

Synchronized multi-person eye-tracking in dynamic scenes

Shreshth Saxena^{1,2} and Lauren Fink^{1,2}

¹ Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Canada. ² Department of Music, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany.

Introduction

- Eye movements are linked to audiovisual perception, attention and aesthetic evaluation.
- Real world applications provide ecological and external validity to eye-tracking studies.
- However, current methods are limited to controlled, single-person setups.

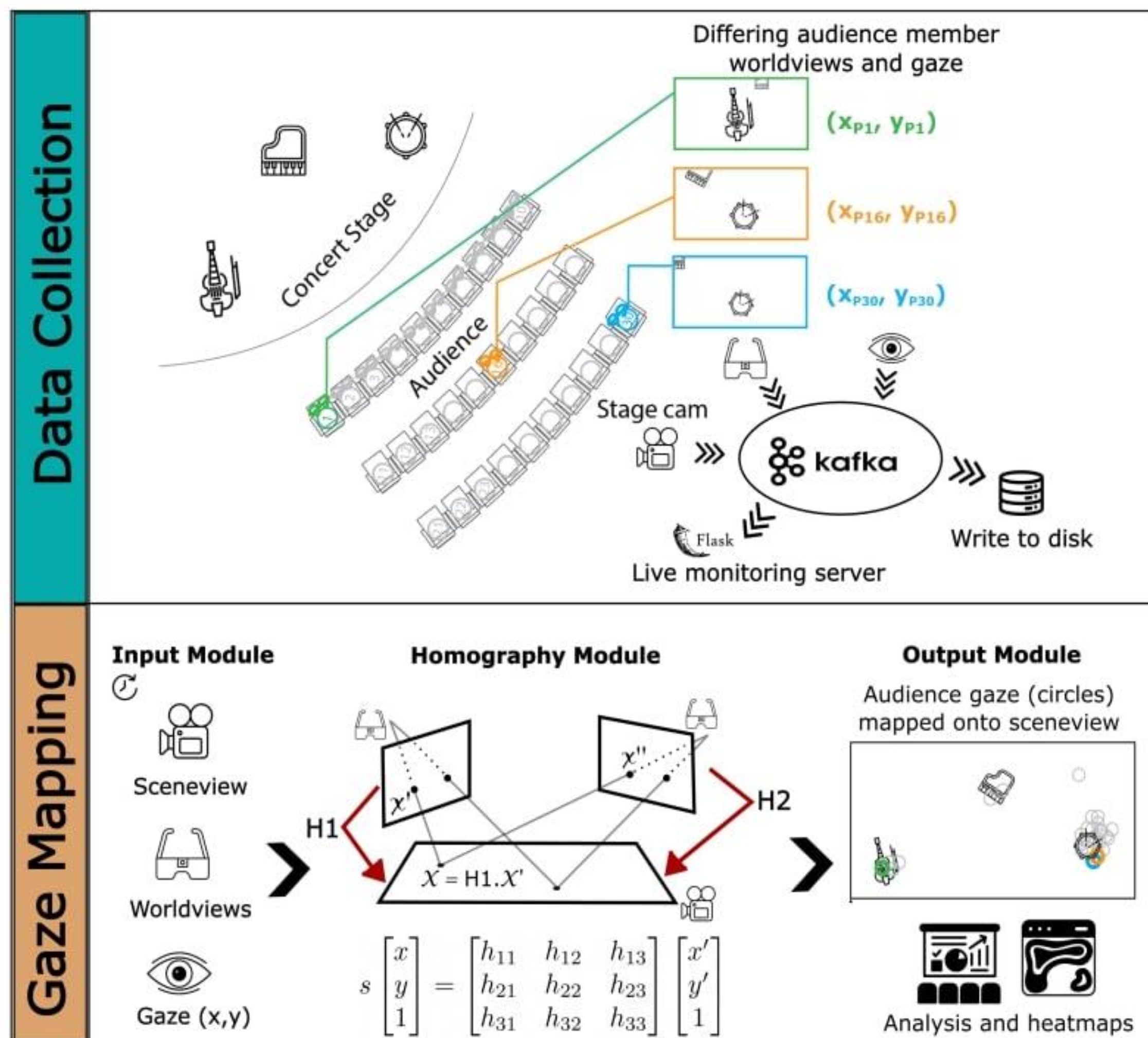
Challenges

- Synchronized data collection from multiple sensors in a shared physical space.
- Joint analysis of egocentric gaze data from differing viewing perspectives.

Current Study

- Validate multi-person eye-tracking in a concert hall setting.
- Stream simultaneous gaze data from eye-tracking glasses worn by audience members.
- Synchronize incoming streams and add timestamps for event triggers.
- Perform homography to map each wearer's egocentric view to a central view of the stage.
- Characterize performance over a comprehensive battery of tasks.

