SF_{\max} - SF_{\min}$, 其中, SF_{\max} 表示空间频率图像 $SF(x,y)$ 中视觉重要区域的最大空间频率值, SF_{\min} 表示空间频率图像 $SF(x,y)$ 中视觉重要区域的最小空间频率值。

(5) VIR 空间频率敏感因子 λ : $\lambda = \eta / \mu$ 。

2.5 立体图像视觉舒适度感知特征融合

本文在 VIR 掩膜图像的基础上, 提取基于 VIR 的视差幅度特征矢量 F_1 、视差梯度边缘特征矢量 F_2 和空间频率特征矢量 F_3 , 然后将特征矢量 F_1, F_2, F_3 连接起来形成一个新的特征矢量, 作为待评价立体图像的特征矢量 X , 即有

$$X = [F_1, F_2, F_3] \quad (12)$$

接着, 通过 SVR 训练得到视觉舒适度预测函数, 基于 SVR 的立体图像视觉舒适度感知特征融合过程主要包含以下两个阶段。

2.5.1 训练阶段 (1) 随机选择立体图像集合中 $(9 \times n) / 10$ 对立体图像构成训练样本数据集合, 记为 Ω_q , $\{X_{k'}, MOS_{k'}\} \in \Omega_q$, 其中, n 表示立体图像集合中立体图像的对数, q 表示训练样本数据集合 Ω_q 中包含的立体图像的对数, $X_{k'}$ 表示训练样本数据集合 Ω_q 中的第 k' 对立体图像的特征矢量, $MOS_{k'}$ 表示训练样本数据集合 Ω_q 中的第 k' 对立体图像的平均主观评分值 (Mean Opinion Score, MOS), $1 \leq k' \leq q$ 。

(2) 构造 $X_{k'}$ 的回归函数 $f(X_{k'})$, $f(X_{k'}) = \sum_{l=1}^q w^T k(X_{k'}, X_{l'}) + b$, 其中, w 为权重矢量, w^T 为 w 的转置矩阵, b 为偏置量, $k(X_{k'}, X_{l'})$ 为径向基

核函数, $k(X_{k'}, X_{l'}) = \exp\left(-\frac{\|X_{k'} - X_{l'}\|^2}{\gamma^2}\right)$, $X_{l'}$ 为训练

样本数据集合 Ω_q 中的第 l' 对立体图像的特征矢量, γ 为核参数, 用于反映输入样本值的范围, 本文取 $\gamma = 54$ 。

(3) 利用支持向量回归对训练样本数据集合 Ω_q 进行训练, 使得经过训练后回归的函数值集合与 MOS 集合之间的误差最小, 拟合得到最优的权重矢量 w^{opt} 和最优的偏置量 b^{opt} , 从而构造最优支持向量回归训练模型。

2.5.2 测试阶段 (1) 立体图像数据库中剩下的 $n/10$

对立体图像构成测试样本数据集合, 记为 Ω_t , t 表示测试样本数据集合 Ω_t 中包含的立体图像的对数。根据支持向量回归训练模型, 对测试样本数据集中每一对立体图像进行测试, 预测得到该子集中每一对立体图像的客观视觉舒适度评价价值。对于该组子集中第 i 对立体图像的客观视觉舒适度评价预测值, 将其记为 $Q_i, Q_i = f(X_i)$, $f(X_i) = \sum_{j=1}^t (w^{opt})^T k(X_i, X_j) + b^{opt}$, 其中, X_i 表示该组子集中第 i 对立体图像的特征矢量。

