
IEEE VR
IEEE COMPUTER SOCIETY
IEEE
27 MARCH - 3 APRIL • 2021 VIRTUAL
Make Virtual Reality diverse and accessible

FixationNet: Forecasting Eye Fixations in Task-Oriented Virtual Environments

Zhiming Hu[1], Andreas Bulling[2], Sheng Li[1], Guoping Wang[1]

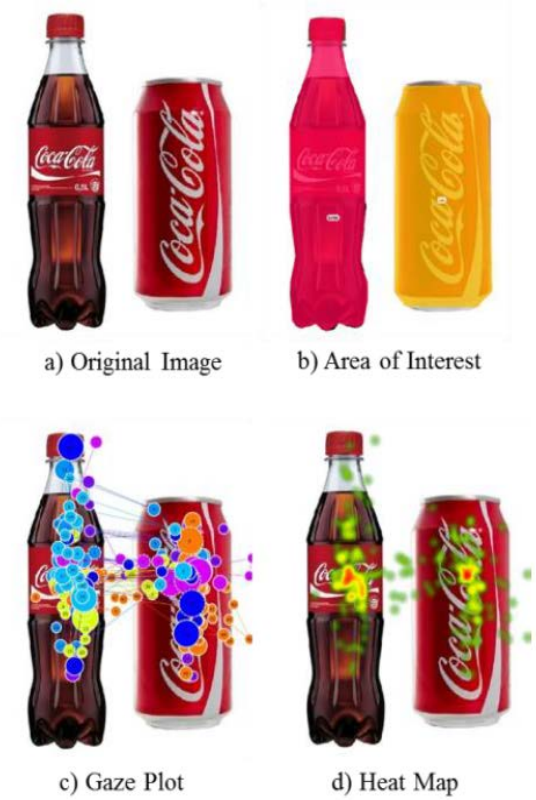
[1]Peking University
[2]University of Stuttgart

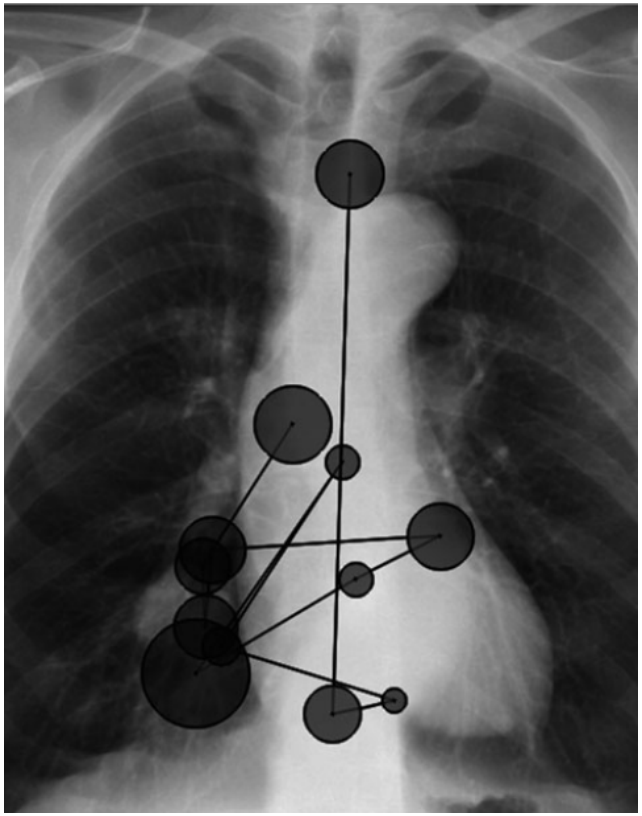
Project URL: cranehzm.github.io/FixationNet

Human Visual Attention

北京大学 PEKING UNIVERSITY

Application of Visual Attention





Marketing Strategy Analysis [Zamani et al. 2016]

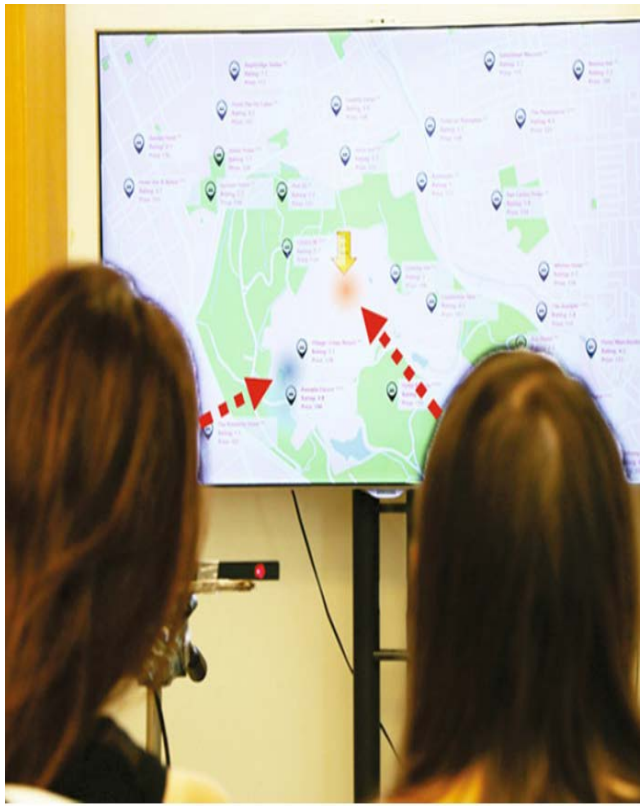
Cognitive Research [Kiefer et al. 2017]

Medical Education [Kok et al. 2017]

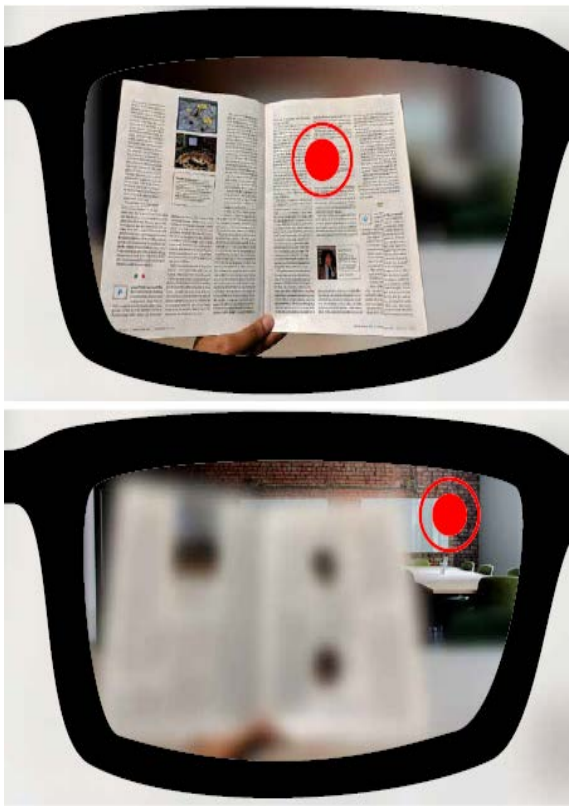
北京大學 PEKING UNIVERSITY

Application of Visual Attention

Gaze-based Interaction
[Pfeiffer et al. 2008]



