

Visual attention-guided augmented representation of geographic scenes: a case of bridge stress visualization

Weilian Li, Jun Zhu, Qing Zhu, Jinbin Zhang, Xiao Han & Youness Dehbi

To cite this article: Weilian Li, Jun Zhu, Qing Zhu, Jinbin Zhang, Xiao Han & Youness Dehbi (17 Jan 2024): Visual attention-guided augmented representation of geographic scenes: a case of bridge stress visualization, International Journal of Geographical Information Science, DOI: [10.1080/13658816.2023.2301313](https://doi.org/10.1080/13658816.2023.2301313)

To link to this article: <https://doi.org/10.1080/13658816.2023.2301313>



Published online: 17 Jan 2024.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



RESEARCH ARTICLE



Visual attention-guided augmented representation of geographic scenes: a case of bridge stress visualization

Weilian Li^{a,b,c,d}, Jun Zhu^a, Qing Zhu^a, Jinbin Zhang^a, Xiao Han^a and Youness Dehbi^d

^aFaculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China; ^bInstitute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany; ^cGuangdong–Hong Kong–Macau Joint Laboratory for Smart Cities, Shenzhen, China; ^dComputational Methods Lab, HafenCity University Hamburg, Hamburg, Germany

ABSTRACT

Efficient geovisualization is beneficial for understanding geospatial phenomena, an important research direction for GISers and Cartographers. However, the current research on geovisualization overemphasizes the visual effects while neglecting the prominent representation of crucial information and failing to consider the user's cognitive workload of information processing. Following the laws of visual perception of the human eyes, this article proposes a visual attention-guided augmented representation approach of geographic scenes that involves area of interest computation, background simplification, and compound graphic variables. Finally, we select bridge stress visualization as a case study for experimental analysis. The experimental results of eye-tracking show that augmented representation could draw the participants' attention to areas of interest in a short time, increasing their duration of fixations and the accuracy of completing given tasks. These findings suggest that our approach can enhance geographic scenes' cognitive efficiency, offers a new idea for the theoretical studies of geovisualization, and holds promising potential for broader application in various geographical phenomena visualization.

ARTICLE HISTORY

Received 20 September 2023
Accepted 28 December 2023

KEYWORDS

Visual attention; augmented representation; geovisualization; 3D GIS; bridge stress

1. Introduction

In the context of the rapid development of digital twins, there is an increasing emphasis on cross-departmental collaboration (Broo et al. 2022; Ma et al. 2022; Wu et al. 2023), in which advanced visualization of geographic scenes can facilitate communication among stakeholders (Kim et al. 2013; Boton 2018; Ma et al. 2021). Taking digital construction as an example, bridge stress simulation plays an essential role in structure safety assessment and risk prevention (Li et al. 2015; Zhou and Zhang 2019). However, the existing representation of stress in finite element analysis (e.g., ANSYS, Midas Civil) leads to tenuous communication among stakeholders (e.g., decision-

makers, bridge engineers, GISers) involved in construction due to excess engineering details and a lack of integrated geovisualization.

Geovisualization generally refers to the visual representation of geoinformation to facilitate thinking, understanding, and knowledge discovery about the geographic environment (MacEachren 2004; Laurini 2017; Li et al. 2020; Degbelo 2022; Zuo et al. 2022). In an era dominated by digital earth, it is still a delicate task to design appropriate visualization that empowers users to perform efficient visual processing of geographic information (Reichenbacher and Swienty 2007; Goodchild et al. 2012; Çöltekin et al. 2018; Li et al. 2019, 2023; Zhu et al. 2023).

Scientific visualization plays a pivotal role in enhancing the effectiveness and readability of geoinformation (Bodum 2005; Li et al. 2019, 2022; Zhu et al. 2024). A critical objective of geovisualization is much more to allow users to quickly locate and easily decode relevant information in geographic scenes, which is not just a simple rendering task, usability and cognition should also be focused (Swienty et al. 2006; Li et al. 2021). However, whether the current research on map-based geovisualization, 3D geovisualization, or immersive geovisualization, they all overemphasize the visual effects while neglecting the prominent representation of crucial information and fail to consider the user's cognitive workload of information processing (MacEachren et al. 2004; Dransch et al. 2010; Dong et al. 2020; Zhang et al. 2020; Badwi et al. 2022). Çöltekin et al. (2019) stated that geovisualization should focus on user-centric thinking both from a theoretical and a practical perspective, and that a critical challenge needs to be addressed is how to improve the understanding of human perceptual and cognitive processes and developing effective solutions (Andrienko et al. 2010). For the task at hand, we expect to propose a novel representation method of geographic scenes that takes into account the user's cognition. In our case, we also want to fill the missing link between the bridge stress simulation and geovisualization to promote the stakeholders' communication.

Human visual scanning of a specific scene is distinguished by two fundamental actions (Reicher et al. 1976; Treisman 1982; Torralba et al. 2006): attention shifting (i.e. saccade) and visual information processing (i.e. fixation). In this process, the eyes will prioritize searching for the area of interest (AOI) and ignore other areas that do not attract visual attention (Jahnke et al. 2008; Fu et al. 2021), which provides a new perspective for studying visual attention-guided geovisualization. From the standpoint of people's cognition, it is essential that the background should not dominate the geovisualization, as it could distract the users' fixations from relevant information. To guide the user's attention to the focal information, the virtual scene's background should be represented as non-salient. In contrast, the relevant information should be processed in a local mode and augmented by a salient visual stimulus.

Following this idea, this article presents a novel approach that utilizes visual attention guidance to enhance the representation of geographic scenes, explicitly focusing on bridge stress visualization as a case study for experimental analysis. The primary objective is to gradually guide the user's attention through background simplification, stress focusing, and visual augmentation, thereby improving their comprehension and perception of the bridge stress scene. Finally, a cognitive evaluation program for

augmented representation using eye tracking is designed to verify the effectiveness of the proposed approach.

The remainder of this article is structured as follows: [Section 2](#) briefly describes the related terms in this article, whereas [Section 3](#) gives insights into the introduced approach. Subsequently, [Section 4](#) introduces the implementation of the prototype system and experiment results. [Section 5](#) presents the discussion. [Section 6](#) finally summarizes the article and gives an outlook for future research.

2. Definition of relevant terms

To improve the reader's comprehension of this article, we have defined and explained the following relevant terms.

1. **Visual attention.** Visual attention is a cognitive mechanism in the human visual system for selecting and focusing on a specific target in the observed scene. In geovisualization, users are required to locate promptly and decode the geoinformation easily to gain visual awareness, which is closely related to visual attention and controlled by the amount of information and stimulus drive (Wolfe and Horowitz 2004; Reichenbacher and Swienty 2007; Robinson 2011).
2. **Area of interest.** AOI refers to a selected sub-region that might interest the user in the displayed stimuli, which provide valuable visual information for completing cognitive tasks (Holland and Lee 2019). In this article, the area where the overlapping area of both eyes is considered the potential AOI.
3. **Augmented representation.** Augmented representation differs from augmented reality (Moore et al. 2020; Zhang et al. 2020). The concept of augmented representation is expanded in this article upon Li et al. (2021), which define it as the utilization of a virtual geographical scene as the primary carrier, incorporating visual attention to achieve the quick localization of the AOI and reduce the interference of irrelevant background information, and additionally using compound graphic variables to enhance visual information saliency, facilitating participants to decode and master critical details quickly.
4. **Bridge stress.** Bridge stress refers to the distribution of internal forces, which describes the magnitude and direction of internal force at a specific location and along a certain cross-section within the bridge.

3. Methodology

[Section 3.1](#) describes the algorithm corresponding to the AOI computation and background simplification. [Section 3.2](#) presents multi-level spatial modeling of bridge stress. [Section 3.3](#) introduces compound graphic variables for the augmentation of visual attention. [Section 3.4](#) designs cognitive evaluation for augmented representation using eye tracking.

3.1. Overlapping vision area computation and background simplification

3.1.1. Overlapping vision area computation considering eye field of view

The observation area of the human eye has a specific range called the field of view (FOV), and the ability to discriminate information is most robust in the overlapping area of both eyes within the FOV, which is considered as AOI in this article. In this context, we present a computational method for determining the AOI considering the user's eye FOV. The primary objective is precisely drawing the user's attention within the AOI. Figure 1 illustrates the fundamental idea behind our method.

We assume that the center point of the eyes is in a horizontal plane with the center point of the monitor, and the distance is denoted as D_s , the pupil distance between the left and right eye as PD , and the screen width and height as S_w and S_h , respectively. The width and height of the sight view are represented as w and h . The width and height of the overlapping vision area are denoted as O_w and O_h . The horizontal and vertical angles of the FOV are represented as α and β , respectively.

