This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

Understanding and Enhancing Stereoscopic 3D Graph Perception with Eye Tracking

Yao Wang, Lin Zhang, Timo Hüttner, Ying Zhang, Wilhelm Kerle-Malcharek, Karsten Klein, Falk Schreiber, Andreas Bulling

***Abstract*—Immersive Analytics (IA) utilises stereoscopic 3D (S3D) graph visualisations in Virtual and Augmented Reality to harness spatial understanding and engagement beyond 2D screen displays. Despite growing research interest, little is known about how users perceive such structures, or how their gaze behaviour relates to task performance. We address this through a series of eye tracking studies on visual analytical tasks over S3D graphs in VR. Our first study, on single-layer graphs, reveals systematic links between gaze behaviour and task performance, with distinct exploration strategies across tasks. We then extend to multilayer graphs and show how their spatial arrangement and complexity reshape gaze behaviour. Finally, by mining the gaze patterns of successful analysts, we enhance the visualisations with targeted visual cues; in a follow-up study, these improve task correctness in the majority of trials. Overall, our findings deepen the understanding of graph perception in immersive environments, from single to complex multilayer structures, and demonstrate how eye tracking can both reveal reading strategies and inform S3D graph design through gaze-derived cues.**

***Index Terms*—Stereoscopic 3D, Graph Visualisation, Eye Tracking, Graph Perception.**

I. Introduction

Graph visualisation in stereoscopic 3D (S3D) is gaining significant attention, with an increasing number of research publications focusing on graphs displayed in S3D and interaction methods that enhance graph analysis through immersive environments. Early studies in the 1990s indicated the beneficial effects of S3D displays on graph analysis performance [1]. Subsequent research has reaffirmed these findings across various tasks and technologies, including community detection [2] and augmented reality (AR) applications [3]. The advent of affordable, high-quality head-mounted display (HMD) systems and the emergence of Immersive Analytics as a research domain have contributed to a substantial increase in studies addressing not only graphs but also diverse abstract data types [4]–[6].

Despite this progress, the extensive design space encompassing technology, graph representation, environmental characteristics, movement, interaction, and navigation has resulted in only a partial understanding of the effects of S3D graph visualisation. Evidence suggests no general superiority of S3D over 2D [2], [7]; performance instead depends on graph characteristics, representation, task type, and interaction support [5], [8]–[11], sometimes trading off accuracy against response time [2], [8], [10], [11]. Consequently, it is crucial to investigate how graph reading and task-solving strategies differ between S3D and 2D and why these differences exist. While these questions are complex and cannot be fully addressed in a single project, this study aims to advance our understanding by examining task-related graph perception in S3D. To explore potential issues, misperceptions, or advantages associated with S3D, we employ eye tracking technology to analyse graph-reading behaviour.

Although recent studies have begun to investigate perception in graph visualisations in traditional 2D setups [12]–[15], research on S3D perception remains limited. Investigating the perception and gaze behaviour of analysts in S3D can help identify potential challenges and obstacles associated with 3D visualisations, thereby contributing to the development of improved quality measures and layout methods specifically tailored for S3D.

Crucially, the collection and analysis of eye-tracking data can provide objective insights into reading behaviour, helping to identify or validate task-solving strategies and comprehension models that are often derived solely from self-reports [16], [17]. An important consideration for S3D is the applicability of existing quality criteria and the potential need for new ones. Although eye tracking has been underutilised in graph visualisations, it has proven effective in investigating readability factors to enhance quality measures, such as edge-crossing angles [18], [19], and in comparing layout effects [20], [21].

Building on these methods, eye tracking is uniquely well suited to graph perception, as it captures *where* and *when* viewers attend without interrupting the task, yielding a time-resolved signal that coarse measures (e.g. accuracy, task completion time) cannot provide. We characterise this gaze behaviour through established metrics: saccadic eye movements reveal the spatial and temporal distribution of attention and act as physiological correlates of cognitive load, where longer, faster saccades indicate increased load [22]. Meanwhile, the task-related fixation ratio captures the proportion of attention towards task-relevant elements rather than distractors. Relating these metrics to task correctness reveals success reading strategies.

The preliminary results of this work appeared in a conference paper at Graph Drawing 2025 [23], which has reported an 18-participant eye tracking study on single-layer S3D graph vi-

Yao Wang, Lin Zhang, Timo Hüttner, and Andreas Bulling are with the University of Stuttgart, Germany. E-mail: yao.wang@vis.uni-stuttgart.de, zhangln@studi.informatik.uni-stuttgart.de, st148236@stud.uni-stuttgart.de, andreas.bulling@vis.uni-stuttgart.de. Ying Zhang, Wilhelm Kerle-Malcharek, Karsten Klein, and Falk Schreiber are with the University of Konstanz, Germany. E-mail: {ying.zhang, wilhelm.kerle, karsten.klein, falk.schreiber}@uni-konstanz.de. Yao Wang and Lin Zhang contributed equally. Yao Wang is the corresponding author.

Manuscript received November 11, 2025; revised September 25, 2026.

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

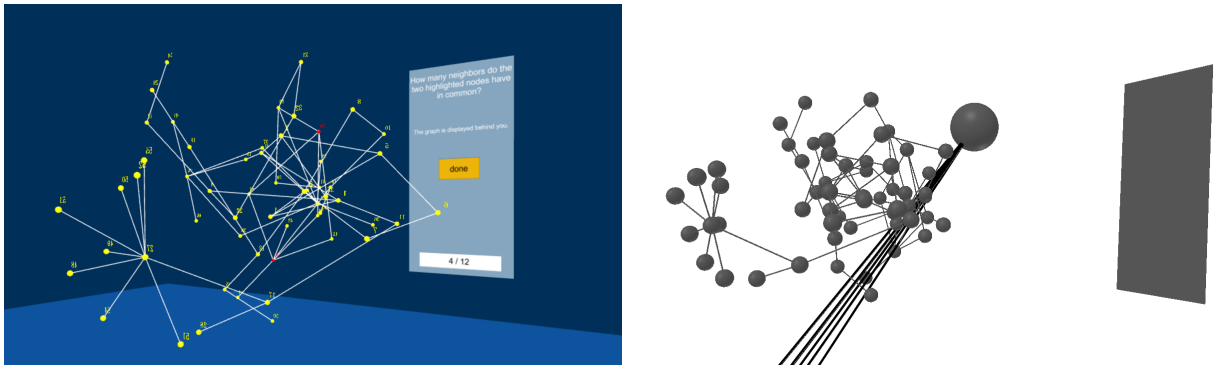

Fig. 1. **Left**: Our immersive environment displaying a 3D graph with highlighted nodes shown in red for the common neighbours task (T_{CN}). The question panel displays the task instruction and progress indicator. **Right**: Visualisation of the perspective reconstruction and ray casting. The large sphere represents the camera position. A cone of nine rays emanates from the camera origin. The black rectangular plane represents the question panel.

sualisation in which participants solved three analytical tasks – shortest path, highest node degree, and common neighbours (Study 1 in this manuscript). It established the first systematic links between gaze behaviour and task performance in immersive analytics, and showed that the gaze behaviours of correct and incorrect responders differ in how they distribute attention across task-relevant and confusion elements. These findings characterised how viewers read single-layer graphs, but left two questions open. First, whether these identified perceptual strategies generalise to more complex graph structures, such as *multilayer networks*, whose nodes are partitioned into several stacked or spatially arranged layers. Second, whether the gaze patterns of successful analysts can be turned into visual cues that actively support task solving.

This article extends the Graph Drawing 2025 conference paper to address these two questions through additional eye tracking studies and a deeper gaze analysis. Study 2 investigates multilayer networks, testing whether the gaze behaviour and strategies observed on single-layer graphs transfer to layered structures, and how their spatial arrangement (2.5D vs. 3D) and complexity (number of layers) reshape attention. Study 3 validates a gaze-informed enhancement: from sequential gaze patterns that differentiate successful from unsuccessful analysts, we derive targeted visual cues and compare the enhanced visualisations against the original ones. The three studies form a progression from understanding the perception of a single graph (Study 1), generalising this understanding to multilayer structures (Study 2), and actively enhancing graph perception with gaze-informed cues (Study 3).

II. Related Work

A. Human Graph Perception

We review prior work on perception and graph visualisation, with a focus on virtual and augmented reality (VR/AR). For broader overviews of eye tracking in 3D, we refer to surveys on gaze interaction in XR [24] and 3D gaze data visualisation [25]. Graph layout strongly affects readability. Eye-movement smoothness depends on edge-crossing angles [18], [19], clustered layouts improve search [20], force-directed layouts suit node/path tasks [21], and tapered edges aid trajectory perception [26]. Other works examine how layout dictates the perception of density [13], [14], stress [27], degree centrality [28], and visual complexity [29].

Eye tracking provides deep insights into the perceptual mechanisms in graph reading. Gaze data reveals reading strategies for graphs on complex surfaces [30], predicts user tasks and cognitive states [15], [31]–[34], and facilitates graph reading via gaze-based highlighting [35]. In immersive analytics, tailoring visualisations to human cognition is critical [36], [37]. Visual cues such as motion, shading, and perspective significantly affect accuracy [7], and eye tracking confirms higher engagement with central 3D view arrangements [38].