(2) 重复以上过程 N 次, 将得到的 N 组客观参数的平均值作为最终该视觉舒适度客观评价模型的评价性能, 本文取 $N=200$ 。

3 实验结果与分析

为验证该客观评价模型的性能, 采用韩国先进科学学院提供的立体图像库作为立体图像样本数据集合^[9]。该立体图像库由 120 幅分辨率为 1920×1080 的立体图像组成, 该数据库包含了各种室内场景和室外场景图像, 其右图像如图 3 所示。同时, 该数据库提供了每幅立体图像的 MOS, MOS 值越大表示立体图像的视觉舒适度越高。该立体图像库的主观测试按照 ITU-R BT.1438 规定的标准进行, 实验中采用 Redrover SDM-400 3 维显示器, 主观实验人数为 18 人, 观看者的观测距离为显示器高度的 3 倍, 测试环境均符合 ITU-R BT.500-11 标准。实验中要求观察者根据自身的生理症状或者疲劳反应 (例如眼睛疲劳, 全身不适, 头痛, 聚焦困难, 恶心等症状) 的严重程度, 对立体图像按照 5 个等级 (5 表示非常舒适, 4 表示舒适, 3 表示一般舒适, 2 表示比较不舒适, 1 表示非常不舒适) 进行主观打分, 然后通过统计分析得到每幅立体图像的 MOS 值。此外, 该测试库采用深度估计软件计算得到右视差图像, 视差分布范围为 $0 \sim 285$ 个像素点。

本文采取 10 折交叉验证方法来评估立体图像视觉舒适度感知特性的预测性能。所谓 10 折交叉验证, 就是随机地将立体图像样本数据集合等分成 10 个互不相交的子集, 将其中的 9 份立体图像子集作为训练样本集, 构造支持向量回归训练模型, 剩余的 1 份立体图像子集作为测试样本集, 根据训练模型, 预测测试样本集中所有立体图像的视觉舒适度客观评价分值 Q 。由于在主观质量评价实验中会引入一系列非线性因素, 为避免非线性因素对客观评价模型性能的影响, 需要将模型的客观评价分值 Q 做非线性拟合, 本文采用 5 参数 Logistic 函数进行非线性拟合, 从而预测得到客观模型预测值 MOS_p 。



图3 立体图像数据库右视点图像

$$MOS_p = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(Q - \beta_3))} \right) + \beta_4 Q + \beta_5 \quad (13)$$

其中, 参数 $\beta_1, \beta_2, \beta_3, \beta_4$ 和 β_5 通过训练得到。

然后计算得到客观预测值 MOS_p 与平均主观评分值 MOS 的相关性来度量客观模型的性能, 在本文中采用 Pearson 相关系数 (Pearson Linear Correlation Coefficient, PLCC), Spearman 相关系数 (Spearman Rank-order Correlation Coefficient, SRCC), Kendall 相关系数 (Kendall Rank-order Correlation Coefficient, KRCC) 和均方根误差 (Root Mean Squared Error, RMSE) 4 组客观参量, 其中, PLCC, RMSE 反映客观评价模型预测值的准确性, SRCC 和 KRCC 反映客观评价模型预测值的单调性和一致性; PLCC, SRCC, KRCC 越大, RMSE 越小说明客观评价预测值与平均主观评分值的相关性越好。

3.1 客观性能评价

表 1 给出了本文方法与文献[9]和文献[12]方法在相同测试库上的各项性能指标, 由表中数据可以看出, 本文方法的 PLCC 在 0.87 以上, SRCC 在 0.83 以上, 均方根误差 RMSE 接近 0.39, 均优于文献[9]和文献[12]的各评价指标。此外, 表 1 还给出了采用本文所提取的视差幅度特征矢量 F_1 、视差梯度边缘特征矢量 F_2 、空间频率特征矢量 F_3 不同特征组合的评价性能指标。由表中数据可以看出, 只采用两个特征矢量组合得到的评价预测值与主观评分值间的相关性并不是最佳的, 并且视差幅度特征矢量 F_1 对评价性能的影响比视差梯度边缘特征矢量

表1 不同方法视觉舒适度客观评价模型评价性能指标比较

特征	PLCC	SRCC	KRCC	RMSE
$X = [F_1, F_2]$	0.8354	0.7969	0.6091	0.4418
$X = [F_1, F_3]$	0.8560	0.8057	0.6132	0.4132
$X = [F_2, F_3]$	0.7543	0.7093	0.5231	0.5283
文献[9]	0.7661	0.7834	0.5956	0.5172
文献[12]	0.8464	0.8178	0.6251	0.4285
本文方法	0.8736	0.8346	0.6472	0.3914

