计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200500031

虚拟现实环境下基于眼动跟踪的导航需求预测与辅助

朱晨爽　程时伟
浙江工业大学计算机科学与技术学院　杭州 310023

摘　要　针对复杂虚拟现实场景中传统导航方法对用户支持不足和用户沉浸感较低等问题，文中提出了基于梯度提升迭代决策树的二分类模型，利用用户在虚拟现实环境中使用辅助导航时需要的眼动数据，来分析和预测用户在任务过程中是否需要辅助导航。根据用户的注视序列对该模型进行评估，得到用户需求判定方法的平均精确率和准确率分别为77.6%与77.2%。此外，文中借助所设计模型实现了一个导航辅助原型系统，通过对用户的眼动数据进行分类，来自动呈现导航辅助界面。实验结果表明，与传统的永久性辅助导航方法相比，新提出的自适应辅助导航具有更好的用户体验。

关键词：虚拟现实；眼动跟踪；人机交互；导航
中图法分类号　TP391

Prediction and Assistance of Navigation Demand Based on Eye Tracking in Virtual Reality Environment

ZHU Chen-shuang and CHENG Shi-wei
School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

Abstract　In order to solve the problems of insufficient user support and low user immersion in traditional navigation methods in complex virtual reality scenes, this paper proposes a binary classification model based on the gradient boost decision tree, which uses eye movement data before and after the user's need for auxiliary navigation in the VR environment to analyze and predict whether the user needs navigation during the task. The model is evaluated according to the user's gaze sequence, and the average precision and accuracy of the user demand judgment method are 77.6% and 77.2%, respectively. In addition, we implemente a navigation aid prototype system. By classifying the user's navigation requirement based on eye movement data, the user interface of the prototype system can automatically present the map for navigation. Experimental results showed that, compared with the traditional permanent assisted navigation method, the adaptive assisted navigation proposed in this paper can provide better user experience.

Keywords　Virtual reality, Eye tracking, Human computer interaction, Navigation

1　引言

在VR环境中，用户常常会失去方向感，从而导致不适，严重影响了沉浸式体验。而开发人员并不完全了解这种用户需求，通常只是将导航指南设计为小地图、箭头或覆盖路径等形式来呈现给用户。但是，这些形式往往会占用显示器上相当大的可见范围，导致在视觉上对用户没有吸引力或过度分散用户的注意力。

本文在已有研究[1]的基础上，通过分析用户的眼动数据来预测用户对辅助导航的需求。用户在虚拟现实场景中完成搜索任务，同时记录他们在任务期间的眼动数据，进而训练眼动行为分类器，并预测用户在虚拟环境下对导航辅助功能的需求程度。在此基础上，将需求预测模型应用于相同场景的不同任务中，设计自适应辅助导航功能，以在虚拟环境中有效地为用户提供导航服务，并通过用户实验进行了验证。

2　相关工作

眼动注视模式可以对人类行为和认知进行分析。例如，Vlcek等[2]研究了注视动作与空间导航之间的关系，并将这一发现用于研究阿尔茨海默氏症患者在迷路时的表现。Laura等[3]发现了不同的导航指南形式与其注视模式之间的关系。上述研究主要在二维空间基于眼动数据进行分析。近年来，研究人员也逐步在三维环境下利用眼动数据开发相关的交互应用。例如，Kevin等[4]提出了在虚拟现实环境中模拟人眼注视，以提高用户与虚拟人物之间的交互沉浸感，开发了一个以用户与虚拟人物交互为核心的眼动跟踪交互叙事框架，该框架能够对玩家的注视作出反应，模拟人与人之间的真实交流场景，实验结果证明用户体验得到了提升。此外，

到稿日期：2020-05-08　返修日期：2020-07-17　本文已加入开放科学计划(OSID)，请扫描上方二维码获取补充信息。
基金项目：国家自然科学基金(61772468)
This work was supported by the National Natural Science Foundation of China(61772468).
通信作者：朱晨爽(759240316@qq.com)

Boyd 等[5]探讨了在沉浸式虚拟现实环境中使用眼神进行面对面交流的方式对自闭症儿童的影响。

眼动跟踪技术也可用于 VR 环境中的导航。例如，Kasahara 等[6]的研究通过眼动跟踪技术分析人们的第一视角注视行为，结果表明，人们可以通过改善自己的注视行为来理解自己在复杂环境下自身位置与他人位置的空间关系。Gandrud 等[7]利用用户的注视和头部动作来推断用户在虚拟现实环境中的导航方向。Lee 等[8]也提取此类数据，比较了在导航过程中玩家在第一人称视角中的视觉注意力与在第三人称视角中作为化身的视觉注意力的不同之处。Saha 等[9]在 VR 环境中使用嵌入路径的导航助手将用户引导到目的地。Alghofaili 等[1]将用户在 VR 中导航的画面投影到屏幕上，并对投影的二维眼动数据进行分析，进而预测用户行为，实现了用户在虚拟现实环境中的自适应导航。

