

文章编号 1004-924X(2024)04-0595-14

虚拟现实视觉诱发晕动症时空多特征评价

董奇峰¹, 郁梅^{1*}, 蒋志迪², 鲁子昂¹, 蒋刚毅¹

(1. 宁波大学 信息科学与工程学院, 浙江 宁波 315211;

2. 宁波大学 科学技术学院 信息工程学院, 浙江 宁波 315212)

摘要:沉浸式的虚拟现实体验中视觉诱发晕动症(Visually Induced Motion Sickness, VIMS)是影响虚拟现实系统发展与应用的一个重要问题。现有的基于视觉内容的评价方案大多考虑的要素不够全面,对运动信息的提取较为简单,且少有考虑时域上的突变对晕动症的影响。针对上述问题,提出了虚拟现实中视觉诱发晕动症时空多特征评价模型;基于立体全景视频中空域、时域信息来设计视觉诱发晕动症评价模型,采用更符合人眼感知的加权运动特征,并考虑了立体全景视频时域的突变信息设计了特征提取方式。所提出模型分为预处理模块、特征提取模块及时域聚合与回归模块。预处理模块用于视口提取和光流图、视差图、显著图的估计。特征提取模块包含前背景加权运动特征提取、基于变换域的视差特征提取、空间特征提取及时域突变特征提取。时域聚合后通过支持向量回归得到 VIMS 评价分数。实验结果表明,该模型在立体全景视频数据库 SPVCD 上的预测结果与平均主观意见分的皮尔逊线性相关系数为 0.821、斯皮尔曼相关系数为 0.790、均方根误差为 0.489。该模型取得了优良的预测性能,验证了所提出特征提取模块的有效性。

关键词:虚拟现实;立体全景视频;晕动症;运动感知;突变特征

中图分类号:TP391.41 **文献标识码:**A **doi:**10.37188/OPE.20243204.0595

Spatiotemporal multi-feature evaluation of visually induced motion sickness in virtual reality

DONG Qifeng¹, YU Mei^{1*}, JIANG Zhidi², LU Ziang¹, JIANG Gangyi¹

(1. Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China;

2. College of Information Engineering, College of Science and Technology Ningbo University,
Ningbo 315212, China)

* Corresponding author, E-mail: yumei@nbu.edu.cn

Abstract: Visually induced motion sickness (VIMS) in immersive virtual reality experience is an important problem that impedes the development and applications of virtual reality systems. Most of the existing assessment methods based on visual content are not comprehensive enough, and the extracted features of motion information are relatively simple, and the influence of abrupt changes in time domain on motion sickness is rarely considered. To solve these problems, a spatio-temporal multi-feature assessment model for VIMS in virtual reality was proposed. The spatial and temporal information of stereoscopic panoramic video was used to design the assessment model of VIMS. The weighted motion features more in line with human perception were adopted, and the feature extraction method was designed considering the temporal

收稿日期:2023-09-02;修订日期:2023-10-17.

基金项目:浙江省自然科学基金项目(No. LY21F010003);国家自然科学基金项目(No. 61871247)

mutation information of stereoscopic panoramic video. The proposed model was divided into preprocessing module, feature extraction module and time domain aggregation and regression module. The preprocessing module was used to extract the viewport images and estimate the optical flow map, disparity map and saliency map. The feature extraction module included foreground-background weighted motion feature extraction, disparity feature extraction based on transform domain, spatial feature extraction and time domain abrupt change feature extraction. VIMS evaluation scores were finally obtained through time domain aggregation and support vector regression. The experimental results show that the PLCC, SROCC and RMSE of the proposed model are 0.821, 0.790 and 0.489, respectively when tested on the stereoscopic panoramic video database SPVCD. The model achieves excellent prediction performance which verifies the effectiveness of the proposed feature extraction module.

Key words: virtual reality; stereoscopic panoramic video; motion sickness; motion perception; abrupt change feature

1 引 言

虚拟现实(Virtual Reality, VR)技术作为数字世界元宇宙的重要基础,在教育、专业培训、医疗和娱乐等领域有着广泛应用^[1]。然而,阻碍VR进一步发展的一个重要因素是沉浸式VR体验可能对用户造成视觉诱发晕动症(Visually Induced Motion Sickness, VIMS)^[2],也被称为网络病(Cybersickness)、模拟器病(Simulator Sickness)等^[3]。其主要表现为视觉不适、迷失方向甚至恶心呕吐等^[4]。当用户沉浸式观看VR时,其视觉感知到的是虚拟运动而非物理运动,前庭系统无法感知沉浸式显示器视觉诱导的自我运动^[5],这种视觉和前庭感官的不匹配诱发了晕动症。让用户预知所要观看的VR内容潜在的晕动程度有助于避免晕动症的发生、保障用户的健康。因此,评价VR内容可能诱发的晕动症等级是一个亟待解决课题。

目前,主流的VIMS评价方式可分为基于主观问卷、生理信号、视觉内容及三者之间结合的方式。主观问卷方式,如模拟器疾病问卷(Simulator Sickness Questionnaire, SSQ)^[6]等,耗时长且受限于VR内容数量,难以实时应用。基于生理信号的评价方式依赖于佩戴式仪器测量结果,如脑电图、皮肤电反应、胃电图、心率等^[7],对受试者易产生除视觉内容外的不舒适感。因此,建立有效的不依赖于主观或生理信号的客观VIMS评价模型对VR内容生成与体验质量改善有十分重要意义。

基于视觉内容的VIMS评价研究尚不成熟,文献[8]建立了一个包括36个虚拟场景的VR病数据库,利用交互模型提出了一个VR疾病预测器框架。文献[9]建立了名为网络病参考的数据库,包含两个参考场景并考虑多因素创建了52个不同的VR场景;提出了一种两阶段网络架构实现神经表征和时空表征。文献[10]为了优化VR病与VR沉浸感,建立了包含10个参考虚拟现实场景的VR-SP数据库,并通过两种运动类型和四种水平速度将参考场景拓展10种变化。同时,设计时空旋转帧差图的统计模型预测VIMS,通过视觉活动、内容特征捕捉VR内容沉浸感。文献[4]建立了包含52个不同内容属性的VR场景,统计各项因素对晕动程度的重要程度。文献[11]构建了10种类型两种帧速率的VR疾病数据库,并提出了一种新的生理融合深度网络,通过内容刺激和生理反应来估计个体VR疾病。文献[12]提出的SACA Net分别从VR内容和人体生理信号中提取症状特征以预测晕动症。文献[13]重点考虑VR内容中的异常运动,建立的全景视频测试数据库包含3种运动模式,9个场景;所提网络模型将原始视频和生成视频之间的差异投影至主观评分空间来预测VIMS。文献[14]针对视频拍摄中相机抖动引起的VR疾病,建立了包含20个不同程度相机抖动的全景视频数据库,将VR视频场景方向和速度的变化作为特征预测VR疾病。文献[15]侧重于预测VR疾病的主要症状的水平,而非VR疾病的总体程度;建立了包含20个基准视频,4种不同帧速率,共80个

视频的全景视频数据库。文献[16]设计了内容刺激引导器以模仿人类感受VR疾病。文献[17]考虑人类视觉来学习信息的特征,并模拟VR内容被接受时的学习过程,通过虚拟内容学习重构器和疾病评分预测器预测分数。上述工作所涉及的虚拟现实VIMS评价的研究均为基于非立体全景视频内容,因而未考虑人眼立体感知特性。