Consequently, the sight view width w can be calculated using Equation 1:

$$w = 2D_s \times \tan \frac{\alpha}{2} \quad (1)$$

Similarly, the sight view height h can be calculated using Equation 2:

$$h = 2D_s \times \tan \frac{\beta}{2} \quad (2)$$

According to the size of the sight view and the pupil distance PD , the overlapping vision area can be calculated using Equation 3:

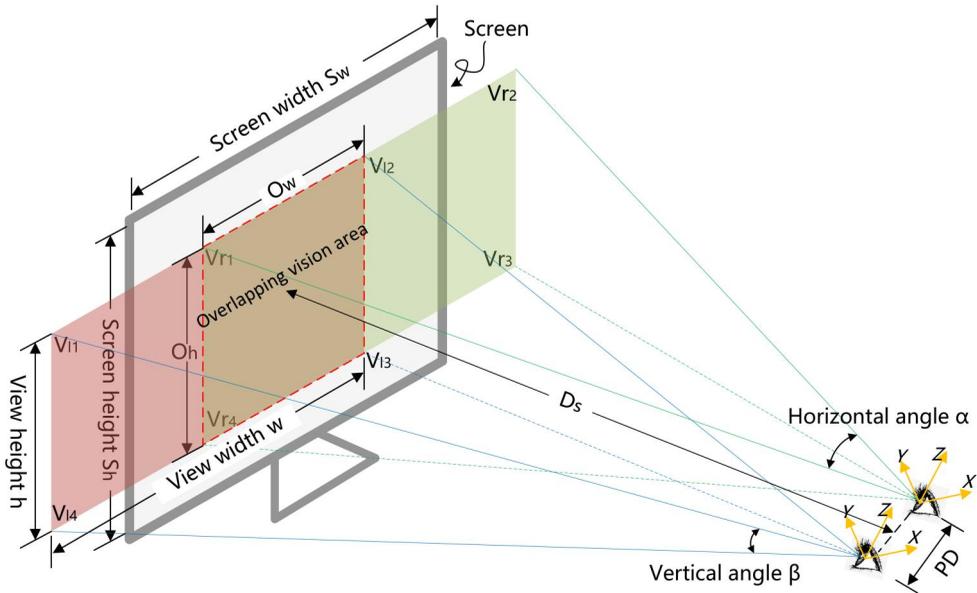


Figure 1. Schematic diagram of the AOI based on the human eye FOV.

$$\begin{cases} O_w = w - PD \\ O_h = h \end{cases} \quad (3)$$

Thus, the AOI is the rectangular area with O_w as the width and O_h as the height. The proportion of the AOI to the screen can be calculated based on the screen size S_w and S_h .

3.1.2. Background simplification based on Gaussian blur algorithm

There are many image blur algorithms, such as Gaussian, Kawase, Dual, etc. However, Gaussian blur simulates the optical effect of defocusing and has high computational efficiency, which enables the creation of a smooth and natural appearance of blurred areas similar to that observed by human eyes in a short time (Zhang and Ma 2019). In this context, we employ the Gaussian blur algorithm to minimize the influence of the background information by simplifying the details outside the AOI and keeping relevant information as much as possible in the visual foreground. The algorithm is described as follows:

$$G(i, j) = \frac{1}{2\pi\sigma^2} e^{-(i^2+j^2)/2\sigma^2} \quad (4)$$

$G(i, j)$ denotes the weight matrix of the Gaussian convolution kernel, (i, j) indicates the kernel position, and σ is the standard deviation of the Gaussian distribution.

$$Z(i, j) = \frac{G(i, j)}{\sum_{u=-n}^{u=n} \sum_{v=-n}^{v=n} G(i + u, j + n)} \quad (5)$$

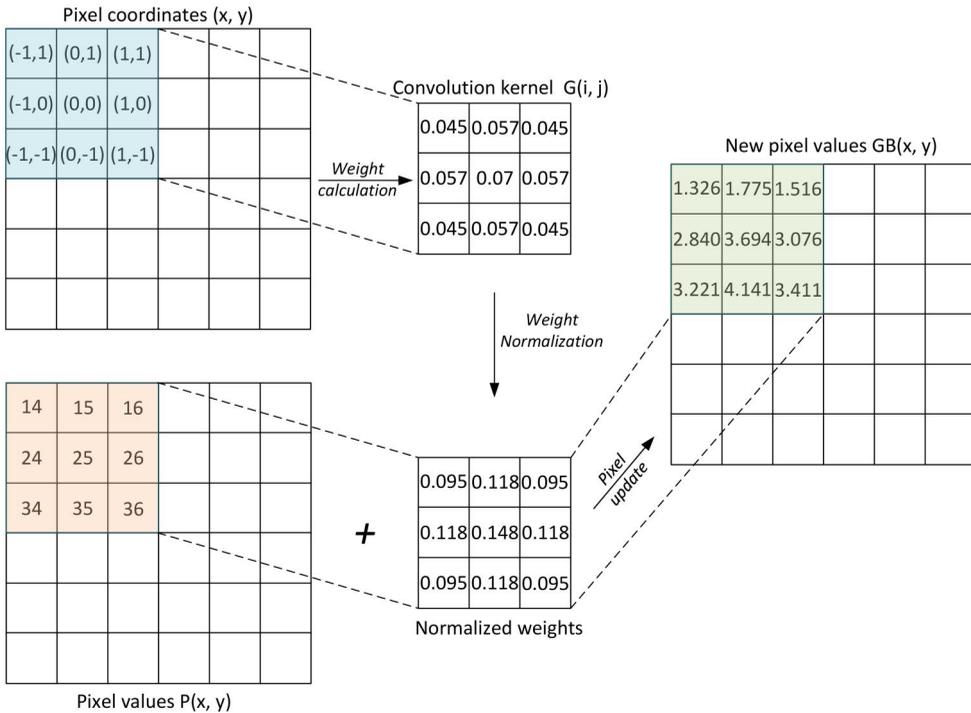


Figure 2. An example of Gaussian blur calculation.

$Z(i,j)$ represents the normalized weight matrix, and n denotes the convolution kernel radius. If $n = 1$, the convolution kernel is of size $3 * 3$.

$$GB(x,y) = P(x,y) * Z(i,j) \quad (6)$$

$GB(x,y)$ denotes the updated pixel values, and the $P(x,y)$ indicates the original pixel values. Figure 2 shows a Gaussian blur calculation with $\sigma = 1.5$ and $n = 1$.

3.2. 3D Representation of bridge stress in the AOI

The bridge stress results are spatially modeled from the whole to a local level to help stakeholders' step-by-step understanding of the stress distribution while also improving the bridge scene's rendering efficiency. The main idea is shown in Figure 3.

3.2.1. Line stress representation

In line stress representation, we simplify the bridge to a rod-like structure comprising multiple units that store essential data, including node positions and stress values. The node ID serves as a crucial identifier for retrieving information regarding the node's position, stress values, and topologic relationships, and they are connected to represent the stress distribution of the whole bridge in 3D space, as shown in the upper right of Figure 3.

3.2.2. Surface stress representation

To accurately reflect the stress conditions of various bridge components, we propose an automatic modeling method of surface stress based on cross-section lofting, drawing inspiration from parametric modeling (Zhu et al. 2015). Figure 4 shows the automatic modeling process of surface stress, and we aim to generate surface stress based on line stress automatically.

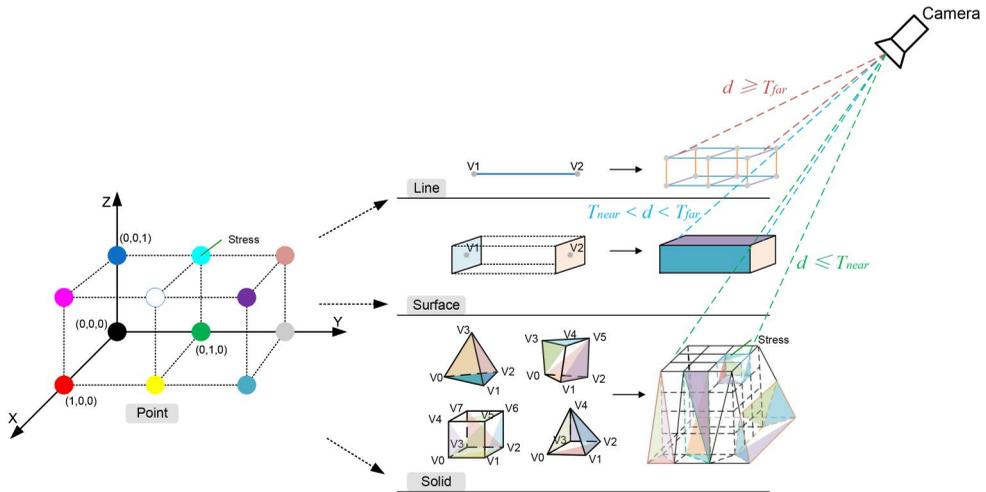


Figure 3. The spatial modeling and representation of stress simulation results. The left shows the position and stress of the sampling points, and the right shows three different stress modeling approaches from the line, the surface, to the solid. In addition, the different levels of the stress model will be dynamically scheduled according to the viewpoint.