本文在上述研究的基础上，提出了基于梯度提升迭代决策树(Gradient Boost Decision Tree，GBDT)的二分类模型，利用模型来预测用户在虚拟环境中是否需要辅助导航，并相应地在虚拟现实场景中显示地图，为用户指示方向。

3 导航辅助需求预测模型

建立一种关于眼动注视模式的分类器，来预测用户对导航辅助功能的实际需求。使用该分类器，尝试向用户自适应地显示辅助导航工具。因此，首先获取 VR 环境中用户的眼动数据，然后利用 GBDT 算法训练分类器以预测用户的导航需求。

3.1 VR 环境中的眼动跟踪技术

本文采用基于瞳孔中心-角膜反光点(Pupil Center Cornea Reflection，PCCR)的方法来获取眼动数据，在获得眼动数据的 PCCR 特征向量后，将特征向量数据导入虚拟场景中，并将其赋值给虚拟现实中的射线功能模块，实现了在虚拟场景中用射线代替眼动视线；然后，通过射线在虚拟场景中与场景物体的碰撞来获取其三维坐标数据。本文首先采用五点标定法，即让当前用户在虚拟现实场景中进行五点标定，如图 1 所示，图中屏幕上依次会出现标定点(绿色圆点)和用户眼动注视点(白色圆点)。然后，从中心位置 1 开始标定，依次按编号 1—5 的顺序进行眼动标定，用户视线跟随着标定点依次移动，每当标定点到达目标位置后，需要视线保持注视标定点一段时间(本文设置为 2 s)，同时记录用户的瞳孔中心和普尔钦斑中心，从而计算得到 PCCR 向量。最后，将每个标定点在虚拟现实场景中的映射坐标和相应的 PCCR 向量作为眼动映射模型的训练数据[10-11]。

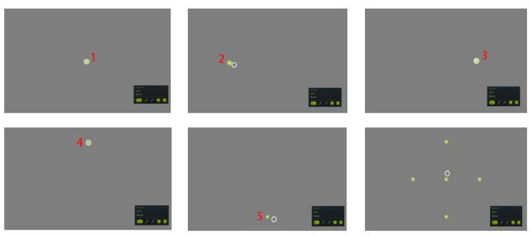


图 1 五点标定示意图(电子版为彩色)

Fig. 1 Five-point calibration diagram

将获取到的 PCCR 特征向量赋值给虚拟现实中的射线模块，将获得特征向量的射线记为 X，此时在虚拟现实场景中射线 X 即为被试者的眼动视线方向，当检测到射线 X 与虚拟现实场景中的场景内容发生碰撞时，得到发生碰撞点的位置 P，再通过拟合方法计算出用户在三维空间中的注视点坐标[12]。

如图 2 所示，*origin* 为在世界坐标中射线 X 的起点，***direction*** 为射线的方向矢量，*distance* 为射线长度，即起点到碰撞点 P 的距离，根据这些信息就可以得到碰撞点 P 相对于起始点 *origin* 的三维坐标。

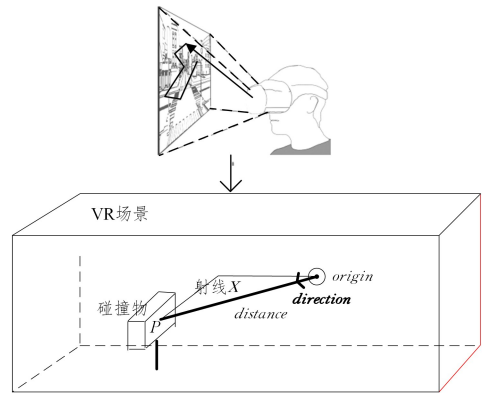


图 2 VR 场景中的坐标点数据

Fig. 2 Coordinate point data in VR scene

虚拟现实场景刷新率为 30 fps，因此将移动窗口的大小设置为 30 个时间帧每秒。对注视样本进行预处理，生成用于分类的最终序列，随后使用 10 s 内获得大小为 300 个时间帧的窗口，保证精确度与最短预测窗口之间的最佳效益。根据实验设计提到的辅助导航功能来标记窗口注视序列，被试者按下提示按钮表明他们需要导航辅助。因此，在导航标记之前将窗口宽度(或 300 个时间帧)内的所有窗口注视序列视为“迷路”，标记为“0”(需要辅助导航)。在任务的前半阶段，用户将按照给定路径从起点出发的窗口注视序列视为“不迷路”，标记为 1。

3.2 眼动数据获取

记录被试者注视场景时的眼动注视点的三维坐标等数据，并将它们作为有标记的注视序列送至分类器进行处理。具体地，招募了 4 名年龄在 22～26 岁之间的研究生参加实验。如图 3 所示，每位被试者都佩戴装有眼动模块的 HTC VIVE 虚拟现实头盔。图 4 给出了 HTC VIVE 头盔中的眼动模块，以及该模块所提取的被试者眼动数据和向被试者呈现的虚拟现实场景示例。

