Abstract—Understanding the focus and visual scanning behavior of users during a collaborative activity in a distributed environment can be helpful in improving users’ engagement. Eye tracking measures can provide informative cues to understanding human visual search behavior. In this study, we present a distributed eye-tracking system with a gaze analytics dashboard. This system extracts eye movements from multiple participants utilizing common off-the-shelf eye trackers, generates real-time traditional positional gaze measures and advanced gaze measures such as ambient-focal coefficient $K$, and displays them in an interactive dashboard. We evaluate the proposed methodology by developing a gaze analytics dashboard and conducting a pilot study to (1) investigate the relationship between $K$ with collaborative behavior, and (2) compare it against the User Experience Questionnaire (UEQ) benchmark. Our results show that groups that spent more time had more ambient attention, and our dashboard has a higher overall impression compared to the UEQ benchmark.

Index Terms—Data Visualization, Eye Tracking, Multi-user, Information Retrieval

I. INTRODUCTION

The spread of the COVID-19 pandemic has forced many organizations and individuals to transform their activities online, including collaborative activities and communication with participants across multiple geographical regions. Understanding the visual attention of users in an online collaborative environment can be helpful in identifying important visual cues, interpreting visual information, and navigating user interaction effectively. Eye tracking techniques such as shared gaze visualization enables remote collaborators to use non-verbal cues which can improve communication and collaboration in online environments [1]. The advancement of eye tracking technology allows us to use different eye movement measures to analyze the quality of collaborative interactions [2], [3].

Eye movement measures including fixation, saccades, micro-saccades, and pupil diameter have been widely used to assess human visual attention during a task [4]–[7]. Even though these eye-trackers perform well for single-user studies, they lack the scalability for multi-user studies mainly because they cannot track more than one person. Eye-tracking studies have been single-user studies [8]–[11] often conducted with a participant in isolated environments primarily due to eye-trackers being unable to track more than one person [12], [13].

Distributed eye-tracking is a phenomenon that uses eye-tracking measures from multiple users in online collaborative tasks in real time [14]. Previous studies on real-time distributed eye-tracking systems focused on joint visual attention using gaze measures such as gaze positions, pupil diameters, and fixations and saccade-related measures [14]–[17]. We focus on gaining further insights into visual scanning behavior in distributed eye tracking systems using real-time advanced gaze measures. In this study, we integrate a real-time advanced gaze measure, namely, Ambient/Focal Attention with Coefficient $K$ [6] to distributed eye-tracking systems. Additionally, we utilize real-time traditional positional gaze measures such as fixation duration, saccade duration, and saccade amplitude in this system. A short demo of the proposed work is available at https://youtu.be/20LzU9NmF4o

The main contributions of this work are:

1) We introduce a distributed multi-user eye-tracking system with advanced gaze measures and traditional positional gaze measures.

2) We display advanced gaze measures along with traditional positional gaze measurements in an interactive dashboard in real-time.

3) We demonstrate the utility of the proposed methodologies through a prototype.

II. RELATED WORK

Shared gaze visualization provides non-verbal cues in remote collaborative tasks allowing the users to see the gaze information of their remote partner [3]. Research has been conducted to investigate the effect of shared gaze visualization on collaborative tasks such as learning [18], [19], programming [19]–[21], co-writing [22], meeting [16], puzzle solving [23], game playing [1], and visual search [24], [25]. Alternatively, gaze visualization can be used to analyze eye-tracking measures of multiple users in collaborative tasks. Despite the increased adaptation of shared gaze visualization in collaborative environments, very limited studies have been conducted in real-time gaze analytics visualization during collaborative tasks. In our prior work [14], we introduced an analytics dashboard that provides real-time visualizations of individual and aggregate measures in the distributed multi-user eye-tracking system.