Collaborative System
[Zhang et al. 2017]

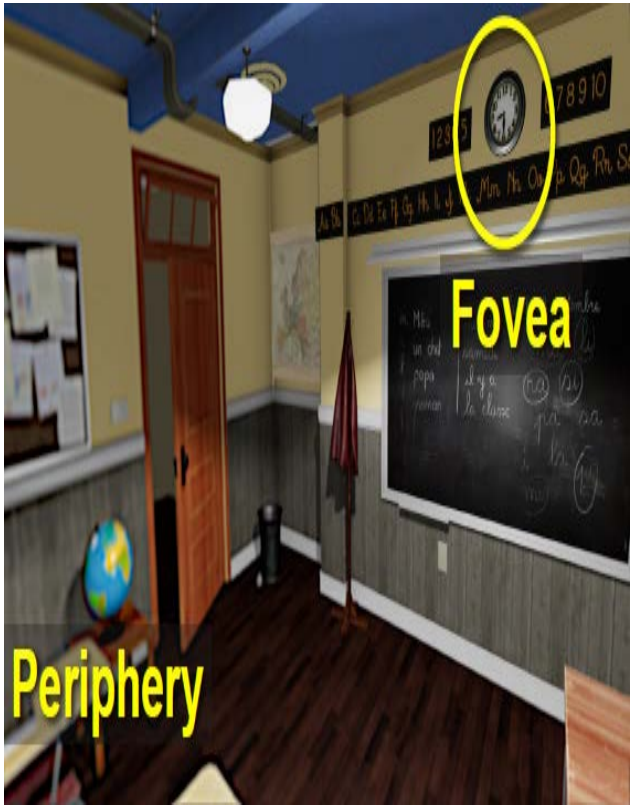


Gaze-contingent Eyeglasses
[Padmanaban et al. 2019]

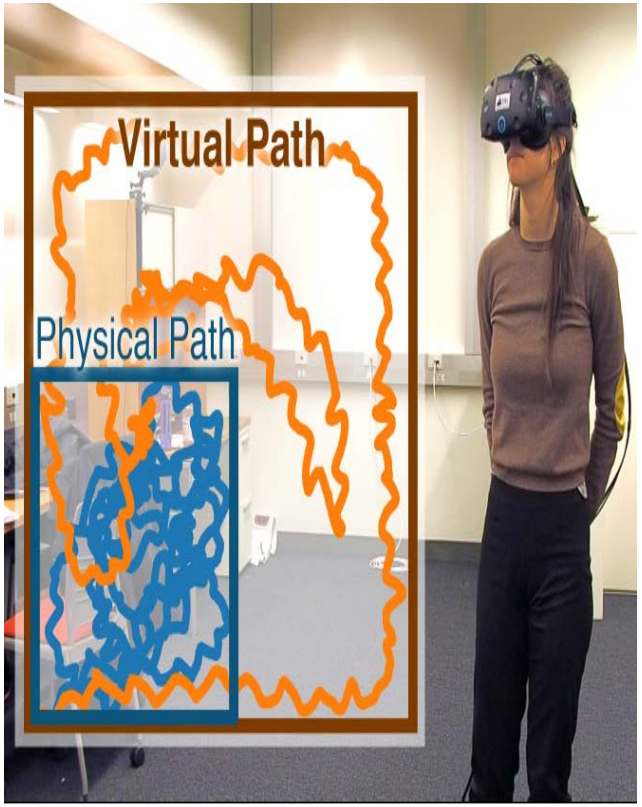
Background

北京大學 PEKING UNIVERSITY

Application of Visual Attention in VR



Gaze-contingent Rendering
[Patney et al. 2016]



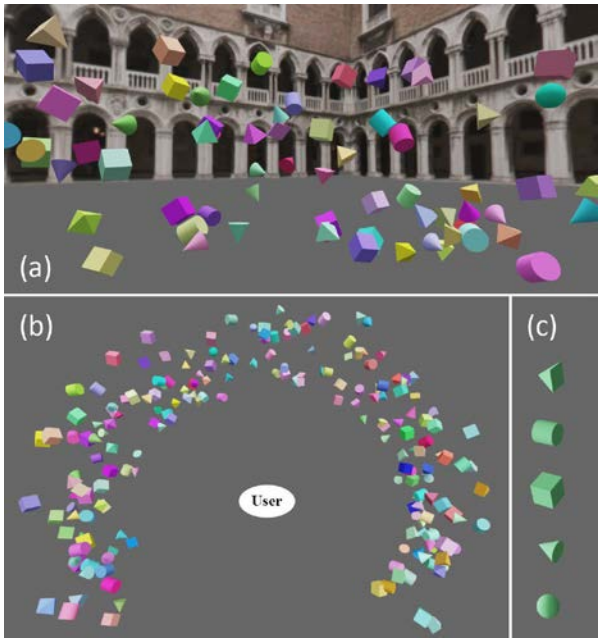
Redirected Walking
[Sun et al. 2018]

Layout Optimization
[Alghofaili et al. 2019]

北京大學 PEKING UNIVERSITY

Application of Visual Attention in VR

VR Content Design
[Sitzmann et al. 2018]



Gaze Guidance
[Grogorick et al. 2017]

LOD Management
[Lee et al. 2009]

Background

北京大學 PEKING UNIVERSITY

Research Goals

- Analyze and reveal the characteristics of users' task-oriented visual attention in virtual reality
- Forecast (temporally predict future) eye fixations based on the characteristics of visual attention