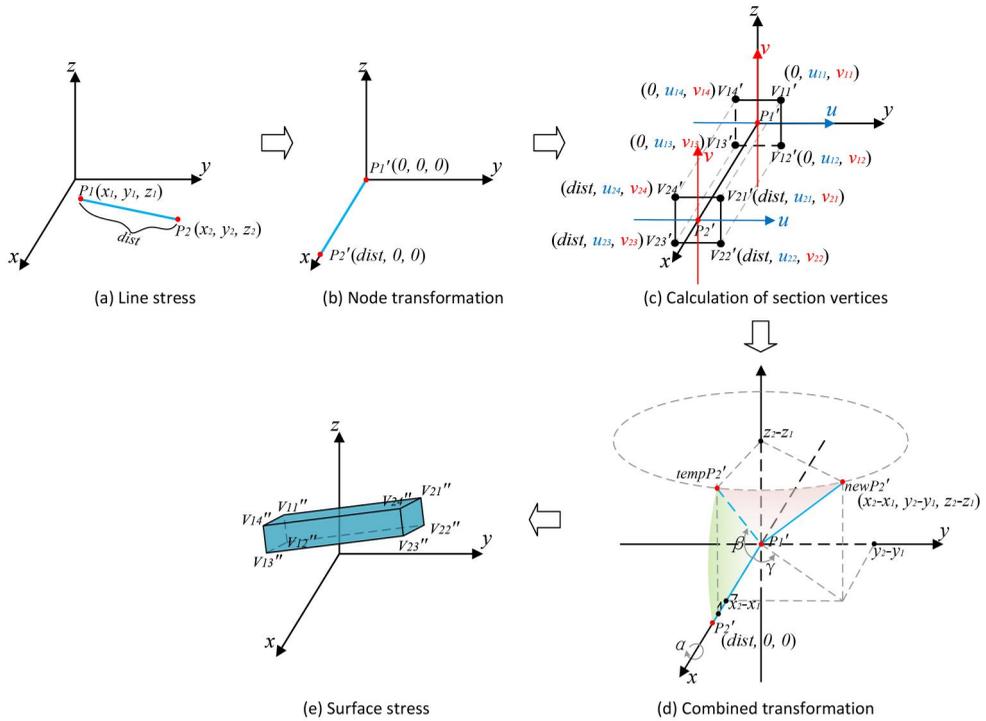


Figure 4. The automatic modeling process of surface stress based on cross-section lofting.

Assuming there is line stress segment P_1P_2 (Figure 4(a)), which can be transformed to $P_1'P_2'$ (Figure 4(b)) in a 3D space. Then, we can calculate the vertices of each node based on the cross-section information, and the example of section type in Figure 4(c) is a quadrilateral.

Subsequently, the combined transformation shown in Figure 4(d) is used to restore the surface stress from $P_1'P_2'$ to its initial orientation $P_1'newP_2'$, where the transformation process involves $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$, and $\Delta z = z_2 - z_1$. The specific procedure is as follows:

- Rotate $P_1'P_2'$ around the y -axis by β to obtain $P_1'tempP_2'$, i.e. sweep through the green sector in Figure 4(d).

$$\beta = \arcsin \frac{\Delta z}{dist} \quad (7)$$

Rotate $P_1'tempP_2'$ around the z -axis by γ to obtain $P_1'newP_2'$, i.e. sweep through the purple sector in Figure 4(d).

Table 1. The rotation angles for particular cases.

Number	α	β	γ
1	0	0	0
2	0	0	90
3	0	0	$\arctan\left(\frac{\Delta y}{\Delta x}\right)$
4	0	0	0
5	0	90	0
6	0	$\arcsin\left(\frac{\Delta z}{\text{dist}}\right)$	90
7	0	$\arcsin\left(\frac{\Delta z}{\text{dist}}\right)$	0

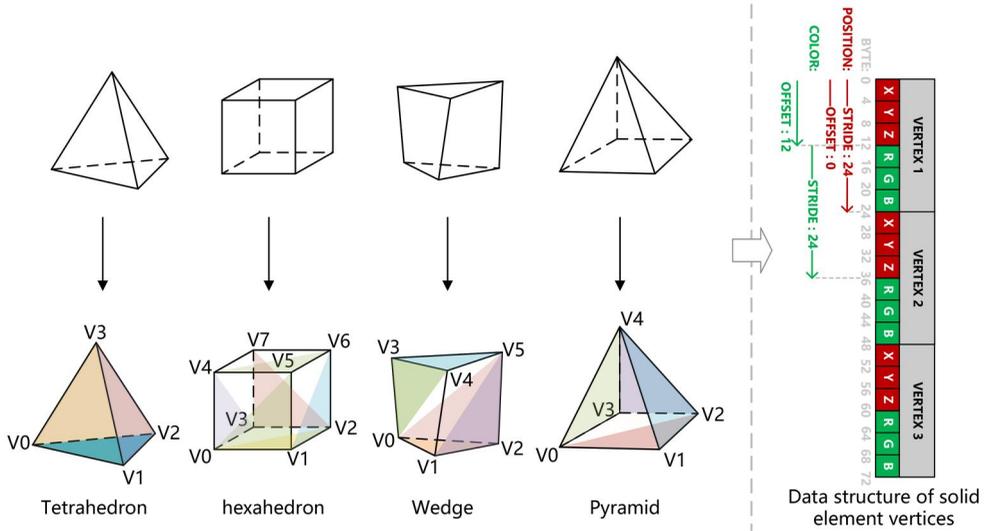


Figure 5. Solid stress modeling of bridges.

$$\gamma = \arctan \frac{\Delta y}{\Delta x} \tag{8}$$

For the particular case where $\Delta x, \Delta y,$ and Δz are 0, the rotation angles are as shown in Table 1.

Translating $P1'newP2'$ back to its initial position (Figure 4(e)). A closed surface is formed based on the topological relationship between vertices, creating a 3D stress model for the bridge surface.

3.2.3. Solid stress representation

Simplifying the bridges into a rod-like structure for stress analysis is insufficient to handle irregular bridge components and accurately analyze the local details of the bridge. This study also considered four types of solid elements: tetrahedron, hexahedron, wedge, and pyramid. The vertices of these solid elements are stored according to the

data structure shown on the right side of Figure 5. Subsequently, triangular facets are generated based on vertex indices and stress values.

3.3. Compound graphic variables for the augmentation of visual attention

In visual brain areas, the posterior parietal cortex (PPC) and inferior parietal cortex (IPC) located on the 'where' and 'what' pathways are responsible for the location and semantic decoding of relevant information for decision-making (Ungerleider 1995; Swienty et al. 2006). Likewise, geovisualization is about bottom-up visual stimulation and sends signals to visual brain areas to tell the user when, where, and what is happening.

Visual variables influence visual attention and cognitive process by controlling changes in stimulus material, which include time, space, and attribute in dimensions. However, basic visual variables have low-level visual characteristics (Li et al. 2020), while multidimensional associations and combinations among these underlying variables enable the creation of more complex compound graphic variables (Chen et al. 2021), enhancing the graphical encoding of geoinformation. The basic