F_2 和空间频率特征矢量 F_3 更大, 而结合 F_1, F_2 和 F_3 得到的评价性能最佳, 说明本文的特征描述方法有效。

3.2 各特征对性能的影响

图 4(a), 图 4(b)和图 4(c)分别给出了只采用两种特征矢量组合得到的客观评价与 MOS 值的散点图, 从图中可以看出, 只采用视差梯度边缘特征矢量 F_2 与空间频率特征矢量 F_3 组合得到的散点图最分散, 与主观评价值之间的吻合度最低, 而采用视差幅度特征矢量 F_1 与视差梯度边缘特征矢量 F_2 组合或者视差幅度特征矢量 F_1 与空间频率特征矢量 F_3 组合得到的散点图较为集中, 说明视差幅度特征矢量 F_1 对评价性能的影响最大。图 4(d), 图 4(e)和图 4(f)分别给出了文献[9]和文献[12]以及本文方法的散点图, 从图中可以看到, 本文评价方法的客观评价值 Q 与 MOS 值的散点图较为集中, 更加符合人眼的主观感知。

4 结束语

本文根据人眼的视觉注意力机制, 提出了一种

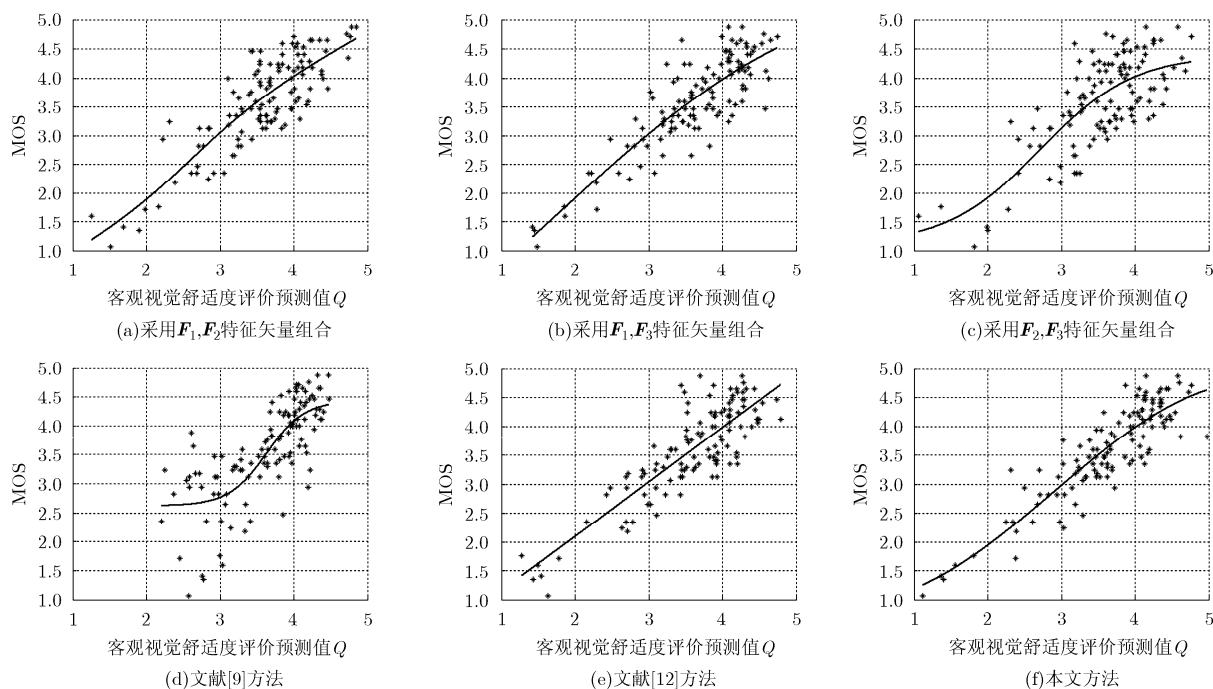


图4 采用不同方法得到的客观视觉舒适度评价预测值与MOS值的散点图

基于视觉重要区域的立体图像视觉舒适度客观评价方法。该方法通过利用图像显著图模型和视差信息得到视觉重要区域，然后提取视觉重要区域的视差幅度特征，视差梯度边缘特征和空间频率特征，作为反映立体图像舒适度的感知特征信息，并通过支持向量回归预测得到客观评价价值。实验结果表明，该方法能够精确地预测人眼对立体图像的视觉舒适程度。接下来，我们将在本文的基础上，考虑建立更为精确的立体显著图模型，从而进一步提高预测的准确性。