北京大學 PEKING UNIVERSITY

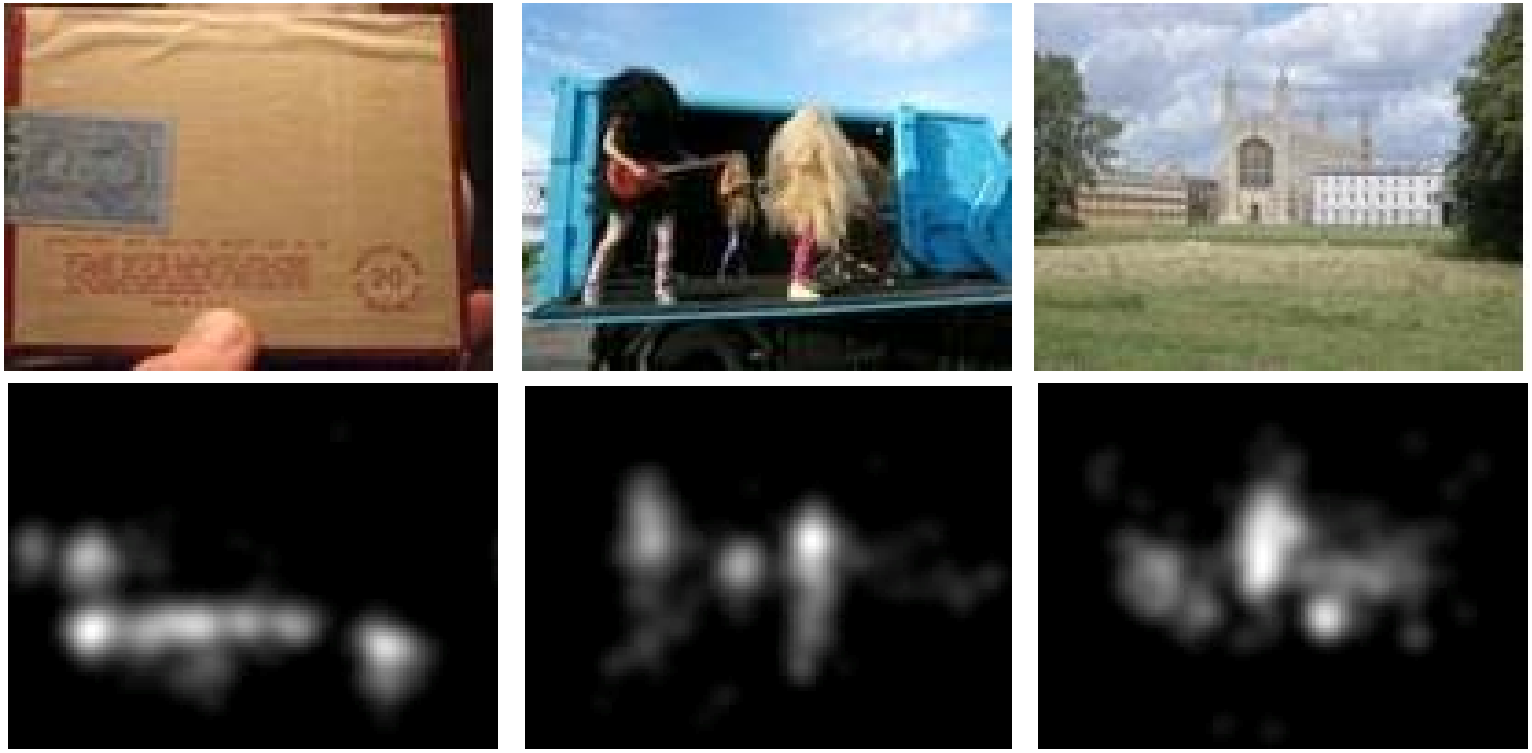
Salient Object Detection

Top: Original Images; Bottom: Salient Objects

[1] https://mmcheng.net/msra10k/

北京大學 PEKING UNIVERSITY

Saliency Prediction

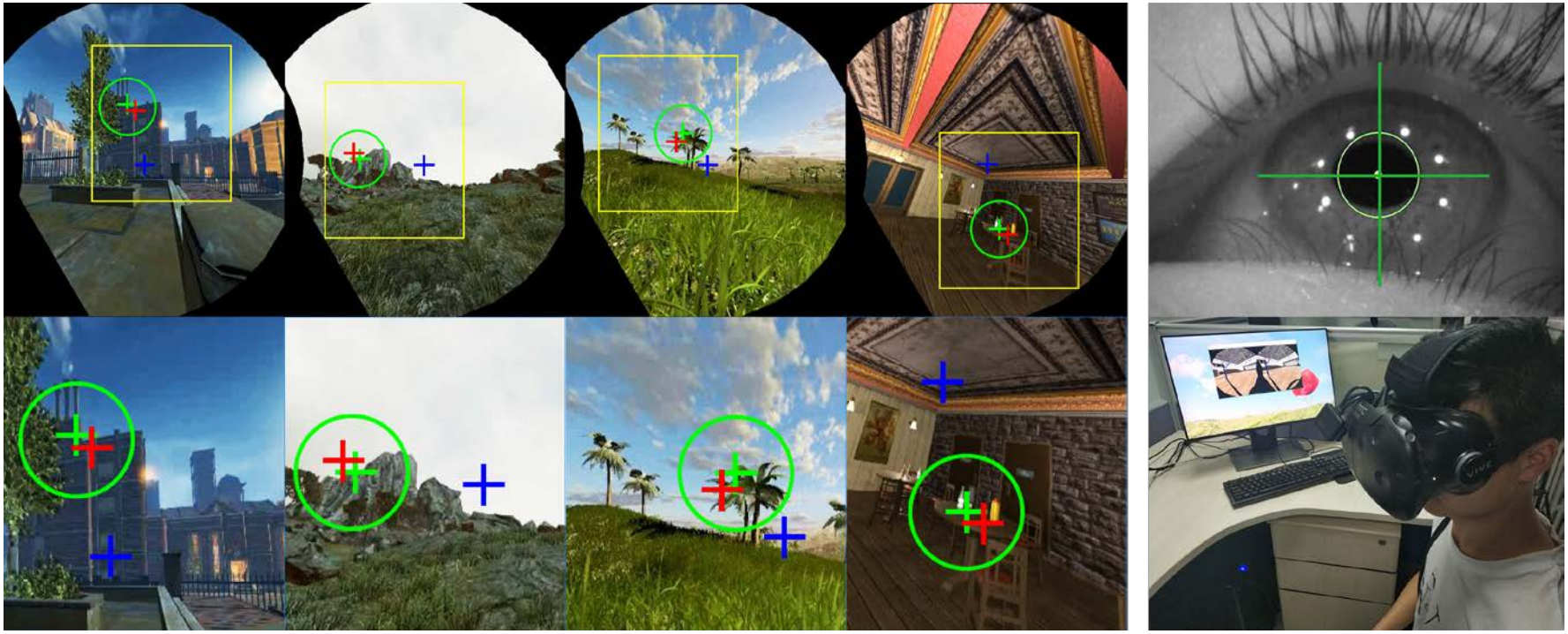


Top: Original Images; Bottom: Saliency Maps

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

北京大學 PEKING UNIVERSITY

Visual Attention Prediction in VR

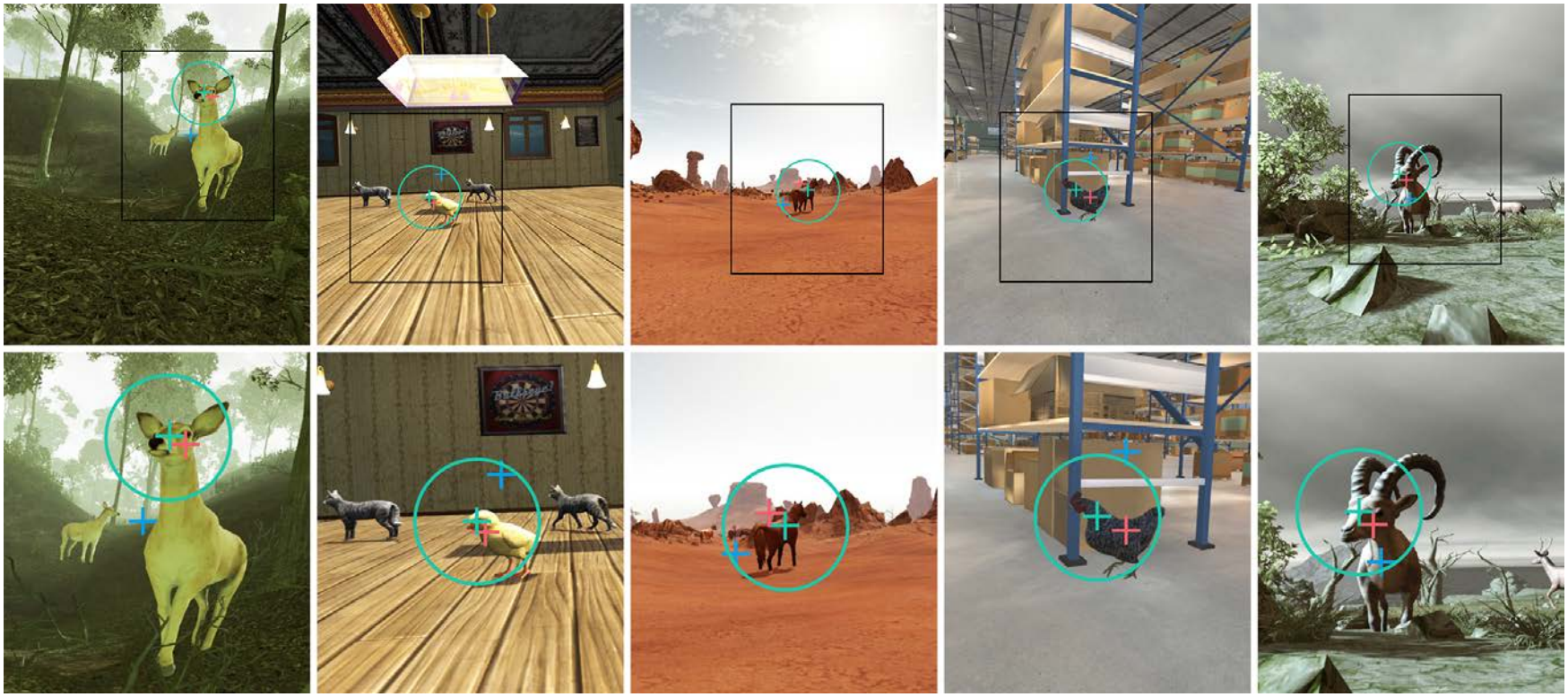


Gaze Prediction in Static Free-Viewing Virtual Environments
[Hu et al. 2019]

Related Work

北京大學 PEKING UNIVERSITY

Visual Attention Prediction in VR



Gaze Prediction in Dynamic Free-Viewing Virtual Environments
[Hu et al. 2020]

Related Work

北京大學 PEKING UNIVERSITY

Our Work *vs.* Previous Work

