A Cognitive Mechanism–Driven Optimization Framework for Visual Communication Based on Visual Information Processing

Xuejie Qin, Jinrui Li*, Jiufang Mei, Ling Jin

Harbin Institute of Information Technology, Harbin, China

*2413613533@qq.com

Abstract. Visual communication aims to efficiently convey information by guiding user attention and reducing cognitive effort; however, existing design practices largely rely on subjective experience and lack systematic integration with visual information processing mechanisms. To address this gap, this study proposes a cognitive mechanism–driven optimization framework that links key cognitive processes, including visual attention, visual search, and cognitive load, to operable design variables such as saliency, visual hierarchy, information density, and text–image integration. A comparative experiment was conducted across infographic reading, poster information extraction, and interface guidance tasks, using eye-tracking metrics (e.g., Time to First Fixation and scanpath length), behavioral performance, and NASA-TLX scores for evaluation. The results show that the proposed framework significantly improves attention allocation, reduces cognitive load, and enhances task efficiency compared with baseline and single-strategy approaches, demonstrating its effectiveness as a measurable and empirically validated method for evidence-based visual communication design.

Keywords: Visual communication; Visual attention; Cognitive load; Eye-tracking; Design optimization.

1 Introduction

Visual communication aims to convey information efficiently by guiding users' visual attention and supporting rapid understanding. However, with increasing information complexity, many designs still rely on subjective experience rather than systematic cognitive principles, leading to unclear attention guidance, inefficient visual search, and increased cognitive load.

Existing research has explored several approaches to improving visual communication. Saliency-based methods (e.g., Itti and Koch) focus on predicting attention distribution based on bottom-up visual features, but provide limited guidance for design optimization. Layout- and hierarchy-based approaches enhance structural organization, yet often lack explicit connections to underlying cognitive mechanisms. In

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_23

addition, cognitive load theory (Sweller et al.) offers important principles for reducing mental effort, but remains largely conceptual and is not directly translated into operable design variables. As a result, existing approaches are fragmented and address isolated aspects of visual communication.

Therefore, despite these efforts, there remains a lack of a unified and mechanism-driven framework that integrates visual attention, cognitive load, and perceptual organization into an operable design methodology.

To address this gap, this study proposes a cognitive mechanism–driven optimization framework based on visual information processing. Building upon prior research, the proposed framework translates cognitive mechanisms into actionable design variables and integrates multiple factors into a unified system. Unlike traditional predictive models, it emphasizes design-oriented optimization and incorporates an evaluation–feedback loop based on eye-tracking and behavioral measures, enabling iterative refinement of design solutions..

2 Literature Review

2.1 Visual Attention and Saliency in Visual Design

Visual attention plays a central role in visual information processing. Existing studies distinguish between bottom-up attention driven by visual saliency (e.g., color, contrast, orientation) and top-down attention guided by task goals [1]. Classical computational models, such as the Itti model, as well as recent deep learning-based approaches, have been widely applied to predict fixation distribution and identify visually salient regions in images and interfaces [2][3].

In the context of visual communication design, saliency principles are commonly used to highlight key elements and guide user attention. However, most existing studies primarily focus on attention prediction rather than design-oriented optimization, and typically treat saliency as an isolated factor [5]. Moreover, they often neglect the interaction between saliency and higher-level cognitive processes, such as task-driven attention and information integration. This limitation reduces their effectiveness in complex visual environments, where multiple cognitive mechanisms jointly influence user behavior [6].

2.2 Cognitive Load and Design Optimization

Cognitive Load Theory provides an important theoretical framework for understanding how information presentation affects mental effort. Prior research has shown that excessive visual complexity, inefficient layout, and separated text–image representations can increase extraneous cognitive load, thereby reducing comprehension efficiency [4]. To address these issues, design principles such as information segmentation, redundancy reduction, and spatial integration have been proposed to improve usability and learning outcomes.

In addition, recent studies have explored optimization approaches in visual communication, including eye-tracking–based evaluation and data-driven layout refinement . These methods highlight the value of empirical feedback in assessing design effectiveness. However, they are often task-specific and fragmented, lacking a systematic framework that connects design variables with underlying cognitive processes and performance outcomes.

To summarize, although existing studies have provided important insights into visual attention and cognitive load, they are typically developed in isolation and remain either predictive or conceptual in nature. There is still a lack of an integrated, mechanism-driven approach that systematically links cognitive processes to operable design variables and optimization strategies. This gap motivates the present study, which aims to bridge cognitive theory and visual communication practice through a unified and empirically validated framework.