参考文献

- [1] Lambooi M, Ijsselsteijn W A, and Heynderickx I. Visual discomfort and visual fatigue of stereoscopic displays : a review[J]. *Journal of Imaging Science and Technology*, 2009, 53(3): 030201-1-030201-14.
- [2] Kim D, Choi S, Sohn K, *et al.* Effect of vergence-accommodation conflict and parallax difference on binocular fusion for random dot stereogram[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(5): 811-816.
- [3] Jung Y J, Sohn H, and Ro Y M. Visual discomfort visualizer using stereo vision and time-of-flight depth cameras[J]. *IEEE Transactions on Consumer Electronics*, 2012, 58(2): 246-254.
- [4] Jung Y J, Kim D, Sohn H, *et al.* Subjective and objective measurements of visual fatigue induced by excessive disparities in stereoscopic images[J]. *SPIE*, 8648, 2013: DOI: 10.1117/12.2005818.
- [5] Tam W J, Speranza F, Yano S, *et al.* Stereoscopic 3D-TV: visual comfort[J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 335-346.
- [6] Patterson R. Human factors of stereo displays: an update[J]. *Journal of the Society for Information Display*, 2009, 17(12): 987-996.
- [7] Schor C, Wood I, and Ogawa J. Binocular sensory fusion is limited by spatial resolution[J]. *Vision Research*, 1984, 24(7): 661-665.
- [8] 王勤, 王琼华, 刘春玲. 视差与空间频率对自由立体显示器观看舒适度的影响[J]. *光电子·激光*, 2012, 23(8): 1604-1608.
Wang Qin, Wang Qiong-hua, and Liu Chun-ling. Effects of parallax and spatial frequency on visual comfort in autostereoscopic display[J]. *Journal of Optoelectronics*, 2012, 23(8): 1604-1608.
- [9] Lambooi M, Ijsselsteijn W A, and Heynderickx I. Visual discomfort of 3-D TV: assessment methods and modeling[J]. *Displays*, 2011, 32(4): 209-218.
- [10] Choi J H, Yun M J, Kim A R, *et al.* Visual comfort measurement for 2D/3D converted stereo video sequence[C]. 2012 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON 2012), Zurich, Switzerland, 2012: 1-4.
- [11] Kim D and Sohn K. Visual fatigue prediction for stereoscopic image[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(2): 231-236.
- [12] Sohn H, Jung Y J, Lee S, *et al.* Predicting visual discomfort using object size and disparity information in stereoscopic

- images[J]. *IEEE Transactions on Broadcasting*, 2013, 59(1): 28-37.
- [13] Itti L, Koch C, and 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.
- [14] Hou X D and Zhang L. Saliency detection: a spectral residual approach[C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Minnesota, USA, 2007: 1-8.
- [15] Harel J, Koch C, and Perona P. Graph-based visual saliency[J]. *Advances in Neural Information Processing Systems*, 2007, 19: 545-552.
- [16] Wang J L, Silva M, Callet P L, *et al.* A computational model of stereoscopic 3D visual saliency[J]. *IEEE Transactions on Image Processing*, 2013, 22(6): 2151-2165.
- [17] Percival A. The zone of comfort: predicting visual discomfort with stereo displays[J]. *Journal of Vision*, 2011, DOI: 10.1167/11.8.11.
- [18] 马恒, 贾传炎, 刘爽. 基于空间频率的图像融合方法[J]. *计算机工程与应用*, 2005, 41(31): 215-218.
- Ma Heng, Jia Chuan-ying, and Liu Shuang. Image fusion based on spatial frequency[J]. *Computer Engineering and Applications*, 2005, 41(31): 215-218.
- [19] Sohn H, Jung J Y, Lee S, *et al.* Korea Advanced Institute of Science and Technology. IVY Lab Stereoscopic Image Database[OL]. <http://ivylib.kaist.ac.kr/demo/3DVCA/3DVCA.htm>, 2013.
- 姜求平: 男, 1991年生, 硕士生, 立体图像质量及视觉舒适度评价、图像处理.
- 邵枫: 男, 1980年生, 副教授, 研究方向为多媒体信息通信、图像处理与视频编码.
- 蒋刚毅: 男, 1964年生, 教授, 研究方向为多媒体信息通信、图像处理与视频编码.
- 郁梅: 女, 1968年生, 教授, 研究方向为多媒体信息通信、图像处理与视频编码.