Eye-tracking measures such as joint visual attention have been used in the literature to understand users’ collaborative behaviors in different interactive tasks [14], [26], [27]. The measures of joint visual attention have been developed for specific contexts in collaborative tasks [3]. For instance,
DisETrac [14] uses the distance from the centroid of the gaze position to the gaze position of a particular user as the joint attention measure. With-me-ness [27] was developed to measure joint visual attention by aggregating entry time, first fixation duration, and the number of revisits. These eye-tracking studies which utilize joint visual attention measures in gaze visualization have focused on pupillary information and traditional gaze metrics such as fixations and saccade-related metrics [14], [18]. In this work, we are extending the DisETrac to support advanced gaze measures along with traditional positional gaze measures for collaborative interactions. We integrate dynamic Ambient/Focal attention with coefficient \( \mathcal{K} \) with the gaze analytics dashboard allowing us to examine visual search behaviors during collaborative tasks.

### III. METHODOLOGY

We use the distributed eye-tracking setup proposed DisETrac [14] for our experiments, comprising two main components for eye-tracking; (1) data acquisition and transmission, (2) aggregation and visualization. Similar to DisETrac study, we sample data from common off-the-shelf eye trackers using the vendor API/SDK. Then we transmit data to an MQTT broker through a public network. MQTT broker is a message server that facilitates communication between publisher and subscriber clients.

In our setup, we acquire the gaze position of each user on the screen \((x,y)\) and the pupil dilation of each user, along with confidence estimates as determined by the vendor software. During the transmission of the data, we add an originating timestamp and a sequence number to facilitate synchronization and the reconstruction of the original sequence of messages. To ensure the comparability of transmitted data, we periodically perform manual clock synchronization using Network Time Protocol (NTP).

At the processing end, we subscribe to the eye-tracking data streams of the MQTT broker and use them to compute eye-tracking measures. We utilize user identifier information to distinguish and compute eye-tracking measures for each user, which we then use to compute aggregate measures. For our computations, we use Real-Time Advanced Eye Movements Analysis Pipeline (RAEMAP), [28]–[30] an eye movement processing library. Finally, we present the data to a proctor through an interactive dashboard. The overall architecture of our setup is shown in Figure 1.

#### A. Real-time Gaze Measures

We use RAEMAP to compute real-time gaze measures in two steps generating, (1) traditional positional gaze measures for each user, and (2) advanced gaze measures for each user and the group. In the first step, we start by forming sliding time windows by aggregating incoming data using the timestamps for each user. For each time window, we identify fixations, periods where the gaze remains stationary, and saccades, where the gaze shifts rapidly [31]. Then we compute fixation duration \((d)\) for each fixation, saccade amplitudes \((\alpha)\), the distance corresponding to the shift in gaze, and saccade duration. In the second step, we use results to generate an alternative version of Ambient/Focal Attention Coefficient \( \mathcal{K} \), an indicator of visual search behavior for each window. Instead of using global statistical information as in \( \mathcal{K} \), we use statistical information in each time window. For this purpose, we propose a windowed coefficient, defined for the \( i \)-th fixation in a time window \( w \) as,

\[
\mathcal{K}_w^{i} = \frac{d_i - \bar{d}_{w,d}}{\sigma_{w,d}} - \frac{\alpha_{i+1} - \bar{\alpha}_{w,a}}{\sigma_{w,a}}
\]

where \( \bar{d}_{w,d}, \sigma_{w,d}, \bar{\alpha}_{w,a}, \sigma_{w,a} \) represent statistical information corresponding to the window, \( \alpha_{i+1} \) saccadic amplitude preceding the fixation, and \( d_i \) the duration of the fixation. We use the average in the presence of multiple fixations during the selected time window \((w)\).

