

虚拟现实环境中用户视觉注意 的分析与预测

胡志明

北京大学

导师：汪国平

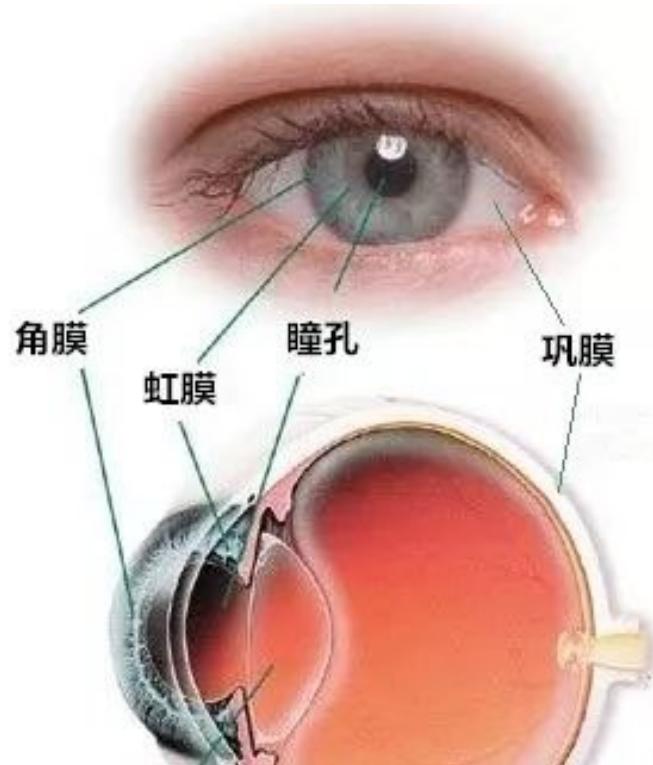
jimmyhu@pku.edu.cn

<https://cranehzm.github.io>

- 研究背景
- 相关工作
- 基于眼动头动协调性的注视预测模型
- 基于卷积神经网络的注视预测模型
- 任务驱动虚拟现实环境中的注视预测模型
- 工作总结

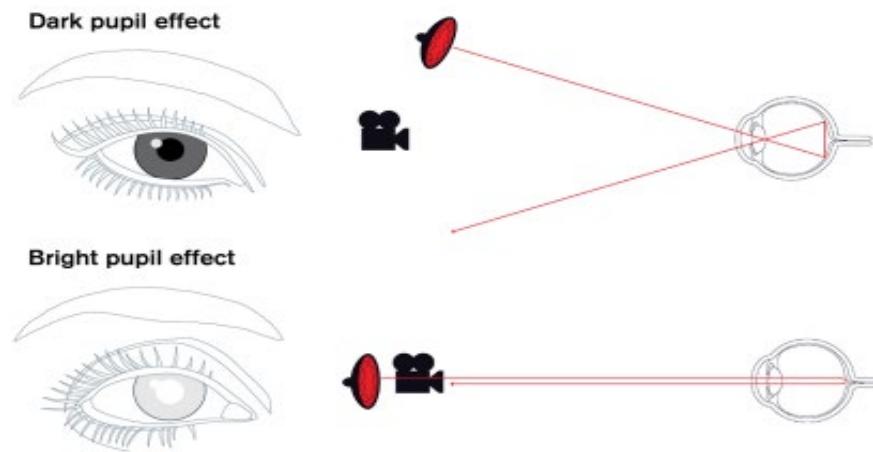
人眼的结构

- 瞳孔：眼球中间黑色圆孔
- 虹膜：瞳孔周围有颜色的部分
- 巩膜：眼白
- 角膜：眼球最前方透明组织



瞳孔角膜反射法原理

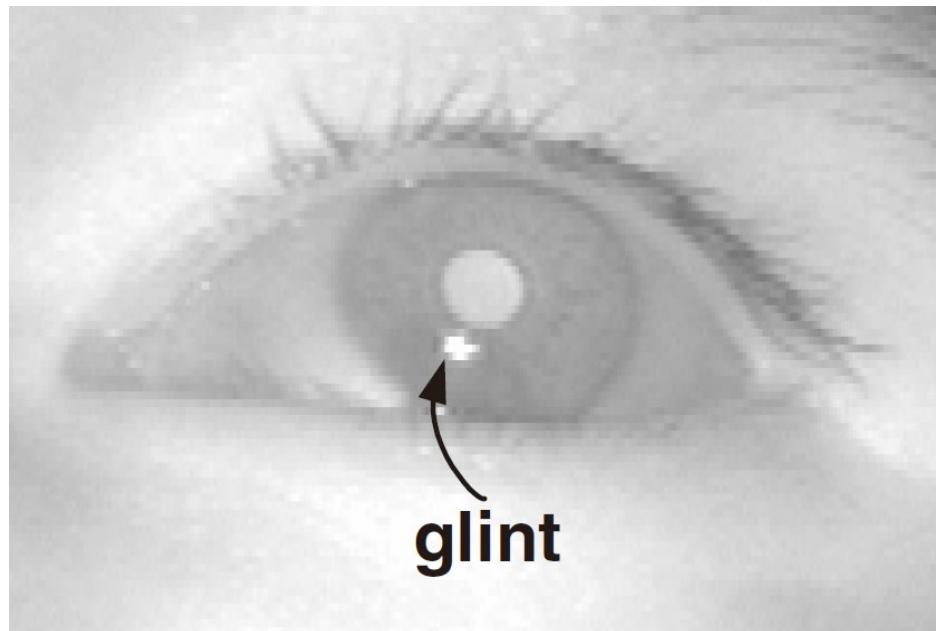
- 暗瞳追踪：光源放置在离成像设备较远的位置(不在同一条光学路径上)，产生瞳孔比虹膜暗的效果
- 明瞳追踪：光源与成像设备在同一条光学路径上，使瞳孔出现发亮的效果



[<https://www.tobiipro.cn>]

瞳孔角膜反射法原理

➤ 角膜反射



[Zhu et al. 2005]