文献[18]建立了一个包含19个场景的立体全景视频内容及其相应的晕动程度评级的数据集,手工提取特征训练机器学习算法。文献[19]建立了目前包含基准场景最多的立体全景视频数据库,共包含116个不同的立体全景视频内容。作者采用深度学习的方式考虑人类视觉系统的注意力机制和视网膜的多通道特性,对虚拟环境中的感官冲突以及双目立体感知的三阶段过程进行仿真建模;并在时间池化模块利用LSTM(Long Short-Term Memory)模拟人眼的时间记忆效应。然而,基于视觉内容的评价方案大多考虑的要素不够全面,对运动信息的提取较为简单,同时少有考虑时域上的突变对晕动症的影响。

针对上述问题,本文从人眼感知特性和视频时空特征的角度考虑,提出了一种基于时空多特征的虚拟现实视觉诱发晕动症评价模型。实

验结果表明,该模型在立体全景视频数据库SPVCD上取得了优良的预测性能,预测结果与人眼所感知的晕动程度有较好的一致性。

2 提出的方法

本文从人眼感知特性和视频时空特征的角度考虑,提出了一种基于时空多特征的虚拟现实视觉诱发晕动症评价模型,其框架如图1所示。该模型分为预处理模块、特征提取模块及时域聚合与回归模块。预处理模块利用左、右视点的中心视口 I_L, I_R 求取光流图 F 、显著图 S 和视差图 D 用于特征提取。运动特征提取模块用于前背景加权运动特征 M 和加速度特征 U 的提取;视差特征提取模块用于视差强度特征 E 和视差梯度能量特征 G 提取;空域感知特征提取模块包含一阶色度亮度特征 H_1, V_1 ,二阶色度亮度特征 H_2, V_2 及空间感知信息特征 SI 的提取。时域突变特征提取模块分别对所提取的前背景加权运动特征 M 、视差强度特征 E 、一阶色度亮度特征 H_1, V_1 计算最大运动突变特征 $C_{\max, M}$ 、运动突变特征和 $C_{\text{sum}, M}$ 、最大视差强度突变特征 $C_{\max, E}$ 、视差强度突变特征和 $C_{\text{sum}, E}$ 、最大色度、亮度突变特征 $C_{\max, H}$, $C_{\max, V}$,色度、亮度突变特征和 $C_{\text{sum}, H}$, $C_{\text{sum}, V}$ 。最终,将提取的帧级特征进行时域聚合并对其进行突变

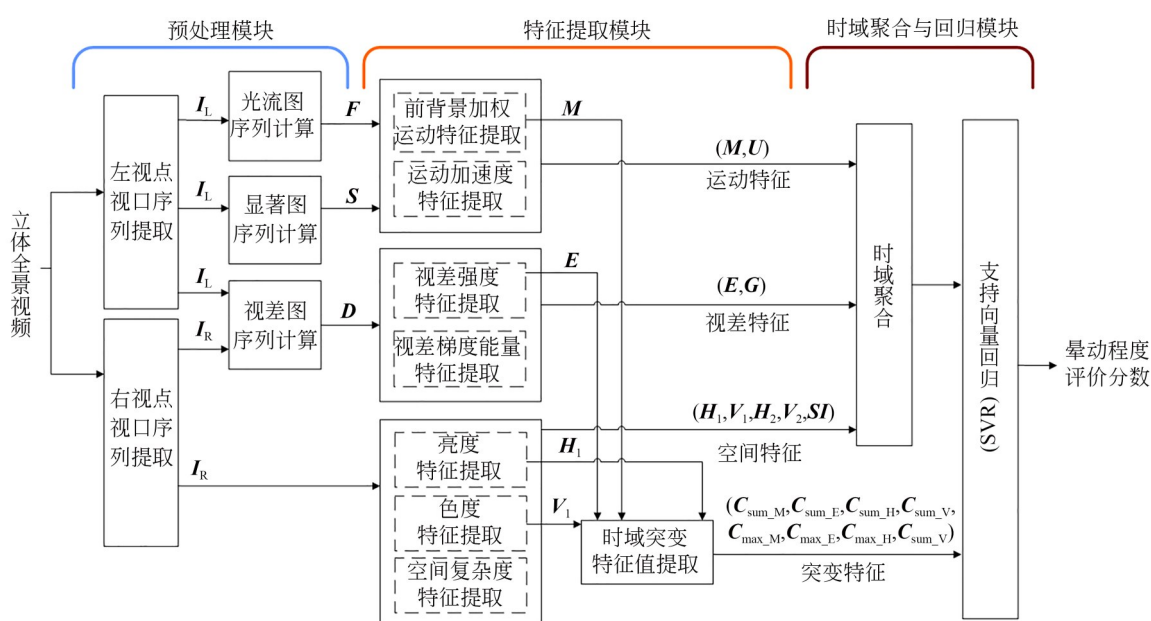


图1 视觉诱发晕动症评价模型

Fig. 1 Assessment model of visually induced motion sickness

特征采用支持向量回归(Support Vector Regression, SVR)得到 VIMS 评价价值。

2.1 预处理

在立体全景视频数据库 SPVCD 的主观评分实验中,为避免头部旋转而非视频内容引起晕动,受试者只对每个视频的前中央视口进行评分。因此,本文将原始的立体全景视频序列投影为立方体投影格式,并取正面投影面作为中央左视口、右视口。对获取的左、右视口 I_L, I_R 通过文献[19]中的方法计算得到显著图 S ,并通过文献[20]中的方法计算得到视差图 D 与光流图 F ,用于后续的特征提取。

2.2 运动特征

与真实物理运动不同,用户沉浸式观看 VR 时所体验到的运动为虚拟运动,而非物理运动;因此,用户的前庭系统无法感知由沉浸式显示器视觉诱导的自我运动^[22]。这种视觉和前庭感官的不匹配可能诱发晕动症。虽然感官冲突是一个不能直接测量的参数,但究其根本是由虚拟环境中的视觉刺激(即视觉运动)驱动的^[23]。视觉所感知到的运动强度的大小在一定程度上反映了视觉内容潜在的晕动程度^[18]。

2.2.1 前背景加权运动特征提取

不同的视觉信息对人眼感知具有不同贡献度^[24]。通常在我们的认知中认为只有视觉关注区域中物体的剧烈运动会对观看者造成晕动,其实不然,用户感知到的视觉诱导的自我运动主要是由周边视场中看到的视觉运动引起^[5]。在所能观看到的视场大小一定时,周边视场的运动特性将对视觉感知产生更大的影响^[25]。因此,本文采用背景和前景目标运动强度的加权和来描述动态运动特征。权重由前景目标与背景之间的运动强度和运动复杂度决定,以此反映运动前景目标和场景变化对晕动程度的影响:

$$M_k = \omega_B M_{kB} + \omega_O M_{kO}, \quad (1)$$

$$M_{kB} = F_k \times (1 - S_{bk}), \quad (2)$$

$$M_{kO} = F_k \times S_{bk}, \quad (3)$$

其中: M_k, M_{kB} 和 M_{kO} 分别为第 k 帧的加权运动强度、背景运动强度和前景目标运动强度。 M_{kB} 和 M_{kO} 分别通过二值化后的显著图 S_{bk} 和非显著图 $1-S_{bk}$ 加权光流图 F_k 得到, ω_B 和 ω_O 分别为背景运动向量强度和前景目标运动向量强度的权重,满

足条件: ω_B 和 $\omega_O \in [0, 1], \omega_B + \omega_O = 1$ 。

$$\omega_B = \omega_1 \times \omega_2, \quad (4)$$

$$\omega_1 = \frac{N_{kB}}{N_k}, \quad (5)$$

$$\omega_2 = \frac{\sum_{i=1}^{N_{kB}} M_{kB}(i)}{\sum_{i=1}^N M_{kO}(i) + M_{kB}(i)}, \quad (6)$$