B. Graph Visualisation in Virtual and Augmented Reality

Immersive graph visualisation is an active field [4]–[6], with research targeting spatial arrangement [8], navigation [39], viewpoint selection [40], [41], visual encoding [42], and domain applications [43]–[45]. AR enhances immersion and spatial perception [46], while egocentric rendering with adaptive layouts improves response times [10], [11]. S3D projections aid community detection [2] and reveal distinct perceptual trade-offs across 2D, 2.5D, and 3D layouts [8]. Regarding visual encoding, straight edges are preferred over curved ones [47], and combining stereo with kinetic depth cues improves graph comprehension by up to an order of magnitude compared to 2D or stereo-only viewing [1], [3], [48], [49].

Despite these advances, prior work has evaluated immersive graph designs largely using coarse metrics such as task accuracy and completion time (subsection II-B), leaving the temporal dynamics of human graph perception in 3D environments under-explored (subsection II-A). We address this gap by introducing an AOI-based visual analysis of 3D gaze data to reveal how task demands and structural characteristics shape visual attention over time. Furthermore, we mine the gaze transitions of successful analysts and re-inject them into the environment as visual cues, directly linking measured perception to the enhancement of S3D graph visualisations.

III. Study Design and Setup

A. Graph Selection

To evaluate diverse structural properties, we included three undirected single-layer graphs and three undirected multilayer

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

graphs. We restricted graph sizes to 50–75 nodes to balance interpretability and task difficulty, minimising 3D occlusion without inducing cognitive overload [50].

The single-layer graph set comprised two real-world networks from the Konect collection and one synthetic graph, with 3D layouts generated via D3-force[1]. $\mathbf{Graph}_{ic}$ (Iceland[2]) is a social contact graph containing 75 nodes and 114 edges. $\mathbf{Graph}_{ed}$ (Edit-bmwikiquote[3]) uses the largest connected component of a bipartite edit graph, containing 56 nodes and 65 edges. $\mathbf{Graph}_{sy}$ is a synthetic Watts-Strogatz graph of 50 nodes and 100 edges generated via the OGDF library [51]. These graphs were deliberately selected to provide variation across key topological metrics (mean degree, diameter, mean clustering coefficient, mean shortest path): Graph_{ic} (3.04, 6, 0.286, 5.20), Graph_{ed} (2.32, 11, 0, 3.19), and Graph_{sy} (4.00, 6, 0.07, 2.86). All single-layer graphs were bounded within a $1\,m^3$ volume.

For the multilayer graphs, we utilised datasets from prior work [8][4], generated via a modified OGDF algorithm. Each graph consisted of either 3 or 7 layers (30 nodes and 45 edges per layer). We evaluated two spatial arrangements: **2.5D** (layers positioned side by side) and **3D** (layers distributed across a spherical surface surrounding the participant). A standard 2D layout was excluded, because previous research has shown its inferior performance in immersive environments [8].

B. Task Selection

Our study tasks contain three commonly performed graph perceptual tasks [52]:

- Task shortest path (T_{SP}): *How long is the shortest path between the two highlighted nodes (in number of edges)?* Two randomly selected nodes are highlighted with a red sphere. The participants must search for the shortest path. All participants see the same highlighted nodes for each graph. The length was controlled between 3 and 5 to ensure a feasible but non-trivial task complexity in line with previous studies. The selected range allows us to derive solution patterns. The participants should select the answer on the panel (buttons 0 to 9).
- Task highest node degree (T_{HD}): *Which node has the highest number of edges to other nodes?* We ensure that in every graph a unique node exists with the highest degree. For our study, the highest node degree ranged from 7 to 24. The participants must search for the highest-degree node and enter the node label on the answer panel.
- Task common neighbour (T_{CN}): *How many common neighbours do the two highlighted nodes share?* All participants see the same highlighted nodes for each graph. For our study, the answer ranged from 3 to 5. The participants should count common neighbours and select the answer on the panel.

For the multilayer graphs in Study 2, we adopted six analytical tasks from Feyer et al. [8], grounded in the same taxonomy [52]. To maintain comparability with Study 1, our primary analysis focuses strictly on the two shared task types: $\mathbf{T}_{SP}$ (shortest path spanning across different layers) and $\mathbf{T}_{HD}$ (highest cross-layer degree). To preserve task difficulty under layered arrangements, T_{HD} graphs were specifically adjusted to always yield two equally valid highest-degree solutions [8]. The four remaining layer-specific tasks are provided in the supplementary material.

C. Study Design

1) VR Environment: The VR environment was developed using the WebXR Device API[5] and Babylon.js[6] (source code available at https://doi.org/10.18419/DARUS-5259). Participants accessed the application via the immersive mode of the native web browser on a Meta Quest 3 HMD, which maintained a minimum refresh rate of 50 Hz to ensure smooth rendering. Eye tracking data were recorded at 200 Hz using an integrated Pupil Labs NeonXR add-on. To maintain consistent interaction design and visual encoding with prior work [8], [53], nodes (diameter 0.01) were labelled with ascending integers starting from 0. Task-relevant nodes were highlighted in red (#FF0000), irrelevant nodes in yellow (#FFFF00), and all edges were rendered as white lines by the Line System function in Babylon.js (see Figure 1).

The virtual scene simulated a 5×5 metre room with dark blue walls. Graphs were centred at a height of 0.5–1.5 metres, enabling participants to freely walk and explore the structures from multiple perspectives. A virtual question-and-answer panel positioned on one wall served as the sole user interface. Participants interacted with this panel via controller-based raycasting, following a standardised pre-experiment training session to ensure baseline familiarity with the VR controls.

2) Experiment Protocol: Each session commenced with a 5-point eye-tracking calibration, requiring participants to fixate on sequentially presented crosses (centre, then four corners; 1 s display, 1 s interval). Participants then entered the virtual room, where they had unlimited time to read the prompt on the question panel. Clicking the start button revealed the graph in the centre of the room. Participants could freely explore the structure and, once confident, return to the panel to input and confirm their answer. Responses could be revised before confirmation, but revisiting previous tasks was prohibited.

Before the main experiment, participants completed three think-aloud tutorial tasks (one per task type), guided by the experimenter, advancing only after correctly solving all tasks. For the subsequent nine main tasks, task types and graph stimuli were counterbalanced using a Latin square design to prevent consecutive identical graphs or task types.

D. Taxonomy of Task-related Graph Elements

We treat nodes and edges in the graph as Areas of Interest (AOIs) as follows:

[1]https://d3js.org/d3-force
[2]http://konect.cc/networks/iceland/
[3]http://konect.cc/networks/edit-bmwikiquote/
[4]https://doi.org/10.18419/DARUS-3387
[5]https://immersiveweb.dev/
[6]https://www.babylonjs.com/

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

Highlighted nodes are the nodes in red, which are two nodes in the shortest path (T_{SP}) and identify common neighbours task (T_{CN}), but the highest degree task (T_{HD}) contains no highlighted nodes.

Task-relevant elements vary according to the specific task conditions:

- T_{SP}: The nodes and edges compose the shortest path(s) between the two highlighted nodes. When multiple shortest paths exist, all constituent elements are included.
- T_{HD}: The nodes include the node with the highest degree, and the edges include all edges connecting to these highest-degree nodes.
- T_{CN}: The nodes represent the common neighbours shared by both highlighted nodes, and the edges include all connections between these common neighbours and the highlighted nodes.

Confusion elements represent potentially misleading graph components that could interfere with task completion:

- T_{SP}: Confusion nodes and edges comprise all nodes and edges belonging to the second shortest path(s) between the highlighted nodes. When multiple second-shortest paths exist, all constituent elements are included.
- T_{HD}: Confusion nodes include all nodes with the second-highest degree, while confusion edges encompass their connections. When multiple nodes share the second-highest degree, all such nodes and their edges are classified as confusion elements.
- T_{CN}: Confusion nodes encompass all neighbours of either highlighted node (including both common and non-common neighbours), with confusion edges representing all connections between these neighbours and the highlighted nodes.

Other elements include all remaining nodes and edges that do not fall into the above categories. When a graph element could be assigned to multiple categories, we applied the following hierarchical priority: highlighted nodes $>$ task nodes/edges $>$ confusion nodes/edges $>$ other elements.

E. Data Processing

The image plane of the scene camera measured 1,600×1,200 pixels, covering a 103°×77° field of view [54]. We logged the VR headset's position and rotation, along with timestamps and event markers (e.g. task start/done), to segment the collected gaze data. Eye fixations were identified using an enhanced Identify by Velocity Threshold (I-VT) algorithm for head-mounted eye tracking [55]. Timestamps between the VR system and the eye tracker were synchronised by the five-point fixation pattern from the calibration phase (see Section III-C2). Participants' gaze data were analysed by reconstructing the study scene using the Python Trimesh library[7]. To account for eye tracking inaccuracies and foveal imprecision [56], [57], a cone-based ray-casting method was applied [58], recording all intersected AOIs and their distances from the ray origin. See supplementary material for details on gaze-ray casting.