- Prediction Goal

Eye Fixations *vs.* Salient Objects, Saliency Maps

- Scene

Immersive Virtual Environments *vs.* Images, Videos

Task-Oriented Situations *vs.* Free-Viewing Conditions

北京大學 PEKING UNIVERSITY

Contributions

- Propose a novel learning-based fixation prediction model (*FixationNet*)
- Analyze and reveal the characteristics of users' task-oriented visual attention in VR
- Build a task-oriented VR eye tracking dataset

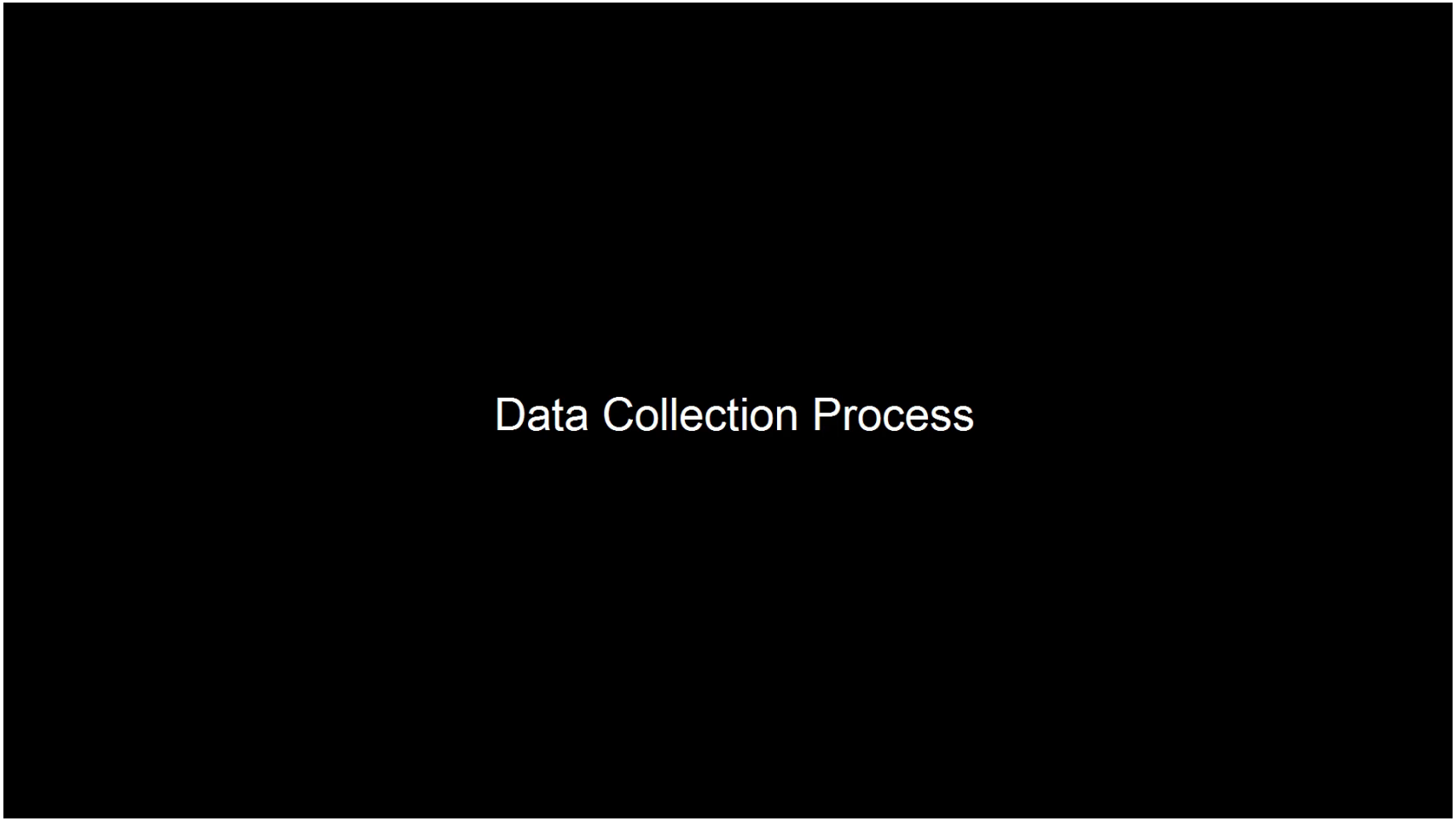
Data Collection

北京大学 PEKING UNIVERSITY

- Participants: 27 users (15 male, 12 female, ages 17-32)
- Stimuli: four immersive virtual environments
- Apparatus: HTC Vive, eye tracker
- Procedure: visual search task
- Data: VR content, task-related objects, eye fixations, head movements