其中: N_{kB}, N_k 分别为第 k 帧背景部分和视口中运动强度非零的像素数, ω_1 表示运动密度,反映运动的空间分布。该值越大,运动分布越广,越关心背景中所发生的运动。 ω_2 表示背景运动强度,反映背景运动能量。该值越大,背景所包含的运动能量越高。

作为实例,图 2 给出了前背景加权运动特征提取过程,图 2 为 SPVCD 中两个不同运动类型场景其中一帧的中心视口图,上图为序列 47 (MOS: 4.09) 的第 53 帧中心视口,下图为序列 18 (MOS: 1.95) 的第 296 帧视口,图 2(b)~图 2(e) 分别为两视口所对应光流图、显著图、背景运动强度图、目标运动强度图。其中,序列 47 第 53 帧的背景运动强度权重 ω_B 为 0.28,前景目标运动强度权重 ω_O 为 0.72;序列 18 第 296 帧 ω_B 为 0.88, ω_O 为 0.12。采用背景与前景目标运动加权的方式,对于序列 47 此类前景目标运动而背景几乎不运动的视频序列,用户会更关注其前景目标的运动状态。而对于序列 18 此类前景目标与背景都处于运动状态的序列,前景目标会受到关注,同时更需关注背景的运动状态,因为此时背景的运动状态与视觉诱导的自我运动具有更高的关联性,是影响晕动程度的重要因素。

图 3 展示了本文所提出的前背景加权运动特征与 SPVCD 中 70 个(数据库共包含 116 个视频序列)视频序列平均主观得分差异(Differential Mean Opinion Score, DMOS)之间的关系。DMOS 越大代表视频在主观实验中晕动程度越严重,反之,晕动程度越轻。SPVCD 中采用平均主观意见分(Mean Opinion Score, MOS),由于运动强度越大往往所对应视频序列的 MOS 越低,因此,此处采用 DMOS 用于展示本文所提前背景加权运动特征与主观评价之间的一致性。由图 3 可知,本文所提的前背景加权运动特征与 SPVCD 中大部分的视频序列的主观评分存在较高的一致性,对晕动程度预测具有较好的表征能力。

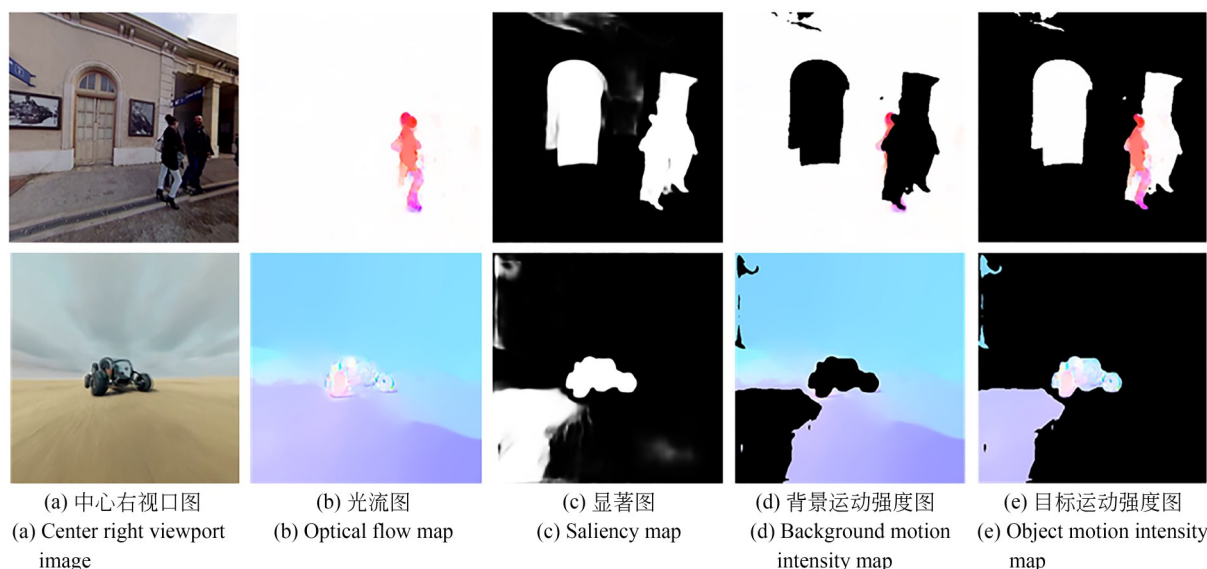


图2 前背景加权运动特征提取

Fig. 2 Foreground-background weighted motion feature extraction

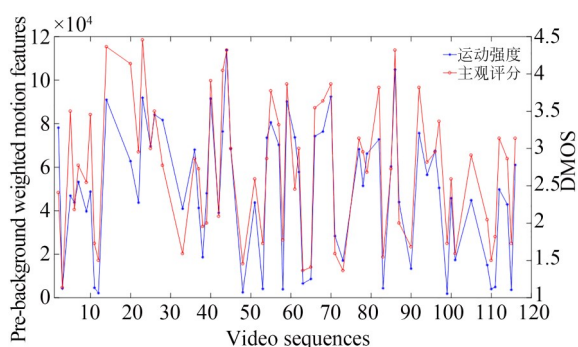


图3 前背景加权运动特征与SPVCD部分视频序列DMOS之间的关系

Fig. 3 Relationship between pre-background weighted motion features and DMOS of partial videos in SPVCD

2.2.2 加速度特征

对于连续帧的视频序列来说,若未发生加速度等异常运动,那么连续帧间的光流变化很小;反之则连续帧之间就会出现较明显的光流变化^[26]。基于该思想,通过相邻帧之间的速度变化差异计算序列的加速度,当视频场景中出现加速或减速时,两帧之间会出现较为明显的速度差别。对由预处理得到的光流图像 F ,计算得到每一帧内像素点对应的水平和垂直位移,并根据水平、垂直位移矩阵和帧间间隔计算每一帧上每一个像素点的速度。通过相邻两帧的速度差表示加速度,得到基于帧间差的加速度特征。根据水

平、垂直位移和帧间隔 n 计算第 k 帧速度 V_k :

$$V_k = \frac{1}{N} \sum_{i=1}^N \sqrt{\left(\frac{\Delta x_{i,k}}{n}\right)^2 + \left(\frac{\Delta y_{i,k}}{n}\right)^2}, \quad (7)$$

$$U_k = \frac{|V_k - V_{k-1}|}{n}, \quad (8)$$

其中: $\Delta x_{i,k}, \Delta y_{i,k}$ 分别为每个像素点的水平与垂直位移量, N 为视口中的总像素数, n 为两帧之间间隔, n 取1。通过计算相邻两帧图像的速度差的绝对值表示第 k 帧的加速度特征 U_k 。

2.3 双目感知特征

立体视觉中的神经机制是双目视差,在观看自然世界时,人眼的焦点调节和会聚是相互协作联合运动的,焦点的调节始终跟随着会聚的变化,通常二者是一致的。然而在观看人工制作的立体视觉内容时,调节和会聚之间就存在产生冲突的可能性。本文从视差图的多尺度二维DCT系数中推导出晕动信息。将输入的视差图划分为4种不同数量的图像块,包括 $1 \times 1, 4 \times 4, 8 \times 8, 16 \times 16$ 。利用多尺度二维离散余弦变换将不同大小的块变换到离散余弦变换域,得到离散余弦变换系数。基于DCT系数的统计特性提取两种类型的多尺度特征。