[7]https://trimesh.org/

IV. Understanding Graph Perception

A. Hypothesis

Our study aims to understand the key factors influencing users' visual attention and how this behaviour reflects the answer correctness of visual analytical tasks in S3D graph visualisations. All study data and analysing scripts are available at https://doi.org/10.18419/DARUS-5950. As human visual attention is influenced by both bottom-up (graph) and top-down (task) factors [36], [37], we hypothesise the effects:

- H1: Graph significantly influences users' gaze behaviour during graph perception in S3D. Since graph characteristics comprise multiple structural measures, H1 is further divided into the following sub-hypotheses:
  - H1.1: Graph arrangement
  - H1.2: Graph complexity
- H2: Task type significantly influences users' gaze behaviour during graph perception in S3D.
- H3: Users' gaze behaviour is significantly correlated with answer correctness during graph perception in S3D.

Our hypotheses follow from prior evidence. H1 reflects bottom-up control of attention: graph structure and spatial arrangement are known to shape readability and scanning behaviour in both 2D screen [13], [14], [21] and immersive settings [8], [11], so we expect the arrangement (H1.1) and complexity (H1.2) of multilayer graphs to influence gaze. Viewing strategies have been shown to differ systematically across graph tasks [30], [35], leading us to expect task-driven control of gaze behaviour (H2). Because more efficient, goal-directed scanning has been linked to better outcomes in graph reading [18], [20], we expect gaze behaviour to correlate with answer correctness (H3).

We break the hypotheses down under these gaze metrics: a) saccade velocity, b) saccade length, and c) task-related fixation ratio, i.e. the ratio of fixation counts landing on task-related elements. These metrics are informative for graph reading: saccade velocity and saccade length capture, respectively, how fast and how far the gaze moves, serving as physiological indicators of search effort and cognitive load (slower, shorter saccades signal more localised, effortful inspection), while the task-related fixation ratio directly measures whether attention reaches the elements that matter for the task rather than distractors. Used together, they let us separate *how* a graph is scanned from *where* attention lands, and relate both to performance. As such, every hypothesis consists of three sub-hypotheses, such as H1.1(b), which states that graph arrangements significantly affect saccade length during graph perception in S3D. Since graph arrangement and complexity are less meaningful in single-layer graphs, H1 was tested only on multilayer graphs.

B. Study 1: Single-layer Graphs

We conducted our user study at the local university with the approval of the ethics committee. We recruited 18 participants (11 male, 7 female) between 21 and 28 years old ($\mu = 25.2$, $\sigma = 2$ years). All 18 participants had no visual impairment but limited VR experience. Participants were compensated 10€ for the session, which included approximately five minutes

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

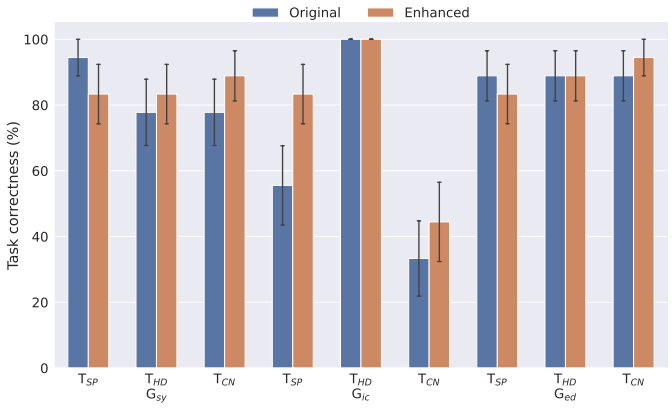

Fig. 2. Average task correctness across task types and graphs. Error bars indicate the standard error. Original: the single-layer graphs in Study 1, Enhanced: the gaze-enhanced graphs in Study 3 (Section VI).

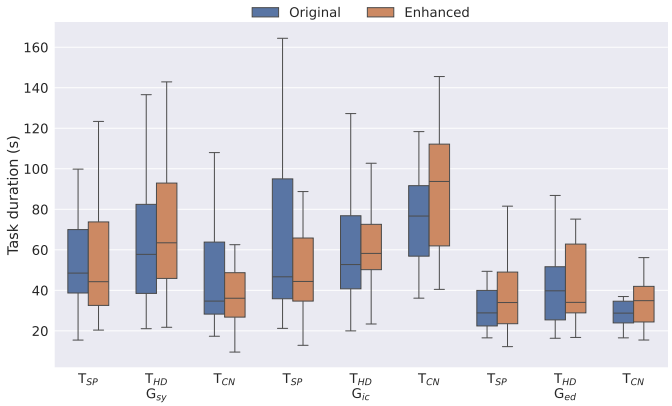

Fig. 3. Task duration for the original (Study 1) and enhanced (Study 3) conditions across various task types and graph datasets. The boxes span the first to third quartiles, with the median indicated by the centre line.

for setup and calibration, and three minutes for tutorial tasks. Using a counterbalanced Latin square design, each participant completed nine main tasks (three task types $\times$ three graphs).

1) Task Performance Overview: Participants achieved a mean correctness of 78.4% (Figure 2): 79.6%, 88.9%, and 66.7% for T_{SP}, T_{HD}, and T_{CN}, and 83.3%, 63.0%, and 88.9% for Graph_{sy}, Graph_{ic}, and Graph_{ed}. The mean completion time was 53.85 s ($\sigma = 33.28$ s; Figure 3), with per-graph means of 57.44 s, 69.20 s, and 34.93 s for Graph_{sy}, Graph_{ic}, and Graph_{ed}; participants spent most time on T_{HD} (57.45 s) and least on T_{SP} (50.55 s).

2) Gaze Behaviour: Overall, participants exhibited a mean saccade velocity of 27.28°/s ($\sigma = 5.00$°/s) and a mean saccade length of 7.89° ($\sigma = 1.45$°). Task type had no significant main effect on either saccade velocity ($F = 1.601, p = 0.205$) or saccade length ($F = 1.210, p = 0.301$), rejecting H2(a) and H2(b) for single-layer graphs.

The overall mean fixation count was 145.65 ($\sigma = 84.95$), remaining relatively stable across tasks: T_{SP} (149.11), T_{HD} (139.61), and T_{CN} (148.24). However, because Graph_{ic} elicited significantly more fixations ($\mu = 181.98$) than Graph_{sy} ($\mu = 154.19$) and Graph_{ed} ($\mu = 100.80$), we normalised the data by analysing the *task-related fixation ratio*. A two-way ANOVA on this ratio revealed significant main effects for graph ($F = 22.758, p < 0.001$) and task ($F = 36.233, p < 0.001$), as well as a significant interaction effect ($F = 7.063, p < 0.001$). Post-hoc Tukey tests confirmed significant pairwise differences between T_{SP} and T_{CN} in both Graph_{sy} and Graph_{ic}, and between T_{SP} and T_{HD} in Graph_{ed}, confirming H2(c).

Correlations between answer correctness and gaze metrics (saccade velocity, length, and task-related fixation ratio) were not significant for T_{SP} and T_{HD}, partially rejecting H3(a–c). However, under T_{CN}, answer correctness correlated significantly with saccade velocity ($r_s = 0.401, p = 0.003$), saccade length ($r_s = 0.386, p = 0.004$), and task-related fixation ratio ($r_s = -0.353, p = 0.009$), confirming H3(a–c) exclusively for this task type. Furthermore, T_{CN} correctness exhibited a strong negative correlation with task duration ($r_s = -0.492, p < 0.001$). This indicates that for common-neighbour tasks, prolonged visual search did not guarantee success, suggesting that participants who struggled exerted considerable but ultimately ineffective cognitive effort.

C. Study 2: Multilayer Graphs

To investigate the generalisability of the findings in single-layer graphs, multilayer graphs were employed to investigate graph arrangements and complexity. 11 participants were recruited (7 male, 4 female), with ages ranging from 23 to 58 years ($\mu = 34.5$, $\sigma = 13.2$ years). None of the participants had taken part in the single-layer graph study. For vision problems, two participants reported astigmatism, and one reported presbyopia. They were compensated with 15€ for their participation. The tasks only include T_{SP} and T_{HD}, as the common neighbour task (T_{CN}) was found to be easy for multilayer graphs in a pilot study. T_{SP} was adjusted to exclude nodes in the same layer as the highlighted nodes. For T_{HD}, the graphs were adjusted to always have 2 solutions [8]. See supplementary material for the full report on multilayer graphs.

1) Task Performance Overview: Across the analysed tasks, mean task correctness was 56.8% for T_{SP} and 72.7% for T_{HD}, and mean completion time was 50.57 s for T_{SP} and 73.31 s for T_{HD}. However, arrangement effects were task-dependent. 3D clearly outperformed 2.5D for T_{SP} (72.7% vs. 40.9%), whereas T_{HD} was virtually unaffected (72.7% vs. 72.7%). Across all six tasks of the full study, the 3D advantage (80.1% vs. 72.4%) stemmed mainly from T_{SP} and cross-layer pattern comparison, and did not hold for within-layer comparison tasks. Likewise, correctness decreased from 3 to 7 layers on average (80.1% vs. 72.4%) and for T_{HD} (86.4% vs. 59.1%), but not for T_{SP} (50.0% vs. 63.6%).

The similar T_{HD} means conceal an arrangement $\times$ complexity interaction: 2.5D correctness dropped from 100% (11/11 correct) at 3 layers to 45.5% (5/11 correct) at 7 layers, while 3D remained at 72.7% (8/11 correct). This likely reflects viewing geometry: three side-by-side layers fit into a single outside-in view, whereas seven span a wider field of view with long cross-layer edges passing in front of intermediate layers; the surrounding 3D arrangement keeps all layers at a similar distance. Consistently, 2.5D-7L

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

yielded the lowest confidence and the longest completion time for T_{HD} (82.43 s). Given 11 responses per cell, we report this interaction descriptively. The full per-condition breakdown of correctness and completion time is reported in the supplementary material.