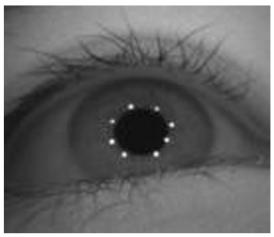


图 3 用户佩戴 HTC VIVE 虚拟现实头盔

Fig. 3 Users wear HTC VIVE virtual reality helmet



(a)HTC VIVE 头盔中的眼动模块



(b)眼动特征点



(c)佩戴头盔后看到的虚拟现实场景

图 4 示例图片

Fig. 4 Sample picture

为了收集训练分类器的眼动数据集，本文利用 Unity 设计了一个三维虚拟城市，包括模拟商场、路边商店、花园广场和路标等[13]，其布局如图 5 所示。每位被试者在虚拟城市中通过导航来完成搜寻任务，后台程序在此期间获取他们的眼动数据。

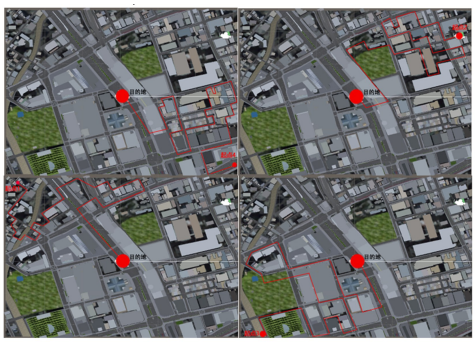


图 5 虚拟城市的布局以及 4 个不同起点的轨迹路线

Fig. 5 Layout of VR city and trajectory of 4 different starting points

如图 6 所示，在任务的前半阶段，每位被试者都将被随机安排在虚拟城市的 4 个预定位置，将这 4 个预定位置作为他们的起始点，并要求他们按照虚拟场景中的给定路径到达指定位置[14]。在此期间，系统将向被试者展示路径以帮助他们完成任务，还将收集被试者认为“不需要辅助导航”时的眼动数据。

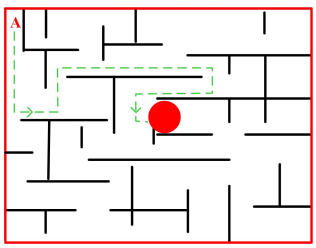
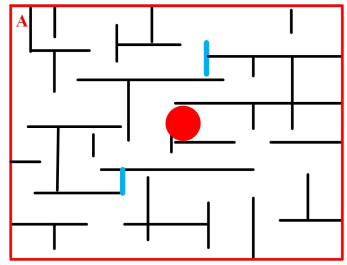


图 6 被试者跟随给定路线(虚线部分)前往目的地

Fig. 6 Participants followed the given route (dotted line) to their destination

被试者根据上述要求完成训练任务并熟悉路线之后，将进入任务的后半阶段。在此阶段开始前，被试者返回前半阶段的起始点位置。与前半阶段不同的是，该阶段将不再展示给定路径，这意味着被试者只能通过按下控制器上的按钮来触发小地图导航，并且，如果触发小地图导航，则被试者需要等待至少 30 s 才能再次触发导航。这是为了约束被试者，避免其频繁使用提示，并确保被试者在最需要时才借助导航辅助功能，从而降低误报的可能性。此外，为了消除学习效应、保持任务难度，并在任务执行过程中获得足够多的眼动数据，在被试者回到起始点的过程中，系统在原先给定的路径上随机放置了一些屏障(见图 7)。



注：粗体线条为随机放置的屏障

图 7 用户需从目的地返回起点

Fig. 7 Subject is asked to return from the destination to the starting point

由于虚拟城市的布局较复杂，任务难度提高。同时，添加了具有时间间隔的小地图导航功能和城市道路屏障，以确保系统能够提取到被试者处于认为“需要辅助导航”时的眼动数据。

实验总共收集了 16 组注视序列，每位被试者完成了 4 个不同起点的任务。被试者在每次任务的注视序列上的标记汇聚成数据集，分别形成了 5 400 个标记为“0”和“1”的数据样本。为了创建训练集，提取了 4 320 个样本用于创建平衡数据集(包含 80%的“0”类数据和相同数量的“1”类数据)。通过未包括在训练集中的剩余 20%的“1”类数据，以及包括在训练集中相同数量的随机选择的“0”类数据来创建测试集。其中，采集到的某一未迷路阶段 10 s 的数据如表 1 所列(坐标数据为相对用户虚拟空间位置的局部坐标)。

表 1 眼动数据样本

Table 1 Example of eye tracking data

	第 1 帧	第 2 帧	…	第 30 帧	迷路标记
1 s	(25.3,11.2,34.5)	(22.6,16.4,31.2)	…	(35.3,24.6,10.7)	1
2 s	(32.2,22.4,13.6)	(30.3,24.4,14.5)	…	(26.4,21.5,24.8)	1
⋮	⋮	⋮	⋮	⋮	⋮
10 s	(19.3,18.7,15.5)	(14.7,11.6,18.5)	…	(17.3,16.4,16.2)	1