2.3.1 视差强度特征

双眼视差过大是诱发视觉不适的主要原因之一,比起直接采用整张视差图像素的均值作为视觉不适特征,DCT系数能够更好地表达图像

中的特征信息^[27]。对分块后的视差图进行多尺度二维离散余弦变换,每个块的2D-DCT系数由DC系数和AC系数组成,其中DC系数表示块的平均强度。因此,使用位于DCT块左上角的DC系数代表整块视差强度特征 E 。将所有块对应的DC系数之和作为整个视差图的视差强度特征 E 。因此,第 k 帧中心视口图所对应的视差强度特征 E_k 计算为:

$$E_k^g = \sum_j DC(I_j), \quad (9)$$

其中, g 表示DCT块的尺度级别,对于 $g=1,2,3,4$ 分别对应于块的数量 $M \times N=1 \times 1, 4 \times 4, 8 \times$

$$G_k^g = \frac{1}{J} \sum_{m=1}^M \sum_{n=1}^N (1/4(E(m+1,n) + E(m-1,n) + E(m,n+1) + E(m,n-1)) - E(m,n)), \quad (10)$$

$$E(m,n) = \sum_x \sum_y [A^{mn}(x,y)]_{M \times N}, \quad (11)$$

其中: $E(m,n)$ 表示位于空间位置 (m,n) 的DCT块 A^{mn} 的能量, g 表示DCT块的尺度级别, J 代表视差图中DCT块的数量。

2.4 空域感知特征

2.4.1 亮度、色度特征

人类视觉系统对颜色和亮度的感知非常敏感,在观看视觉内容时颜色分布的不均匀以及过亮或过暗的内容都会在观看过程中产生不舒适的体验感,进而产生晕动。

图像的颜色分布主要集中在低阶矩中,其中,一阶矩反映图像明暗程度、二阶矩描述了颜色整体分布范围。一阶矩是计算所有像素点的均值,二阶矩是计算所有像素点的方差。将RGB颜色空间转化到与人类描述彩色方式更为一致的HSV颜色空间,并在HSV颜色空间分别提取左右视口一、二阶的亮度矩和色度矩。以色度分量为例,公式如下:

$$H_{1,k} = \frac{1}{P} \sum_x \sum_y H_k(x,y), \quad (12)$$

$$H_{2,k} = \sqrt{\frac{1}{P} \sum_{xy} (H_k(x,y) - H_{1,k})^2}, \quad (13)$$

其中: H_k 为第 k 帧HSV颜色空间中的色度分量, $H_{1,k}$ 为第 k 帧的色度一阶矩, $H_{2,k}$ 为第 k 帧色度二阶矩, (x,y) 为色度分量的坐标, P 为视口大小。亮度分量一、二阶矩 $V_{1,k}, V_{2,k}$ 由相同方式求得。

2.4.2 空间感知信息特征

空间感知信息(SI, Spatial-perceptual Infor-

8, 16×16 ; I_j 为分块视差图中第 J 块, $J=M \times N$; $DC(\cdot)$ 表示取DC系数。

2.3.2 视差梯度能量特征

空间方向上的快速视差变化同样容易引起辐辏调节冲突,视差梯度能量特征是为了捕捉压缩域中相邻块之间的局部视差能量变化。首先,使用整个块的DCT系数之和来表示每个视差块的能量,并计算当前DCT块与其上、下、左、右块之间的能量差来获得视差梯度能量。因此,尺度级别为 g 的第 k 帧中心视口图所对应特征 G_k^g 计算为:

mation)量化了视频序列中存在的空间细节的复杂程度,它随着视觉画面内容丰富程度增加而增加。对于虚拟运动强度相似的两个视频序列,更高的晕动程度可能是由于更高的空间感知信息所造成的。空间感知信息通过计算梯度向量得到。首先,利用Sobel算子计算得到梯度向量,对于每个视频帧计算梯度向量幅度的标准差。第 k 帧的空间感知信息计算公式为:

$$SI_k = \text{std}[\text{Sobel}(I_k)]. \quad (14)$$

以图4中两序列为例,图4(a)的上下图分别



(a) 序列3和序列53第1帧中心视口图

(a) Central viewport of the first frame of Video 3 and Video 53

(b) 对应梯度计算结果

(b) Corresponding gradient image

图4 SPVCD中相似序列空间感知信息比较

Fig. 4 Comparison of SI of similar videos in SPVCD

为SPVCD中序列3和序列53的中心视口图,图4(b)的上下图分别为所对应梯度计算结果。

序列3的MOS值为4.86,平均前背景加权运动特征 M 为4.276;序列53的MOS值为4.27, M 为4.084。通过计算,序列3的空间感知信息特征 SI_k 的均值为0.04,而序列53的空间感知信息特征 SI_k 的均值为0.10。当两个视频序列的运动强度相近且均未出现场景切换以及大幅度的视差变化时,空间感知信息成为影响视觉舒适度的因素之一。画面内容丰富的视频序列相比于画面内容简单的视频序列往往更容易在观看时产生晕动的感受,但相比于运动所造成的不舒适,空间感知信息的增大所造成的晕动程度较为有限。

2.5 时域聚合

上述的三类特征均为对待评价的视频序列的每一帧进行提取所得,称为帧级特征。通过时域聚合将帧级特征计算为用以描述视频序列整体晕动程度的特征。对帧级的前背景加权运动特征 M_k 、加速度特征 U_k 、空间感知信息特征 SI_k ,3个帧级特征在时域上进行平均值、最大值、中值、方差聚合,得到对应的12维立体全景视频中心视口序列特征;对4个尺度的帧级视差强度特征 E_k 、视差梯度能量特征 G_k ,在时域上进行平均值聚合得到对应的8维立体全景视频中心视口序列特征;对帧级的一阶亮度、色度特征 $V_{1,k}$, $H_{1,k}$,二阶亮度、色度特征 $V_{2,k}$, $H_{2,k}$,在时域上进行平均值聚合得到相应的8维立体全景视频中心视口序列特征。最终通过时域聚合共获得28维特征作为模型的部分特征用于评价立体全景视频的晕动症。将用于时域聚合的帧级特征记为 F_k ,在时域上实现平均值、最大值、中值、方差聚合的公式如下:

$$F_{\text{avg}} = \frac{1}{K} \sum_{k=1}^K F_k, \quad (15)$$

$$F_{\text{max}} = (F_k)_{\text{max}}, \quad (16)$$

$$F_{\text{mid}} = (F_k)_{\text{median}}, \quad (17)$$

$$F_{\text{var}} = \frac{1}{K} \sum_{k=1}^K (F_k - F_{\text{avg}})^2, \quad (18)$$

其中: F_{avg} , F_{max} , F_{mid} 和 F_{var} 分别为帧级特征时域平均值、最大值、中值和方差聚合结果。 K 为视频序列帧数,SPVCD中序列统一为 $K=400$ 。 $(\cdot)_{\text{max}}$ 为取最大值, $(\cdot)_{\text{median}}$ 为取中值。

2.6 时域突变特征

视觉内容的运动强度一定程度反映了视觉前庭不匹配的程度,在沉浸式的视觉环境中,运动的平滑度对VIMS的影响也很大^[28]。快速、突然的运动变化在运动强度一定的基础上更容易导致视前庭不协调,从而引发晕动症。为了衡量快速、突然的运动变化,并同时延伸考虑视差、亮度、色度的突然变化对观看者可能引起的晕动程度,本文提取时域突变特征用于评价此类因素所带来的VIMS。

对得到的帧级前背景加权运动特征 M_k 、视差强度特征 E_k 、一阶色度、亮度特征 $H_{1,k}$, $V_{1,k}$ 分别进行时域突变特征的提取。首先,对帧级特征在时域上做窗口 W 大小为16的滑动平均处理,以减少在特征图提取过程以及帧级特征计算过程中所带来的误差。滑动平均后的前背景加权运动特征值 $\bar{M}_k = 1/W \sum_{k=k^+}^{k+W} M_k$,并求取帧级特征值的极值(局部最大、最小值)。图5展示了滑动平均处理前后SPVCD序列32的前背景加权运动特征值。