2) Gaze Behaviour: The mean saccade length was 8.30° for T_{SP} and 7.64° for T_{HD}, and the mean saccade velocity was 26.17°/s for T_{SP} and 22.44°/s for T_{HD}. See the per-condition means of all gaze metrics in the supplementary material.

To test our hypotheses on graph arrangements (H1.1), complexity (H1.2), and task types (H2), we conducted a three-way ANOVA on gaze metrics. No significant effects of graph arrangements and complexity were found on saccade velocity and saccade length (all $p > 0.05$), rejecting H1.1(a-b) and H1.2(a-b). Both H1.1(c) and H1.2(c) are confirmed with significant effects of graph arrangement (F(arrangement) = 20.947, $p < 0.001$) and graph complexity (F(complexity) = 4.538, $p = 0.036$). Task type had a significant effect on saccade velocity for T_{SP} and T_{HD}, F(task) = 4.77, $p = 0.032$, confirming H2(a). The effect of task type on saccade length was not significant, F(task) = 1.64, $p = 0.205$, rejecting H2(b). For T_{SP}, the mean fixation count was 116.77 ($\sigma = 60.85$), while for T_{HD} it was 135.09 ($\sigma = 74.84$).

Similar to Study 1, we instead analysed the task-related fixation ratio. ANOVA revealed no significant effect of task (F(task) = 2.675, $p = 0.106$), thus rejecting H2(c). However, a significant interaction was found between arrangement and task (F(arrangement:task) = 4.227, $p = 0.043$), whereas no other significant interactions were found ($p > 0.1$). A post-hoc Tukey test indicated that for T_{SP}, the 3D arrangement yielded a higher proportion of task-related fixation ratio than the 2.5D arrangement. No significant difference between arrangements was observed for T_{HD} ($p = 0.317$). In T_{SP}, no significant effects of task correctness were found on saccade velocity, saccade length, and task-related fixation ratio, rejecting H3(a-c). For T_{HD}, only the task-related fixation ratio significantly correlated with correctness (Spearman's $r_s = 0.386$, $p = 0.010$), accepting H3(c), but rejecting H3(a-b).

V. Gaze Behaviour Analysis

We first adapted two complementary visualisation techniques to capture both the temporal dynamics and spatial distributions of gaze behaviour, illustrating how participants explored the graphs while solving the task. Then, we analysed their viewing strategies by mining gaze patterns, providing qualitative insights into users' task-solving strategies.

A. Temporal Distribution

We employed uncertainty-aware scarf plots [59] to analyse the temporal evolution of participants' gaze behaviour. Each vertical bar represents a fixation event, with its width encoding duration. Coloured segments within the bars denote the intersected AOIs (nodes and edges), mapped to the eight-category classification defined in Section III-E. The implementation details are provided in the supplementary material.

The scarf plot (Figure 4) revealed consistent spatio-temporal patterns. Fixations on task-related AOIs typically occurred only after initial attention fell on highlighted nodes and confusion-associated regions. Furthermore, task-relevant AOIs systematically shifted from the background (top) to the foreground (bottom) over time (green block moving to bottom across consecutive fixations). This indicates that participants actively adjusted their viewpoints to minimise spatial distance and reduce visual clutter, thereby enhancing the visibility of relevant structures. The primary takeaway is therefore both temporal and depth-related: successful task solving is preceded by a deliberate spatial reorganisation that progressively brings relevant elements into the foreground. Recognising that the early dominance of confusion elements signals the exact moment viewpoint adjustment is most needed directly motivates the gaze-derived visual cues introduced in Section VI.

B. Spatial Distribution

To analyse the spatial distribution of visual attention, we aggregated fixation data across all participants for each task-graph combination to compute AOI hit frequencies. Heatmaps and implementation details are provided in the supplementary material. The resulting saliency distributions reveal that participants adapted their viewing strategies to specific task requirements while remaining strongly guided by the underlying topological features. Specifically, fixations consistently concentrated on structurally significant areas, including high-degree hubs, bridging nodes, and tightly interconnected clusters. This pattern indicates that viewers actively leveraged the graph's inherent organisation to efficiently isolate task-relevant information, rather than relying on exhaustive visual search.

We also observed anomalously high saliency on a small subset of long, task-irrelevant edges spanning large portions of the graph volume. This inflation is a methodological artifact rather than a perceptual phenomenon. Because our cone-based raycasting registers an intersection if any segment of an edge falls within the gaze cone, long edges crossing multiple sightlines accumulate incidental hits, whereas short, localised edges do not. We therefore attribute this elevated saliency strictly to edge-level intersection accounting. This limitation could be mitigated in future pipelines by spatially partitioning long edges into discrete subregions (Section VII-D).

C. Gaze Pattern Mining

We analysed participants' viewing strategies using a differential pattern mining approach [60], which identifies systematic gaze differences between correct and incorrect responders, revealing attention patterns linked to task correctness.

We focused on 2-pattern activities, which capture the sequential precedence $A \succ C$ between two AOIs, indicating that some fixations on element A occurred before some fixations on element C within a viewing sequence. Since each fixation may intersect multiple graph elements, each item in a sequence represents not a single AOI, but an array of candidate AOIs identified by gaze ray casting. To handle this uncertainty, we employed a weighted approach when computing pattern support. If fixation f_i intersects k_i AOIs and fixation f_j intersects k_j AOIs (f_i precedes f_j), each of the $k_i \times k_j$ possible 2-patterns receives a weight of $\frac{1}{k_i \times k_j}$. Following Schlieder et

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

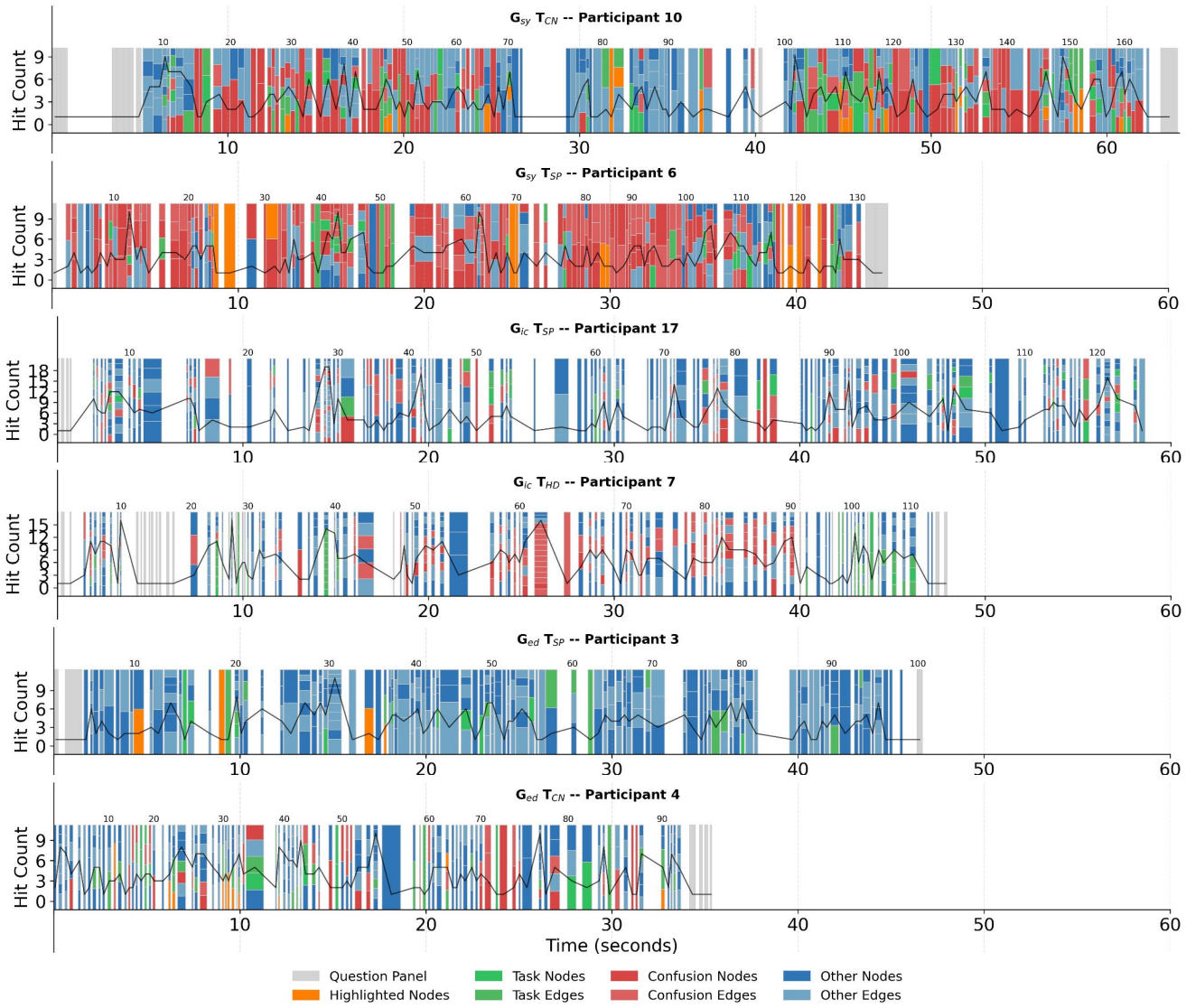

Fig. 4. Sample timeline (scarf-plot) visualisations of gaze behaviour, enlarged to full width for legibility. Each vertical bar represents a fixation event, with width indicating duration and coloured blocks representing the AOIs that intersected the eye-gaze ray; colours follow the eight-category AOI scheme of Section III-E. Within each block, objects are stacked from closest (bottom) to farthest (top).

al. [60], we computed two support metrics for each pattern: *s-support* (sequence support) measures the fraction of sequences containing the pattern, while *i-support* (instance support) measures the average weighted occurrence count of the pattern within sequences.