Unlike \( \mathcal{K} \), \( w \mathcal{K} \) requires only the gaze details of the window \( w \) for the computation. We can progressively calculate the coefficient using a sliding window as data arrive for processing by compromising including the global context in the computation. For experiments beyond the window \( w \), we generate an aggregate coefficient through the average of all coefficients across the set of time windows \((W)\), providing a summary of attention during the experiment defined as,

\[
\mathcal{K}_W = \frac{1}{|W|} \sum_{w \in W} w \mathcal{K}
\]

For a multi-user environment, we extend the coefficient by defining the group coefficient as the average across all the users, either in a specific time window \((U \mathcal{K})\) or for the entire experiment \((U \mathcal{K})\). Our study uses \( U \mathcal{K} \) for visualizations and \( U \mathcal{K} \) when comparing group performance.

\[
U \mathcal{K} = \frac{1}{|U|} \sum_{u \in U} U \mathcal{K} \quad \text{and} \quad U \mathcal{K} = \frac{1}{|U|} \sum_{u \in U} U \mathcal{K}
\]

However, our modifications to the definition of \( \mathcal{K} \) do not affect the interpretation of the values. Similar to \( \mathcal{K} \), the windowed coefficients \( s \mathcal{K}, \mathcal{K} \), \( s \mathcal{K}, \mathcal{K} \), \( U \mathcal{K} < 0 \) indicate ambient visual scanning, while positive coefficients suggest focal processing.

#### B. Gaze Analytics Dashboard

The gaze analytics dashboard provides a detailed real-time visualization of (1) advanced gaze measures for each user \((\mathcal{K})\) and the group \((U \mathcal{K})\), and (2) traditional positional gaze measures for each user for the ongoing experiment (see Figure 2). Further, this dashboard provides more interactive functionalities to monitor, analyze, and control the gaze measure visualizations. The gaze analytics dashboard has four main key components as illustrated in Figure 3.

1) **Tabs**: Tabs allow the proctor to switch between the views of different gaze measure types. The views of two types of gaze measures that are designed in the dashboard (advanced gaze measures and traditional positional gaze measures) are shown in Figure 2.
Architecture of the proposed distributed eye tracking system for visual attention. Here, we use common off-the-shelf eye trackers to collect data from multiple users. Then we transmit the eye-tracking data to the MQTT broker. Next, we calculate real-time traditional positional gaze measures and real-time advanced gaze measures by passing the data through RAEMAP [29]. Finally, we stream the gaze measures to the gaze analytics visualization dashboard.

2) **Play/Pause Control:** As the gaze measures are visualized in real-time charts (data streaming charts), they automatically update themselves after every $n$ second. Hence, this play/pause control allows the proctor to pause the real-time charts and replay as necessary.

3) **Gaze Measures:** Real-time visualization of gaze measures calculated during the user experiment.

4) **Controls:** The control widgets include box zoom, wheel zoom, save, and reset.

### C. User Study

We conducted a pilot user study comprising ten participants (6M, 4F) and evaluated their attention in a collaborative activity. We conducted the study as physically isolated pairs (chosen randomly) collaborating online. The participants were graduate students in Computer Science and aged between 25-35 years. All the participants had normal or corrected-to-normal vision. We selected an online collaborative Jigsaw puzzle-solving activity comprising a 50-piece jigsaw puzzle pieces (see Figure 4).

We used identical computer setups for each user comprising of desk-mounted GazePoint GP3 eye tracker, a 23.8-inch screen (1920x1080). The eye trackers operated at 60 Hz, and our setup sampled data at 30 Hz from the eye trackers. We hosted the MQTT broker and the analytics dashboard on another two computers connected through the public network. Considering that each session lasted less than 10 minutes, we synchronized all the devices only once at the beginning of each session.

Each session started with a proctor calibrating each eye tracker using the standard 9-point calibration and manually testing the accuracy of the calibration. Then, the proctor presented a similar jigsaw puzzle as in the activity, explained the controls in the user interface, and allowed users to familiarize themselves with the activity. Meantime, we started the transmission, processing, and visualizations to ensure proper data flow. Once everything was in order, we presented the puzzle activity to the users and recorded the experimental data. During the experiment, we collected gaze location data from the eye trackers and formed advanced gaze measures upon reception at the gaze analytic dashboard. When forming advanced measures, we used a window of $w = 3000$ms, sliding at each 300ms. Further, we measured the time each pair took to complete the task.