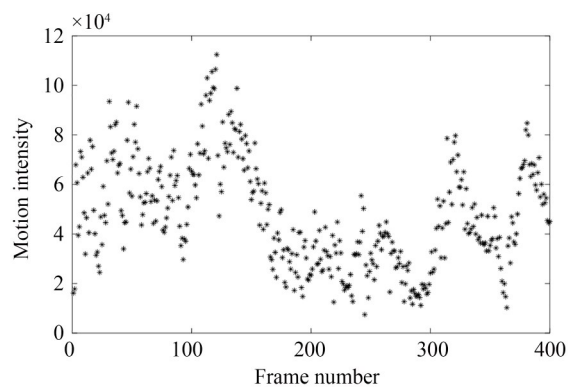
以运动强度突变为例,在一段时间内运动强度变化越大,所造成的晕动程度越大,在统计上表示为相邻的两极值点所在直线的斜率的绝对值越大,所对应运动强度变化越快。简化算法默认在两个极值点间各点之间的斜率值保持不变。在相同斜率变化下,所持续时间越长,所造成的运动程度也会越大,因此通过计算相邻的两极值点所在直线的斜率与两极值点之间的距离的比率表示第 m 段运动突变值:

$$C_{m,M} = \begin{cases} \frac{D_{m,M}}{k_{m,\text{max}} - k_{m,\text{min}}} & d_m > 8 \\ 0 & d_m \leq 8 \end{cases}, \quad (19)$$

$$D_{m,M} = K \sqrt{d_M^2 + d_k^2}, \quad (20)$$

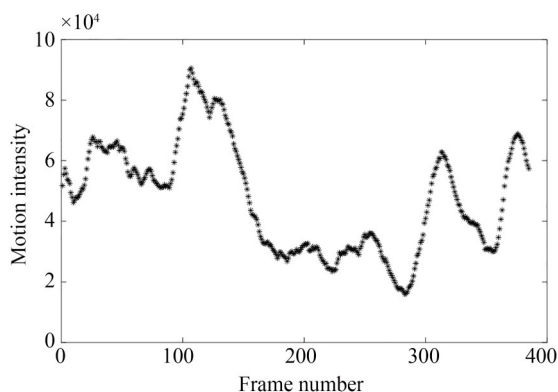
$$K = \left| \frac{M_m^{\text{max}} - M_m^{\text{min}}}{k_m^{\text{max}} - k_m^{\text{min}}} \right|, \quad (21)$$

其中: $D_{m,M}$ 代表线性归一化后第 m 组局部最大运动特征值 $\max(M_m)$ 和最小运动特征值 $\min(M_m)$ 之间的欧氏距离, d_M 与 d_k 用于表示最大、最小特征值之间的强度差以及帧间隔。 K 表示两相邻极大极小运动特征值点之间的斜率的绝对值。 $C_{m,M}$ 为第 m 组运动强度突变值, $k_{m,\text{max}}$, $k_{m,\text{min}}$ 分别表示局部最大运动特征值和最小运动特征值所对



(a) SPVCD中序列32原始运动特征

(a) Original motion features of Video 32 in SPVCD



(b) 序列32滑动平均后的运动特征

(b) Sliding averaging motion feature of Video 32

图5 帧级运动特征值滑动平均处理

Fig. 5 Frame level motion feature value through sliding averaging processing

应的帧数,即两者之差 d_m 表示一次运动强度突变所持续的时间间隔。当一次突变所持续的时间间隔过短时($d_m < 8$)默认对用户所造成的晕动感可以忽略不计。

最终对一个视频中得到的多段运动强度突变值求取总和 C_{sum_M} 以及最大值 C_{max_M} ,作为表示一个视频序列中由于运动突变所造成的晕动程度的运动突变特征值。图6(a)和图6(b)展示了SPVCD中序列32和序列39的前背景加权运动特征值与运动突变特征值结果。序列32的MOS=1.59,属于晕动程度较大的视频序列;序列39的MOS=4.00,属于晕动程度较小的视频序列。通过实验结果可以发现,两个序列的前背景加权运动特征平均值接近,但序列32的运动突变总和($C_{\text{sum}_M}=37.50$)与运动突变最大值($C_{\text{max}_M}=$

6.17)大于序列39的运动突变总和($C_{\text{sum}_M}=22.28$)与运动突变最大值($C_{\text{max}_M}=3.80$)。由此证明对于整体运动强度接近的视频序列,存在更多运动变化的视频序列更容易造成VIMS。

利用上述相同方法求得视频序列的视差强度突变特征值 C_{sum_E} , C_{max_E} ,亮度突变特征值 C_{sum_V} , C_{max_V} ,色度突变特征值 C_{sum_H} , C_{max_H} 。最终,提取通过时域聚合的28维特征以及8维时域突变特征,共36维特征用于立体全景视频VIMS评价。

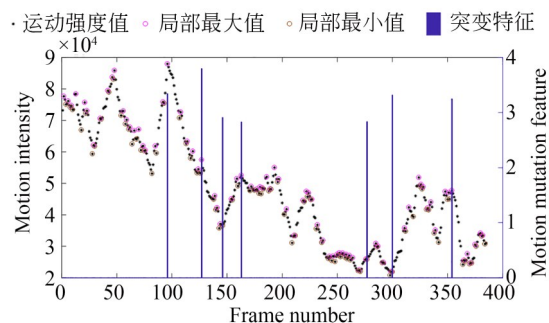
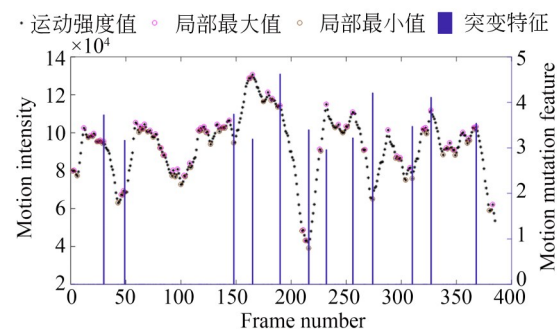
(a) 序列32的运动突变特征值分布MOS=1.59, M_k 平均值=47 607, $C_{\text{sum}_M}=37.50$, $C_{\text{max}_M}=6.17$ (a) Original motion features of Video 32 in SPVCD MOS=1.59, M_k average=47 607, $C_{\text{sum}_M}=37.50$, $C_{\text{max}_M}=6.17$ (b) 序列39的运动突变特征值分布MOS=4.00, M_k 平均值=47 984, $C_{\text{sum}_M}=22.28$, $C_{\text{max}_M}=3.80$ (b) Original motion features of Video 39 in SPVCD MOS=4.00, M_k average=47 984, $C_{\text{sum}_M}=22.28$, $C_{\text{max}_M}=3.80$

图6 运动强度相似序列运动突变特征比较

Fig. 6 Comparison of motion mutation feature in videos with similar M_k

2.7 回归模型

采用SVR作为映射函数对特征进行回归以得到最终的客观评价分数。考虑一组训练数据 $\{(x_i, y_i), \dots, (x_l, y_l)\}$,其中 $x_i \in F$ 为提取的VIMS

感知特征, y_i 是相应的 MOS。给定参数 $C > 0$ 和 $\varepsilon > 0$, SVR 的标准形式表示为:

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} w^T w + C \left\{ \sum_{i=1}^l \xi_i + \sum_{i=1}^l \xi_i^* \right\}, \quad (22)$$

$$\text{subject to } w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i, \quad (23)$$

$$y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i^*, \quad (24)$$

$$\xi_i, \xi_i^* \geq 0, i = 1, \dots, l, \quad (25)$$

其中, $K(x_i, x_j) = \varphi^T \varphi$ 为核函数。本文使用核函数为 $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ 的径向基函数。

3 实验结果与分析

所提出的预测模型在 SPVCD 立体全景视频数据库上进行了测试。立体全景视频数据库 (SPVCD)^[19] 共包含 116 个时长为 20 s 的立体全景视频, 其中既包含自然场景, 也包含人工合成场景。该数据库涵盖丰富的场景类型、运动类型, 包含各种可能造成不同程度晕动的立体全景视频。SPCVCD 数据库部分视频首帧如图 7 所示。

3.1 整体性能分析

为了检验模型的预测性能, 本文在 SPVCD 立体全景视频数据库上进行了实验。数据库中的 116 个视频, 80% 用于训练, 20% 用于测试, 并确保用于训练与测试的视频序列互不重叠。采用皮尔逊线性相关系数 (Pearson Linear Correlation Coefficient, PLCC)、斯皮尔曼相关系数 (Spearman Rank-order Correlation Coefficient, SROCC)、均方根误差 (Root Mean Square Error, RMSE) 三个性能指标来衡量所提出模型的性能, 计算公式如下:

$$\text{PLCC} = \frac{\sum_{i=0}^{N-1} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=0}^{N-1} (X_i - \bar{X})^2} \sqrt{\sum_{i=0}^{N-1} (Y_i - \bar{Y})^2}}, \quad (26)$$

其中: N 表示样本个数, X_i 和 Y_i 分别表示第 i 个样本的客观模型预测值和主观评分值, $\bar{X}$ 和 $\bar{Y}$ 分别表示客观模型预测均值和主观评分均值。PLCC 是用于衡量两个变量之间线性关系强度的指标, 其取值范围为 $-1 \sim 1$ 。当 PLCC 接近 1 时, 表示两个变量之间存在强正线性关系。

$$\text{SROCC} = 1 - \frac{6 \cdot \sum_{i=0}^{N-1} (R_{X_i} - R_{Y_i})^2}{N \cdot (N^2 - 1)}, \quad (27)$$

其中: R_{X_i} 和 R_{Y_i} 分别表示主观评分值和客观模型预测值按相同顺序排序后, 第 i 个分数在各自序列中的索引。SROCC 用于衡量两个变量之间的单调关系强度, 当其接近 1 时, 表示两个变量之间存在强正单调关系。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2}. \quad (28)$$

RMSE 是衡量预测值与实际值差异的指标, 数值越小代表预测值与实际值越接近。

同时, 为了验证本文所提出模型的性能以及各特征的有效性, 表 1 给出了本文所提出的 VIMS 评价模型以及去除其中一个类型特征后在 SPVCD 数据库上的性能指标, 并比较了将本文所提出的加权运动特征提取方法改为仅使用简单非加权的显著区域光流统计提取运动特征后的性能指标。为了尽量保证实验中训练集与测试集的随机性, 对 SPVCD 立体全景视频数据库共进行 1 000 次 8:2 比例的训练-测试集划分, 并取 1 000 次结果的 RMSE 的中值所对应模型的预测结果作为最终预测性能指标。

通过表 1 实验结果可以发现, 在去除任何一个类型的特征后模型性能均出现了不同程度的下降, 说明在沉浸式视觉内容观看过程中视觉内容的运动强度、视差、色彩类型、时域突变等都是影响晕动症程度的要素。其中, 在不包含运动特征时, 模型性能出现了最大程度的下降, 这说明在观看 VR 内容的过程中, 视觉内容中的运动强度对晕动症的产生起着主导作用, 这也印证了产生晕动症的主要原因为视觉和前庭感官感知运动的不匹配。同时, 视差特征的缺失对模型性能产生的影响最小, 一方面的原因可能是相比于立体图像、立体视频中备受关注的视差特征, 在立体全景视频中由于观看方式的改变, 沉浸式的观看使得受试者更多地受到运动等更为主观的因素的影响; 另一方面, 由于观看方式的不同, 立体全景视频采用头戴式显示器的观看方式, 容易发生在立体图像、视频中的焦点调节和会聚冲突现象有所减少。通过对比不同的运动特征提取, 相比于以往简单的对显著区域的光流进行统计以表示视觉内容的运动强度, 采用本文所提出的显

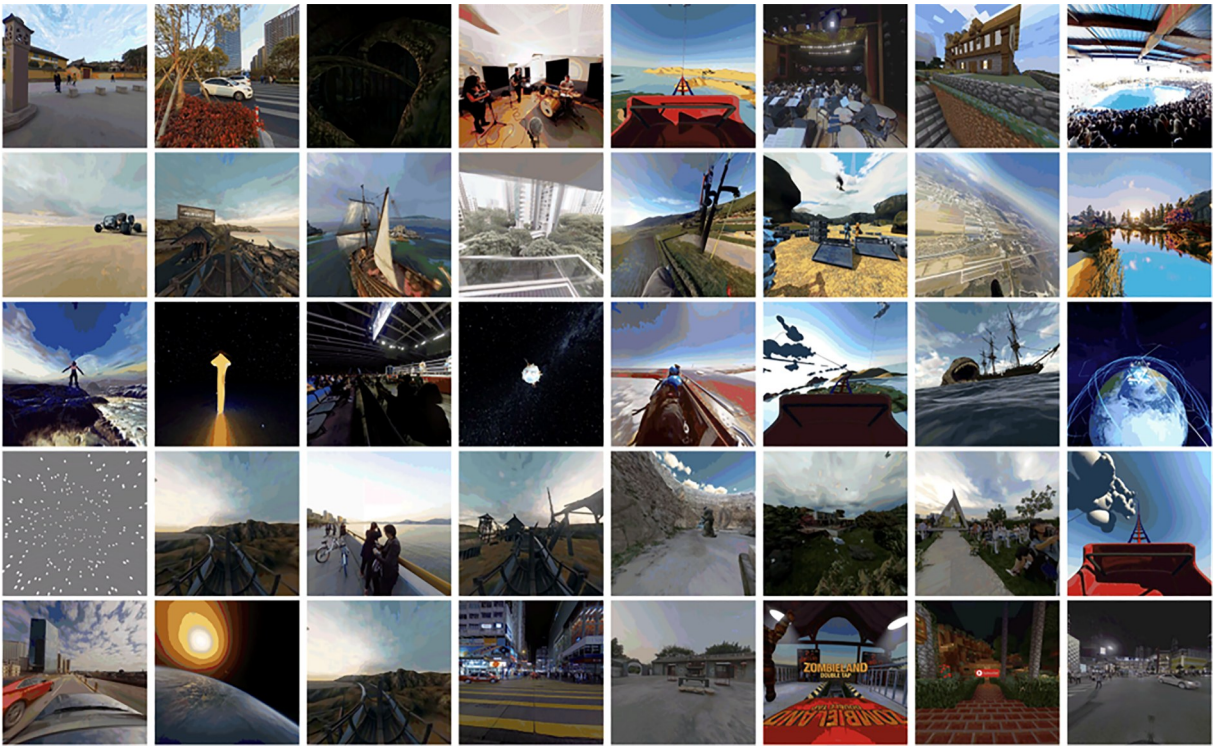


图7 SPVCD数据库部分视频示例
Fig. 7 Sample videos in SPVCD database

表1 不同特征集在 SPVCD上测试的性能指标
Tab.1 Performance indicators of different feature sets when tested on SPVCD