1) Differential Pattern Identification: We partitioned the viewing sequences into two groups by task correctness: the *pos* group and the *neg* group. We then applied sequential pattern mining (SPM), which identifies *differentially frequent* patterns satisfying two criteria: (1) *s-support* exceeds 0.5 in at least one group; and (2) either the pattern appears in only one group, or *i-support* differs significantly between groups as determined by a Student's T-test ($p < 0.05$). Patterns fulfilling both criteria were classified as *pos*-patterns, distinctive of successful task performance; *neg*-patterns, associated with incorrect responses; and *both*-patterns, occurring in both groups with significantly different frequencies.

The differential i-support, defined as $\Delta i = i_{pos} - i_{neg}$, serves as the primary metric for ranking patterns. Patterns with large positive Δi values indicate viewing sequences that successful participants employed more frequently, highlighting critical attention transitions underlying effective graph-reading strategies. Conversely, patterns with large negative Δi values reveal potentially inefficient viewing behaviours. We mine gaze patterns from both groups and output the differentially frequent 2-patterns with corresponding labels (see supplementary material for the Algorithm).

Since pattern mining relies on the same conical raycasting, the long-edge artifact also enters the viewing sequences. However, its effect on pattern identification is limited. Unlike the pooled saliency maps, the differential analysis compares the *pos* and *neg* groups on identical graph geometry. Inflation affecting both groups equally, whether as an offset or a proportional scaling, preserves the sign of Δi and leaves the t-test unchanged, since the t-statistic is invariant to a common shift and scale; inflation unrelated to group membership only

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

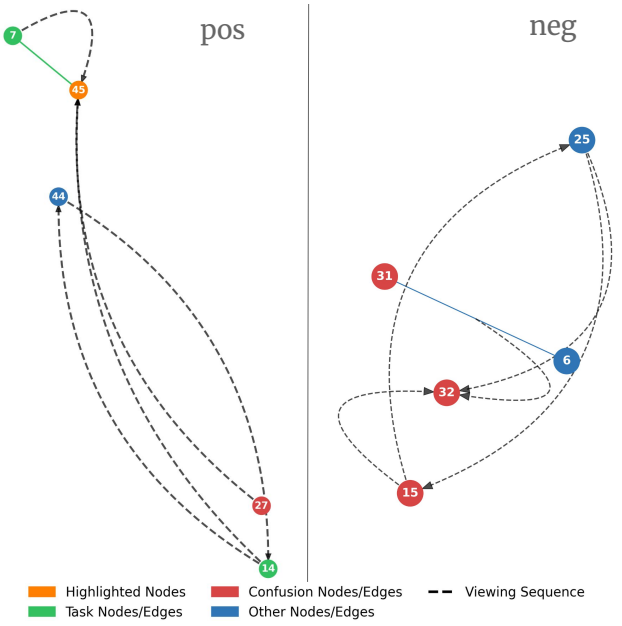

Fig. 5. Precedence state diagrams for the Graph$_{ed}$ under task shortest path, showing the top 5 differentially frequent patterns for the *pos*-group (left) and *neg*-group (right).

adds variance. Moreover, incidental hits pair a long edge with varying partners and rarely form a consistent 2-pattern with *s-support* above 0.5. Residual effects may amplify $|\Delta i|$ and thus the ranking of edge-involving patterns, but cannot introduce distractors into the Study 3 cues, which are restricted to task-relevant elements (Section VI).

2) Pattern Selection: Representative patterns were selected from the *pos*- and *neg*-groups to characterise their respective reasoning strategies. For the *pos*-group, we prioritised differentially frequent patterns assigned to the *pos* or *both* groups with positive differential i-support ($\Delta i > 0$), selecting the five patterns with the highest Δi values. If fewer than five patterns met these criteria, all qualifying patterns were included. The selection procedure for the *neg*-group followed the same principle, prioritising patterns from the *neg* or *both* groups with negative differential i-support ($\Delta i < 0$), and selecting the five with the lowest Δi values.

3) Precedence State Diagrams: To visualise the identified patterns, we constructed precedence state diagrams for each task condition, separately illustrating the *pos* and *neg* groups. In these diagrams, nodes and edges are projected onto the XY-plane. Black dashed arrows denote temporal precedence relations, indicating the sequential order of visual attention. The colour coding of AOIs follows the category definitions in Section III-E.

4) Shortest Path Task Strategies: We demonstrate the viewing strategies revealed by differential pattern mining using Graph$_{ed}$ under the shortest path task as an exemplar. The precedence state diagrams (Figure 5) reveal clear strategic differences between successful and unsuccessful participants in the shortest path task. The *pos* group showed a goal-directed approach focused on task-relevant elements. Dominant patterns including <Node 14, Node 44>, <Node 44, Node 14>, <Edge 7-45, Node 45>, <Node 14, Node 45> indicate a path-tracing strategy, where participants sequentially examined nodes and edges forming the shortest path. Patterns involving intermediate nodes reflect methodical verification of path continuity, while the circle between Node 14 and Node 44 suggests revisiting key nodes to confirm the shortest path.

In contrast, the *neg* group exhibited a less focused search strategy dominated by attention to confusion elements. The most prominent patterns – from Node 25, 15, and Edge 6-31 to Node 32 – predominantly involve confusion nodes (Nodes 15, 31, and 32) that belong to alternative, non-optimal paths. The high differential i-support values (ranging from -55.19 to -70.19) indicate that unsuccessful participants devoted substantially more attention to these misleading elements. The circus between Node 25 and Node 15 suggests participants attempted to verify connections between confusion and unrelated nodes, potentially exploring alternative paths without successfully identifying the optimal solution. This dispersed search strategy, which failed to distinguish between optimal and suboptimal paths, represents a misidentification of task-relevant elements.

VI. Gaze-based Graph Enhancing

A. Study Design

To validate the effectiveness of the patterns identified through differential mining, we conducted a follow-up experiment (Study 3) with 18 new participants (15 male, 3 female) between 23 and 29 years old ($\mu = 26.06$, $\sigma = 1.90$). Visual cues were designed based on task-relevant patterns from the *pos* group. Apart from the inclusion of visual cues, the experimental setup, apparatus, stimuli, and task conditions were identical to the initial study (Study 1). Participants performed the same nine trials (3 graphs $\times$ 3 task types) with identical instructions.

The experimental procedure replicated the protocol from Study 1 (Section III-C2). Participants completed the identical five-point calibration of the eye tracker, three mandatory tutorial tasks, and nine counterbalanced main trials. The sole experimental difference was the addition of gaze-derived visual cues to the graphs; all underlying structures, highlighted target nodes, and correct answers remained identical. Following the trials, participants completed a closing questionnaire to provide the Likert-scale ratings and free-text feedback.

Visual cues were constructed by selecting patterns that exclusively involved task-relevant graph elements. Specifically, we filtered patterns where both elements belonged to the categories of highlighted nodes, task nodes, or task edges as defined in Section III-E. The selection prioritised differentially frequent patterns from the *pos* or *both* groups with positive differential i-support ($\Delta i > 0$), extracting the top 3 patterns ranked by Δi for each task configuration. When fewer than 3 differentially frequent patterns met these criteria, we relaxed the differential frequency constraint and selected from all eligible patterns based on Δi values.

From the selected patterns, we extracted all unique graph elements (nodes and edges) and used these as the basis for visual guidance. The visual cue implementation augmented

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

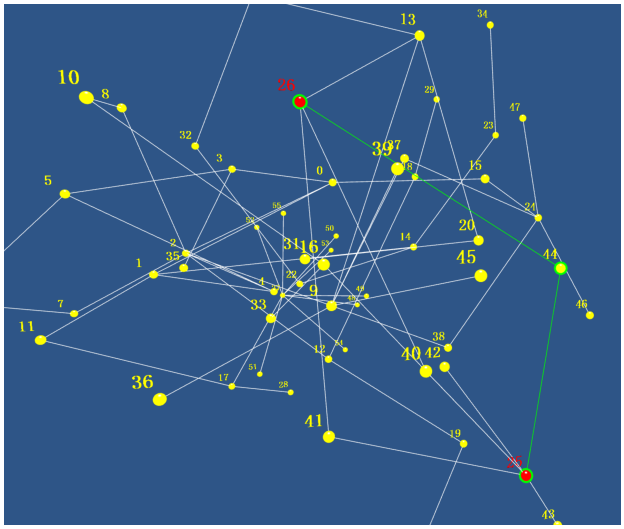

Fig. 6. The gaze-enhanced visual representation under G_{ic}, T_{CN}. Cued task-relevant nodes and edges, derived from the gaze patterns of successful (*pos*-group) analysts, are emphasised with green outlines.

the original 3D graph visualisation by rendering green outlines around the identified task-relevant graph elements, i.e. a green halo around each cued node and a green casing along each cued edge, leaving colour, size, and position of all other elements unchanged so that the cue added emphasis without altering the underlying layout (see Figure 6).