Using the same set of participants, we evaluated our proposed gaze analytics dashboard and compared it against DisETrac [14]. For the evaluation, we used UEQ [32], a fast and reliable questionnaire to measure the User Experience of interactive products. For each participant, we presented both dashboards with simulated data. Once the participants have used both dashboards, we provided them with the UEQ and asked them to provide feedback regarding their experience with each dashboard. To avoid the sequence effect, the two dashboards were presented in random order per participant.

### IV. RESULTS

#### A. Latency Analysis

Similar to previous studies, we computed the latency by computing the delay between the transmission from the originating device to the destination dashboard in our system. We considered all eye-tracking data messages received during the experiment for the computation, assuming the effect of clock drifts to be negligible. Our results (see Table I) indicate that our setup transmitted data with a mean latency of 407 ms and average maximum latency of 994 ms in a public network. This indicate that our approach can notify a proctor on changes on average $wK + 407$ ms, where $d$ represents the duration of the last fixation. To emulate potential real-world conditions, we did not adjust the quality of service parameters of the network to prioritize our data.

#### B. Ambient/Focal Attention Analysis

To demonstrate the potential utility of the proposed windowed coefficient of attention, we investigated the relationship
Fig. 2: Visualizations of gaze measures in the analytics dashboard. Left: Traditional positional gaze measures, Right: Advanced gaze measures.

![Dashboard Visualization](image.png)

Fig. 3: Layout of the gaze analytics dashboard illustrating key components

![Dashboard Layout](image.png)

Table I: Data latency (gaze and pupil data) during the experiment.

<table>
<thead>
<tr>
<th>Session</th>
<th>Mean Latency (ms)</th>
<th>Max Latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>394 ± 235</td>
<td>973</td>
</tr>
<tr>
<td>2</td>
<td>408 ± 301</td>
<td>976</td>
</tr>
<tr>
<td>3</td>
<td>398 ± 313</td>
<td>1035</td>
</tr>
<tr>
<td>4</td>
<td>421 ± 330</td>
<td>1001</td>
</tr>
<tr>
<td>5</td>
<td>414 ± 341</td>
<td>1019</td>
</tr>
<tr>
<td>Mean</td>
<td>407 ± 308</td>
<td>994</td>
</tr>
</tbody>
</table>

Table II: Ambient/Focal Attention with Coefficient ($U_W K$) during the experiment.

<table>
<thead>
<tr>
<th>Session</th>
<th>Attention Coefficient ($U_W K$)</th>
<th>$\sigma$</th>
<th>Total time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.0515</td>
<td>0.4307</td>
<td>261</td>
</tr>
<tr>
<td>2</td>
<td>-0.0350</td>
<td>0.4386</td>
<td>174</td>
</tr>
<tr>
<td>3</td>
<td>-0.0575</td>
<td>0.4757</td>
<td>207</td>
</tr>
<tr>
<td>4</td>
<td>-0.0350</td>
<td>0.3273</td>
<td>168</td>
</tr>
<tr>
<td>5</td>
<td>-0.0996</td>
<td>0.4451</td>
<td>365</td>
</tr>
</tbody>
</table>

C. Dashboard Evaluation

We used the UEQ Data Analysis Tool, which uses T-Test [33] with 95% confidence interval to analyze the UEQ
responses. The 26 items in the UEQ are categorized into six scales (see Table III) that cover a comprehensive impression of user experience. We compared the scale means of the two dashboards as depicted in Figure 5. Our analysis did not show a statistically significant difference between the gaze analytics dashboard and DisETrac dashboard for the UEQ scales with $\alpha = 0.05$ (see Table IV).