Software Design Pattern



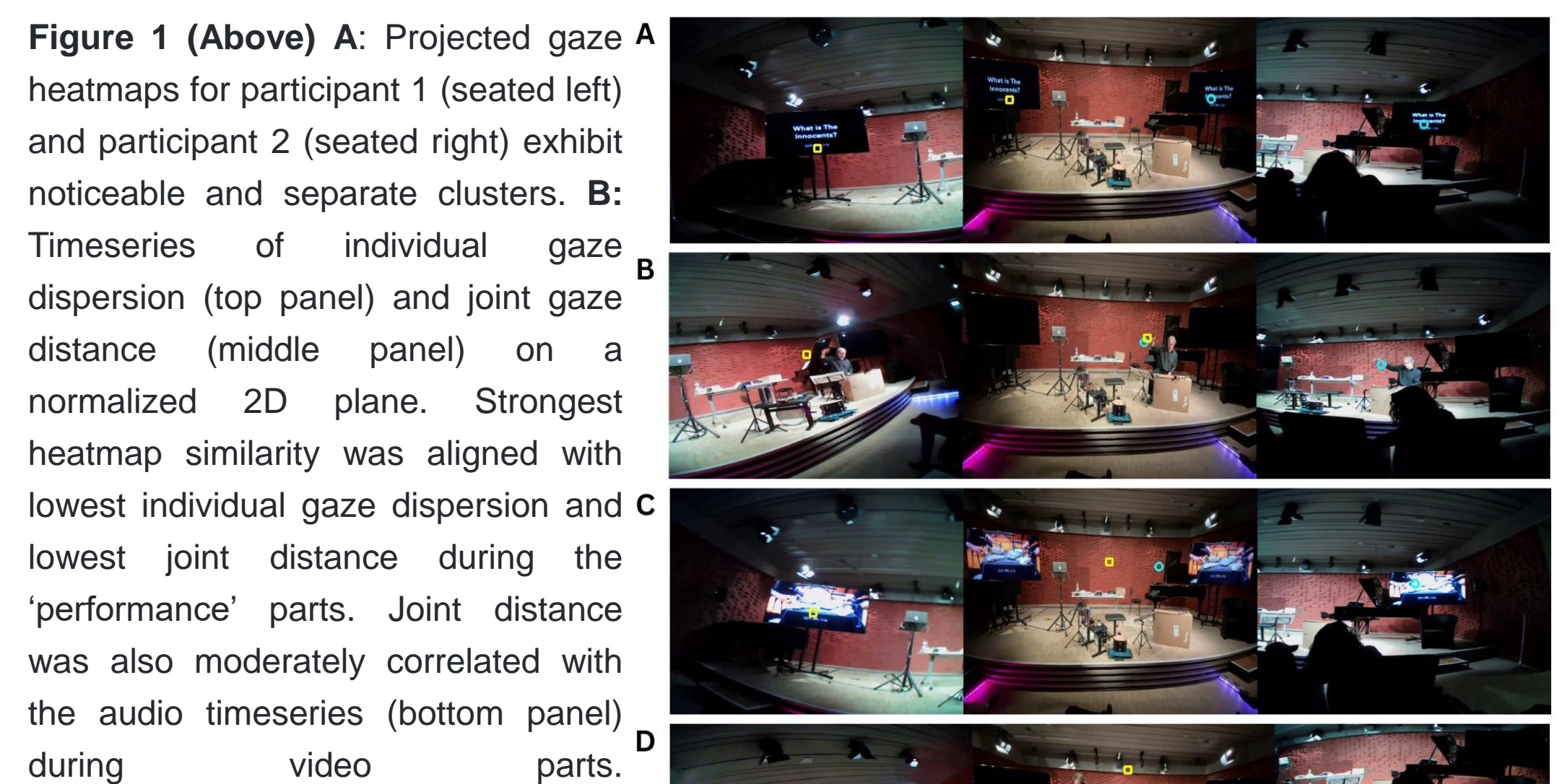
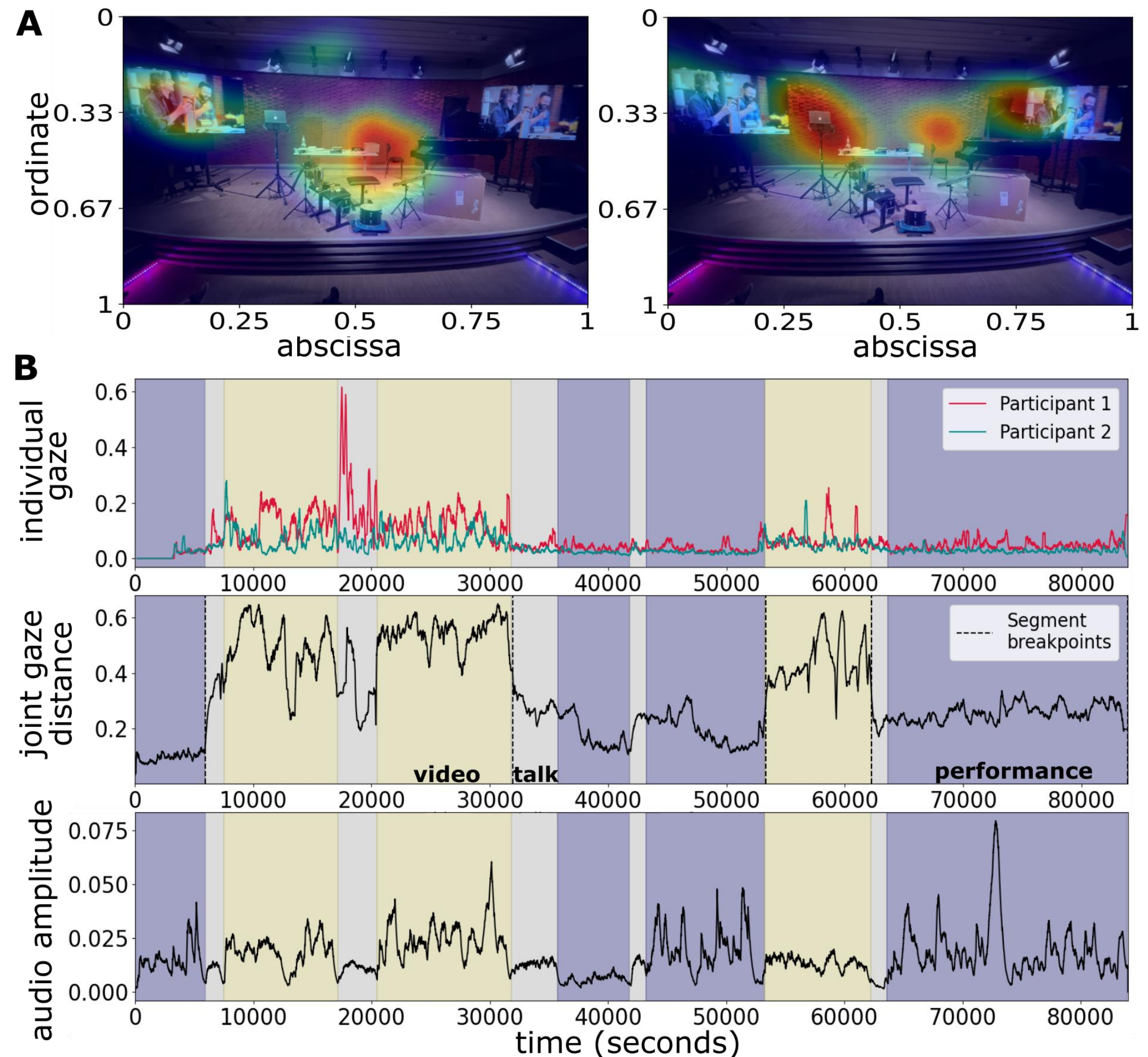
- Gaze and worldview recorded from the viewer's perspective.
- Sceneview recorded from a stationary stage camera placed centrally.
- Data Collection layer initiates synchronized recording from all devices.
- Incoming gaze, worldviews, and sceneview are efficiently streamed using Apache Kafka.
- Deep-learning-based feature mapping² and homography projects gaze to sceneview.
- Transformed gaze is analyzed and plotted in the output module.