3.3 需求预测模型

本文使用梯度提升迭代决策树 GBDT 对用户在虚拟环境中的眼动注视模式进行分类。GBDT 使用分类与回归树(Classification and Regression Tree，CART)，其解决的核心问题是：在每轮迭代时，利用真实标签减去弱分类器的输出结

果得到残差[15]。本文将使用逻辑回归的对数似然损失函数，将结果的概率值与真实概率值的差值作为残差进行拟合。

梯度提升迭代决策树分类的主要思想为：每一次建立模型的方向是之前建立模型损失函数的梯度下降方向。损失函数越大，说明模型越容易出错。如果模型能够让损失函数持续地下降，则说明模型在不停地改进，而最好的方式就是让损失函数在其梯度方向上下降[16]。

对数似然损失函数是在概率估计基础上进行定义的，用于评估分类器的概率输出。通过舍弃错误的分类结果，来实现对分类器准确度的量化。最小化对数损失基本等价于最大化分类器的准确度。为了计算对数损失，分类器必须提供输入的每个类别的概率值。因此，首先计算逻辑回归的对数损失函数[17]，逻辑回归的预测函数为：

$$h_\theta(x)=\frac{1}{1+e^{-\theta^T x}} \tag{1}$$

函数 $h_\theta(x)$ 的值有特殊的含义，它表示结果取 1 的概率，因此对于输入 x 分类结果为类别“1”和类别“0”的概率分别为：

$$P(Y=1|x;\theta)=h_\theta(x) \tag{2}$$

$$P(Y=0|x;\theta)=1-h_\theta(x) \tag{3}$$

推导出逻辑回归的对数损失函数为：

$$P(Y=y|x;\theta)=(h_\theta(x))^y\ (1-h_\theta(x))^{1-y} \tag{4}$$

通过取似然函数，并在同一 θ 处取得极值，使用梯度下降法求得最佳参数 θ，从而得到对数损失函数 $L(\theta)$ 为：

$$L(\theta)=\sum_{i=1}^{N}[y_i\log h_\theta(x_i)+(1-y_i)\log(1-h_\theta(x_i))] \tag{5}$$

由于 GBDT 让损失函数沿着梯度方向下降，因此需要利用损失函数的负梯度在当前模型的值作为算法中残差的近似值来拟合一个回归树。为了让损失函数尽可能地收敛到最优解，算法在每轮迭代时都需要拟合损失函数在当前模型下的负梯度值，以计算得到最佳残差拟合值。

根据上述对数损失函数可以得到最佳残差拟合值，单个样本 (x_i,y_i) 的损失函数 $L(\theta)$ 为：

$$L(\theta)=-y_i\log\hat{y}_i-(1-y_i)\log(1-\hat{y}_i) \tag{6}$$

设 $\hat{y}_i=h_\theta(x)$ 是逻辑回归预测的结果，第 M 步迭代之后当前学习器为 $F(x)=\sum_{m=0}^{M}h_m(x)$。代入式(6)可得到新的损失函数 $L(y_i,F(x_i))$，这样就能计算第 m 棵树对应的损失函数的负梯度，即伪残差。

$$r_{m,i}=y_i-\hat{y}_i \tag{7}$$

计算得到各个叶子节点的最佳残差拟合值为 $c_{m,j}$。因此，梯度提升迭代决策树二分类算法的完整过程如下：

初始化第一个弱学习器 $F_0(x)$：

$$F_0(x)=\log\frac{P(Y=1|x)}{1-P(Y=1|x)} \tag{8}$$

其中，$P(Y=1|x)$ 是训练样本中 $y=1$ 的比例，利用先验信息来初始化学习器。

对于建立 M 棵分类回归树 $m=1,2,\cdots,M$：

对于 $i=1,2,\cdots,N$，根据上述得到的式(7)计算第 m 棵树对应的响应值 $r_{m,i}$：

$$r_{m,i}=y_i-\frac{1}{1+e^{-F(x_i)}} \tag{9}$$

对于 $i=1,2,\cdots,N$，利用回归树拟合数据 $(x_i,r_{m,i})$，得到第 m 棵回归树，其对应的叶子节点区域为 $R_{m,j}$，其中 $j=1,2,\cdots,J_m$，且 J_m 为第 m 棵回归树叶子节点的个数。