**TABLE III: Scales of User Experience Questionnaire.**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attractiveness</td>
<td>Overall impression of the product. Do users like or dislike the product?</td>
</tr>
<tr>
<td>Perspicuity</td>
<td>Is it easy to get familiar with the product? Is it easy to learn how to use the product?</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Can users solve their tasks without unnecessary effort?</td>
</tr>
<tr>
<td>Dependability</td>
<td>Does the user feel in control of the interaction?</td>
</tr>
<tr>
<td>Stimulation</td>
<td>Is it exciting and motivating to use the product?</td>
</tr>
<tr>
<td>Novelty</td>
<td>Is the product innovative and creative? Does the product catch the interest of users?</td>
</tr>
</tbody>
</table>

The UEQ tool offers a benchmark that helps to interpret the results and the benchmark relies on a number of studies concerning different products [32]. We compared the results obtained for our gaze analytics dashboard with the benchmark to gain insight into the user experience quality of our visualization dashboard compared to typical products in the market.

As illustrated in Figure 6, the gaze analytics dashboard has “excellent” results in “Attractiveness” compared to the benchmark. Moreover, the gaze analytics dashboard introduced in this work shows “good” results in “Efficiency”, “Stimulation”, and “Novelty” scales which is 75% better than the results in the benchmark data set. However, for the “Perspicuity” and “Dependability” scales, our gaze analytics dashboard is better than 25% of the results in the benchmark which indicates as “below average”.

The results of the dashboard evaluation does not indicate a significant different between the UEQ scales in our dashboard compared to the existing DisETrac dashboard (see Table IV). We further compared the UEQ results obtained for gaze analytics dashboard against the benchmark. The results indicate that the overall impression of our interactive dashboard is in the range of the 10% best results. However, the results of UEQ indicate that our dashboard is difficult to get familiar with and learn how to use compared to the average results in the benchmark. We believe having eye-tracking specific measures in our dashboard caused this low score as the majority of evaluators are not eye-tracking experts. We mainly focused on data visualization and analysis aspects in our dashboard rather than data security. Hence, we observed that “Dependability” scale results of our dashboard is below the average of the benchmark. The “Dependability” is interpreted in the sense that the interaction is save and controllable by the user. However, according to UEQ analysis, our dashboard has provided users with exciting and motivating experiences, allowed users with less effort, and caught users’ interests compared to 75% of results in the benchmark data set.

**V. DISCUSSION**

In this study, we implemented a real-time ambient/focal attention coefficient $K$ and extended the concept to distributed multi-user eye tracking systems. Even though we demonstrate the utility through a pilot study, our approach requires further validation to determine the potential usage in analyzing user behaviors. Moreover, our study did not investigate defining the ideal window size ($w$) and remains unexplored. However, our pilot study revealed that the time a group takes to complete a puzzle is related to the ambient visual scanning behavior quantified by $U^W K$. Our results indicate that groups that spent more time had more scanning of the screen and searching behavior. Considering that jigsaw puzzle solving requires the participants to identify and match pieces based on their visual characteristics (color, shape, texture), we presume a relationship exists between the ambient scanning behavior and the finding of a matching piece.

A trivial approach to determine the effective and efficient means of computing variations of $K$ would be to conduct a comprehensive set of user studies encompassing different combinations of user behaviors. However, this approach could be costly and time-consuming. Alternatively, we can use synthetic data or re-stream data from previous experiments [34] to investigate the broad spectrum of possibilities for variations of $K$.

The results of the dashboard evaluation does not indicate a significant different between the UEQ scales in our dashboard compared to the existing DisETrac dashboard (see Table IV).
VI. CONCLUSION

In this study, we presented distributed eye-tracking system with real-time advanced gaze measures. 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 with advanced gaze measures. We presented the real-time gaze measures through an interactive dashboard. In the future, we plan to improve through the incorporation of other advanced gaze measures such as the Real-Time Index of Pupillary Activity (RIPA) [7] and Gaze Transition Entropy in a multi-user distributed environments. Further, we plan to integrate real-time scan-path visualizations in our dashboard by streaming user viewpoints.

ACKNOWLEDGMENT

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

REFERENCES