Stimuli

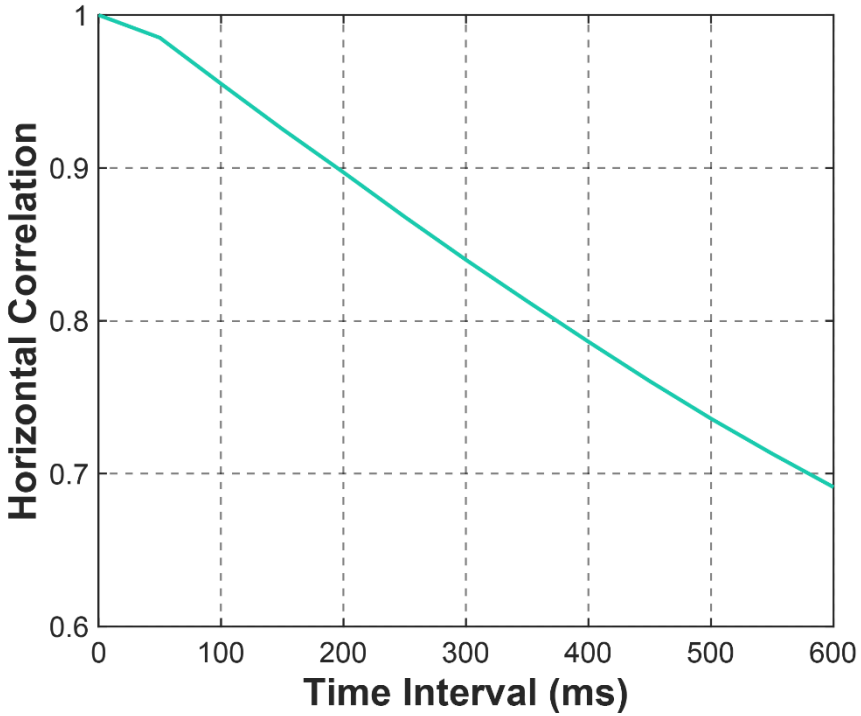
Data Collection Process

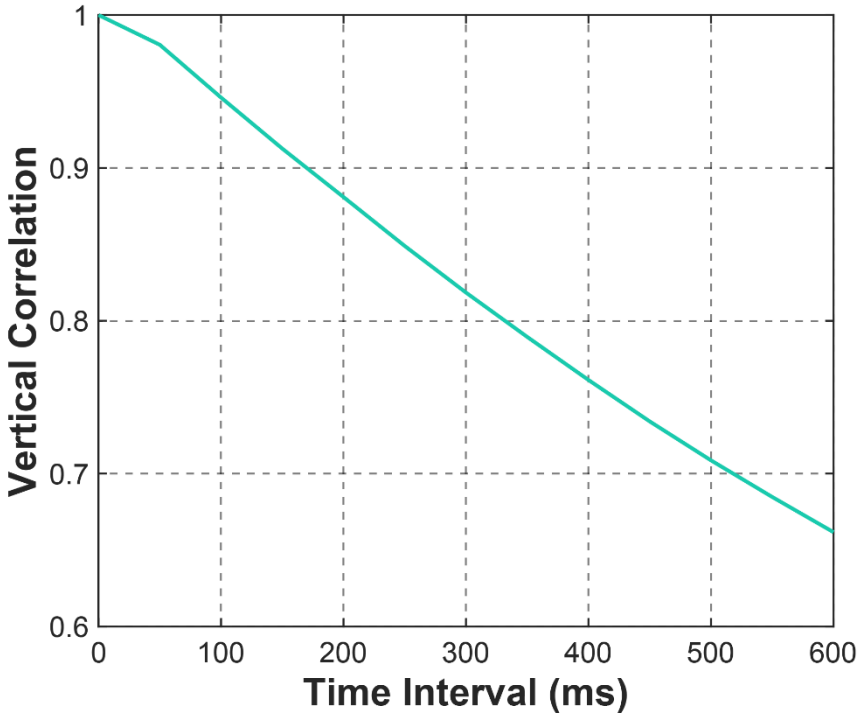


Analysis of Eye Fixation

北京大學 PEKING UNIVERSITY

Fixation-Gaze Correlation





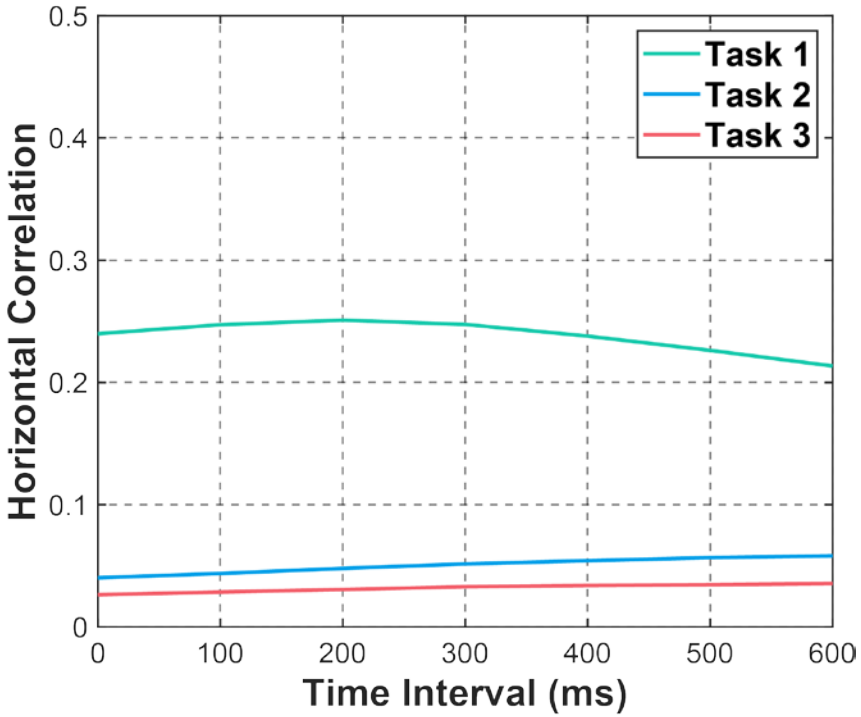
Fixation-gaze correlations in the horizontal (left) and vertical (right) directions