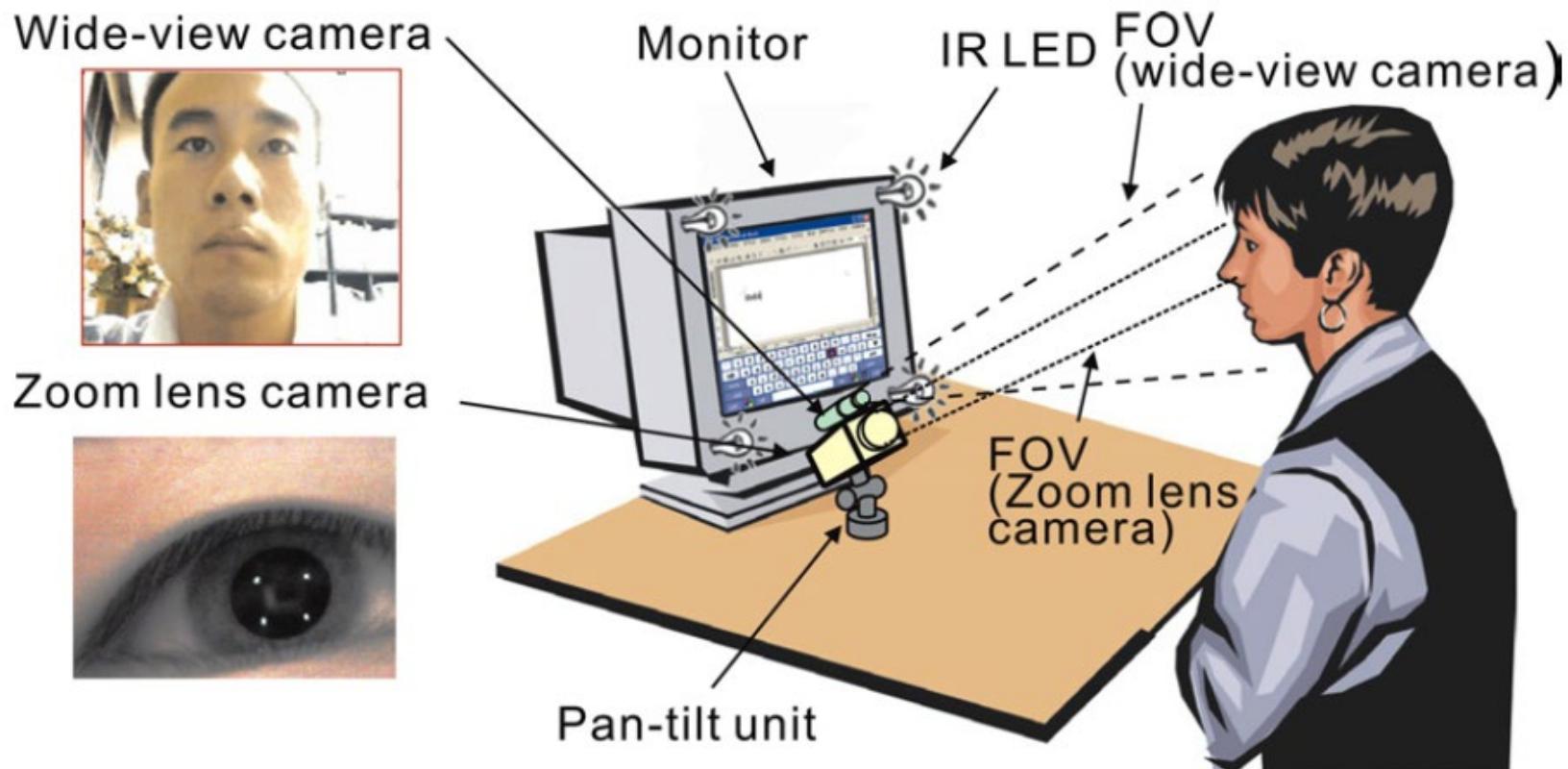
瞳孔角膜反射法原理

➤ 角膜反射



[Duchowski 2007]

瞳孔角膜反射法示意图



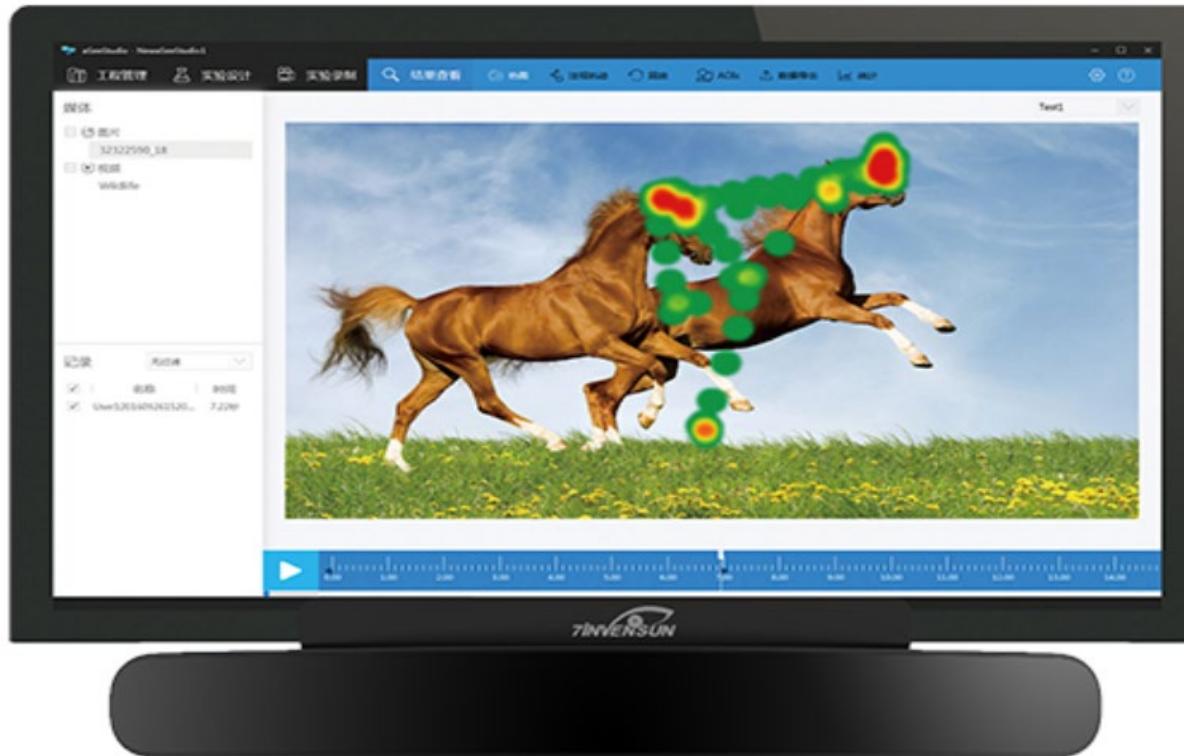
[Yoo et al. 2005]

眼镜式眼动仪



[<https://pupil-labs.com/>]

桌面式眼动仪



[<https://www.7invensun.com/>]

VR眼动仪



[<https://pupil-labs.com/>]

常见应用



a) Original Image



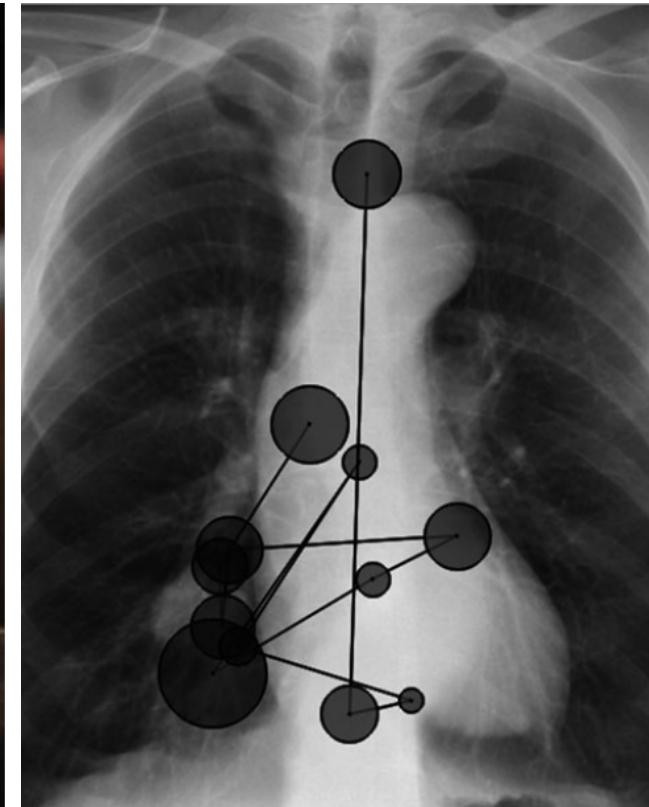
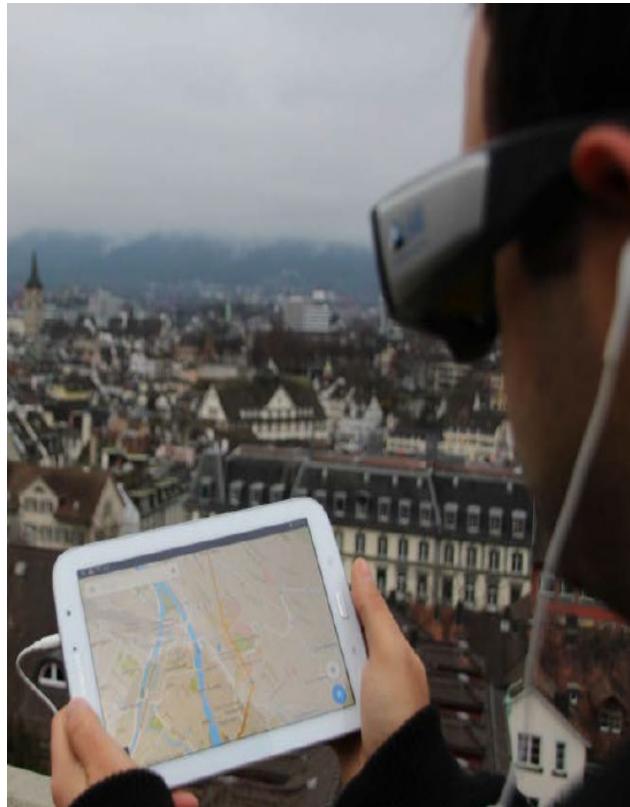
b) Area of Interest



c) Gaze Plot



d) Heat Map



营销策略分析
[Zamani et al. 2016]