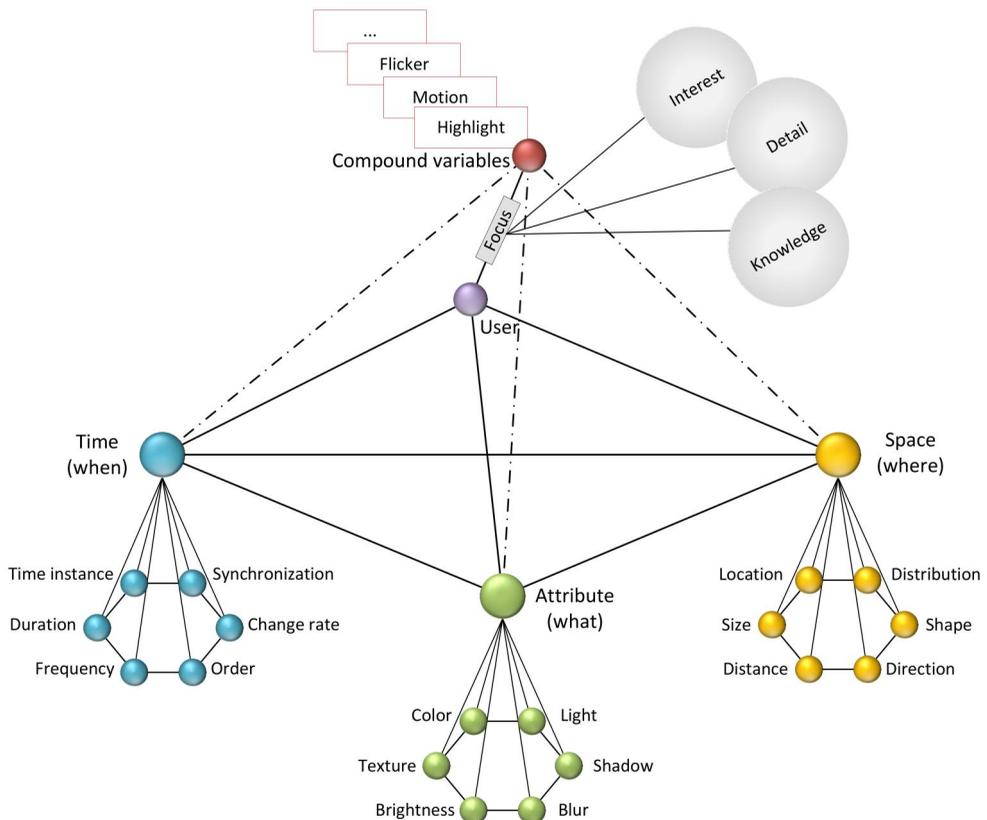


Figure 6. Visual variable dimensions and compound graphic variables for the augmentation of visual attention. The conceptual framework of compound graphic variables was inspired by Li et al. (2020) and Chen et al. (2021).

principle is shown in Equation 9:

$$F(e) = \{T(e, t_i), S(e, s_i), A(e, a_i)\} \bowtie I \quad (9)$$

Where F denotes the compound graphic variables; T refers to visual variables with the ability to describe temporal characteristics; S represents the spatial semantic characteristics of the visual elements; A reflects the change in the appearance of the visual elements; I indicates the interplay and combination of diverse visual variables.

Specifically, the idea for the compound graphic variables to enhance visual attention is shown in Figure 6. The association of basic visual variables in time, space, and attribute dimensions forms compound graphic variables, which stimulate the user's attention, help them focus on more details and ultimately decode the geoinformation into knowledge. For example, duration, shape, and orientation can be combined to form a new compound graphic variable names' motion', which provokes high neural responses and is detected faster than static attributes.

3.4. Cognitive evaluation for augmented representation using eye-tracking

Attention can shift without eye movements, but users cannot move their eyes without a concomitant shift of attention (Swienty 2008). In this context, we opt for eye-tracking to evaluate the cognitive efficiency of the augmented representation.

Figure 7 depicts a workflow comprising three primary components: experiment preparation, procedure design, and result analysis. In the experiment preparation phase, essential tasks involve determining the purpose of the eye tracking study, selecting the appropriate software environment (e.g. Unity3D), and hardware equipment (e.g. Tobii Pro Nano). Additionally, participant selection and preparation of comparative materials for the study are essential aspects. Subsequently, the procedure for the eye tracking experiment is designed, wherein participants are invited to observe the test materials and respond to predetermined questions while data is continuously recorded. Finally, relevant effectiveness and efficiency indices are chosen to qualitatively and quantitatively evaluate the experiment results.

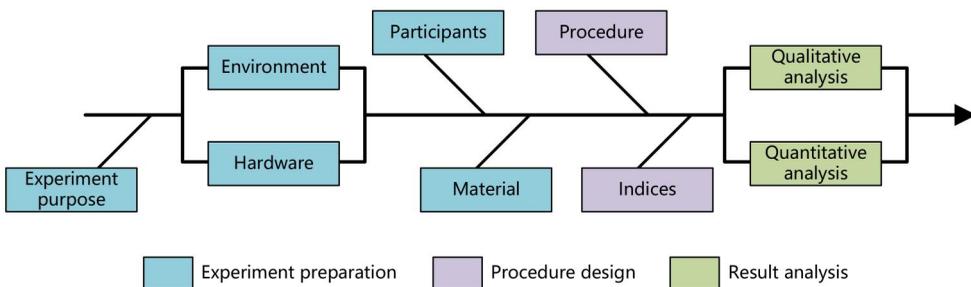


Figure 7. Experiment design of eye tracking for augmented representation.



Figure 8. The interface of the prototype system. The prototype system comprises Function Menu, Stress LOD, Layer management, and 3D representation. The Function Menu expertly governs the system, overseeing tasks like frame rate monitoring, gauss blur, and geo-analysis. The parameters of the vision and the screen size can also be set in the Function Menu. The Stress LOD module is responsible for loading stress data, while the Layer Management controls the visibility and renders the order of layers. 3D representation zone visualizes the bridge stress and the corresponding geoinformation.

4. Experiment analysis

4.1. Study area and data processing

In this article, a mega suspension bridge under construction located in Luding ($29^{\circ}55'32''\text{N} \sim 29^{\circ}56'03''\text{N}$, $102^{\circ}13'18''\text{E} \sim 102^{\circ}14'07''\text{E}$) was selected as the case for experiment analysis. The stress analysis was performed in Midas Civil under the supervision of bridge specialists. The finite element model for the whole bridge contains 12,970 nodes and 12,436 cells, with a more refined tower section featuring 10,834 nodes and 97,331 cells.

We restructured the stress simulation results according to node type, position, topologic relationships, cell type, and stress value. A more lightweight GeoJSON format was used for data storage and exchange. Additionally, the Drone acquired the remote sensing images and digital elevation model (DEM) required for geographic scene construction and were further processed into tiles.

4.2. Prototype system implementation

Using WebGL technique, we successfully developed a prototype system with B/S architecture. This system enabled us to perform visual attention-guided augmented representation of bridge stress in virtual geospatial space, as shown in Figure 8. The server side was built with NodeJS v16.18.0, and the interface was designed using HTML5, CSS, and JavaScript within the Vue framework. An open-source library Cesium.js v1.73 was used to construct the 3D virtual scene and visualize the stress information.

The prototype system was tested on Google Chrome 96.0.4664.110 and ran on Lenovo Legion R9000P2021H. The processor was an AMD Ryzen 7 5800H with Radeon Graphics, 16 GB memory, and NVIDIA GeForce RTX 3060 Laptop GPU 6 GB. In addition, Tobii Pro Nano was employed to carry out the subsequent eye tracking experiment.

4.3. Augmented representation of bridge stress

4.3.1. Background simplification and bridge semantic division

As shown in Figure 9, the previously mentioned AOI computation and Gaussian blur algorithm have been applied to simplify the background information outside the AOI

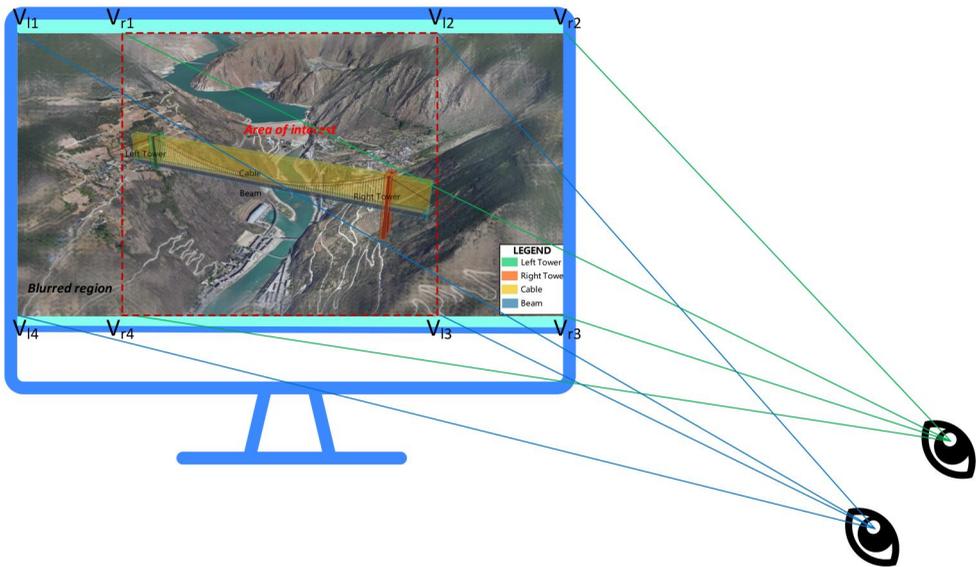


Figure 9. Background simplification and bridge semantic division.

