DisETrac: Distributed Eye-Tracking for Online Collaboration

Bhanuka Mahanama  
Old Dominion University  
Norfolk, VA, USA

Mohan Sunkara  
Old Dominion University  
Norfolk, VA, USA

Vikas Ashok  
Old Dominion University  
Norfolk, VA, USA

Sampath Jayaratna  
Old Dominion University  
Norfolk, VA, USA

Abstract
Coordinating viewpoints with another person during a collaborative task can provide informative cues on human behavior. Despite the massive shift of collaborative spaces into virtual environments, versatile setups that enable eye-tracking in an online collaborative environment (distributed eye-tracking) remain unexplored. In this study, we present DisETrac, a versatile setup for eye-tracking in online collaborations. Further, we demonstrate and evaluate the utility of DisETrac through a user study. Finally, we discuss the implications of our results for future improvements. Our results indicate promising avenue for developing versatile setups for distributed eye-tracking.

CCS Concepts
• Human-centered computing → Interaction techniques; Collaborative interaction; • Information systems → Specialized information retrieval; Users and interactive retrieval.

Keywords
Eye Tracking, Multi-user, Information Retrieval

ACM Reference Format:

1 Introduction
Human gaze and pupillary information have a wide range of applications, from human-computer interaction [Palinko et al. 2016; Papoutsaki et al. 2017] to psychology [De Silva et al. 2019; Mahanama et al. 2022a; Michalek et al. 2019] and behavioral science research [Jayawardena et al. 2020, 2019; Mahanama et al. 2021]. While many studies are single-user studies, often conducted in isolated environments, studies on collaborative behaviors have become increasingly popular with the advances in eye-tracking techniques.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHIIR ’23, March 19–23, 2023, Austin, TX, USA
© 2023 Copyright held by the owner/author(s).
https://doi.org/10.1145/3576840.3578292

2 Related Work
Eye tracking can be used primarily to convey cues (often termed social cues) for initiating joint attention in a collaborative environment. The technique commonly referred to as gaze sharing (or shared gaze), involves visualizing the gaze positions of the collaborators indicating their attention. The shared gaze visualizations have proven to improve performance in collaborative environments such as meetings [Langner et al. 2022], writing [Kütt et al. 2019], search [Zhang et al. 2017], puzzle solving [D’Angelo and Gergle 2016], and code review [Cheng et al. 2022]. Even though it is straightforward to identify and react to cues (often termed social cues) such as gaze direction, online collaborations often require these social cues to be explicitly visualized in the medium of interaction.

The advent of the COVID-19 pandemic transformed many collaborative activities into online interactions with participants spanning multiple geographical regions. Similar to co-located collaborations, one can use eye-tracking in online collaborative tasks to either investigate [Langner et al. 2022; Špakov et al. 2019] or convey cues for collaboration [Langner et al. 2022], which can be termed distributed eye-tracking. These distributed eye-tracking systems leverage the capability of eye-tracking data (gaze and pupillary) to provide insight into human cognition processes [Jayawardena et al. 2022] such as attention [Mahanama et al. 2022a]. As a result, distributed eye tracking plays a pivotal role in understanding human behavior in online collaborations.

Despite the popularity and importance of eye tracking in online collaborations, there is a lack of studies that propose and evaluate the versatility of a generalized setup. In our study, we attempt to address the issue by introducing a generalized setup for distributed eye tracking. A short demo of the proposed work is available at https://youtu.be/yJli4VGbrHA.

The contributions of our study are as follows,

1. Introduce DisETrac: a distributed multi-user eye-tracking utilizing off-the-shelf eye-trackers
2. Demonstrate the utility of DisETrac through a prototype and evaluate the setup
3. Discuss our observations during the study and potential applications.

[Kütt et al. 2019]. These studies examine the behavior of co-located participants engaging in collaborative activities by utilizing eye tracking to convey gaze or for analyzing collaborative behavior [Broz et al. 2012; Kütt et al. 2019; Zhang et al. 2017].
Alternatively, eye tracking can be used to evaluate joint attention between participants of a collaborative task. Prior studies on joint attention tracking include virtual-human interactions via immersive experiences [Kim and Mundy 2012; Swanson and Siller 2014] and human-human social interactions [Ambrose et al. 2020]. These virtual-human interactions are based on the virtual reality paradigm using Heads-Up Display systems with eye-tracking optics and virtual avatar [Kim and Mundy 2012]. Similar studies have incorporated social video viewing tasks [Swanson and Siller 2014] using desktop eye trackers. Social interactions during human-human interactions [Ambrose et al. 2020] have used wearable eye-tracking devices to track participants simultaneously during an interactive task using two mobile eye-tracking devices.