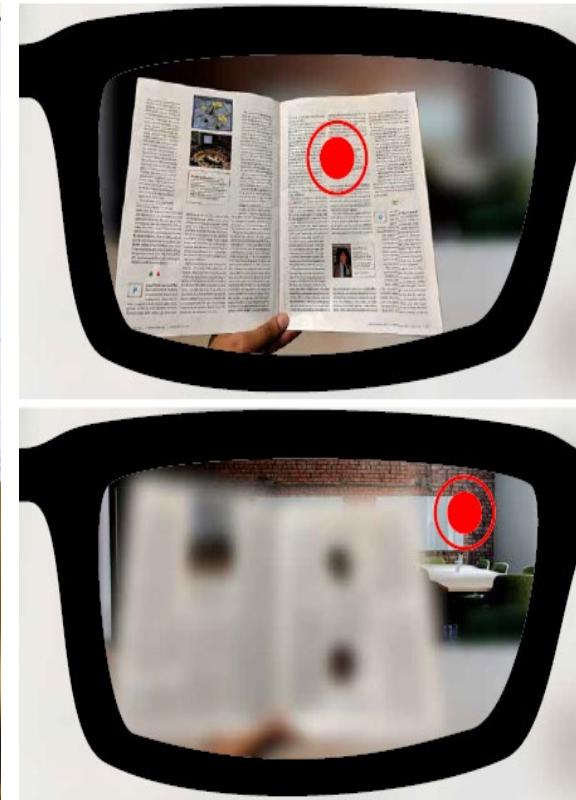
感知研究
[Kiefer et al. 2017]

医学教育
[Kok et al. 2017]

研究背景：眼动追踪应用



常见应用



眼动交互

[Pfeiffer et al. 2008]

协作系统

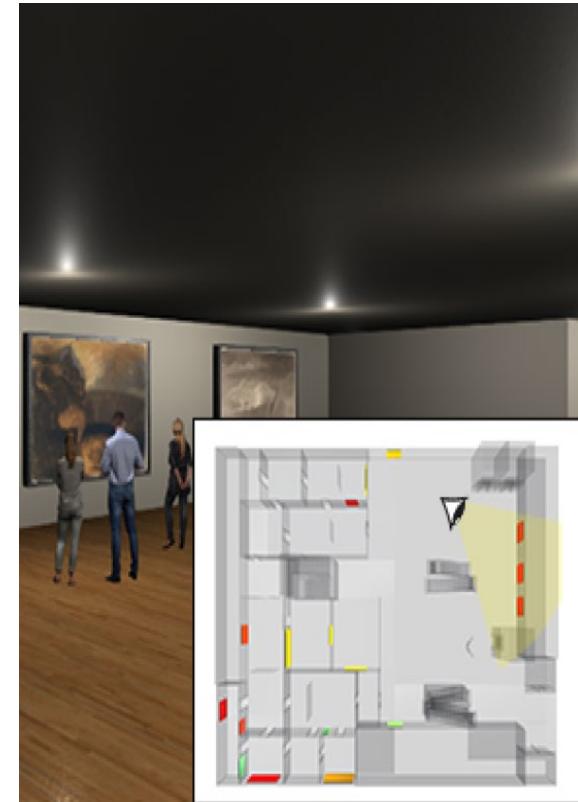
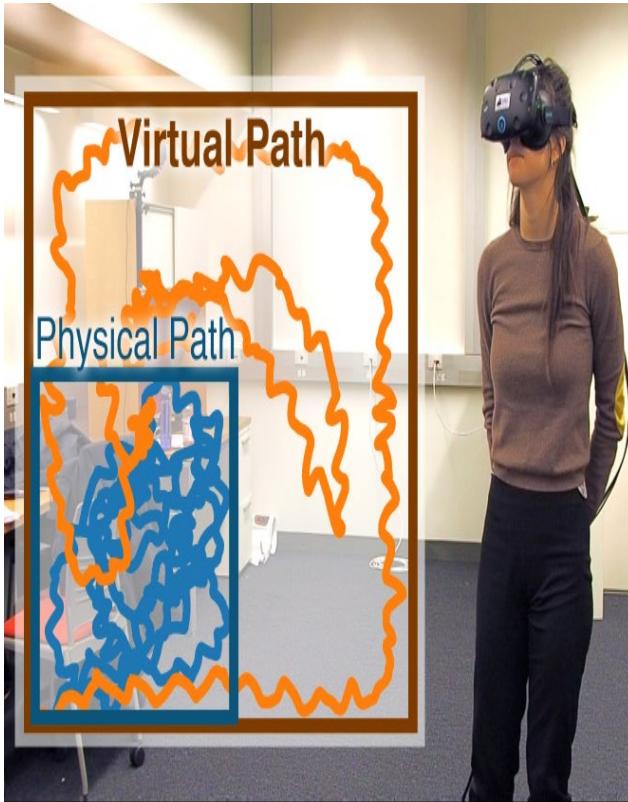
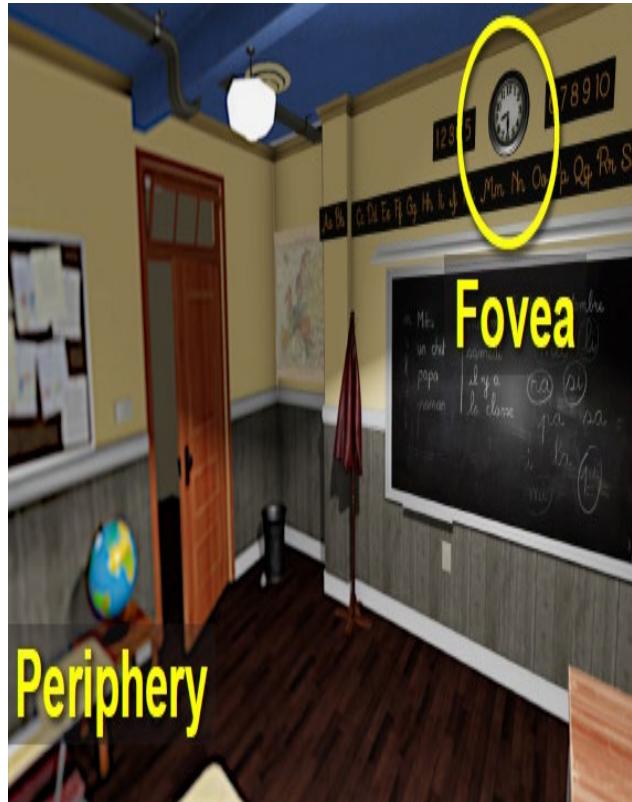
[Zhang et al. 2017]

自动调焦眼镜

[Padmanaban et al. 2019]