3 Methodology

3.1 Overall Framework Design

As shown in Fig. 1, this study proposes a cognitive mechanism–driven framework that models visual design as an information-processing process. The framework includes six layers: input, cognitive processing, variable extraction, optimization, output, and evaluation–feedback. Infographics, posters, and interfaces are analyzed to identify cognitive bottlenecks before optimization.

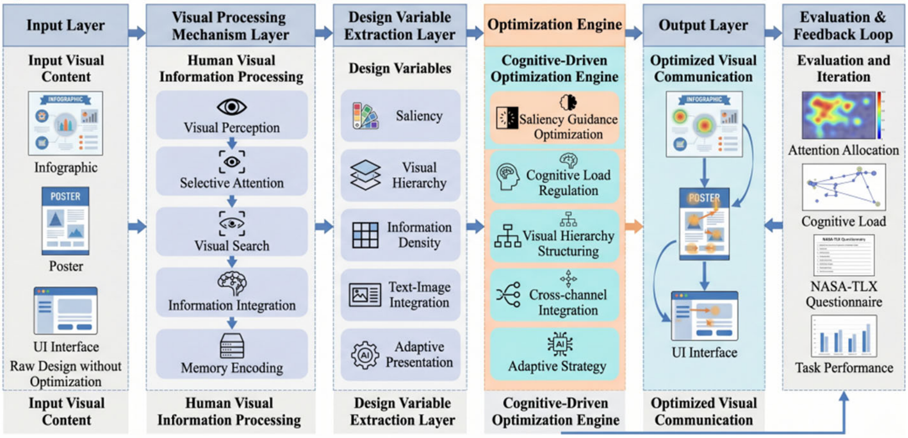

Fig. 1. Cognitive-Driven Visual Design Optimization Framework.

The workflow includes four steps: task definition, problem diagnosis, design optimization, and iterative evaluation. Weak saliency, poor hierarchy, excessive information density, and separated text-image relations are first identified, then corresponding strategies are applied. The overall optimization target is defined as:

$$Q = \alpha A - \beta L + \gamma P \quad (1)$$

where Q is design quality, A is attention allocation efficiency, L is cognitive load, and P is task performance.

3.2 Design Variable Extraction and Cognitive Mapping

Five groups of design variables were derived from visual cognition theory: saliency, visual hierarchy, information density, text–image integration, and adaptive presentation. These variables are directly manipulable in design practice and correspond to specific cognitive processes. Saliency was implemented through contrast, local emphasis, and directional cues, while visual hierarchy was controlled by font size, spacing, alignment, and position. Information density was defined as the number of visual elements per unit area, and text–image integration as the spatial proximity between related elements.

A rule-based mapping was established between cognitive bottlenecks and design operations. Weak attention was addressed by increasing contrast and reducing competing elements. Disordered scanpaths were improved using Gestalt grouping principles (e.g., proximity, similarity). High cognitive load was mitigated through content segmentation and removal of redundant elements, while separated text and graphics were integrated by reducing spatial distance.

To improve reproducibility, key variables were semi-quantitatively controlled. Contrast was increased by approximately 20–40% relative to the baseline. Local emphasis was achieved by enlarging target elements (1.2–1.5×), increasing color saturation, or adding directional cues. Information density was reduced by 15–30%, and text–image distance was minimized to enhance perceptual grouping.

3.3 Cognitive-Driven Optimization Engine

The core module is the cognitive-driven optimization engine in Fig. 1, which integrates five strategies: saliency guidance, cognitive load regulation, visual hierarchy structuring, cross-channel integration, and adaptive strategy. Each design sample was first divided into Areas of Interest (AOIs), including primary targets, secondary information, and decorative regions. Then, different parameters were adjusted according to task type. For example, poster tasks emphasized titles and call-to-action elements, infographic tasks reduced dense clusters through modular layout, and interface tasks combined controls with concise labels and directional cues.

The optimization process was semi-structured rather than fully automatic. Adobe Illustrator and Figma were used to build controlled design variants, while saliency and hierarchy were manually adjusted according to predefined rules. Visual load was modeled as:

$$L_v = \sum_{i=1}^{n} w_i x_i \tag{2}$$

where L_v is visual load, x_i is the value of a design factor, and w_i is its weight. Reducing unnecessary L_v helped lower extraneous cognitive load.