特征	PLCC	SROCC	RMSE
不含运动特征	0.672	0.599	0.666
不含视差特征	0.803	0.717	0.537
不含空间特征	0.711	0.617	0.640
不含突变特征	0.796	0.720	0.551
非加权运动特征	0.778	0.697	0.571
所有特征	0.821	0.790	0.486

著区域内外加权的运动特征提取方式在性能指标上也得到了提升,由此也说明在观看过程中观看者的晕动程度不仅受显著区域的运动情况所影响,也会受到显著区域外的运动情况所引起的自我运动影响。

3.2 不同统计方式下模型性能对比

考虑到数据库场景的多样性,难以遍历每一种训练-测试集划分结果所对应的模型性能。因此,此处以多种统计方式对模型性能进行验证与对比。本文采用SVR对特征空间进行回归,回归时对数据集进行1 000次的8:2比例训练-测试

集划分,并分别取三个指标所对应的中值作为最终的结果。最终,进行50次实验分别以RMSE、SROCC作为中值取对应的指标结果以及50次实验的指标平均值。对比算法^[19]为深度学习方法,将数据库视频随机划分为50对训练集和测试集在其网络模型上进行训练测试,并选择50次实验结果中RMSE中值所对应的模型代表最终性能。本文统计了对比算法50次实验结果中SROCC中值所对应的模型性能及50次实验结果的均值。表2为不同的统计方式得到的两个模型的性能指标。

表2 不同统计方式下性能对比
Tab.2 Performance comparison under different statistical methods

方法	统计方式	PLCC	SROCC	RMSE
Lu ^[19]	Median RMSE	0.816	0.875	0.740
	Median SROCC	0.796	0.846	0.809
	Average	0.810	0.837	0.763
Proposed	Median RMSE	0.821	0.790	0.486
	Median SROCC	0.825	0.762	0.499
	Average	0.840	0.751	0.489

实验结果表明,在以多种统计方式对模型性能进行比较时,两个模型的各项指标均存在一定程度的波动。这与SPVCD的特点有关,SPVCD包含了目前用于VIMS评价中数量最多的场景类型。因此,不同的训练、测试场景选择对模型预测结果的影响较大。本文所提出的方法在三种统计方式中PLCC和RMSE均优于对比算法,SROCC低于对比算法。这与对比算法网络训练时以最高SROCC值选取最优网络模型参数存在一定关联。

3.3 跨数据库性能对比

为验证所提出模型的泛化性,在斯坦福大学的数据库上进行了跨数据库实验^[18]。Stanford数据库包含19个立体全景视频场景,每个场景持续60 s。帧速率分别为24 fps,25 fps和30 fps,分辨率为1 K,2 K或4 K。以SPVCD数据库中的116个视频序列作为训练集,对Stanford数据库的19个视频序列进行测试,预测得到最终的性能指标。

值得注意的是,在SPVCD数据库中的主观得分越高,晕动病的程度越低。这与Stanford数据库的主观分数相反,Stanford数据库的MOS范围为4.3~41.1,远远超过了SPVCD库的最大MOS(MOS范围:1.318 2~4.863 6)。因此,在跨数据库实验中,首先将Stanford数据库的MOS映射到SPVCD的得分空间。将Stanford数据库的最大MOS对应SPVCD数据库的最小MOS,Stanford数据库的最小MOS对应SPVCD数据库的最大MOS。

表3 所提出模型在Stanford数据库上的总体及对比性能指标

Tab.3 Overall and comparative performance indicators of the proposed model on Stanford database

方法	PLCC	SROCC	RMSE
Lu ^[19]	0.606	0.614	0.713
Proposed	0.678	0.633	0.672

表3显示了所提出的晕动病程度预测模型在Stanford数据库上的跨数据库实验结果以及文献[19]的对比结果。所提出模型在Stanford数据库上的PLCC,SROCC和RMSE分别为0.678,0.633和0.672,取得了较为良好的预测性能,各

项性能指标均高于文献[19],原因可能是文献[19]所提出的方法为深度学习方法,在针对规模较小数据集时,模型容易过拟合导致模型性能不佳。

3.4 相关方法分析

目前对于虚拟现实VIMS评价的研究多是基于非立体内容,而对于立体全景视频的VIMS评价研究还较少。表4给出了两个团队各自数据库的简单介绍。分别为延世大学团队和韩国科学技术院团队。延世大学团队和韩国科学技术院团队均是研究非立体全景内容的VIMS评价,其中,各团队使用数据库不同且多样。截至目前,延世大学团队建立了三个用于VR晕动症研究的主观数据库,但数据库均尚未公开。韩国科学技术院团队建立了四个用于VR晕动症研究的主观数据库,数据库均包含生理信号,可以获取基准视频,但数据库中所包含的基准视频的变体没有直接提供。例如,文献[15]中的数据库包含80个视频,该数据库具有20个基准视频,可以直接获取;而每个基准视频又具有三种变体视频,不可直接获取。

因目前已公开的数据库均是非立体全景视频数据库且不可完全获取,与包含生理信号的方法对比不公平,致使本文所提模型未在其数据库上进行相关实验,暂时不能完全验证本文所提模型的泛化性。延世大学团队和韩国科学技术院团队所提出的模型大多使用了生理信号,这在本文所使用的SPVCD基准数据中是不存在的,因此也无法将其所提模型在SPVCD数据库上进行验证。不同的数据库配置导致设计和建立VIMS评价模型的出发点不同,无法在同一数据库上直接对比模型的性能优劣,因此,表4仅对已有的相关研究结果进行了汇总。由于相关数据库不同,其结果仅用于对客观VIMS评价模型的分析,不能直接通过表4中指标值来比较判断客观评价模型间的优劣。只有文献[19]的方法与本文方法同在SPVCD数据库上进行了性能测试。表4中,本文所提模型得到的性能指标采用与其相同的模型选取方式:对数据库随机划分进行测试并选取RMSE为中值所对应的模型结果作为最终的性能指标。对比结果发现本文所提出的模型在PLCC,RMSE两项指标上取得了更优异的指

表 4 所提方法与其他模型的结果汇总

Tab. 4 Summary of results the proposed method and other model

方法	所用数据库	是否包含 生理信号	视频 个数	是否 立体	PLCC	SROCC	RMSE
H. Oh ^[3]	CYRE database	是	52	非立体	0.602	0.606	/
J. Kim ^[7]	VR sickness database	否	36	非立体	0.724	0.710	/
J. Kim ^[8]	CYRE database	是	52	非立体	0.862	0.861	/
W. Kim ^[9]	VR-SP database	否	100	非立体	0.789	0.721	/
S. Lee ^[10]	VRSA-DB-FR	是	20	非立体	0.830	0.819	7.341
S. Lee ^[11]	VRSA-DB-shaking	是	20	非立体	0.751	0.679	25.373
	VRSA-DB-FR	是	20	非立体	0.801	0.671	22.937
H. G. Kim ^[12]	VR video database	是	9	非立体	0.885	0.882	10.251
S. Kim ^[13]	VRSA-DB-FR	是	20	非立体	0.804	0.698	27.406
H. G. Kim ^[14]	VRPS-DB-FR	是	80	非立体	0.831	0.812	7.112
S. Lee ^[15]	VRSA-DB-shaking	是	20	非立体	0.767	0.706	25.946
	VRSA-DB-FR	是	20	非立体	0.837	0.719	20.350
Lu ^[19]	SPVCD	否	116	立体	0.816	0.875	0.740
Proposed	SPVCD	否	116	立体	0.821	0.790	0.486