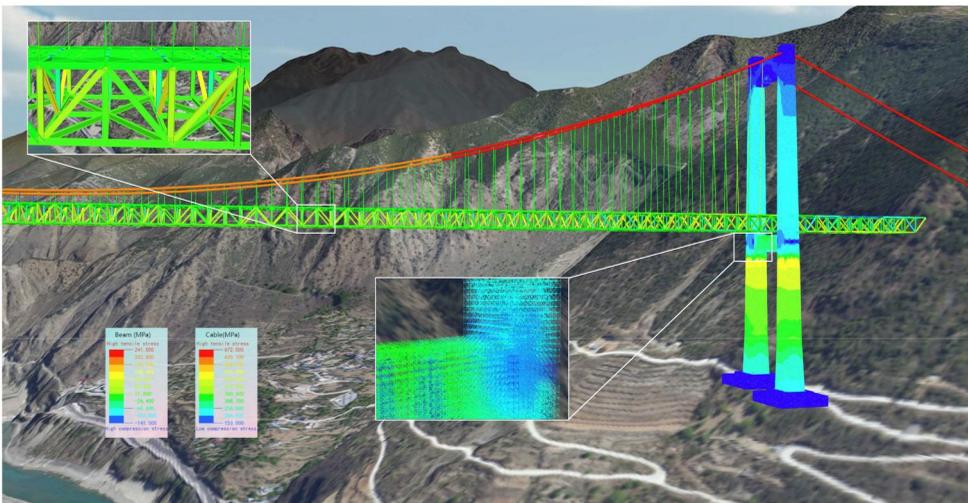


Figure 10. Spatial modeling and representation of bridge stress.

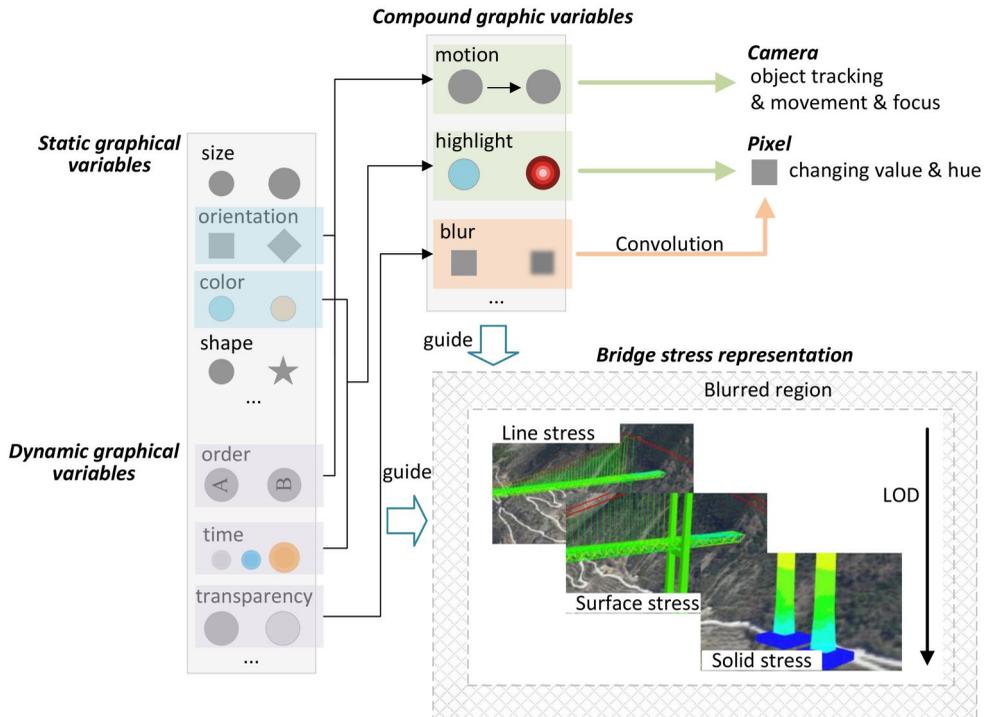


Figure 11. Augmented representation of bridge stress based on compound graphic variables.

to reduce the distraction of non-relevant information. It can be seen that the bridges and the corresponding geoinformation in the AOI were rendered with high fidelity, which allowed users to focus on the AOI and effectively perceive the bridge information. In addition, we divided the bridge structure from a semantic perspective and provided the corresponding annotations, which served to facilitate the users' understanding of the stress distribution of each bridge component subsequently. Bridge semantic division will also be used as part of the material for subsequent cognitive experiments on augmented representation.

4.3.2. Augmented representation of bridge stress in the 3D scene

Figure 10 presents the outcomes of spatial modeling and representation of bridge stress, effectively visualizing stress at all levels and seamlessly coupling it to the 3D geographic scene. The LOD modeling of stress can meet the requirements of bridge construction at different stages but also the rendering efficiency of the virtual scene.

People tend to rely on perceptual salience to extract information, and vivid representation may attract their attention, thus guiding them to focus on the AOI and improving memorability. To visualize bridge stress, we performed multiple convolutions on the pixels outside the AOI using the Gaussian blurring algorithm, thus simplifying the background details.

Regarding the augmented representation of bridge stress, we divided the stress intervals and determined the mapping between stress and color values according to the gradient variation. In addition, compound graphic variables are used to guide

visual attention, as shown in [Figure 11](#). For example, the compound variable ‘motion’ holds significant potential in directing attention since it can trigger the high responsiveness of visual brain areas. Consequently, by combining the variables of orientation and order, motion can be generated to effectively guide individuals’ attention toward the center of the virtual scene. Similarly, combining color and time variables can form a flicker, emphasizing stress areas that require focused attention, thus enhancing the participant’s perception of the area and swiftly drawing their gaze into sharp focus.

4.4. Cognitive analysis of augmented representation

In this section, we analyzed the cognitive efficiency of the augmented representation of bridge stress based on eye-tracking experiments following the process shown in [Figure 7](#).

4.4.1. Experiment preparation

(1) **Experiment purpose.** In the context of bridge stress visualization, this study aims to compare the difference in cognition between traditional geovisualization and visual attention-guided augmented representation.

(2) **Test environment and hardware.** The test environment employed for the study was the bridge stress scene in 3D, running on the Lenovo Legion R9000P2021H. The eye-tracking device was a Tobii Pro Nano with a sampling frequency of 60 Hz. The Tobii Pro Eye Tracker Manager and Tobii Pro Lab were used to record and analyze the observation data. In our experiment setup, we took care that all users had the same environment, e.g. the same screen, table, Ds, etc.

(3) **Participants.** We recruited 30 participants (aged between 21 and 28 years old, 18 males and 12 females) to participate in this study who were GIS or bridge engineering students. They were randomly assigned to groups A (14 participants) and B (16 participants). The participants have a naked or corrected vision of 1.0 or higher, with no color vision deficiency. All participants had 2D map-reading experience, and they could use computers proficiently, but none had prior knowledge of this experiment’s content. For the sake of comparability and simplicity, we assumed a default value of PD (60 mm) for the participants.

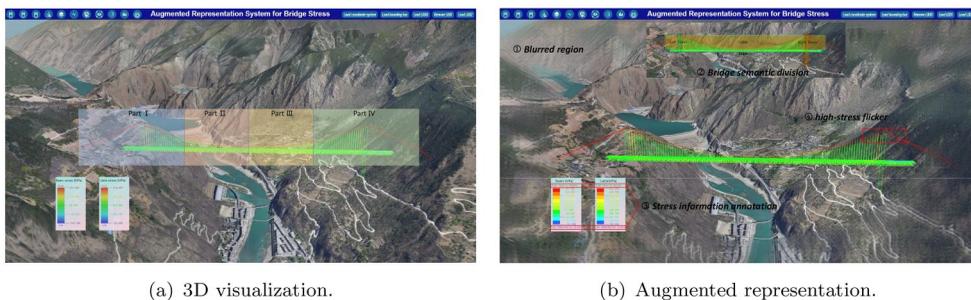


Figure 12. The comparative materials of cognition experiment.

(4) Material. The test material for the eye-tracking experiment is shown in [Figure 12](#). In [Figure 12\(a\)](#), the 3D representation of bridge stress is depicted as scene A, which does not consider the human eye's visual characteristics (Dransch et al. 2010; Zhang et al. 2020; Li et al. 2021). In [Figure 12\(b\)](#), scene B incorporates background blurring, high-stress flicker, and semantic annotation of stress information on the basis of scene A.

4.4.2. Procedure design

(1) Test procedure. The participants in groups A and B were asked to observe scenes A and B, respectively, follow the on-screen prompts, and complete the given questions.

- Among the above four parts, which part has the highest tensile stress?
- Among the above four parts, which part has the highest compressive stress?
- Among the above four parts, which part has the lowest stress?
- In Part IV, which bridge component is subjected to the highest tensile stress?
- In Part IV, which bridge component is subjected to the highest compressive stress?

The above process automatically recorded the eye movement data and response for further analysis.

(2) Analysis indices. The analysis of the eye-tracking data generally includes effectiveness and efficiency. According to Dong et al. (2018), the effectiveness indices include the percentage of fixations in AOIs and the accuracy for the given tasks, while the efficiency indices include the time to first fixation in AOI, the finish time for the given tasks, and the pupil diameter. [Table 2](#) lists the indices and their descriptions.