Utility test

- 2 participants seated at extreme angles in a concert hall (row 1 left & row 2 right).
- Pupil Labs Invisible devices used to record gaze during a percussion concert.
- Homography: SuperGlue² method outperformed standard SIFT+RANSAC.
- Projections were robust to motion artifacts, occlusions and low-textured or dark regions.
- Projected heatmaps show diverging heatmap clusters highlighting varying optimal experiences characterized by varying attentional interests and viewing angles. (Fig. 1)
- Gaze coordination was highest during the "performance" parts of the concert. (Fig. 1)

References

1. Saxena, S., Lange, E. B., & Fink, L. (2023). Deep Learning Models for Webcam Eye-tracking in Online Experiments. *Behavior Research Methods* (accepted, in press). PrePrint, <https://osf.io/42qej/>
2. Sarlin, P.-E., DeTone, D., Malisiewicz, T., & Rabinovich, A. (2020). SuperGlue: Learning Feature Matching with Graph Neural Networks (arXiv:1911.11763). arXiv. <http://arxiv.org/abs/1911.11763>



Next Steps

- N = 4 Pupil Labs Neon devices * 10 groups = 40 Participants
- Participants presented with an eye-tracking task battery proposed in our previous work¹.
- Runtime tests of streaming sample-rate, frame drops and time synchronization.
- Accuracy measures of projected gaze for tasks in the battery.
- Comparison with lab-based and online eye-tracking accuracy evaluated on same tasks.
- Develop Online and Offline modes for the proposed framework.
- Improve precision of projected gaze points by incorporating past temporal information.