标,在 SROCC 上有所不足。文献[19]采用深度学习的方法,同时考虑了双目感知、运动特征等特征,然而对运动特征的考虑较为单一,且未考虑时域上的突变的影响。

4 结 论

本文针对立体全景视频中存在的视觉诱发晕动症这一问题,综合考虑可能导致视觉诱发晕动症的各项因素,设计了基于时空多特征的虚拟

现实中视觉诱发晕动症评价模型。同时,在运动特征提取方面基于视觉感知冲突理论,设计了前背景加权运动特征提取。在时域突变信息度量方面,在时域聚合的基础上,分别对运动信息、视差信息、空域感知信息等特征提取突变特征。在 SPVCD 数据库上,所提方法展现出了与最先进的深度学习方法相当的预测性能。在下一阶段工作中,如何将本文中提取特征方式与深度学习联合,构造自动进行对应特征提取学习的深度神经网络将成为研究的重点。

参考文献:

- [1] KOURTESIS P, AMIR R, LINNELL J, *et al.* Cybersickness, cognition, & motor skills: the effects of music, gender, and gaming experience[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2023, 29(5): 2326-2336.
- [2] HSIAO C Y, KUO C C, LIOU Y A, *et al.* Determining work-rest schedules for visual tasks that use optical head-mounted displays based on visual fatigue and visually induced motion sickness recovery [J]. *International Journal of Environmental Research and Public Health*, 2023, 20(3): 1880.
- [3] YIP S H, SAUNDERS J A. Restricting the distribution of visual attention reduces cybersickness[J]. *Cognitive Research: Principles and Implications*, 2023, 8(1): 18.
- [4] OH H, SON W. Cybersickness and its severity arising from virtual reality content: a comprehensive study[J]. *Sensors*, 2022, 22(4): 1314.
- [5] LOU R D, MÉRIENNE F, SO R H Y, *et al.* Geometric Simplification for Reducing Optic Flow in VR [C]. 2022 *IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. Singapore, Singapore. IEEE, 2022: 682-685.
- [6] KENNEDY R S, LANE N E, BERBAUM K S, *et al.* Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness[J]. *The*

- International Journal of Aviation Psychology*, 1993, 3(3): 203-220.
- [7] GUNA J, GERŠAK G, HUMAR I, *et al.* Influence of video content type on users' virtual reality sickness perception and physiological response [J]. *Future Generation Computer Systems*, 2019, 91: 263-276.
- [8] KIM J, KIM W, AHN S, *et al.* Virtual Reality Sickness Predictor: Analysis of Visual-Vestibular Conflict and VR Contents[C]. 2018 *Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. Cagliari, Italy. IEEE, 2018: 1-6.
- [9] KIM J, OH H, KIM W, *et al.* A deep motion sickness predictor induced by visual stimuli in virtual reality [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(2): 554-566.
- [10] KIM W, LEE S, BOVIK A C. VR sickness versus VR presence: a statistical prediction model[J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 2021, 30: 559-571.
- [11] LEE S M, KIM S, KIM H G, *et al.* Physiological fusion net: quantifying individual VR sickness with content stimulus and physiological response [C]. 2019 *IEEE International Conference on Image Processing (ICIP)*. Taipei, China. IEEE, 2019: 440-444.
- [12] LEE S M, KIM J U, KIM H G, *et al.* SACA Net: Cybersickness Assessment of Individual Viewers for VR Content via Graph-Based Symptom Relation Embedding [M]. *Computer Vision-ECCV 2020*. Cham: Springer International Publishing, 2020: 170-186.
- [13] KIM H G, LIM H T, LEE S M, *et al.* VRSA net: VR sickness assessment considering exceptional motion for 360° VR video [J]. *IEEE Transactions on Image Processing*, 2019, 28(4): 1646-1660.
- [14] KIM S, LEE S M, RO Y M. Estimating VR sickness caused by camera shake in VR videography [C]. 2020 *IEEE International Conference on Image Processing (ICIP)*. Abu Dhabi, United Arab Emirates. IEEE, 2020: 3433-3437.
- [15] KIM H G, LEE S M, KIM S, *et al.* Towards a better understanding of VR sickness: physical symptom prediction for VR contents[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(1): 836-844.
- [16] LEE S M, KIM S, KIM H G, *et al.* Assessing individual VR sickness through deep feature fusion of VR video and physiological response [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(5): 2895-2907.
- [17] LUO S H, REN P, ZHU C F. A model for predicting the level of motion sickness caused by virtual reality content by using attention and effective visual field mechanism [C]. 2023 *9th International Conference on Virtual Reality (ICVR)*. Xianyang, China. IEEE, 2023: 400-407.
- [18] PADMANABAN N, RUBAN T, SITZMANN V, *et al.* Towards a machine-learning approach for sickness prediction in 360° stereoscopic videos [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2018, 24(4): 1594-1603.
- [19] LU Z A, YU M, JIANG G Y, *et al.* Prediction of motion sickness degree of stereoscopic panoramic videos based on content perception and binocular characteristics [J]. *Digital Signal Processing*, 2023, 132: 103787.
- [20] WEI J, WANG S H, WU Z, *et al.* Label decoupling framework for salient object detection [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 13022-13031.
- [21] LIU C, FREEMAN W T, ADELSON E H, *et al.* Human-assisted motion annotation [C]. 2008 *IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, AK, USA. IEEE, 2008: 1-8.
- [22] SUMAYLI Y, YE Y. Motion sickness during roll motion: VR HMD view versus monitor view [J]. *Vibration*, 2023, 6(1): 45-56.
- [23] ROGERS C, RUSHTON S K, WARREN P A. Peripheral visual cues contribute to the perception of object movement during self-movement [J]. *i-Perception*, 2017, 8(6): 2041669517736072.
- [24] 江本赤, 卞仕磊, 史晨阳, 等. 基于色貌尺度相位一致性的全参考图像质量评价 [J]. *光学精密工程*, 2023, 31(10): 1509-1521.
- JIANG B C, BIAN S L, SHI C Y, *et al.* Full reference image quality assessment based on color appearance-based phase consistency [J]. *Opt. Precision Eng.*, 2023, 31(10): 1509-1521. (in Chi-

- nese)
- [25] ZHOU Y, LIANG Y S, LIU W, *et al.* Research on video motion characteristics extraction and description based on human visual characteristics [C]. 2017 *IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. Chongqing, China. IEEE, 2017: 936-940.
- [26] 权巍, 王超, 耿雪娜, 等. 基于运动感知的 VR 体验舒适度研究[J]. 系统仿真学报, 2023, 35(1): 169-177.
- QUAN W, WANG C, GENG X N, *et al.* Research on VR experience comfort based on motion perception [J]. *Journal of System Simulation*, 2023, 35(1): 169-177. (in Chinese)
- [27] ZHOU Y, YU W L, LI Z, *et al.* Stereoscopic visual discomfort prediction using multi-scale DCT Features[C]. *Proceedings of the 27th ACM International Conference on Multimedia*. Nice France. ACM, 2019: 184-191.
- [28] LOHMAN J, TURCHET L. Evaluating cybersickness of walking on an omnidirectional treadmill in virtual reality [J]. *IEEE Transactions on Human-Machine Systems*, 2022, 52(4): 613-623.

作者简介:



董奇峰(1999—),男,浙江宁波人,硕士研究生,2021年于浙大宁波理工学院获得学士学位,主要从事全景视频晕动症评价等方面的研究。E-mail: 792505038@qq.com

通讯作者:



郁梅(1968—),女,江苏无锡人,博士,教授,博士生导师,2000年于韩国 Ajou 大学获得博士学位,主要从事多媒体信号处理与通信、计算成像、视觉感知与编码、图像与视频质量评价等方面的研究工作。E-mail: yumei@nmbu.edu.cn