Irrespective of the role of eye tracking in the joint attention task, the general setup remains common across the experiments. Moreover, we can collectively classify these applications of eye tracking under multi-user eye tracking.

Despite the increasing adoption of eye-trackers in research and for human-computer interaction, very limited studies have been conducted in the realm of multi-user eye-tracking [Mahanama 2022]. Eye-tracking studies have been predominantly single-user studies [Jayawardena et al. 2021; Mahanama et al. 2022b] often conducted with a participant in isolated environments, primarily due to eye-trackers being unable to track more than one person [Mahanama 2022]. Most of the studies in multi-user eye-tracking studies consider focusing primarily on co-located dyads and use eye-tracking data as a medium of sharing social cues [Cheng et al. 2022; He et al. 2021; Kütte et al. 2019; Spakov et al. 2019; Zhang et al. 2017] or investigating the collaborative behavior among the participants [Broz et al. 2012]. Despite the usefulness, none of the studies investigate the feasibility of the experimental setups for practitioners for situations beyond co-located dyads, often requiring the transmission of data across public networks.

In an attempt to develop a setup beyond dyads, Mahanama 2022 proposed a commodity hardware-based solution capable of utilizing the recent advancements in appearance-based eye-tracking [Pathirana et al. 2022; Senarath et al. 2022]. Despite the scalability of the setup beyond dyads, the setup assumes the participants to be co-located. Moreover, the setup fails to leverage the far superior hardware and software specialized in eye-tracking [Pathirana et al. 2022] for generating eye-tracking data. Thus the requirement of a scalable setup with quantitative evaluation remains unexplored for multi-user eye-tracking on beyond-dyad non-co-located experiments.

3 Methodology

In the proposed multi-user distributed eye-tracking system, we identify two main components, (1) eye-tracking metric estimation and transmission, and (2) aggregation and processing. We facilitate communication between the two components through an MQTT broker (see Figure 1). We use common off-the-shelf eye-trackers (e.g., Gazepoint GP3) for the eye-tracking metric estimation task, SDK/API to gather eye-tracking data, and transmit it to the MQTT server. When transmitting eye-tracking data, we add an originating timestamp to facilitate synchronization at the dashboard and a sequence number for reconstructing the original order of messages.

To eliminate the effects of clock drift, we periodically perform manual clock synchronization during the span of time-intensive experiments.

In the aggregation and processing state, we first subscribe to the eye-tracking data streams (i.e., gaze position and pupil dilation). Then we aggregate the messages together based on the attached origin timestamp and the sequence number. Finally, we compute aggregate eye-tracking measures and visualize the data on the dashboard.

3.1 Eye-Tracking Measures

For the simplicity of experiments, we use two types of measures in our proposed system, individual (e.g., gaze position, pupil dilation) and aggregate (joint attention distance). We introduce joint attention distance as the distance from the centroid of the gaze position to the gaze position of a particular user. We compute the joint attention distance (D) of the user x in a group comprising users U as,

$$D_x = \sum_{u \in U} \frac{S_u}{|U|}$$

where $S_u$ is the eye-tracking surface coordinates of the user x. In a surface with a normalized 1-D coordinate system ($S \in [0, 1]$), our measure yields $D \in [0, 1]$, with lower values denoting overlapping gaze positions and higher values indicating deviations from the majority behavior. Similarly, we extend this for surface with 2-D normalized coordinates with an appropriate distance measure. For instance use of Manhattan (L1) distance will yield $D \in [0, 2]$.

3.2 Analytics Dashboard

The analytics dashboard provides a detailed real-time perspective on the ongoing experiment by combining visualizations of individual and aggregate measures. Further, the dashboard offers functionalities to monitor and control the experimental setup. We combine these functionalities in the prototype dashboard through four interactive visual elements (see Figure 2).

1. Control Pane: Monitor and control the experimental setup. Actions include starting/resuming ongoing experiments, exporting experiment data, and monitoring the status of devices connected to the system.
2. Gaze Positions: Real-time visualization of the user experiment surface with color-coded visualizations of the gaze positions of each user.
3. Individual Metrics: Color-coded visualizations of other individual measures for the participants in the experiment.
4. Aggregate Metrics: Visualization of aggregate measures of participants (e.g., joint-attention distance).

3.3 Utility Study

We demonstrate the utility of DisETrac evaluating joint attention in a collaborative environment. We recruited ten participants (4F) for the study and conducted the experiment in physically isolated pairs. We used a primary computer for hosting the MQTT broker and DisETrac dashboard. We provided participants with an identical computer setup (secondary) connected to an online collaboration session, including option for discussion. We used two Gazepoint
Figure 1: Architecture of the Proposed DisETrac Distributed Eye-Tracking System.

Figure 2: DisETrac Distributed Eye-Tracking Dashboard Layout.