研究背景：眼动追踪应用

VR应用

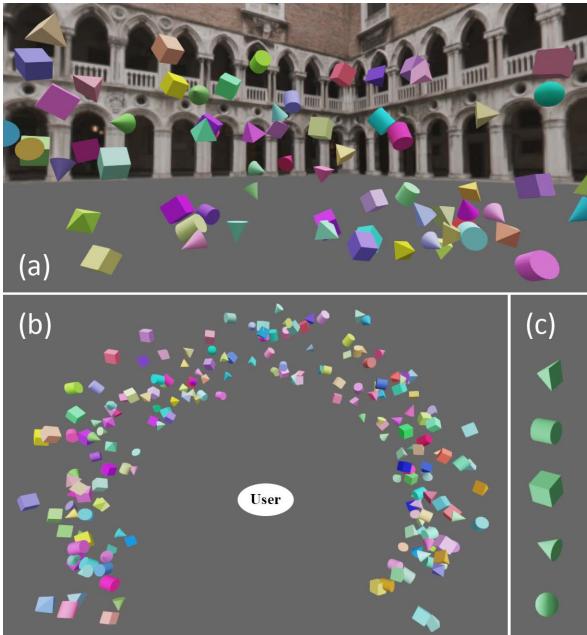
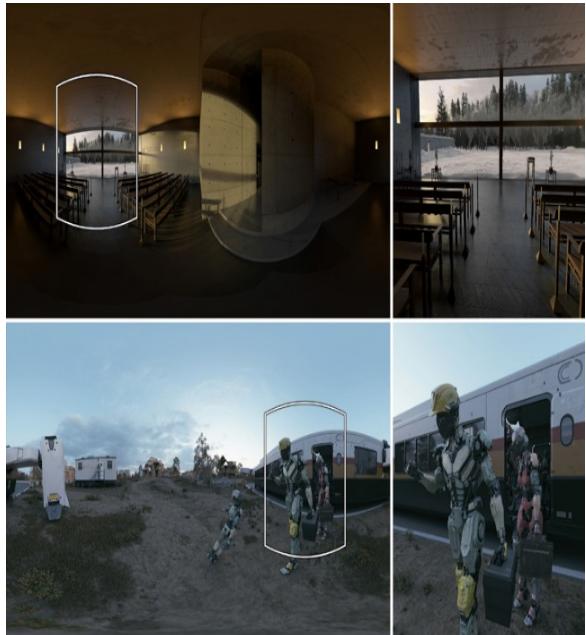


注视点渲染
[Patney et al. 2016]

重定向行走
[Sun et al. 2018]

场景布局优化
[Alghofaili et al. 2019]

VR应用



VR内容设计

[Sitzmann et al. 2018]

注视引导

[Groganick et al. 2017]

细节层次管理

[Lee et al. 2009]

眼动仪的缺陷

➤ 价格昂贵

- Pupil VR/AR (VR/AR眼动仪) 1750欧元
- Pupil Invisible (眼镜式眼动仪) 5500欧元

➤ 校准难度大

- 用户个体差异太大，部分用户难以校准
- 用户眼睛相对位置发生变化可能会导致校准失效

➤ 不能预测用户在未来时刻的注视

- 眼动仪只能提供用户在当前时刻的注视点

虚拟现实环境中注视预测的优势

- 可以直接获取整个场景的全部信息
 - 场景内容信息
 - 三维物体信息：位置、大小、距离等
 - 事件信息：运动物体信息、任务相关物体信息等
- 可以直接获取用户动作信息
 - 头部运动信息
 - 控制器操作信息
 - 手势动作信息

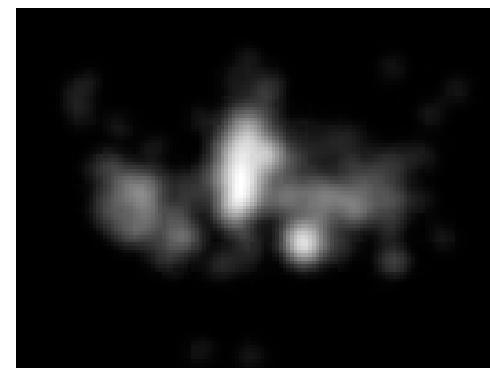
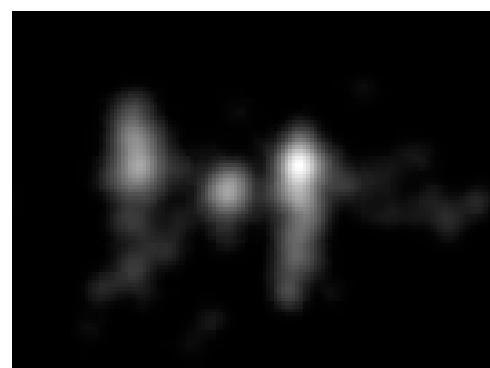
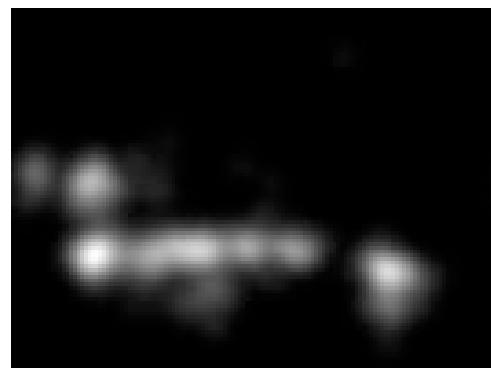
显著性物体检测



上：原始图像；下：显著性物体

[<https://mmcheng.net/msra10k/>]

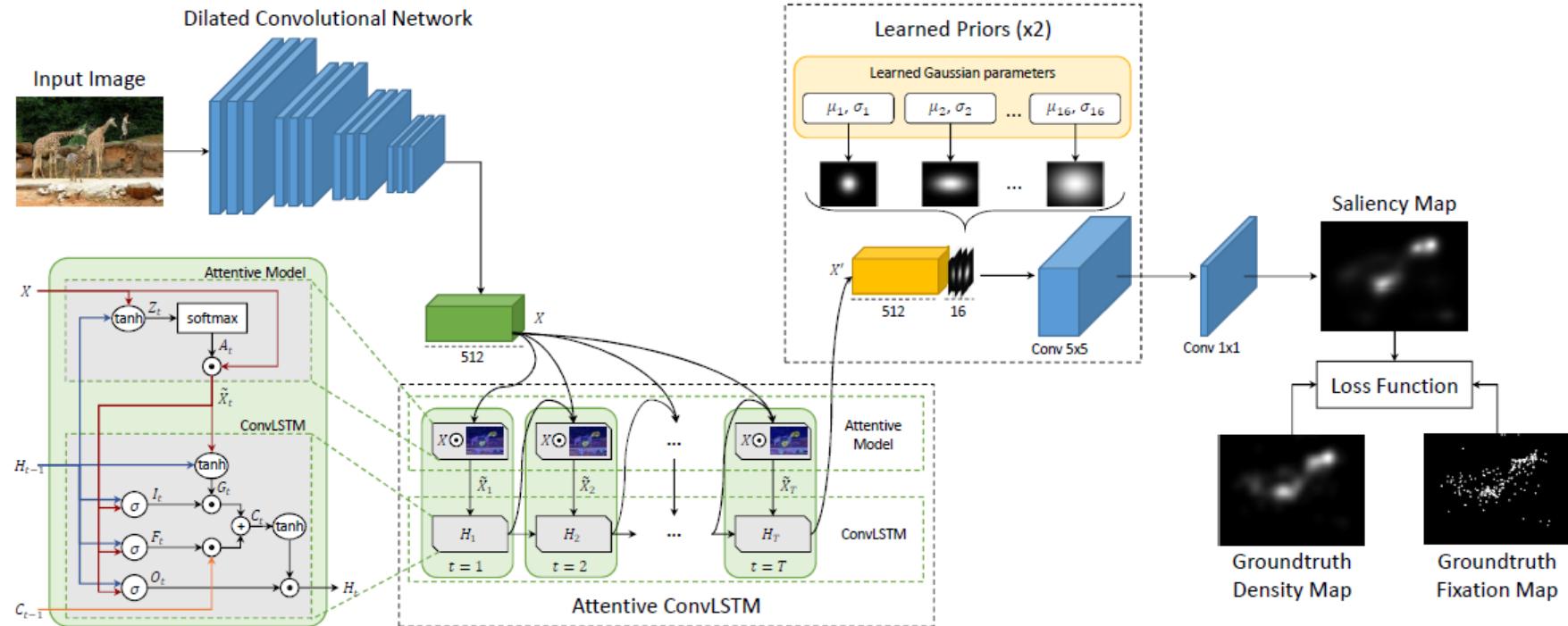
显著性图预测



上：原始图像；下：显著性图

[http://saliency.mit.edu/results_mit300.html]