对于 J_m 个叶子节点区域 $j=1,2,\cdots,J_m$，代入式(7)计算出最佳拟合值 $c_{m,j}$：

$$c_{m,j}=\arg\min_{c}\sum_{x_i\in R_{m,j}}L(y_i,F_{m-1}(x_i)+c) \tag{10}$$

根据最佳拟合值 $c_{m,j}$，可更新 $F_m(x)$ 强学习器：

$$F_m(x)=F_{m-1}(x)+\sum_{j=1}^{J_m}c_{m,j}I(x\in R_{m,j}) \tag{11}$$

得到最终的强学习器的表达式 $F_M(x)$ 为：

$$F_M(x)=F_0(x)+\sum_{m=1}^{M}\sum_{j=1}^{J_m}c_{m,j}I(x\in R_{m,j}) \tag{12}$$

将逻辑回归的公式进行整理，可以得到 $\log\frac{p}{1-p}=\theta^T x$，其中 $p=P(Y=1|x)$，也就是将给定输入 x 预测为正样本的概率。逻辑回归用一个线性模型去拟合 $Y=1|x$ 这个事件的对数概率。二元梯度提升迭代决策树分类算法与逻辑回归思想一样，用一系列的梯度提升树去拟合这个对数概率[18]，其分类模型可以表示为：

$$p=P(Y=1|x)=\frac{1}{1+e^{-F_M(x)}} \tag{13}$$

相比其他算法，梯度提升迭代决策树不仅能筛选特征，还可以很好地处理缺失特征的数据，因为决策树的每个节点只依赖一个特征，如果某个特征不存在，则决策树依然可以进行决策，只是少了一些路径。本文使用对数似然损失函数进行分类，相比其他算法能提供无偏估计的替代损失(surrogate loss)。除此之外，数据不需要归一化，特征的作用只是分裂节点，叶节点的值与特征值的大小无关，是否归一化并不影响叶节点值的大小，也不影响梯度下降的进程。

应用 GBDT 算法对在虚拟环境中的用户的注视模式进行分类。使用 Python 库 Keras 来设置和训练模型的分类，模型参数设置学习率为 0.3，迭代次数为 10，树的深度为 5。在一个基于 GPU 的高性能计算集群节点上进行训练，该节点有两个 4 核 Intel(R)Xeon(R)CPU，处理器基频为 2.40 GHz，RAM 为 32 GB。

4 模型评估

本文计算了分类模型的精确率和 $F1$ 分数，来对分类模型进行评估。其中混淆矩阵用于计算需求判定的精确率，具体定义如表 2 所列。

表 2 需求判定的混淆矩阵说明

Table 2 Explanation of confusion matrix for demand judgment

混淆矩阵	模型判定无需求	模型判定有需求
用户不需要导航	*FP*	*TP*
用户需要导航	*TN*	*FN*

精确率、召回率、准确率的计算表达式如下：

$$precision=\frac{TP}{TP+FP} \tag{14}$$

$$recall=\frac{TP}{TP+FN} \tag{15}$$

$$accuracy=\frac{TP+TN}{TP+FN+FP+TN} \tag{16}$$

$F1$ 分数(F1 score)是精确率(precision)和召回率(recall)的调和平均数，常用于比较模型的分类与识别性能，其计算表达式如下：

$$F1=\frac{2TP}{2TP+FP+FN} \tag{17}$$

本文首先对 4 名被试者的所有眼动数据进行统一的模型预测，得到单统一模型，其模型检验计算需求判定的数据如表 3 所列，其预测效果的数据值都明显差于其他模型，说明不同用户的眼动习惯不同，当对所有用户的眼动数据进行统一需求判定预测时，预测效果不好。

表 3 分类模型的测试结果

Table 3 Test results of classification model

	用户 1	用户 2	用户 3	用户 4	用户平均值	单统一模型	其他文献二维坐标注视分类模型[19]
精确率	0.823	0.805	0.774	0.703	0.776	0.632	0.766
召回率	0.808	0.818	0.740	0.691	0.764	0.614	0.783
准确率	0.817	0.810	0.762	0.700	0.772	0.628	0.719
$F1$	0.815	0.811	0.757	0.697	0.770	0.623	0.775

接着，对每位被试者的眼动数据进行单独的预测，4 名被试者各自的预测模型需求判定数据如图 8 所示。当被试者使用自己的眼动数据进行建模预测时，对比单统一模型，效果得到了显著提升，特别是被试者 1 和被试者 2 的预测效果精确率、召回率、准确率和 $F1$ 分数都高于二维坐标注视数据的分类模型[19]；而被试者 3 的模型预测分类的精确率和准确率高于二维坐标注视点数据的分类模型，而召回率和 $F1$ 分数相对较低；被试者 4 的模型预测分类效果的测试数据值都相对较差。

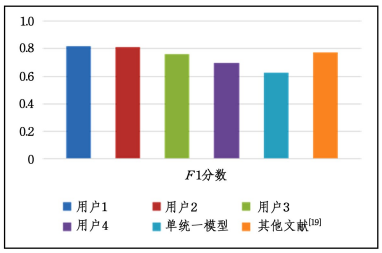


图 8 导航需求预测 $F1$ 分数

Fig. 8 $F1$ score from the prediction of navigation aid

进一步分析发现，在三维虚拟现实场景中，眼动数据模型预测和个人的眼动行为特征关系紧密[20]，不适合应用统一的预测模型。此外，每位被试者各自的模型预测效果普遍较好，但不同被试者的预测效果有较大差异。