4.4.3. Result analysis

(1) Qualitative analysis. [Figures 13\(a\)](#) and [13\(b\)](#) show the eye-tracking heat map of scene A and scene B, respectively. The darker red areas in the heat map indicate the parts that receive more attention. Compared with scene A, it can be seen that scene B's heat map has darker red in the high-stress flicker and the semantic annotation part of the stress information, indicating that these two areas can effectively attract users' visual attention and enhance concentration.

Table 2. Analysis indices.

Category	Indices	Descriptions
Effectiveness	Percentage of fixations in AOIs(PAOIs)	Duration of fixations located within AOIs divided by all reading time
	Accuracy	Number of given questions answered correctly
Efficiency	Time to the first fixation in AOI(TAOIs)	Duration from the beginning of the task to the first fixation located within the AOIs
	Finish time	The average time used to complete given questions
	Pupil diameter	The average pupil size of participants during the experiment

In addition, we depicted the gaze point sequences of all participants, as shown in Figure 14(a) and 14(b). The circles in the figures indicate the participants' gaze points, the numbers indicate the gaze order, and the larger size represents the longer the



Figure 13. The heat map of eye-tracking in the bridge stress scene.

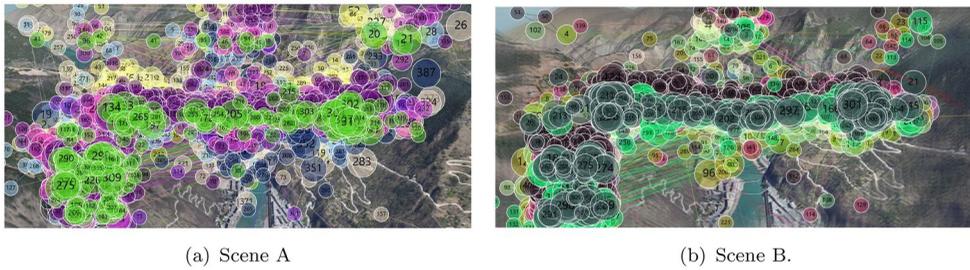


Figure 14. The gaze point sequences of all participants in the bridge stress scene.

Table 3. Shapiro-Wilk for normal distribution test^a.

Indices	Groups	Shapiro-Wilk test		
		sta	df	<i>p</i>
PAOIs	Group A	0.912	14	0.169
	Group B	0.946	16	0.434
TAOIs	Group A	0.814	14	0.008**
	Group B	0.776	16	0.001**
Finish time	Group A	0.912	14	0.167
	Group B	0.912	16	0.125
Accuracy	Group A	0.849	14	0.021*
	Group B	0.787	16	0.002**
Pupil diameter	Group A	0.662	14	0.000**
	Group B	0.872	16	0.029*

^asta = statistic, df = degree of freedom. P stands for the statistical p-value. **p* < 0.05, ***p* < 0.01.

Table 4. Statistical test of the experiment results^b.

Evaluation indices	Descriptive		Inferential					
	Group A M ± SD	Group B M ± SD	T-test			U-test		
			<i>t</i>	df	<i>p</i>	<i>u</i>	<i>z</i>	<i>p</i>
PAOIs	0.41 ± 0.12	0.55 ± 0.12	-3.298	28	0.003**	n/a	n/a	n/a
TAOIs	1.34 ± 1.23	0.65 ± 0.81	n/a	n/a	n/a	59	-2.203	0.028*
Finish time	21.13 ± 5.99	14.42 ± 5.72	3.136	28	0.004**	n/a	n/a	n/a
Accuracy	0.63 ± 0.21	0.84 ± 0.17	n/a	n/a	n/a	49	-2.619	0.010**
Pupil diameter	3.13 ± 0.57	2.82 ± 0.34	n/a	n/a	n/a	61	-2.210	0.034*

^bM = mean, SD = standard deviation, and **p* < 0.05, ***p* < 0.01.

gaze duration. The gaze points are scattered in scene A, and many are gathered in the background area, e.g. the river. In contrast, the gaze points in scene B are mainly distributed in the area surrounding the bridge stress. This suggests that the blurring of background information helps to reduce the interference of non-relevant information, allowing the participants to focus more on the bridge stress information in AOI.

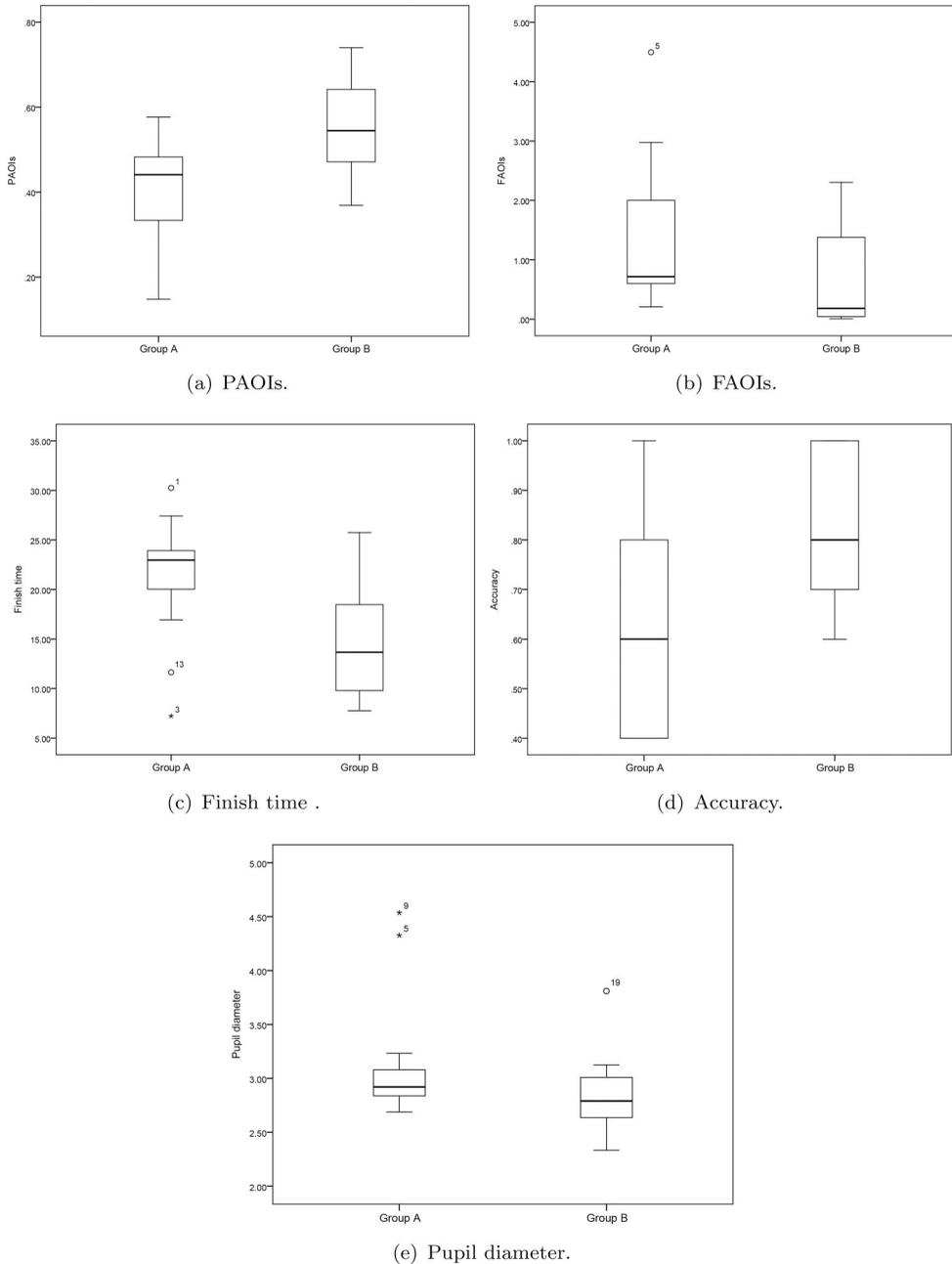


Figure 15. Statistical analysis of the two test groups.

(2) Quantitative analysis. After receiving eye-tracking feedback from participants, we conducted a Shapiro-Wilk test to assess the normal distribution of the results, as presented in Table 3. The statistical results show that the p-values of PAOIs and finish time are greater than 0.05, indicating that the results corresponding to the two indices fit the normal distribution. Conversely, the p-values for TAOIs, accuracy, and pupil diameter are smaller than 0.05, meaning that the data associated with these indices do not conform to a normal distribution.