B. Results on Study 3: Gaze-Enhanced Graphs

We evaluated the visual cues in a between-subject study, comparing the performance of 18 new participants against Study 1 across the identical nine task–graph trials.

Task correctness and completion times for the gaze-enhanced graphs are presented in Figure 2 and Figure 3. Overall mean correctness increased from 78.4% ($\sigma = 41.3$) in the baseline to 83.3% ($\sigma = 37.4$) under the enhanced condition. Across the nine specific trials, correctness improved in five instances, remained unchanged in two, and decreased in two. Both decreases occurred exclusively during the shortest-path task, where correctness dropped from 94.4% (17/18 correct) to 83.3% (15/18 correct) on Graph_{ic}, and from 88.9% to 83.3% on Graph_{sy}. Despite the aggregate upward trend, statistical analysis revealed no significant difference in overall correctness between the baseline and enhanced graphs.

The user study on the gaze-enhanced graphs yielded generally positive results, with participants reporting high satisfaction across most aspects of the experience. On a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), they found the gaze-enhanced graphs effective and easy to interpret. Participants agreed that the visual hints facilitated the identification of important nodes and edges ($\mu = 6.0$, $\sigma = 0.97$) and reduced mental effort during task solving ($\mu = 5.94$, $\sigma = 1.16$). The visual hints were also perceived as clearly distinguishable from regular graph elements ($\mu = 6.11$, $\sigma = 0.76$) and as not introducing visual clutter ($\mu = 6.06$, $\sigma = 1.21$). The participants' qualitative feedback further supported these findings, with comments such as "The visual hints helped identify the starting/potential points" from P12 and P3, and "made it easier to trace the nodes in a dense situation" from P11.

VII. Discussion

A. Graphs, Tasks, and Gaze Behaviour

a) Single-layer graphs: Our findings reveal nuanced effects of task type and answer correctness on gaze behaviour (H2, H3). In single-layer graphs, task type and response correctness did not significantly affect saccadic eye movements, rejecting H2(a,b) and H3(a,b). However, we identified a distinct inverse relationship between graph density and saccadic metrics. As the mean degree of graphs decreases, saccade length and velocity monotonically increase (from 7.53° at 25.94°/s, to 7.80° at 26.39°/s, up to 8.34° at 29.52°/s), respectively for Graph_{sy} (4.0), Graph_{ic} (3.12), and Graph_{ed} (2.32). This trend indicates that denser graphs constrain visual scanning to localised regions, whereas sparser structures encourage broader visual exploration.

Although we formulated no prior hypotheses regarding overall complexity in single-layer graphs, we observed a strong correlation between structural complexity and task-related fixation ratios. Graph_{ic}, exhibiting the highest complexity (75 nodes, 114 edges, 5.2 mean shortest path), elicited the highest task-related fixation ratio (46.64%), compared to the simpler Graph_{sy} (34.11%) and Graph_{ed} (32.13%). The higher cognitive load [61] and visual clutter of complex graphs likely hinder goal-directed search, so users devote more fixations to task-relevant elements to compensate.

b) Multilayer graphs: The task-related fixation ratio was significantly affected by H1.1(c) graph arrangement and H1.2(c) graph complexity. Participants tended to view the 2.5D arrangements from the side rather than entering them, reducing task-related fixations. Moreover, performance in T_{SP} was generally lower in 2.5D, accompanied by a smaller task-related fixation ratio, likely due to the difficulty of perceiving cross-layer shortcuts. Increasing the number of layers from 3 to 7 further decreased the task-related fixation ratio. For T_{HD}, arrangement affected not mean accuracy but its robustness: 2.5D dropped from 100% to 45.5% as layers increased, while 3D remained stable (Section IV-C). Interestingly, participants self-reported higher confidence and had shorter task completion time in T_{SP} (50.57 s) than in T_{HD} (73.31 s), despite lower task correctness in T_{SP}. This suggests that while the shortest-path task appeared easy, verification was challenging.

c) Single-layer vs. Multilayer Graphs: Contrasting the two studies reveals that layered structures fundamentally shift viewing behaviour. In single-layer graphs, attention is governed by density and complexity, constraining exploration to local regions. In multilayer graphs, spatial arrangement dominates: 3D spherical layouts invite users to look *into* the structure (raising task-related fixations and, for T_{SP}, accuracy), while 2.5D layouts encourage passive, outside-in viewing, which remains effective only while all layers are visible together. Furthermore, the confident-but-wrong phenomenon in T_{SP} is highly pronounced in multilayer graphs where cross-layer shortcuts are easily missed.

These findings yield concrete design guidelines: 1) Favour single-layer S3D graph for local topology tasks (e.g. node

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

degree) if density is moderate, as denser single-layer graphs already constrain exploration. 2) Use 3D spherical multilayer for tasks requiring cross-layer integration, as it invites necessary immersive exploration, improving accuracy for cross-layer path tracing (T_{SP}) and keeping cross-layer comparison (T_{HD}) stable as layers increase; 2.5D remains acceptable for within-layer comparisons or graphs with few layers. 3) Limit layer count universally, as added layers measurably dilute attention. 4) Provide verification cues for tasks with high confidence but low accuracy (e.g. T_{SP}) to prevent anchoring.

B. Spatio-temporal Gaze Patterns in Graph Perception

Eye tracking captures rich spatio-temporal data: scarf plots reveal the temporal evolution and sequential ordering of participants' focus, while saliency maps characterise the spatial distribution of attention across graph structures.

Our analysis reveals distinct task-dependent viewing strategies that reflect underlying cognitive processes. In the shortest path task, participants displayed distributed fixation patterns across intermediate vertices, with saliency maps (see supplementary material) showing moderate activation around candidate nodes and scarf plots (see Figure 4) capturing frequent transitions between task-relevant and irrelevant vertices – a breadth-first search strategy comparing multiple routes before identifying the optimal path. In the highest-node-degree task, attention concentrated on structurally prominent nodes (dark blue clusters in the saliency distribution) with fewer scarf-plot transitions outside task-relevant areas, reflecting a targeted search that efficiently isolated high-degree vertices. The common-neighbour task elicited a distinctly localised gaze pattern, with fixations concentrated on the neighbourhoods of highlighted nodes and sparse scarf-plot transitions, indicating a local comparison strategy confined to nearby nodes.

Despite these task-specific strategies, some consistent gaze patterns emerged. In particular, the hub nodes attracted human visual attention across all tasks, implying that participants consistently prioritised structurally significant elements.

C. Gaze-based Graph Enhancement

The gaze-derived cues are not equivalent to revealing the solution, since only elements from the top-ranked *pos*-group patterns are cued. Participants usually see a subset of the task-relevant structure and must still locate the rest and verify the answer themselves. Although the gaze cues were effective overall, correctness decreased for the shortest-path task (Figure 2). Emphasising elements frequently attended by successful participants may have biased attention towards locally salient but suboptimal regions, disrupting the global integration required for path tracing; overly directive guidance is known to narrow attentional scope [42]. The salient cues may also have anchored participants, discouraging exploration of alternative routes. Gaze-based enhancement thus facilitated counting and comparison tasks (e.g. node degree, common neighbours) but hindered shortest-path searching.

These findings call for task-driven, adaptive cue design. Tasks requiring global reasoning, such as path tracing, need context-aware cues that support spatial orientation without dictating solution paths, whereas localised, feature-based tasks benefit from direct cues that guide attention and reduce mental effort. Future gaze-adaptive systems could adjust cue encodings to the task at hand.

D. Limitations and Future Work

Sample Size and Generalisability: Our sample sizes (18, 11, and 18 participants for the three studies) limit statistical power, so non-significant differences (e.g. the post-enhancement accuracy gain) should be read as trends rather than confirmed effects. Moreover, pattern mining and gaze-derived cues were evaluated on specific graph instances and tasks; their generalisability to larger graphs, other topologies, and broader populations remains to be confirmed.

Challenges of Perception Analysis in Graph Visualisations: Analysing gaze in 3D graphs presents two methodological challenges. First, conical raycasting introduces depth ambiguity: a single fixation in a dense region can intersect multiple elements along the line of sight, obscuring the true target. Second, long edges spanning large 3D regions accumulate spurious hits because intersecting any segment flags the entire edge. In our analysis, this mainly affects pooled saliency, as the *pos*/*neg* contrast largely offsets it in pattern mining (Section V-C1); a sensitivity analysis excluding long edges would quantify the residual effect. Mitigating these issues requires sophisticated gaze-depth estimation [62] to pinpoint the attended subregion or nearest intentional target.

Real-time Gaze-based Applications: Our setup relied on offline fixation detection. A natural next step is gaze-responsive graphs that adapt to the analyst's viewing pattern in real time, which requires overcoming low-latency processing, robust pattern detection from short sequences, and gaze uncertainty [63] – constraints avoided by our offline AOI-level analysis. One possible solution is replacing the offline pattern-mining pipeline with saliency models, enabling real-time cue selection without a pre-computed successful-analyst cohort.