GP-3\(^1\) eye-trackers operating at 60 Hz sampled at 30 Hz for extracting eye-tracking data. Then, we connected all devices (2 secondary computers, the primary running dashboard, and the MQTT broker) to the same network and synchronized all devices using the Network Time Protocol (NTP) before each round.

For the collaboration task, we provided the users with an online Jigsaw puzzle\(^2\) with 40 pieces (see Figure 3). Before the experimental task, we allowed the users to familiarize themselves with the controls of the game. We allowed the participants to collaborate in their preferred languages and form game strategies during the

---

\(^1\)https://www.gazept.com/product-category/gp3/

\(^2\)https://www.jigsawexplorer.com/
We primarily test our proposed system’s performance using the latency of the setup, computed using the delay between eye-tracking data generation to the reception at the dashboard. Our results (see Table 1) indicate that the proposed setup captured and transmitted eye-tracking data while maintaining a mean latency of 202.5 ms. Since we use a public network for our experiments, the latencies we report are subjected to delays due to the conditions in the network and the offered quality of service parameters, resembling a real-world scenario. We identify instances of abnormally high latencies (e.g. sessions 3 & 5) to mitigate potential impacts based on the network state at the time of the experiment.

### Table 1: Data latency (gaze and pupil data) during the puzzle-solving task.

<table>
<thead>
<tr>
<th>Session</th>
<th>Mean Latency (ms)</th>
<th>Max Latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>129.6 ± 63.7</td>
<td>462.0</td>
</tr>
<tr>
<td>2</td>
<td>350.3 ± 87.7</td>
<td>833.4</td>
</tr>
<tr>
<td>3</td>
<td>314.9 ± 556.8</td>
<td>4000.7</td>
</tr>
<tr>
<td>4</td>
<td>76.3 ± 425.9</td>
<td>41.0</td>
</tr>
<tr>
<td>5</td>
<td>141.6 ± 410.5</td>
<td>4615.7</td>
</tr>
<tr>
<td>Mean</td>
<td>202.5 ± 308.9</td>
<td>1990.56</td>
</tr>
</tbody>
</table>

We also examined the relationship between the eye-tracking metrics and the overall task by the time taken to complete the puzzle. For the individual measures, we computed the fixational measures (fixation count, fixation duration, and saccadic amplitude) for each participant and computed the average in each session. Our results indicate the existence of a weak correlation between the fixation count and task completion time (correlation coefficient of 0.54).

### Table 2: Average fixation count and duration for experiment sessions.

<table>
<thead>
<tr>
<th>Session</th>
<th>Fixations count</th>
<th>Fixation duration (s)</th>
<th>Saccadic amplitude</th>
<th>Total time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>517</td>
<td>0.707</td>
<td>444</td>
<td>308</td>
</tr>
<tr>
<td>2</td>
<td>496</td>
<td>0.561</td>
<td>451</td>
<td>148</td>
</tr>
<tr>
<td>3</td>
<td>1343</td>
<td>0.517</td>
<td>463</td>
<td>306</td>
</tr>
<tr>
<td>4</td>
<td>588</td>
<td>0.621</td>
<td>387</td>
<td>294</td>
</tr>
<tr>
<td>5</td>
<td>1141</td>
<td>0.619</td>
<td>444</td>
<td>744</td>
</tr>
</tbody>
</table>

We also examined the relationship between the total time taken by each dyad and against the joint attention distance (see Table 3). Our results indicated a weaker correlation between the two measures (correlation coefficient of 0.2).

### 5 Conclusion

In this paper, we present DisETrac, a multi-user eye-tracking framework to demonstrate the utility of distributed eye-tracking. Our setup uses off-the-shelf eye-trackers connected through a public network for providing real-time insights on a multi-user eye-tracking experiment comprising remote participants.

Our objective is to analyze and interpret the attention of students in an online classroom lecture in order to accommodate the instructor’s teaching style to the integrated learning style of multiple students. This scenario is useful during virtual learning and in-person lectures, given that all students are looking at the same screen and from the same viewpoint of joint attention.

A key limitation of our experimental setups is that all participants have identical setups and to be identical in terms of resolution and the content on screen. Experiments on complex screen configurations require additional steps in the aggregation to transform data into comparable metrics. In the future, we will experiment the viability of the non-homogeneous setup [Jayawardana et al. 2022, 2021] and data transfer process [Mahanama et al. 2020].

### Acknowledgments

This work was supported in part by NSF CAREER 2045523. Any opinions, findings and conclusions or recommendations expressed in this material are the author(s) and do not necessarily reflect those of the sponsors.

### References

DisETrac: Distributed Eye-Tracking for Online Collaboration

CHIIR ’23, March 19–23, 2023, Austin, TX, USA

Figure 3: Collaborative Puzzle Solving Activity.