Based on the normal distribution tests' results, we conducted significance tests between two groups using the T-test for PAOIs and Finish time, and the Mann-Whitney U test for TAOIs, accuracy, and pupil diameter, respectively. The mean (M), standard deviation (SD), Median (MD), and Interquartile Range (IQR) were used to analyze participant feedback. As shown in Table 4, the p-value = 0.003, 0.028, 0.004, 0.010, $0.034 < 0.05$ indicated a significant difference between groups A and B in all indices.

As shown in Figure 15(a), the PAOIs increased by 14% when participants observed augmented scenes (Group B, $M = 0.55$, $SD = 0.12$) compared to normal scenes (Group A, $M = 0.41$, $SD = 0.12$), the MD value of group B was greater than group A, and the IQR of group B was approximately distributed in the interval (0.5, 0.6), which suggested that participants have more interests on augmented scenes. As shown in Figure 15(b), the TAOIs of participants on the augmented scene (Group B, $M = 0.65$, $SD = 0.81$) was shorter than the normal scene (Group A, $M = 1.34$, $SD = 1.23$), and the MD value of Group B was close to the lower quartile, indicating that the proposed approach is able to attract users' attention more quickly since the background blurring that reduces the distraction of non-relevant information. The compound graphic variables can stimulate and draw users' visual attention.

In terms of the finish time of given tasks, the participants involved in the augmented scene (Group B, $M = 14.42$, $SD = 5.72$) took less time than the normal scene (Group A, $M = 21.13$, $SD = 5.99$), and the MD value of normal scene is close to the upper quartile (Figure 15(c)), which suggested that the augmented representation proposed in this article can effectively improve the users' ability to read maps or scenes. In contrast, the accuracy of given tasks of the augmented scene (Group B, $M = 0.84$, $SD = 0.17$) increased by 21% compared to the normal scene (Group A, $M = 0.63$, $SD = 0.21$), and many participants were able to answer the questions with 100% accuracy (Figure 15(d)), which indicated that the augmented scene could effectively enhance the participants' understanding and cognition of bridge stress.

Regarding the pupil diameter of participants during the experiment, the group who observed the augmented scene (Group B, $M = 2.82$, $SD = 0.34$) was smaller than the normal scene (Group A, $M = 3.13$, $SD = 0.57$), and the MD value of augmented scene is close to the lower quartile (Figure 15(e)), suggested that the cognitive load of Group B was less than that of Group A during the experiment.

In summary, the analysis results of all indices showed that compared with scene A, scene B can quickly draw the user's attention to focus on the AOI in the 3D scene. There was a significant increase in the accuracy of completing the given task while using less finish time, and the participants had lower cognitive load, indicating that the proposed method can effectively improve the user's cognitive efficiency of geographic scenes.

5. Discussion

In this article, we innovatively introduced visual attention into geovisualization, which is beneficial for people to capture critical information in geographic scenes. Nevertheless, the following points are still worth discussing, and we encourage the readers to apply further creative thinking to tackle the challenges presented in this context.

Regarding background simplification, we actually had two alternatives for implementing the Gaussian blur algorithm in the prototype system. The first option involves directly blurring the screen pixels outside the overlapping vision, as outlined in this article. It is also possible to blur the geographic objects in the scene, which can be performed by reprojection, progressively blurring based on the distance between the objects and the central point. However, it is important to emphasize that both options end up blurring the screen pixels. The difference between them is that the second type of blurring is highly relevant to spatial orientation, distance, and scale. It provides an avenue for investigating the correlations between scene interaction, background blurring, and user perception, which is also a research question worthy of our in-depth study.

According to the statement of Dragoi (2020), the overlapping vision formed by binocular fusion is irregular. However, this article computed a rectangular mask to present the overlapping vision area instead of an irregular mask. Theoretically, the irregular mask formed by binocular fusion can be calculated if we access more parameters of the eyes, such as visual acuity, color vision, and so on. From a practical perspective, obtaining these parameters requires more sensors of visual perception and neuroscience knowledge, which is beyond the scope of this article. Nevertheless, the authors have strategically mentioned this point here and look forward to interdisciplinary collaborations to address this challenge.

In the cognitive evaluation conducted in this study, fixations and pupil diameter were chosen as the eye-tracking analysis indices due to hardware constraints. As Dong et al. (2019) stated, eye movement metrics can be classified into three categories according to their cognitive patterns: information processing, visual search, and cognitive burden metrics. Specifically, fixation and saccade can characterize information processing and visual search processes, respectively. Furthermore, changes in the brain's cognitive load cause the corresponding changes in pupil size. Consequently, pupil size is also commonly used to characterize the level of cognitive load. Therefore, we are coordinating with the cartography lab to introduce a new eye-tracking device to allow more in-depth eye-tracking analysis to address cognitive issues in more complex geographic scenes.

In terms of promoting the augmented representation approach, we took the bridge stress visualization as an example in this article. However, bridge stress is a specialized field with a limited amount of geoinformation, making the amount of stimuli in the eye-tracking experiment relatively small. In this context, we are actively applying the proposed approach to other fields, including but not limited to smart cities and disaster management. Compared with bridge stress, disasters involve more diverse spatial information, which facilitates the design of more prosperous stimulation experiments of eye-tracking. It also throws up a research question of balancing the amount of information within AOI and its context visibility to avoid information overload that increases the cognitive load.

6 Conclusion and outlook

Following the laws of visual perception of the human eyes, this article proposed a visual attention-guided augmented representation approach of geographic scenes. First, related terms, such as visual attention and augmented representation, were defined. Second, we introduced AOI computation and background simplification, 3D representation of bridge stress, and compound graphic variables in detail. Third, focusing on bridge stress visualization as a case study and performed a cognitive analysis of the augmented representation using eye-tracking. We aim to propose a novel geovisualization approach from the visual perception perspective to promote geo-knowledge communication. The main contributions of this article are summarized as follows.

Augmented representation offers a new idea for the theoretical studies of geovisualization. The inherent thinking of detailed geovisualization inevitably increases the cognitive load of visual information processing. However, attention shifting and visual information processing distinguish human visual scanning of a specific scene. In other words, the eyes will prioritize searching for the AOI and ignore other areas that do not attract visual attention, and the salient visual stimulus is beneficial for drawing attention and decoding geoinformation. Leveraging the above visual cognitive mechanism, we computed the AOI considering FOV, blurred the background to reduce the interference of irrelevant information, and introduced the compound graphic variables to provoke high neural responses and draw visual attention. The main objective is to help people in swiftly capture critical information, which was also verified in our eye-tracking experiment on bridge stress visualization. In summary, augmented representation provides new research ideas for efficient geovisualization and could be applied to visualize many geographical phenomena.

In the future, a topic worthy of investigation is the exploitation of the user interaction feedback. This valuable information could be coupled with deep learning, e.g. graph attention networks, for the anticipation and prediction of user focus and preferences for the iterative optimization of geographic scene representations.

Acknowledgements

The authors would like to express their gratitude to all participants in the comparative experiment and the editor and anonymous reviewers for their careful reading and constructive suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This article was supported by the National Natural Science Foundation of China (Grant Nos. U2034202, 42201446 and 42271424), Guangdong Science and Technology Strategic Innovation Fund (the Guangdong–Hong Kong–Macau Joint Laboratory Program, Grant No. 2020B1212030009).

Notes on contributors

Weilian Li is a postdoctoral research associate in the Computational Methods Lab at the HafenCity University Hamburg. He received his Ph.D. in Geomatics Science and Technology at Southwest Jiaotong University in December 2020. Then, he finished his postdoc research stay in the Geoinformation Group at the University of Bonn in October 2023. He has a strong interest in topics related to smart cities, disaster management, and digital twins. He proposed the idea and contributed to this article's methodology, implementation, experiment, and writing.

Jun Zhu is a professor and the deputy dean of the Faculty of Geosciences and Environmental Engineering at Southwest Jiaotong University. He received his Ph.D. in Cartography and GIS from the Chinese Academy of Sciences. His research interests include computer vision, 3D GIS, and virtual geographic environments. He proposed and contributed to the study design, methodology, and review.

Qing Zhu is a professor and the director of the Research Committee in the Faculty of Geosciences and Environmental Engineering at Southwest Jiaotong University. He received his Ph.D. in railway engineering from North Jiaotong University. His research interests include photogrammetry, geographic information systems, and virtual geographic environments. He contributed to the study design, methodology, and review.

Jinbin Zhang is a Ph.D. student in the Faculty of Geosciences and Environmental Engineering at Southwest Jiaotong University. His research interests include geographic knowledge graphs and 3D GIS. He played a crucial role in the literature review and cognitive procedure design.

Xiao Han received his M.S. in Geomatics Science and Technology at Southwest Jiaotong University in July 2023. His research interests include 3D GIS and geovisualization. He provided technical support for the prototype system design and implementation.