VIII. Conclusion

We presented a series of eye tracking studies investigating human perception of S3D graph visualisations in VR. Our spatio-temporal analysis revealed distinct, task-driven viewing strategies, demonstrating that gaze metrics effectively quantify cognitive load and visual complexity. Crucially, by mining the distinct gaze transitions of successful analysts and re-injecting them into the environment as sequential visual cues, we actively improved task correctness across most conditions. Moving from single-layer to multilayer graphs, we demonstrated that spatial arrangement and structural complexity fundamentally reshape human visual attention. These findings yield concrete guidelines for building effective, gaze-aware applications for immersive analytics.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161 (Project A07, A09, and D04).

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

REFERENCES

[1] C. Ware and G. Franck, "Evaluating stereo and motion cues for visualizing information nets in three dimensions," *ACM Trans. Graph.*, vol. 15, no. 2, p. 121–140, Apr. 1996.

[2] N. Greffard, F. Picarougne, and P. Kuntz, "Visual community detection: An evaluation of 2D, 3D perspective and 3D stereoscopic displays," in *Graph Drawing*. Berlin, Heidelberg: Springer, 2012, pp. 215–225.

[3] D. Belcher, M. Billinghurst, S. Hayes, and R. Stiles, "Using augmented reality for visualizing complex graphs in three dimensions," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* IEEE, 2003, pp. 84–93.

[4] M. Kraus, J. Fuchs, B. Sommer, K. Klein, U. Engelke, D. Keim, and F. Schreiber, "Immersive analytics with abstract 3D visualizations: A survey," in *Computer Graphics Forum*, vol. 41. Wiley Online Library, 2022, pp. 201–229.

[5] L. Joos, M. T. Fischer, J. Rauscher, D. A. Keim, T. Dwyer, F. Schreiber, and K. Klein, "Visual network analysis in immersive environments: A survey," *arXiv preprint arXiv:2501.08500*, 2025.

[6] A. Fonnet and Y. Prié, "Survey of immersive analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 2101–2122, 2021.

[7] F. Yang, J. Tompkin, L. Harrison, and D. H. Laidlaw, "Visual cue effects on a classification accuracy estimation task in immersive scatterplots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 12, pp. 4858–4873, 2022.

[8] S. P. Feyer, B. Pinaud, S. Kobourov, N. Brich, M. Krone, A. Kerren, M. Behrisch, F. Schreiber, and K. Klein, "2D, 2.5D, or 3D? An exploratory study on multilayer network visualisations in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 1, pp. 469–479, 2023.

[9] N. Greffard, F. Picarougne, and P. Kuntz, "Beyond the classical monoscopic 3D in graph analytics: An experimental study of the impact of stereoscopy," in *2014 IEEE VIS International Workshop on 3DVis (3DVis)*, 2014, pp. 19–24.

[10] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma, "A study of layout, rendering, and interaction methods for immersive graph visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 7, pp. 1802–1815, 2016.

[11] J. Sorger, A. Arleo, P. Kán, W. Knecht, and M. Waldner, "Egocentric network exploration for immersive analytics," *Computer Graphics Forum*, vol. 40, no. 7, pp. 241–252, 2021.

[12] J. Miller, M. Ghoniem, H.-Y. Wu, and H. C. Purchase, "On the perception of small sub-graphs," in *31st International Symposium on Graph Drawing and Network Visualization (GD 2023)*. Springer, 2023, pp. 213–230.

[13] U. Soni, Y. Lu, B. Hansen, H. C. Purchase, S. G. Kobourov, and R. Maciejewski, "The perception of graph properties in graph layouts," *Computer Graphics Forum*, vol. 37, no. 3, pp. 169–181, 2018.

[14] E. Kypridemou, M. Zito, and M. Bertamini, "The effect of graph layout on the perception of graph properties." *EuroVis (Short Papers)*, pp. 1–5, 2020.

[15] Y. Wang, W. Wang, A. Abdelhafez, M. Elfares, Z. Hu, M. Bâce, and A. Bulling, "SalChartQA: Question-driven saliency on information visualisations," in *the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–14.

[16] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, "State-of-the-Art of Visualization for Eye Tracking Data ," in *EuroVis - STARs*, 2014.

[17] C. Körner, "Sequential processing in comprehension of hierarchical graphs," *Applied Cognitive Psychology*, vol. 18, no. 4, pp. 467–480, 2004.

[18] W. Huang, "Using eye tracking to investigate graph layout effects," in *2007 6th International Asia-Pacific Symposium on Visualization*. IEEE, 2007, pp. 97–100.

[19] Y. Zhang, N. Gröne, G. Liotta, F. Schreiber, and K. Klein, "Investigating crossing perception in 3d graph visualisation," in *33rd International Symposium on Graph Drawing and Network Visualization (GD 2025)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025, pp. 52–1.

[20] B. Sharif and J. I. Maletic, "An eye tracking study on the effects of layout in understanding the role of design patterns," in *2010 IEEE International Conference on Software Maintenance*, 2010, pp. 1–10.

[21] M. Pohl, M. Schmitt, and S. Diehl, "Comparing the Readability of Graph Layouts using Eyetracking and Task-oriented Analysis ," in *Computational Aesthetics in Graphics, Visualization, and Imaging*, 2009.

[22] S. Chen, J. Epps, N. Ruiz, and F. Chen, "Eye activity as a measure of human mental effort in hci," in *the 16th International Conference on Intelligent User Interfaces*, 2011, pp. 315–318.

[23] L. Zhang, Y. Wang, Y. Zhang, W. Kerle-Malcharek, K. Klein, F. Schreiber, and A. Bulling, "Towards a better understanding of graph perception in immersive environments," in *33rd International Symposium on Graph Drawing and Network Visualization (GD 2025)*, 2025, pp. 1–19.

[24] A. Plopski, T. Hirzle, N. Norouzi, L. Qian, G. Bruder, and T. Langlotz, "The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality," *ACM Computing Surveys*, vol. 55, no. 3, Mar. 2022.

[25] V. Sundstedt and V. Garro, "A systematic review of visualization techniques and analysis tools for eye-tracking in 3D environments," *Frontiers in Neuroergonomics*, vol. 3, 2022.

[26] R. Netzel, M. Burch, and D. Weiskopf, "Comparative eye tracking study on node-link visualizations of trajectories," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2221–2230, 2014.

[27] G. J. Mooney, H. C. Purchase, M. Wybrow, S. G. Kobourov, and J. Miller, "The Perception of Stress in Graph Drawings," in *32nd International Symposium on Graph Drawing and Network Visualization (GD 2024)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 320, 2024, pp. 21:1–21:17.

[28] X. Zhao, S. Fu, R. Yang, L. Yang, Y. Chen, J. Zhang, J. Long, F. Zhou, and Y. Zhao, "Investigating visual perception of degree centrality in graph visualization," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2025.

[29] V. Yoghourdjian, D. Archambault, S. Diehl, T. Dwyer, K. Klein, H. C. Purchase, and H.-Y. Wu, "Exploring the limits of complexity: A survey of empirical studies on graph visualisation," *Visual Informatics*, vol. 2, no. 4, pp. 264–282, 2018.

[30] K.-T. Chen, Q. Q. Ngo, K. Kurzhals, K. Marriott, T. Dwyer, M. Sedlmair, and D. Weiskopf, "Reading strategies for graph visualizations that wrap around in torus topology," in *the 2023 Symposium on Eye Tracking Research and Applications*, 2023, pp. 1–7.

[31] T. Kosch, J. Karolus, J. Zagermann, H. Reiterer, A. Schmidt, and P. W. Woźniak, "A survey on measuring cognitive workload in human-computer interaction," *ACM Computing Surveys*, vol. 55, no. 13s, pp. 1–39, 2023.

[32] B. Steichen, G. Carenini, and C. Conati, "User-adaptive information visualization: using eye gaze data to infer visualization tasks and user cognitive abilities," in *the 2013 International Conference on Intelligent User Interfaces*, 2013, pp. 317–328.

[33] Y. Wang, Y. Jiang, Z. Hu, C. Ruhdorfer, M. Bâce, and A. Bulling, "VisRecall++: Analysing and predicting visualisation recallability from gaze behaviour," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, no. ETRA, pp. 1–18, 2024.

[34] X. Jin, Y. Bai, L. Xu, S. Ma, D. Shi, L. Yu, and M. Fan, "Decoding cognitive load: Eye-tracking insights into working memory and visual attention," in *Proceedings of the 2025 Symposium on Eye Tracking Research and Applications*, ser. ETRA '25. New York, NY, USA: Association for Computing Machinery, 2025.

[35] M. Okoe, S. S. Alam, and R. Jianu, "A gaze-enabled graph visualization to improve graph reading tasks," *Computer Graphics Forum*, vol. 33, no. 3, pp. 251–260, 2014.

[36] M. Whitlock, S. Smart, and D. A. Szafir, "Graphical perception for immersive analytics," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2020, pp. 616–625.

[37] Q. Sun, B. Duinkharjav, and A. Patney, "Modeling and optimizing human-in-the-loop visual perception using immersive displays: A review," in *SID Symposium Digest of Technical Papers*, vol. 53. Wiley Online Library, 2022, pp. 190–193.