5 应用实例

5.1 自适应导航辅助系统

在虚拟现实环境下，预测用户何时需要导航是非常有用的。使用辅助导航分类器，能够通过分类结果确定用户是否需要辅助导航。首先收集被试者 10 s 内 300 帧的眼动数据，收集超过 10 s 的数据后，将不断更新收集到的眼动数据，并且不断对新的注视点进行分类。一旦被试者的注视模式被分类为“需要辅助导航”，系统就会通过小地图的形式为被试者呈现辅助导航界面，引导用户到达目的地。

被试者可以使用图 9 给出的“辅助灵敏度档位”来选择他们想要的导航帮助级别。本文设计了三档灵敏度级别：高灵敏度、中灵敏度和低灵敏度。被试者可以在交互界面上选择辅助导航分类器的灵敏度级别，例如图 9 中被试者选择中灵敏度，意味着在 30 s 内若存在 1 个或 2 个注视数据集序列被分类为“迷路”状态时，系统才会为其提供辅助导航。每组数据集的注视序列指从 10 s 内的 300 帧 VR 环境中收集的被试者的眼动数据。在用户实验中，将灵敏度设置为低灵敏度(即 30 s 内有 3 个注视数据集序列被分类为“迷路”时才会显示辅助导航)[21]。

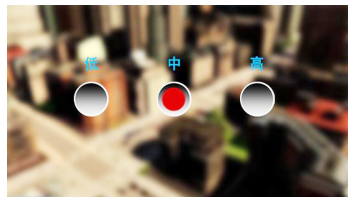


图 9 辅助灵敏度档位

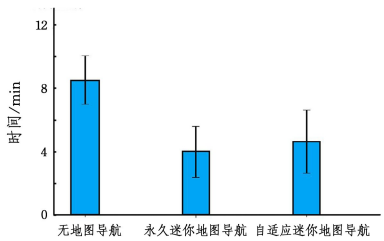
Fig. 9 Auxiliary sensitivity gear

将预测模型应用到虚拟现实场景中得到自适应导航模型，旨在增强用户在虚拟现实环境中的导航体验。本文进行了一项用户研究来验证自适应导航的有效性，以便快速且毫无阻碍地引导用户到达目的地。同时，还将通过预测模型得到的自适应辅助导航与传统的永久性辅助设备(迷你地图)以及无辅助导航进行比较，探索在虚拟现实环境中用户的满意度和沉浸感[22]。

招募 4 名年龄在 22～26 岁之间的健康研究生参与这项研究，所有被试者的视力均为正常或已矫正至正常。采用 HTC VIVE 虚拟现实头戴式显示器与眼动模块。在虚拟城市中，将整个场景分成 3 块区域，让被试者在不同的起点进行实验。要求被试者从起始点分别借助 3 种不同的导航工具寻找路径以到达目的地。由于地标在导航中起着重要作用，因此将场景中的目的地设计成一座塔，其高度足以让人从场景中的任何位置对其进行观察。

5.2 评估与分析

3 种不同的导航方式分别是：由预测模型生成的自适应迷你地图导航、永久迷你地图导航以及无地图导航。在被试者使用不同的导航方式寻找路径以到达目的地的过程中，对比这 3 类导航辅助方式在虚拟环境中对被试者的影响。所有被试者完成任务的平均时间如图 10 所示，可以看到，被试者在使用无地图导航时耗费的时间最长，使用永久迷你地图导航时耗费的时间最短，使用自适应迷你地图导航耗费的时间与使用永久迷你地图导航耗费的时间相近，说明用户使用自适应迷你地图导航完成任务的效率较高。



注：误差线为标准差

图 10 用户在 VR 场景中使用不同导航的平均时间

Fig. 10 Average duration for users with different navigation aids in a VR scene

进一步地，通过眼动热区图分析发现，与永久迷你地图相比，自适应迷你地图不会明显地改变参与者对 VR 场景的注意力(见图 11)。且被试者对于永久迷你地图外部的注视区域的平均注视持续时间短于自适应迷你地图导航的平均注视持续时间；与使用永久迷你地图相比，被试者在使用自适应迷你地图时将更多地探索和关注 VR 场景本身具有的特征[23]。

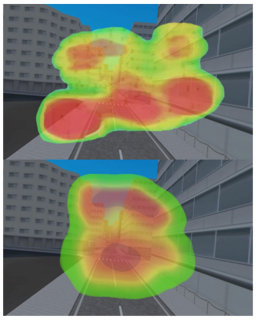


图 11 VR 场景中使用不同导航的眼动热区图

Fig. 11 Eye movement heat map with different navigationin VR scene

被试者在使用永久迷你地图导航时，大部分时间在一直注视迷你地图所在的区域，而将视线转移到辅助导航区域之外的时间非常短，因此对 VR 场景本身不会过多关注。相反，自适应迷你地图使被试者有机会利用自己的启发性认知能力去熟悉 VR 场景中的内容，且不会感到无所适从。此外，被试者在无辅助导航时前往目的地所感受到的阻碍感更强，需要耗费更长的时间去完成任务，虽然这会使得被试者对 VR 场景更加熟悉，但与自适应地图导航相比，仍耗费了太多的时间和精力。