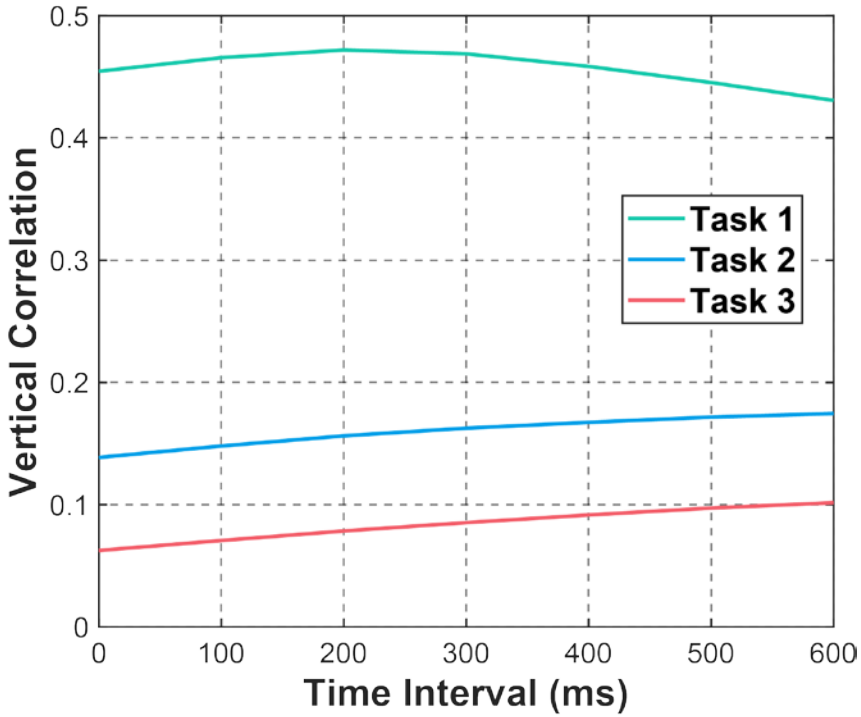
Eye fixations are highly correlated with historical gaze positions

Analysis of Eye Fixation

北京大學 PEKING UNIVERSITY

Fixation-Task Correlation





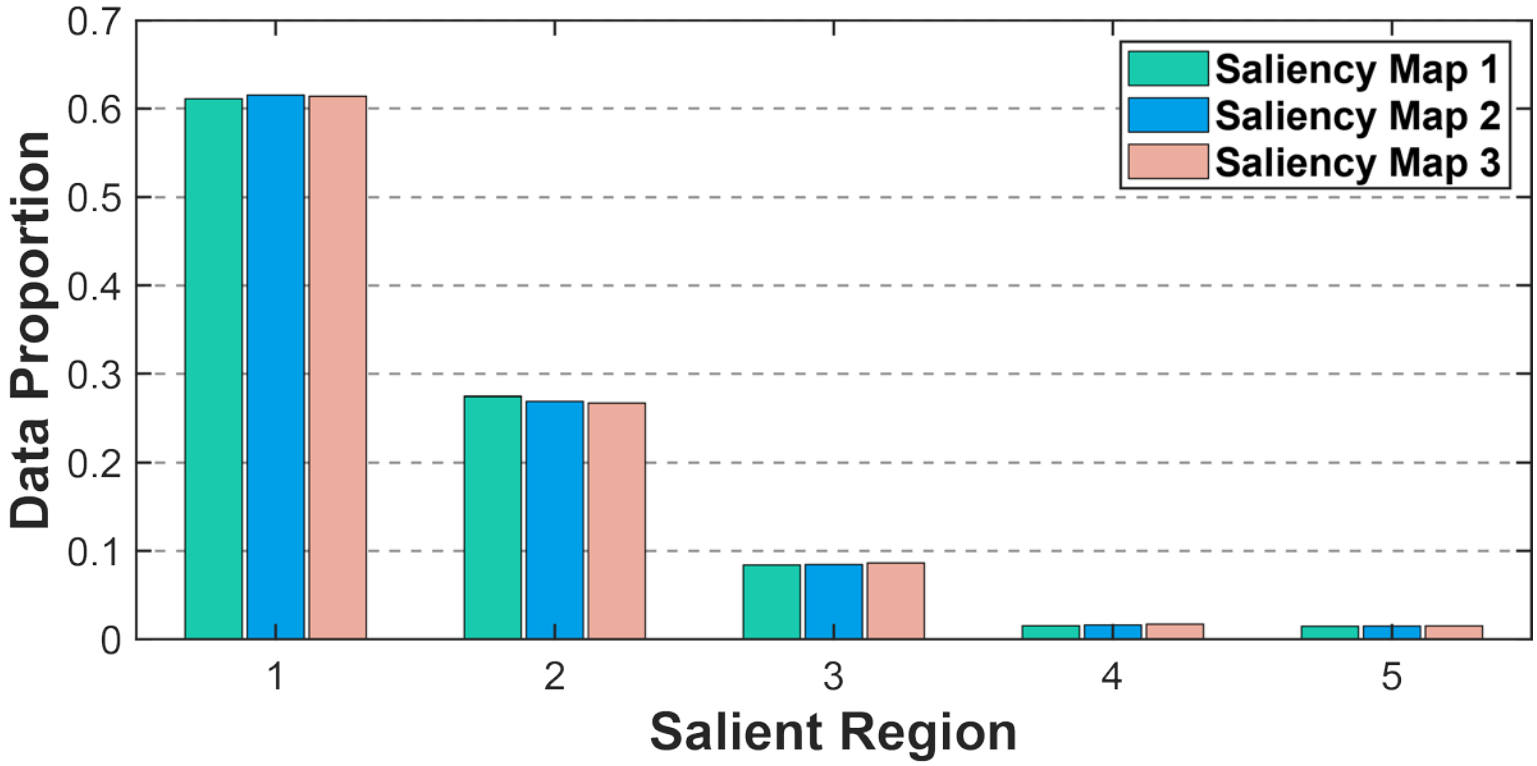
Fixation-task correlations in the horizontal (left) and vertical (right) directions

Fixations are correlated with task-related objects

Analysis of Eye Fixation

北京大學 PEKING UNIVERSITY

Fixation-Saliency Correlation



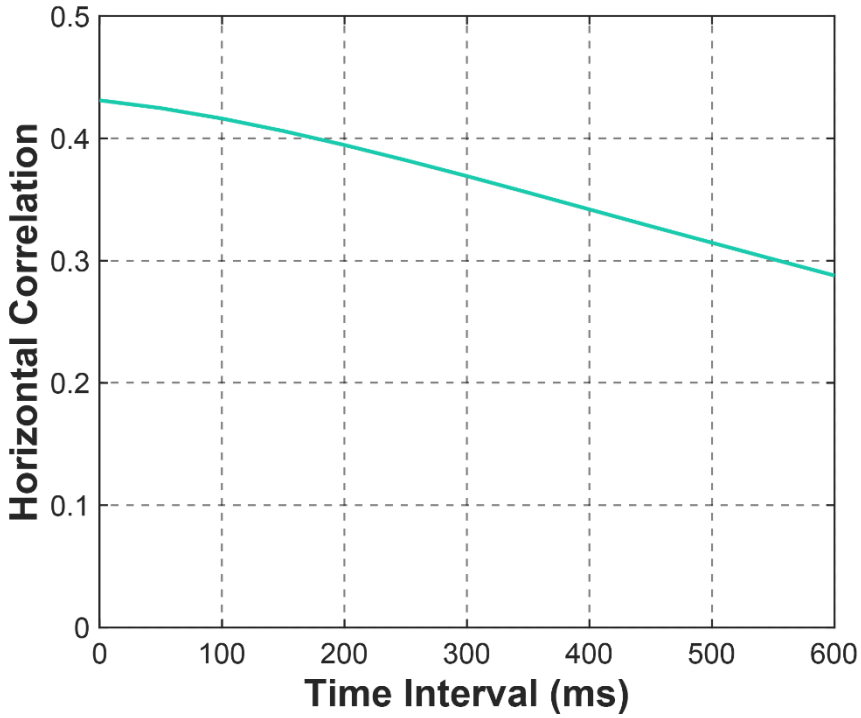
The distribution of users' fixation positions on the salient regions