基于深度学习的显著性预测模型

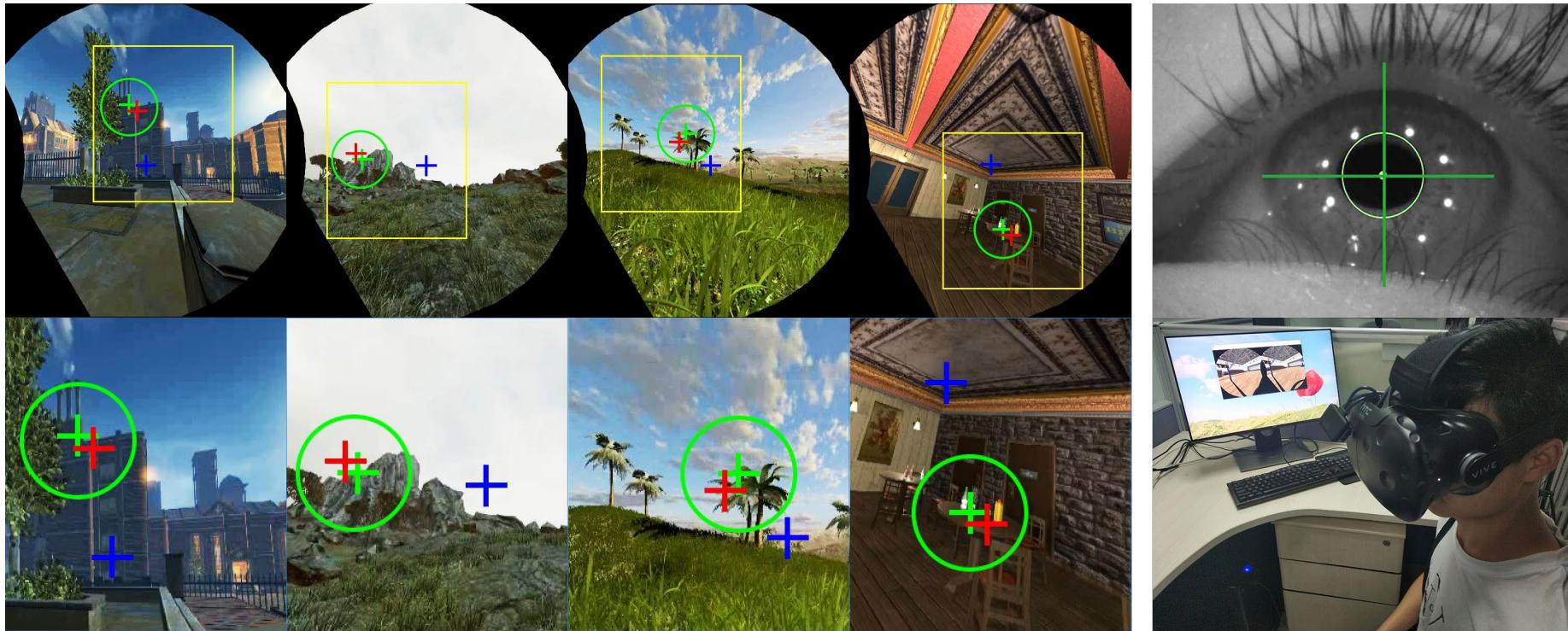


SAM模型
[Cornia et al. 2018]

本文工作 vs. 相关工作

- 预测目标：实时注视位置 vs. 显著性物体、显著性图
- 实验场景：虚拟3D场景 vs. 图像

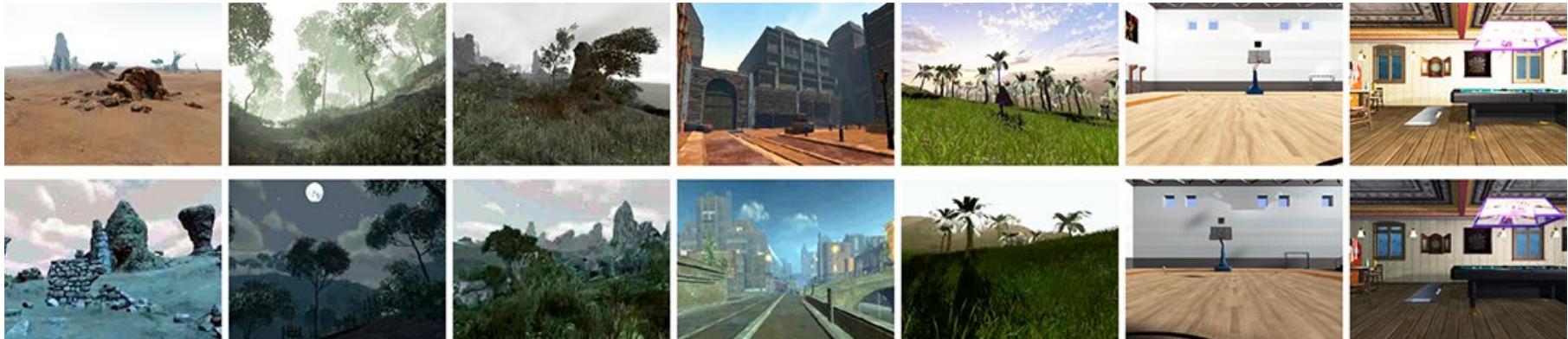
静态虚拟场景注视预测



[Hu et al. TVCG 2019]

数据收集

- 参与者：60名用户（35男，25女，年龄18-36）
- 实验场景：七个静态虚拟场景
- 实验设备：HTC Vive头盔、眼动仪
- 实验过程：自由观察
- 数据：场景内容、注视位置、头部运动



实验场景

眼动头动线性相关性

皮尔逊相关因子 (Pearson's correlation coefficient, PCC)

	Static	Intentional	Sudden	Whole
PCC(v_x)	0.1345	0.5883	0.1511	0.5641
PCC(v_y)	0.1484	0.4969	-0.0906	0.4132

用户注视位置与头部运动之间的相关性

注视位置与头动速度具有很强的线性相关性

SGaze模型

$$x_g(t) = \alpha_x \cdot v_{hx}(t + \Delta t_{x1}) + \beta_x \cdot a_{hx}(t) + F_x(t + \Delta t_{x2}) + G_x(t) + H_x(t)$$
$$y_g(t) = \alpha_y \cdot v_{hy}(t + \Delta t_{y1}) + F_y(t + \Delta t_{y2}) + G_y(t) + H_y(t)$$