[38] M. Tory, M. Atkins, A. Kirkpatrick, M. Nicolaou, and G.-Z. Yang, "Eyegaze analysis of displays with combined 2D and 3D views," in *VIS 05. IEEE Visualization, 2005.*, 2005, pp. 519–526.

[39] A. Drogemuller, A. Cunningham, J. Walsh, B. H. Thomas, M. Cordeil, and W. Ross, "Examining virtual reality navigation techniques for 3D network visualisations," *Journal of Computer Languages*, vol. 56, p. 100937, 2020.

[40] L. Joos, M. T. Fischer, D. A. Keim, and J. Fuchs, "Aesthetic-driven navigation for node-link diagrams in VR," in *the 2023 ACM Symposium on Spatial User Interaction*, ser. SUI '23, 2023.

[41] L. Joos, G. J. Mooney, M. T. Fischer, D. A. Keim, F. Schreiber, H. C. Purchase, and K. Klein, "Show me your best side: Characteristics of user-preferred perspectives for 3d graph drawings," in *33rd International*

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted,
but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2026.3746476

Symposium on Graph Drawing and Network Visualization (GD 2025). Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025, pp. 37–1.

[42] W. Büschel, S. Vogt, and R. Dachselt, "Augmented reality graph visualizations," *IEEE Computer Graphics and Applications*, vol. 39, no. 3, pp. 29–40, 2019.

[43] M. Krüger, Q. Li, T. W. Kuhlen, and T. Gerrits, "A case study on providing immersive visualization for neuronal network data using cots soft- and hardware," in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, 2023, pp. 201–205.

[44] G. Conte, A. Q. Ye, A. G. Forbes, O. Ajilore, and A. Leow, "Braintrinsic: A virtual reality-compatible tool for exploring intrinsic topologies of the human brain connectome," in *Brain Informatics and Health*. Cham: Springer International Publishing, 2015, pp. 67–76.

[45] S. Jaeger, K. Klein, L. Joos, J. Zagermann, M. De Ridder, J. Kim, J. Yang, U. Pfeil, H. Reiterer, and F. Schreiber, "Challenges for brain data analysis in VR environments," in *2019 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 2019, pp. 42–46.

[46] A. B. Marques, V. Branco, R. Costa, and N. Costa, "Immersive unit visualization with augmented reality," *Multimodal Technologies and Interaction*, vol. 7, no. 10, p. 98, 2023.

[47] W. Büschel, J. Chen, R. Dachselt, S. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, and W. Stuerzlinger, "Interaction for immersive analytics," *Immersive Analytics*, pp. 95–138, 2018.

[48] C. Ware and P. Mitchell, "Reevaluating stereo and motion cues for visualizing graphs in three dimensions," in *the 2nd Symposium on Applied Perception in Graphics and Visualization*, ser. APGV '05, Aug. 2005, pp. 51–58.

[49] ——, "Visualizing graphs in three dimensions," *ACM Transactions on Applied Perception*, vol. 5, no. 1, pp. 1–15, Jan. 2008.

[50] V. Yoghourdjian, Y. Yang, T. Dwyer, L. Lawrence, M. Wybrow, and K. Marriott, "Scalability of network visualisation from a cognitive load perspective," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1677–1687, 2020.

[51] M. Chimani, C. Gutwenger, M. Jünger, G. W. Klau, K. Klein, and P. Mutzel, "The open graph drawing framework (ogdf)." *Handbook of Graph Drawing and Visualisation*, vol. 2011, pp. 543–569, 2013.

[52] F. McGee, B. Renoust, D. Archambault, M. Ghoniem, A. Kerren, B. Pinaud, M. Pohl, B. Otjacques, G. Melançon, and T. Von Landesberger, *Visual analysis of multilayer networks*. Springer, 2021, vol. 8.

[53] W. Kerle-Malcharek, S. P. Feyer, F. Schreiber, and K. Klein, "GAV-VR: An Extensible Framework for Graph Analysis and Visualisation in Virtual Reality," in *ICAT-EGVE 2023 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, 2023.

[54] C. Baumann and K. Dierkes, "Neon accuracy test report," *Pupil Labs*, vol. 10, 2023.

[55] M. Drews and K. Dierkes, "Strategies for enhancing automatic fixation detection in head-mounted eye tracking," *Behavior Research Methods*, vol. 56, no. 6, pp. 6276–6298, 2024.

[56] R. Liu, S. Jin, and X. Wu, "Single camera remote eye gaze tracking under natural head movements," in *Advances in Artificial Reality and Tele-Existence*. Berlin, Heidelberg: Springer, 2006, pp. 614–623.

[57] F. B. Narcizo, F. E. D. dos Santos, and D. W. Hansen, "High-accuracy gaze estimation for interpolation-based eye-tracking methods," *Vision*, vol. 5, no. 3, 2021.

[58] A. Kar and P. Corcoran, "Performance evaluation strategies for eye gaze estimation systems with quantitative metrics and visualizations," *Sensors*, vol. 18, no. 9, 2018.

[59] N. Pathmanathan, S. Öney, M. Koch, D. Weiskopf, and K. Kurzhals, "Uncertainty-aware scarf plots," in *the 2025 Symposium on Eye Tracking Research and Applications*, 2025, pp. 1–8.

[60] A. Schlieder, J. Rummel, P. Albers, and F. Sadlo, "Sequential visual cues from gaze patterns: Reasoning assistance for bar charts," in *the 2025 CHI Conference on Human Factors in Computing Systems*, ser. CHI '25, 2025.

[61] A. Martinez-Cedillo, N. Gavrila, A. Mishra, E. Geangu, and T. Foulsham, "Cognitive load affects gaze dynamics during real-world tasks," *Experimental Brain Research*, vol. 243, no. 4, p. 82, 2025.

[62] M. Koch, N. Pathmanathan, D. Weiskopf, and K. Kurzhals, "How deep is your gaze? Leveraging distance in image-based gaze analysis," in *the 2024 Symposium on Eye Tracking Research and Applications*, 2024, pp. 1–7.

[63] Y. Wang, M. Koch, M. Bâce, D. Weiskopf, and A. Bulling, "Impact of gaze uncertainty on aois in information visualisations," in *2022 Symposium on Eye Tracking Research and Applications*, 2022, pp. 1–6.

IX. Biography Section

Yao Wang is a postdoctoral researcher at the University of Stuttgart, Germany. He received the B.Sc. and the M.Sc. degrees from Peking University, China. He received the PhD degree in Information Technology and Electrical Engineering from the University of Stuttgart in 2024. His research interests include human-AI interaction, data visualisation, computer vision, and extended reality.

Lin Zhang is a master's student at the University of Stuttgart, Germany. He received the B.Eng. degree in Software Engineering from Beijing University of Posts and Telecommunications, China, in 2024. His research interests include human-computer interaction and virtual reality.

Timo Hüttner is a master's student at the University of Stuttgart, Germany. He received his B.Sc. degree in Software Engineering from the University of Stuttgart in 2021. His research interests include human-computer interaction, virtual reality and generative AI.

Ying Zhang is a PhD student at the University of Konstanz, Germany. She received her M.Sc. in Computer and Information Science from the University of Konstanz in 2020. Her research explores the utilization of design space in immersive analytics environments, including human perception, multimodal data representation, and interaction techniques for movement and network data.

Wilhelm Kerle-Malcharek is a PhD student at the University of Konstanz, Germany. He received his M.Sc. in Computer and Information Science from the same university in 2022. His primary research focus is the use of immersive spaces in Archaeology. His research includes digitizing monuments and creating virtual spaces where these digitizations can be used for further analysis. Further research includes network analysis in immersive environments and the analysis of animal behaviour data.

Karsten Klein is a postdoctoral researcher in the Life Science Informatics group at the University of Konstanz, Germany, where he manages the Immersive Analytics Lab and works on the design and implementation of visual and immersive analytics approaches for complex data from application areas. A main research focus of his is on network analysis and visualisation and the underlying graph theoretic principles. Karsten Klein received his Ph.D. degree in computer science from TU Dortmund, Germany, and pursued postdoctoral studies at the University of Sydney and Monash University, Australia, before joining the University of Konstanz.

Falk Schreiber is full professor and head of the Life Science Informatics group at the University of Konstanz. He worked at universities and research institutes in Germany and Australia and is adjunct professor at Monash University Melbourne. His main interests are network analysis and visualisation, immersive analytics as well as multi-scale modelling of processes, all with a focus on life science applications from metabolic processes to collective behaviour.

Andreas Bulling is full professor of Computer Science at the University of Stuttgart, Germany, where he directs the research group "Collaborative Artificial Intelligence". He received the MSc degree in Computer Science from the Karlsruhe Institute of Technology, Germany, in 2006 and the PhD degree in Information Technology and Electrical Engineering from ETH Zurich, Switzerland, in 2010. Andreas Bulling was a Feodor Lynen and Marie Curie Research Fellow at the University of Cambridge, UK, and a Senior Researcher at the Max Planck Institute for Informatics, Germany. His research interests include computer vision, machine learning, and human-AI interaction.

Authorized licensed use limited to: UNIVERSITAETSBIBL STUTTGART. Downloaded on October 24,2026 at 04:09:36 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted,
but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.