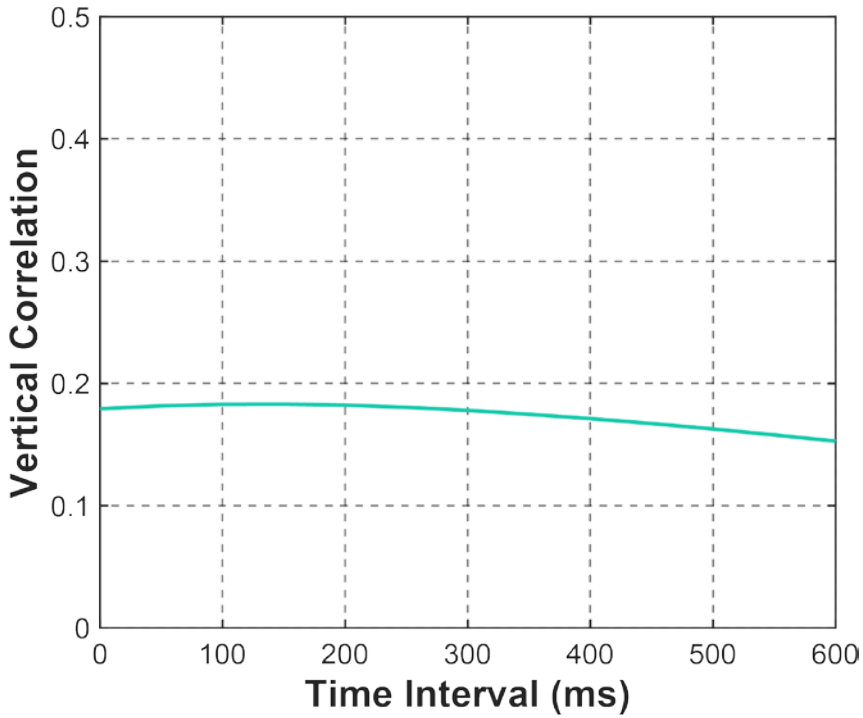
The fixation positions are mostly located in the regions with high saliency values

Analysis of Eye Fixation

北京大學 PEKING UNIVERSITY

Fixation-Head Correlation

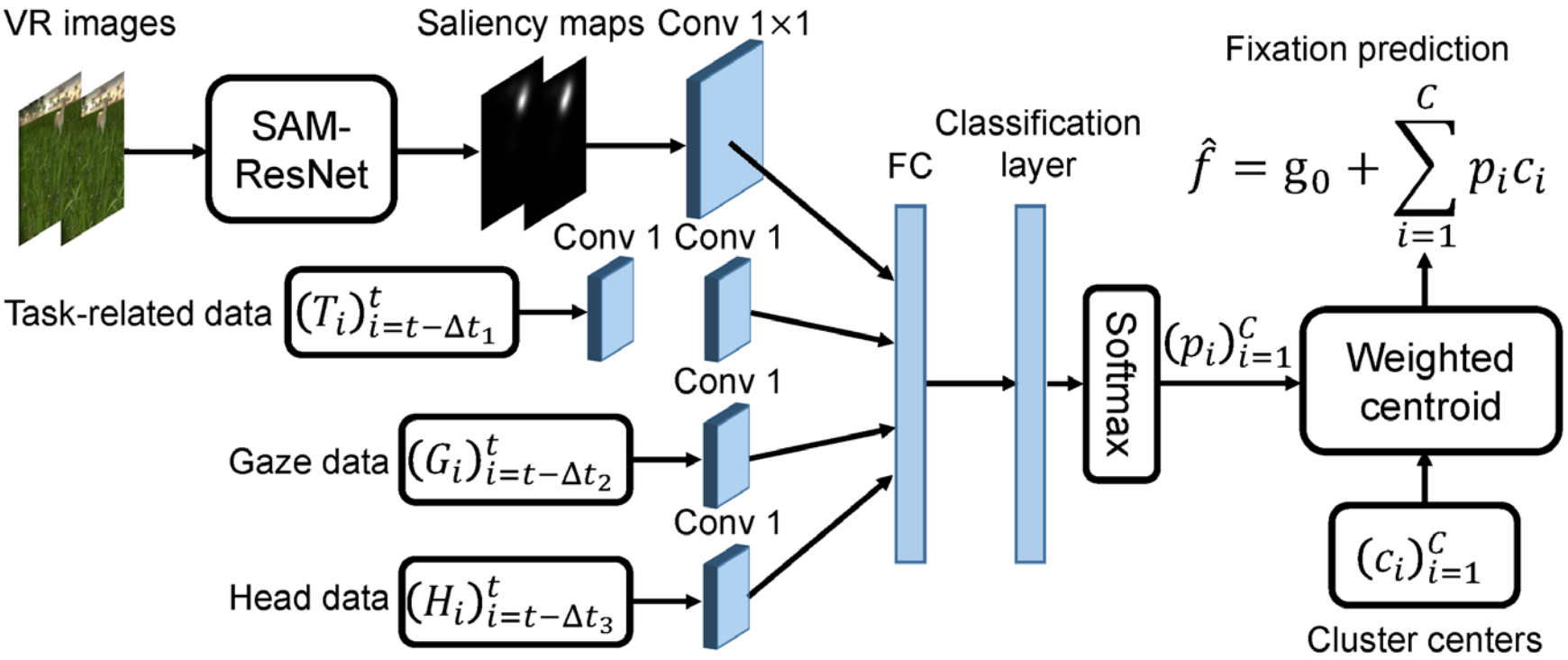




Fixation-head correlations in the horizontal (left) and vertical (right) directions