Youness Dehbi is a professor and the Chair of the Computational Methods Lab at the HafenCity University Hamburg. He received his Ph.D. degree in the Institute of Geodesy and Geoinformation from the University of Bonn in 2016. His research focuses on urban reasoning, analytics, pattern recognition, and machine learning for geoinformation and statistical relational learning. He contributed to the literature review, study design, and review.

Data and codes availability statement

The data and codes that support the findings of this study can be found online at <https://doi.org/10.6084/m9.figshare.24167775.v2>

References

- Andrienko, G., *et al.*, 2010. Geova(t)-geospatial visual analytics: focus on time. *International Journal of Geographical Information Science*, 24 (10), 1453–1457.
- Badwi, I., Ellaithy, H.M., and Youssef, H.E., 2022. 3D-GIS parametric modelling for virtual urban simulation using CityEngine. *Annals of GIS*, 28 (3), 325–341.
- Bodum, L., 2005. Modelling virtual environments for geovisualization: a focus on representation. In: J. Dykes, A. MacEachren, and M.-J. Kraak, eds. *Exploring geovisualization*, international cartographic association. Oxford: Elsevier, 389–402.
- Boton, C., 2018. Supporting constructability analysis meetings with immersive virtual reality-based collaborative bim 4d simulation. *Automation in Construction*, 96, 1–15.
- Broo, D.G., Bravo-Haro, M., and Schooling, J., 2022. Design and implementation of a smart infrastructure digital twin. *Automation in Construction*, 136, 104171.
- Chen, Y., *et al.*, 2021. A framework of pan-maps: facilitating a unification of maps and map-likes, in: Proceedings of the ICA, Copernicus GmbH. pp. 1–8.

- Çöltekin, A., Janetzko, H., and Fabrikant, S.I., 2018. Geovisualization. In: J.P. Wilson, ed. The geographic information science & technology body of knowledge (2nd quarter 2018 edn). <https://doi.org/10.22224/gistbok/2018.2.6>.
- Çöltekin, A., et al., 2019. Where are we now? re-visiting the digital earth through human-centered virtual and augmented reality geovisualization environments. *International Journal of Digital Earth*, 12 (2), 119–122.
- Degbelo, A., 2022. Fair geovisualizations: definitions, challenges, and the road ahead. *International Journal of Geographical Information Science*, 36 (6), 1059–1099.
- Dong, W., et al., 2019. New research progress of eye tracking-based map cognition in cartography since 2008. *Acta Geogr. Sin*, 74, 599–614.
- Dong, W., et al., 2018. Using eye tracking to evaluate the usability of flow maps. *ISPRS International Journal of Geo-Information*, 7 (7), 281.
- Dong, W., et al., 2020. How does map use differ in virtual reality and desktop-based environments? *International Journal of Digital Earth*, 13 (12), 1484–1503.
- Dragoi, V., 2020. Visual processing: Eye and retina. URL: <https://nba.uth.tmc.edu/neuroscience/m/index.htm>. neuroscience Online: An Electronic Textbook for the Neurosciences, provided by Department of Neurobiology and Anatomy, McGovern Medical School at UTHealth.
- Dransch, D., Rotzoll, H., and Poser, K., 2010. The contribution of maps to the challenges of risk communication to the public. *International Journal of Digital Earth*, 3 (3), 292–311.
- Fu, L., et al., 2021. Tunnel vision optimization method for vr flood scenes based on gaussian blur. *International Journal of Digital Earth*, 14 (7), 821–835.
- Goodchild, M.F., et al., 2012. Next-generation digital earth. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (28), 11088–11094.
- Holland, J.L., and Lee, S., 2019. Internet of everything (ioe): Eye tracking data analysis. In: P.J.S. Cardoso, ed. *Harnessing the Internet of Everything (IoE) for Accelerated Innovation Opportunities*. Hershey, PA: IGI Global, 215–245.
- Jahnke, M., et al., 2008. Non-photorealistic rendering on mobile devices and its usability concerns. In: H. Lin and M. Batty, eds. *CD-proceedings virtual geographic environments-an international conference on development on visualization and virtual environments in geographic information science*. Hong Kong: Science Press, 164–177.
- Kim, C., Son, H., and Kim, C., 2013. Automated construction progress measurement using a 4d building information model and 3d data. *Automation in Construction*, 31, 75–82.
- Laurini, R., 2017. *Geographic knowledge infrastructure: applications to territorial intelligence and smart cities*. Paris: Elsevier.
- Li, H., et al., 2015. Bridge stress calculation based on the dynamic response of coupled train-bridge system. *Engineering Structures*, 99, 334–345.
- Li, W., et al., 2023. Immersive virtual reality as a tool to improve bridge teaching communication. *Expert Systems with Applications*, 217, 119502.
- Li, W., et al., 2021. An augmented representation method of debris flow scenes to improve public perception. *International Journal of Geographical Information Science*, 35 (8), 1521–1544.
- Li, W., et al., 2022. Three-dimensional virtual representation for the whole process of dam-break floods from a geospatial storytelling perspective. *International Journal of Digital Earth*, 15 (1), 1637–1656.
- Li, W., et al., 2019. A fusion visualization method for disaster information based on self-explanatory symbols and photorealistic scene cooperation. *ISPRS International Journal of Geo-Information*, 8 (3), 104.
- Li, Y., et al., 2020. Semantic visual variables for augmented geovisualization. *The Cartographic Journal*, 57 (1), 43–56.
- Ma, Z., et al., 2021. Activity-based process construction for participatory geo-analysis. *GIScience & Remote Sensing*, 58 (2), 180–198.
- Ma, Z., et al., 2022. Customizable process design for collaborative geographic analysis. *GIScience & Remote Sensing*, 59 (1), 914–935.
- MacEachren, A.M., 2004. *How maps work: representation, visualization, and design*. New York: Guilford Press.

- MacEachren, A.M., et al., 2004. Geovisualization for knowledge construction and decision support. *IEEE Computer Graphics and Applications*, 24 (1), 13–17.
- Moore, A., et al., 2020. Comparative usability of an augmented reality sandtable and 3d gis for education. *International Journal of Geographical Information Science*, 34 (2), 229–250.
- Reichenbacher, T., and Swienty, O., 2007. Attention-guiding geovisualization. In: Proceedings of the 10th AGILE International Conference on Geographic Information Science, 8th–11th May, Aalborg University, Denmark, Citeseer.
- Reicher, G.M., Snyder, C.R., and Richards, J.T., 1976. Familiarity of background characters in visual scanning. *Journal of Experimental Psychology. Human Perception and Performance*, 2 (4), 522–530.
- Robinson, A.C., 2011. Highlighting in geovisualization. *Cartography and Geographic Information Science*, 38 (4), 373–383.
- Swienty, O., 2008. Attention-Guiding Geovisualisation. Ph.D. thesis. Technische Universität München.
- Swienty, O., Zhang, M., and Reichenbacher, T., 2006. Attention guiding visualization of geospatial information. In: *Geoinformatics 2006: Geospatial Information Technology*, 28–29 October 2006. Wuhan, China: SPIE, 642101.
- Torralba, A., et al., 2006. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological Review*, 113 (4), 766–786.
- Treisman, A., 1982. Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology. Human Perception and Performance*, 8 (2), 194–214.
- Ungerleider, L.G., 1995. Functional brain imaging studies of cortical mechanisms for memory. *Science*, 270 (5237), 769–775.
- Wolfe, J.M., and Horowitz, T.S., 2004. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews. Neuroscience*, 5 (6), 495–501.
- Wu, J., et al., 2023. A dynamic holographic modelling method of digital twin scenes for bridge construction. *International Journal of Digital Earth*, 16 (1), 2404–2425.
- Zhang, G., et al., 2020. An efficient flood dynamic visualization approach based on 3d printing and augmented reality. *International Journal of Digital Earth*, 13 (11), 1302–1320.
- Zhang, Z., and Ma, X., 2019. Lane recognition algorithm using the Hough transform based on complicated conditions. *Journal of Computer and Communications*, 07 (11), 65–75.
- Zhou, X., and Zhang, X., 2019. Thoughts on the development of bridge technology in China. *Engineering*, 5 (6), 1120–1130.
- Zhu, J., et al., 2023. The impact of spatial scale on layout learning and individual evacuation behavior in indoor fires: single-scale learning perspectives. *International Journal of Geographical Information Science*, 38 (1), 77–99.
- Zhu, J. et al., 2024. A knowledge-guided visualization framework of disaster scenes for helping the public cognize risk information. *International Journal of Geographical Information Science*. <https://doi.org/10.1080/13658816.2023.2298299>
- Zhu, J., et al., 2015. A procedural modelling method for virtual high-speed railway scenes based on model combination and spatial semantic constraint. *International Journal of Geographical Information Science*, 29 (6), 1059–1080.
- Zuo, C., et al. (2022). Multiscale geovisual analysis of knowledge innovation patterns using big scholarly data. *Annals of GIS*, 28 (2), 197–212.