x_g, y_g : 注视位置

v_{hx}, v_{hy}, a_{hx} : 头动速度及加速度

F_x, F_y : 内容

G_x, G_y : 任务

H_x, H_y : 其他因素

$\alpha_x, \alpha_y, \beta_x$: 头动速度及加速度的线性影响系数

$\Delta t_{x1}, \Delta t_{y1}$: 眼动头动延迟

眼动头动线性相关性

眼动头动延迟

模型评估

其他方法：Center, Mean, Saliency

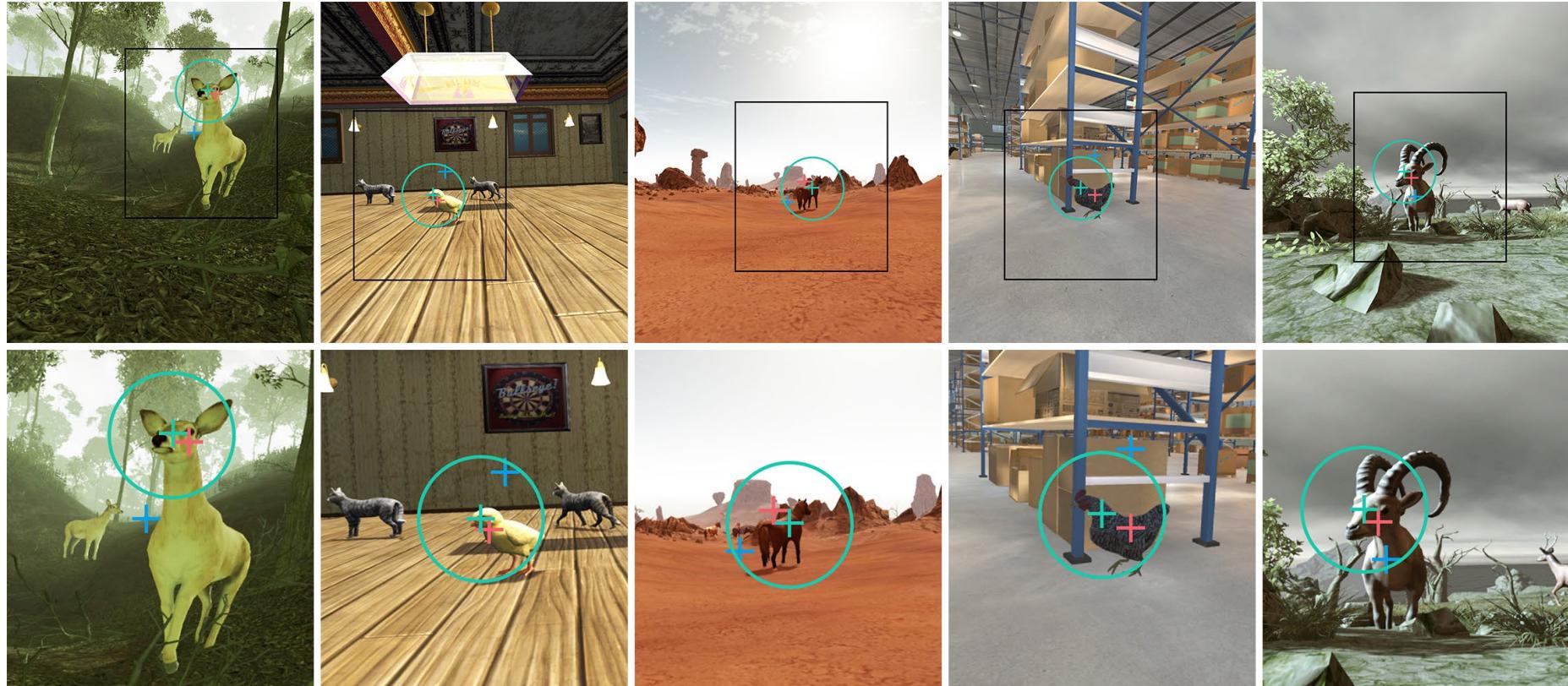
评价指标：视角误差

	SGaze	Mean	Center	Saliency
Mean Error	8.52°	10.93°	11.16°	21.23°
	SGaze	Mean	Center	Saliency

SGaze模型与其他方法的预测误差均值

SGaze模型的预测效果具有显著的提升

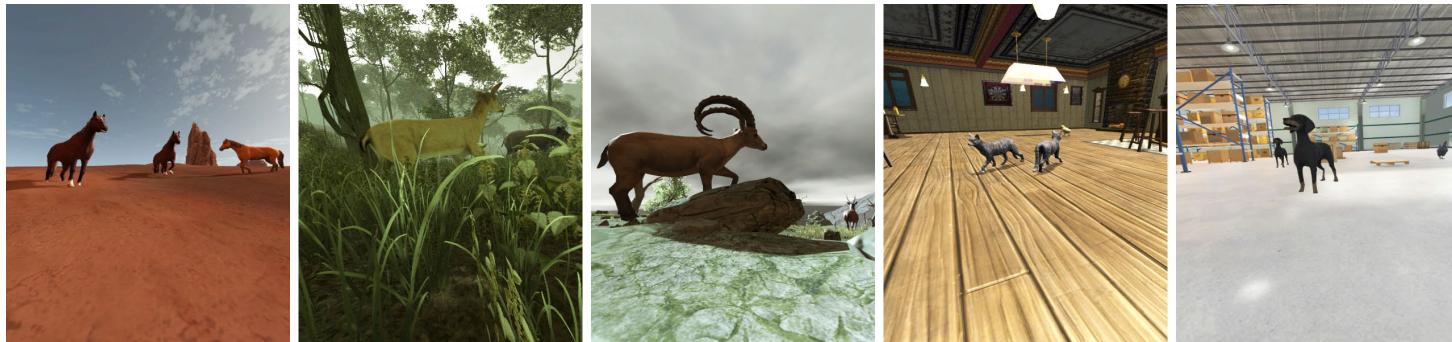
动态虚拟场景注视预测



[Hu et al. TVCG 2020]

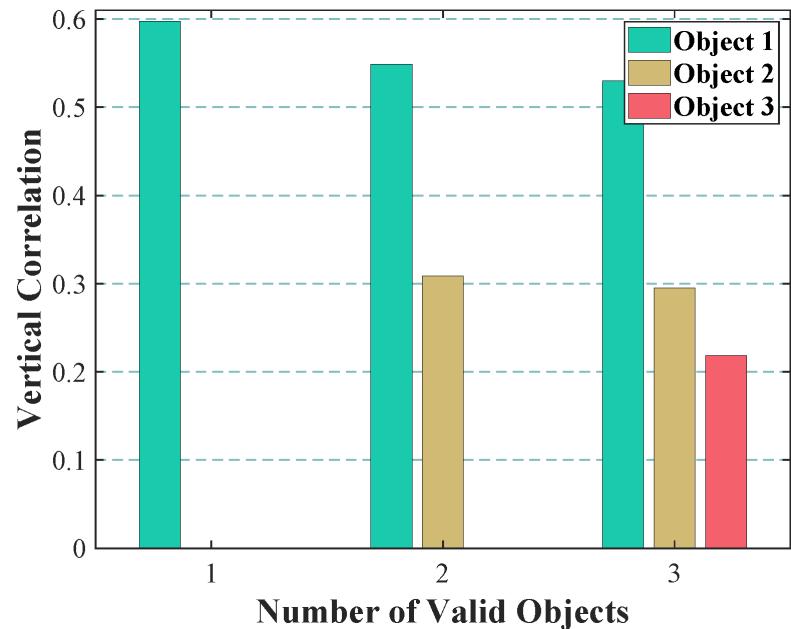
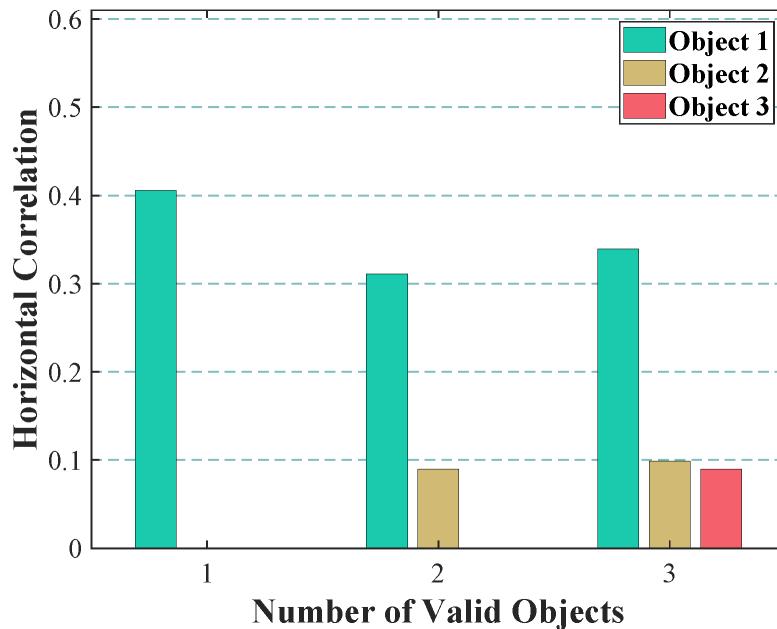
数据收集

- 参与者：43名用户（25男，18女，年龄18-32）
- 实验场景：五个动态虚拟场景
- 实验设备：HTC Vive头盔、眼动仪
- 实验过程：自由观察
- 数据：场景内容、动态物体位置、注视位置、头部运动