Fixations have correlations with head velocities

FixationNet Model

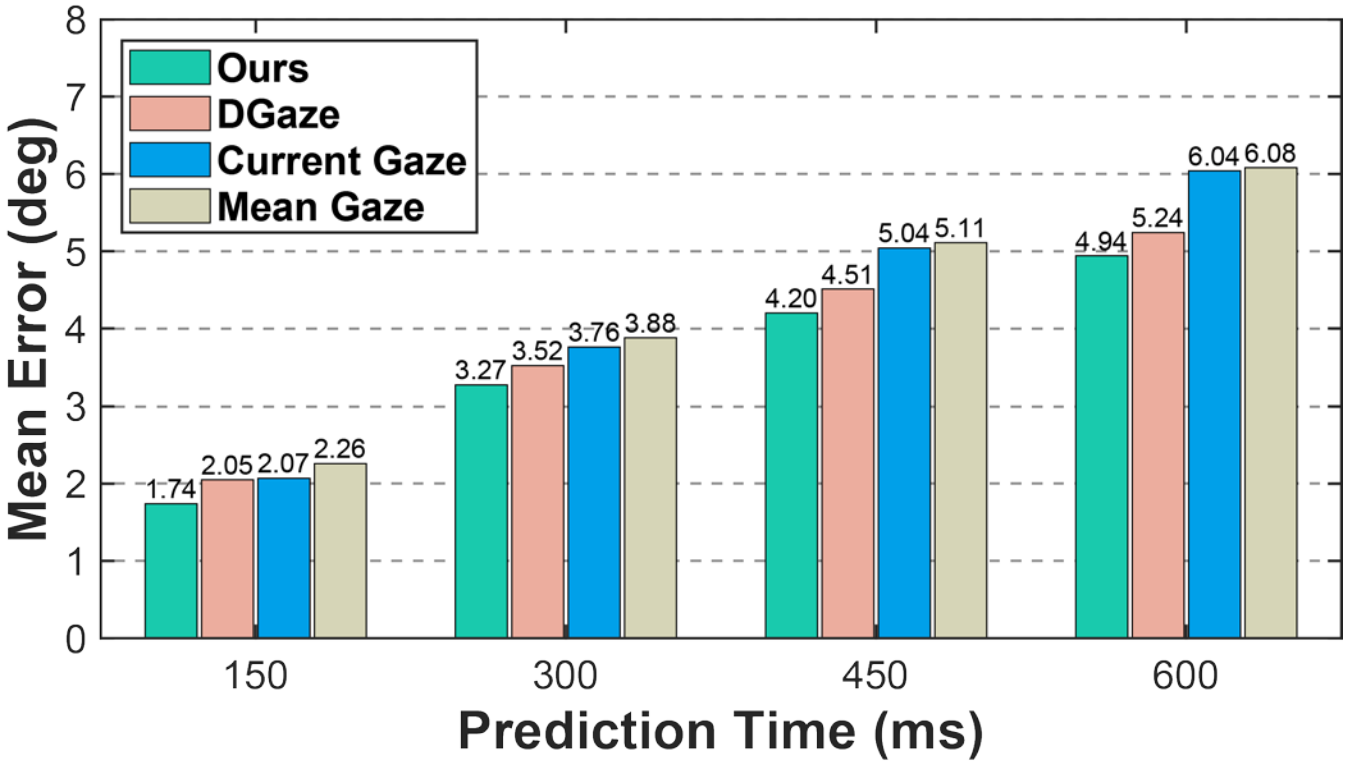


Architecture of FixationNet model

Experiments and Results

北京大學 PEKING UNIVERSITY

Performance in Task-Oriented Situations



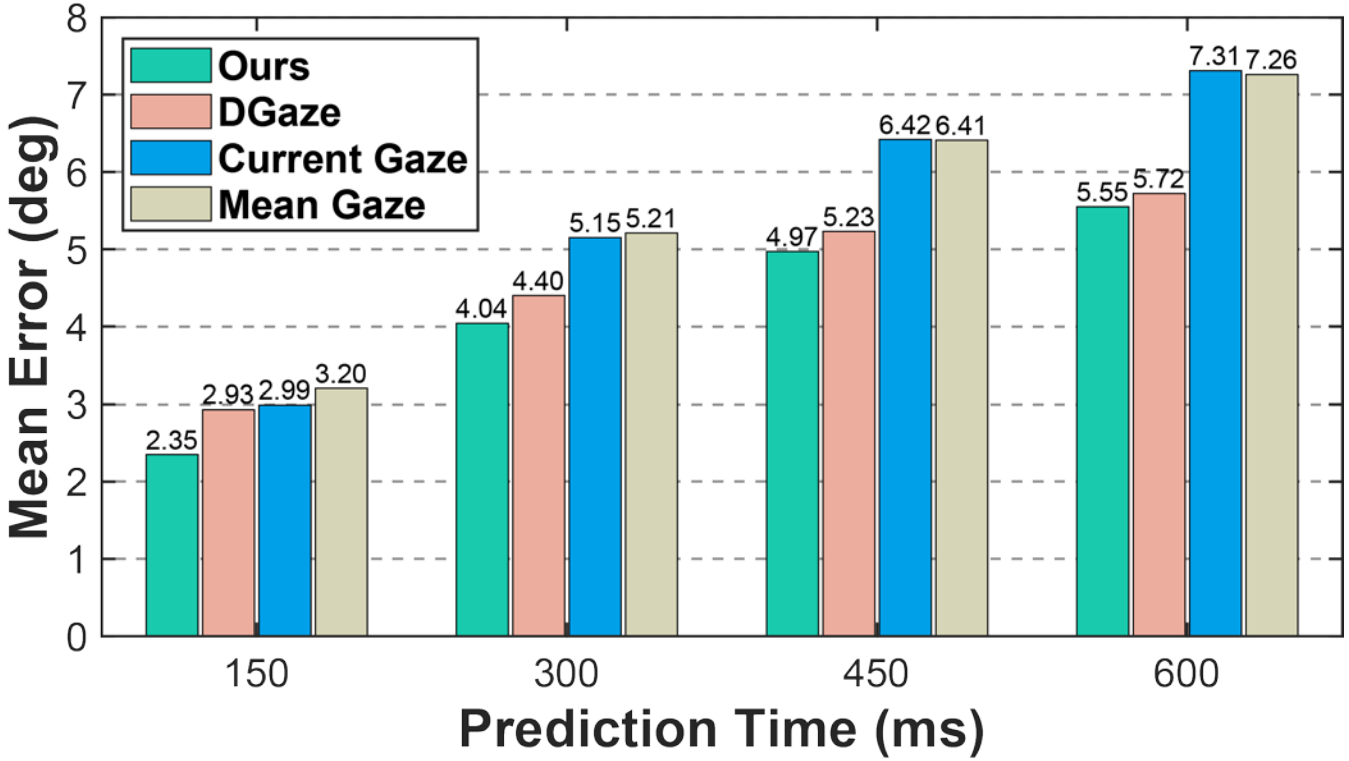
Performances in task-oriented situations at different time intervals

FixationNet outperforms other methods at different prediction times

Experiments and Results

北京大學 PEKING UNIVERSITY

Performance in Free-Viewing Conditions



Performances in free-viewing conditions at different time intervals

FixationNet outperforms other methods at different time intervals

Discussion

北京大學 PEKING UNIVERSITY

Limitations

- Our dataset is restricted to visual search task
- Our model employs pre-computed cluster centers
- The influence of sound is not considered in our model

Discussion

北京大学 PEKING UNIVERSITY

Future Work

- Overcome the limitations
- Improve our model by considering other factors related to users' visual attention
- Explore the problem of forecasting users' long-term eye fixations
- Convert our model to other systems like AR and MR systems

Thank you