此外，本文还比较了被试者在熟悉 VR 场景前后的眼动数据。图 12 是被试者在同一地点、处于“迷路”和“不迷路”情况下的眼动数据热区图。当被试者处于 A1 情形时，分类模型判定被试者需要导航辅助；当被试者处于 A2 情形时，分类模型则判定其无需导航辅助。通过对比发现，A1 中的被试者眼动数据分布较为分散，分布范围也较大；而 A2 中被试者的数据分布则较为集中。类似地，B1 和 B2 也是被试者在同一地方分别需要导航辅助和不需要导航辅助时的眼动热区图。同样可以发现，B1 中的被试者眼动数据分布较为分散，所涉及的场景内容也较多；而 B2 中被试者更多地只关注路面。

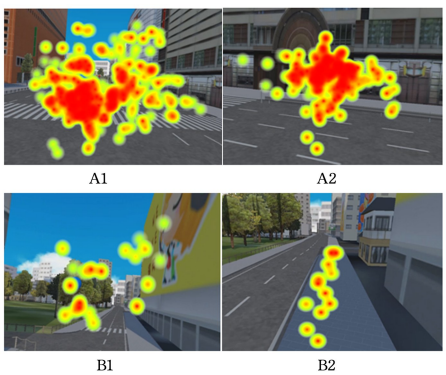


图 12 用户在“迷路”和“不迷路”的情况下在同一地点的眼动热区图

Fig. 12 Users' heat maps from the same place when they are judged as “get lost”(A1, B1) or “not get lost”(A2, B2)

由图 12 可见，当用户在陌生的 VR 场景中迷路时，会更多地搜索和关注标志性建筑、路牌等地标，因此眼动注视点分布分散且分布范围较大[24]；相反，在熟悉的地区时(不迷路)，由于被试用户对路线和地形已经较为熟悉，只需对所在 VR 场景做一个简单的判断就能知道该如何到达目的地，因此眼动注视点数据分布就会较为集中，分布范围也较小。

5.3 主观反馈

任务结束后，向每位被试者发放问卷，以评估在 3 种不同导航辅助模式下被试者的主观体验[25]。问卷中的问题包括：1)导航过程中是否经常会遇到阻碍？2)“迷路”时自适应辅助导航能否及时弹出？其中评分为 1 分(分别代表两个提问中“总是遇到阻碍”和“没有及时弹出”)至 5 分(分别代表两个问题中“从未遇到障碍”和“弹出非常及时”)。具体评分结果如表 4 和表 5 所列。

表 4 是否遇到阻碍主观评价

Table 4 Result of the question “Do you meet obstacle”

主观评价	1 分者/人	2 分者/人	3 分者/人	4 分者/人	5 分者/人	平均值/分
自适应辅助导航	0	0	0	2	2	4.50
永久辅助导航	0	0	0	1	3	4.75
无导航	1	3	0	0	0	1.75

表 5 是否及时弹出导航用户评价

Table 5 Result of the question “Does the navigation aid pop up on time”

主观评价	1 分者/人	2 分者/人	3 分者/人	4 分者/人	5 分者/人	平均值/分
自适应辅助导航	0	0	0	2	2	4.5

被试者对没有地图导航的体验评分集中在 1 分和 2 分，说明他们在 VR 场景中没有地图作为导航辅助时主观体验较差；而使用永久迷你地图导航和自适应迷你地图导航的评分集中在 4 分和 5 分，这表明他们在这两种导航模式下的主观体验更好。

结束语 本文研究了使用注视模式对 VR 环境中用户对导航辅助的需求进行分类，同时设计了一个模型，该模型能够预测用户何时需要导航辅助，并使用这些预测来自适应地帮助他们找到正确的路径。最后，本文通过用户研究评估和验证了所提模型的有效性，揭示了其在 VR 应用领域的潜力。

本文在进行特征选择时，眼动数据只选取了注视点坐标，包含的特征信息还不够全面。在后续的研究中，可以提取更多、更合适的特征进行分析，以提高分类准确率。

参 考 文 献

[1] ALGHOFAILI R, SAWAHATA Y, HUANG H K, et al. Lost in Style: Gaze-driven Adaptive Aid for VR Navigation[C]// CHI Conference. Association for Computing Machinery, New York, NY, USA, 2019: 1-12.

[2] VLCEK K, LACZO J. Neural correlates of Spatial Navigation changes in Mild Cognitive Impairment and Alzheimer's Disease [J]. Front Behav Neurosci, 2014(8): 89.