实验场景

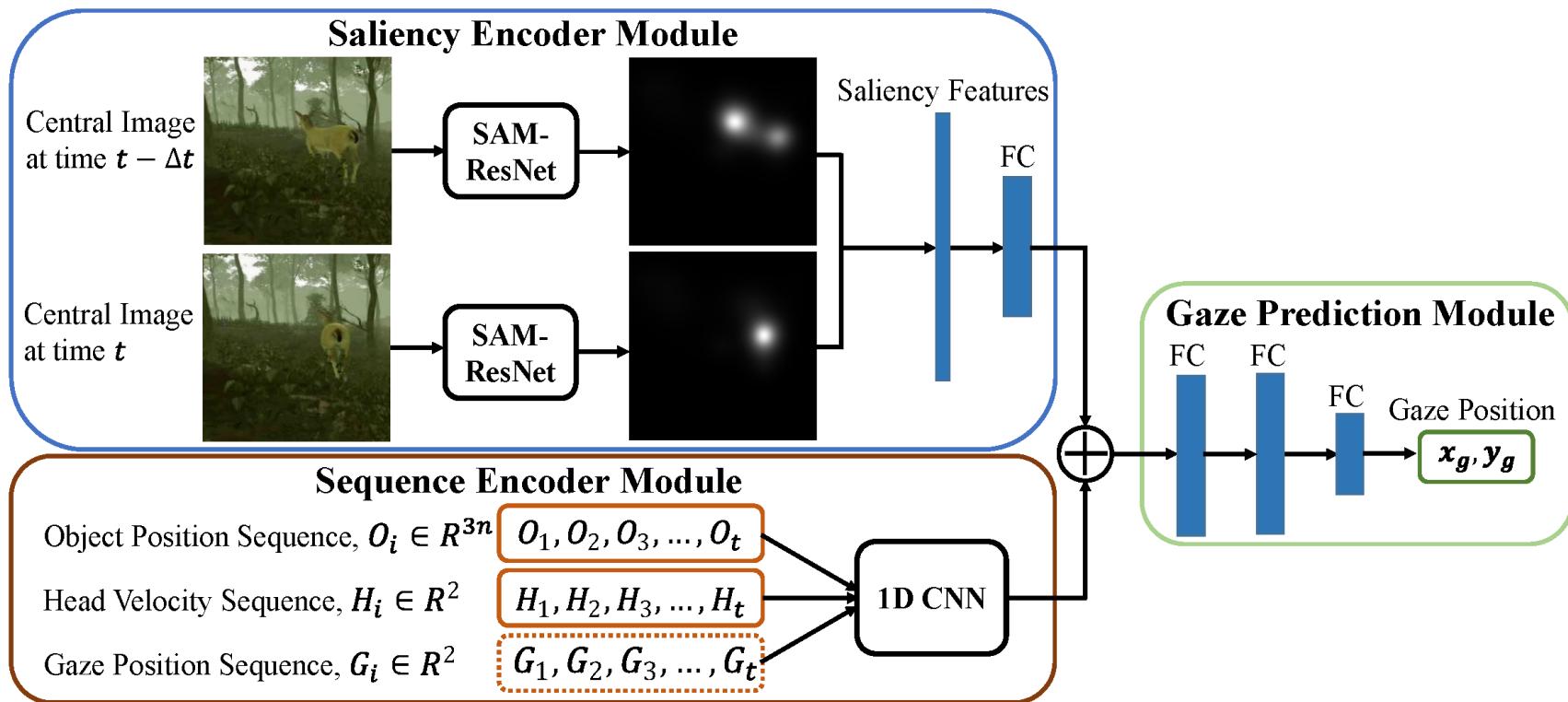
注视与动态物体相关性



注视位置与动态物体位置的相关性
左：水平方向；右：竖直方向

注视与动态物体具有相关性

DGaze模型



DGaze模型框架

模型评估

		DGaze	SGaze	Mean	Center	Object
Dynamic	Mean Error	7.11°	9.11°	10.04°	12.46°	13.25°
Static	Mean Error	7.71°	8.52°	10.93°	11.16°	

DGaze 模型与其他方法在动态和静态数据集上的预测效果

DGaze模型在动态和静态数据集上均有显著的提升

预测效果

Realtime gaze prediction results

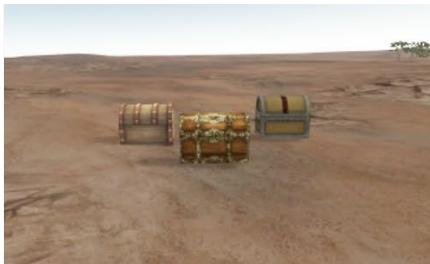
任务驱动虚拟场景注视预测



[Hu et al. TVCG 2021 (Best Journal Nominees)]

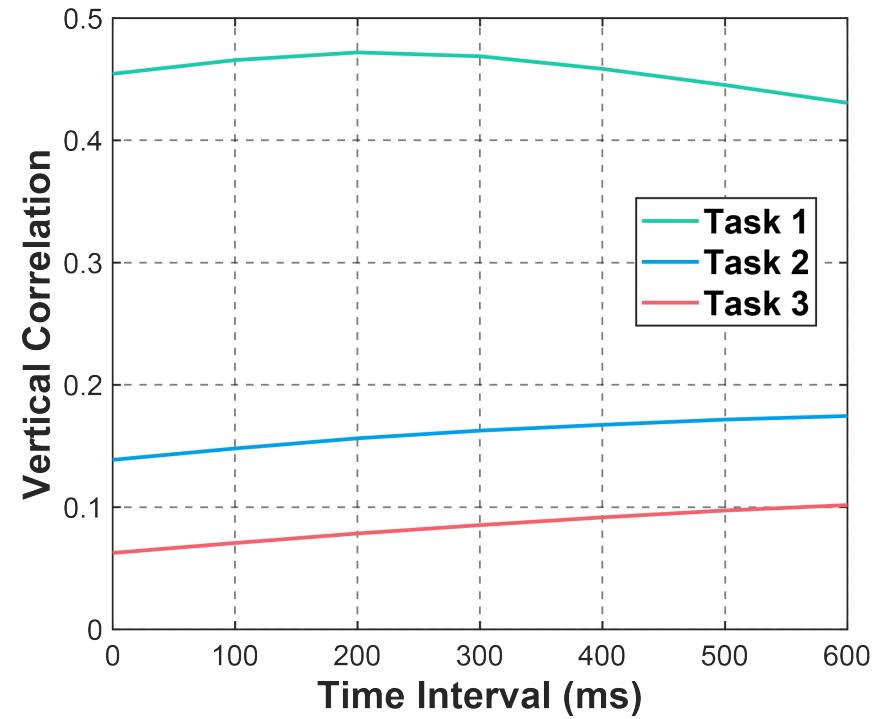
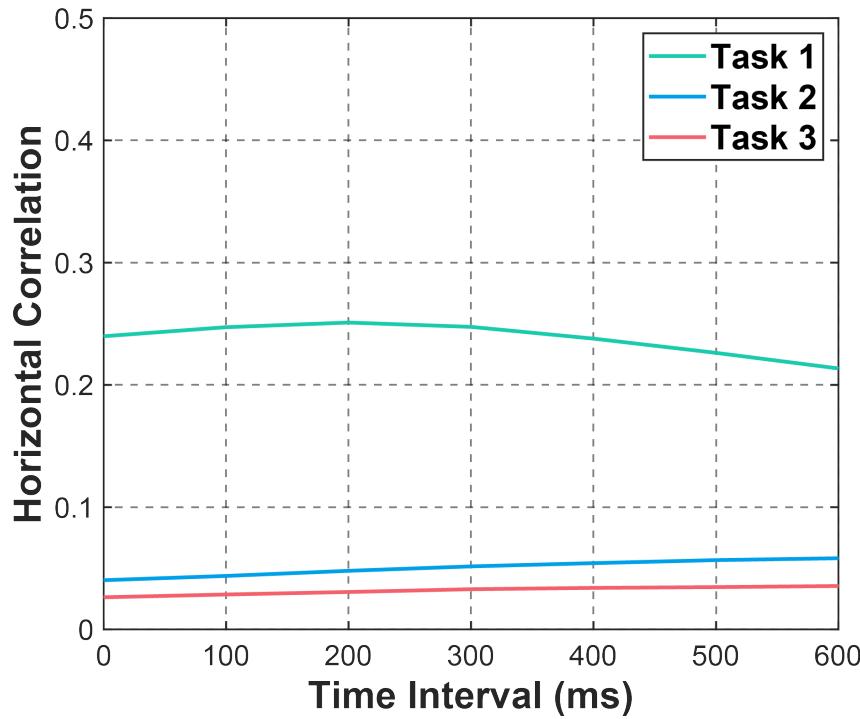
数据收集

- 参与者：27名用户（15男，12女，年龄17-32）
- 实验场景：四个任务驱动虚拟场景
- 实验设备：HTC Vive头盔、眼动仪
- 实验过程：视觉搜索任务
- 数据：场景内容、任务相关物体信息、注视位置、头部运动



实验场景

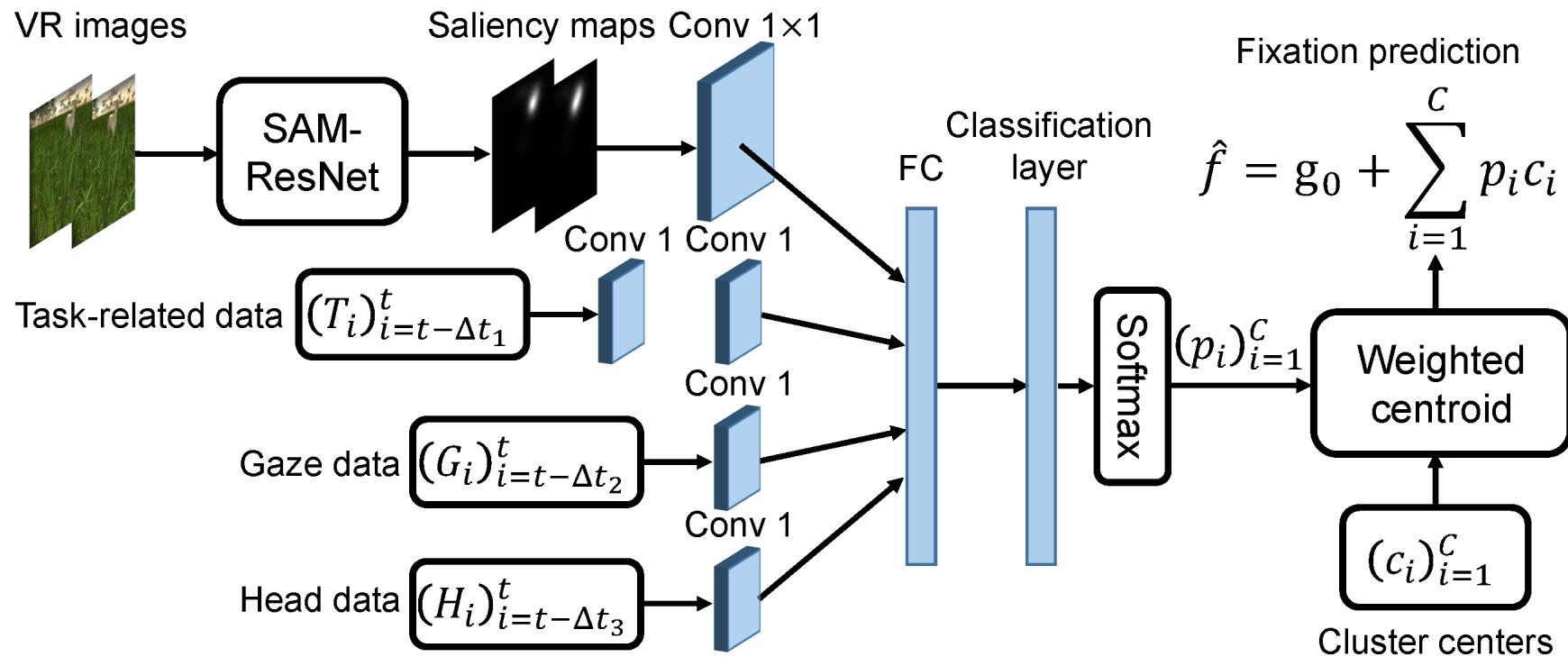
注视-任务相关性



注视位置与任务相关物体的相关性
左：水平方向；右：竖直方向

注视与任务具有相关性

FixationNet模型



FixationNet模型框架

注视预测网络

$$\hat{f} = g_0 + \sum_{i=1}^c p_i c_i$$

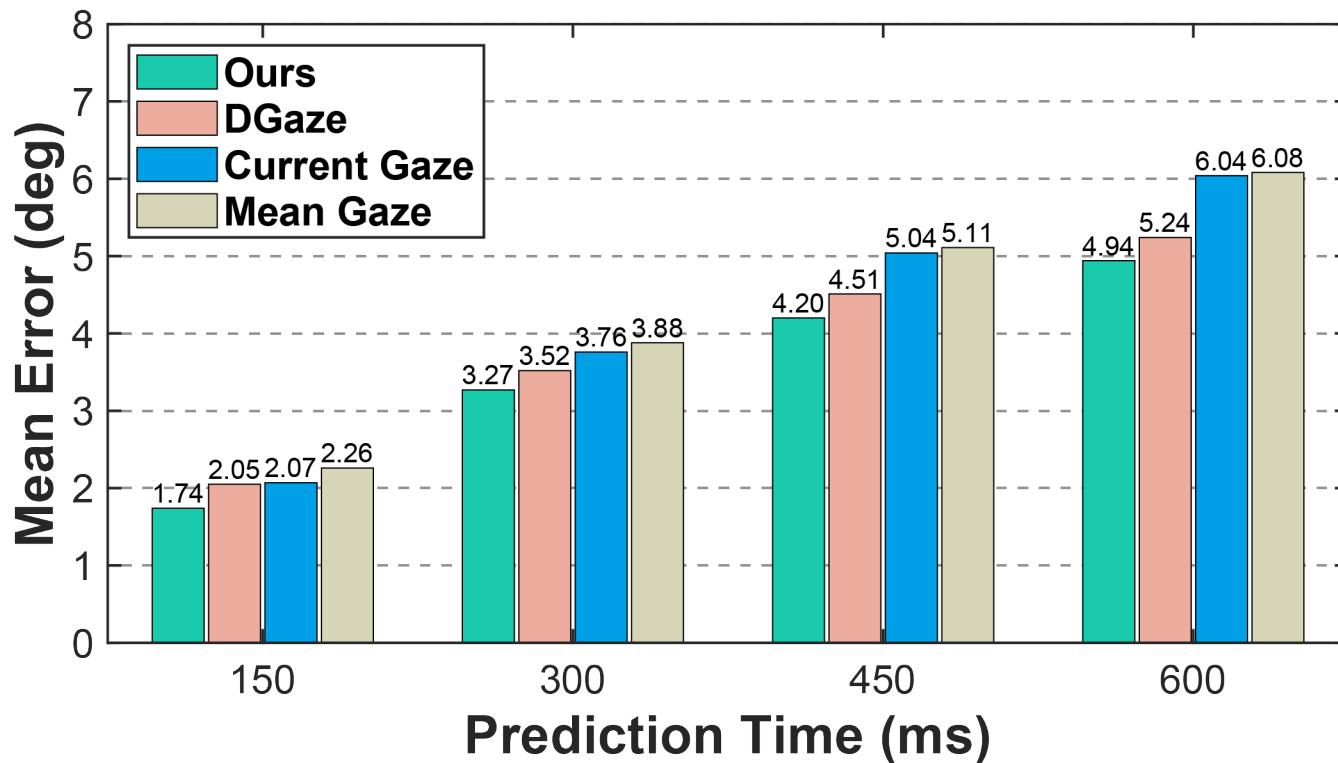
\hat{f} : 预测的未来时刻注视位置

g_0 : 用户在当前时刻的注视位置

c_i : 聚类中心的位置

p_i : 聚类中心的权重

模型评估



任务驱动场景中，不同时间间隔下模型的预测表现

FixationNet在不同预测时间下均具有最好的表现

- 眼动追踪技术在多个领域具有重要的应用
- 现有眼动追踪方案具有价格昂贵、校准难度大、不能预测未来注视等缺陷
- 静态虚拟场景注视预测：SGaze模型
- 动态虚拟场景注视预测：DGaze模型
- 任务驱动虚拟场景注视预测：FixationNet模型

提问环节

谢谢