[3] LAURA P,DE LUCA M,RAFFAELLA N,et al. Navigational Style Influences Eye Movement Pattern during Exploration and Learning of an Environmental Map[J]. Frontiers in Behavioral Neuroscience,2016(10):e43620.

[4] KEVIN S,PAI Y S,KUNZE K. Virtual gaze:exploring use of gaze as rich interaction method with virtual agent in interactive virtual reality content[C]// Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology. ACM,2018:130.

[5] BOYD L A E,GUPTA S,VIKMANI S B,et al. vrSocial:Toward Immersive Therapeutic VR Systems for Children with Autism[C]// The 2018 CHI Conference. Association for Computing Machinery. New York,NY,USA,2018:1-12.

[6] KASAHARA S,ANDO M,SUGANUMA K,et al. Parallel Eyes:Exploring Human Capability and Behaviors with Paralleled First Person View Sharing[C]// CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery,New York,NY,USA,2016:1-2.

[7] GANDRUD J,INTERRANTE V. Predicting destination using head orientation and gaze direction during locomotion in VR [C]// Proceedings of the ACM Symposium on Applied Perception. ACM,2016:31-38.

[8] LEE J,CHEON M,MOON S E,et al. Peripersonal Space in Virtual Reality:Navigating 3D Space with Different Perspectives [C]// Symposium on User Interface Software & Technology. Association for Computing Machinery,New York,NY,USA,2016:207-208.

[9] SAHA D P,KNAPP R B,MARTIN T L. Affective feedback in a virtual reality based intelligent supermarket[C]// Adjunct Acm Joint Conference on Pervasive & Ubiquitous Computing. ACM,2017:646-653.

[10] ZHOU J B. Study and Implementation on the Algorithm of Eye Tracking System[D]. Xi'an:Xidian University,2010.

[11] WU Y Y. Research on the Technology of Eye Tracking[D]. Xi'an:Xidian University,2009.

[12] MIAO L F. Geometry Modeling and Rendering of Point Set Surface[D]. Hangzhou:Zhejiang University,2005.

[13] ZHU M,WAN L,ZHANG Y Y,et al. From 3D Model to Holographic Model:Application of Intelligent Model in Urban Planning-Taking the Urban Planning Assistant System of Chaohu City as an Example[J]. Journal of Anhui Jianzhu University,2019,27(5):36-43.

[14] LIN G L,CAO Y J,BAO Y F. Research on Roaming and Interaction of Ancient Architecture Based on the Integration of 3D Modeling and VR Technology-Taking Jiayuguan Fortress as an Example[J]. Journal of Lanzhou University of Arts and Science (Natural Science Edition),2020,34(1):71-75.

[15] QU H. Using XGBoost to Identify Structures in Time Series [J]. Journal of Jiujiang Vocational and Technical College,2019 (3):80-81.

[16] TANG M Z,ZHAO Q,LONG W,et al. Fault Detection Based on Gradient Boosting Decision Tree for Wind Turbine Gearbox [J]. Hunan Electric Power,2019,39(6):56-60,64.

[17] WANG Y,LI W,WU K H,et al. Application of Fusion Model of GBDT and LR in Encrypted Traffic Identification[J]. Computer and Modernization,2020(3):93-98.

[18] FU F,JIANG J,SHAO Y,et al. An experimental evaluation of large scale GBDT systems[J]. Proceedings of the VLDB Endowment,2019,12(11):1357-1370.

[19] LIU Y Q,LIU Z Y,ZHOU K,et al. Predicting Search User Examination with Visual Saliency[J]. ACM Sigir Forum,2016 (6):619-628.

[20] CHU X F,ZHU C Y. Design and Implementation of 3D Virtual Roaming System[J]. Mechanical and Electrical Information, 2020(6):87-89.

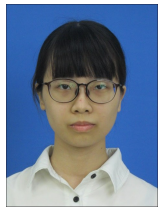
[21] SANCHES C L,KISE K,AUGEREAU O. Japanese Reading Objective Understanding Estimation by Eye Gaze Analysis [C]// UbiComp. ACM,2017:121-124.

[22] AGEEV M,LAGUN D,AGICHTEIN E. Improving search result summaries by using searcher behavior data[C]// Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. Association for Computing Machinery,New York,NY,USA,2013:13-22.

[23] FABRI M,MOORE D,HOBBS D. Mediating the Expression of Emotion in Educational Collaborative Virtual Environments:An Experimental Study[J]. Virtual Reality,2004,7(2):66-81.

[24] LOPES P,ION A,BAUDISCH P. Impacto:Simulating Physical Impact by Combining Tactile Stimulation with Electrical Muscle Stimulation[C]// Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology. Association for Computing Machinery,New York,NY,USA,2015:11-19.

[25] DEY A,PIUMSOMBOON T,LEE Y,et al. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay[C]// CHI Conference on Human Factors in Computing Systems. ACM,2017:4045-4056.



ZHU Chen-shuang, born in 1995, postgraduate. Her main research interests include virtual reality and so on.