3.4 Experimental Validation and Iterative Refinement

To validate the framework, comparative experiments were conducted on infographic reading, poster information extraction, and interface guidance tasks. For each task, baseline and optimized versions were created under identical semantic content, ensuring that only design strategies differed. Participants performed goal-oriented tasks while gaze data were recorded using a 60 Hz eye tracker. Key indicators included Time to First Fixation (TTFF), fixation duration in AOIs, scanpath length, scanpath entropy, and pupil diameter variation. Task completion time, response accuracy, and NASA-TLX scores were also collected.

The results were used to evaluate improvements in attention allocation, cognitive load, and task efficiency, and were fed back into the framework for iterative refinement. Specifically, if TTFF remained high, contrast was increased or competing elements were reduced; if scanpath entropy was high, grouping and alignment were adjusted; if NASA-TLX scores were high, information density was reduced or text–image integration was improved. This process was repeated until no significant improvement was observed, forming an iterative, data-driven design optimization procedure.

4 Experiments and Comparative Analysis

4.1 Experimental Setup

A controlled comparative experiment was conducted using three representative tasks: infographic reading, poster information extraction, and interface guidance. For each task, baseline and optimized versions were created under identical semantic content and display conditions, ensuring that only design strategies differed.

A total of 36 participants with normal or corrected-to-normal vision were recruited. Eye-movement data were recorded using a 60 Hz eye tracker after standard calibration. Measures included Time to First Fixation (TTFF), fixation duration, scanpath length, scanpath entropy, and pupil diameter variation, along with task completion time, response accuracy, and NASA-TLX scores. A within-subject design with randomized stimulus order was used to reduce sequence effects and improve reliability.

4.2 Main Experimental Results

Table 1. Main experimental results under different design conditions

Method	TTFF (ms)	Scanpath Length (px)	Accuracy (%)	Task Time (s)	NASA-TLX
Baseline	1250	1840	78.6	14.8	62.4
Saliency-based	1015	1652	83.1	13.2	56.8
Layout-based	1098	1587	84.5	12.9	54.9
Cognitive Load-based	1124	1610	85.2	12.5	52.7
Proposed Framework	**842**	**1395**	**89.6**	**10.8**	**46.3**

As shown in Table 1, the proposed framework achieved the best overall performance in attention guidance, cognitive load reduction, and task efficiency. Compared with the baseline, it reduced TTFF, scanpath length, and NASA-TLX scores while improving response accuracy. It also outperformed saliency-, layout-, and cognitive load–based methods, demonstrating more consistent improvements across all indicators.

These findings indicate that the framework enables faster target identification, shorter visual search paths, and lower subjective workload, with particularly strong advantages in complex tasks such as infographics and interface guidance..

4.3 Discussion

The results indicate that the effectiveness of the proposed framework stems from integrating multiple cognitive mechanisms rather than relying on a single design strategy. While saliency-based methods improve initial attention capture, they are limited when information structure is unclear. Layout- and cognitive load-based methods enhance organization and reduce processing burden, but are less effective in guiding immediate fixation toward key targets. By combining saliency guidance, hierarchy optimization, and information integration, the proposed framework achieves more consistent improvements across all indicators. These findings suggest that effective visual communication depends on the coordinated regulation of attention, structure, and cognitive load, and that multi-dimensional optimization is better suited for complex design tasks.

5 Conclusion

This study proposed a cognitive mechanism–driven optimization framework for visual communication based on visual information processing. By integrating visual attention, cognitive load, and visual organization, the framework translates cognitive principles into an operable design methodology with measurable variables and clear optimization rules. Experimental results across multiple tasks demonstrate that the proposed method outperforms conventional and single-strategy approaches in attention guidance, task efficiency, and workload reduction. The findings highlight the importance of coordinated cognitive strategies in visual communication and provide a practical reference for evidence-based design. Future work will extend the framework to dynamic media, personalized adaptation, and automated optimization.

References

1. Itti L, Koch C. Computational modelling of visual attention. Nature Reviews Neuroscience, 2001, 2(3): 194-203. https://doi.org/10.1038/35058500
2. Borji A. Saliency prediction in the deep learning era: Successes and limitations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 679-700. https://doi.org/10.1109/TPAMI.2019.2935715

3. Bylinskii Z, DeGennaro E M, Rajalingham R, et al. Towards the quantitative evaluation of visual attention models. Vision Research, 2015, 116: 258-268. https://doi.org/10.1016/j.visres.2015.04.007
4. Sweller J, Ayres P, Kalyuga S. Cognitive Load Theory. New York: Springer, 2011.
5. Ware C. Information Visualization: Perception for Design (4th Edition). Cambridge, MA: Morgan Kaufmann, 2021. https://doi.org/10.1016/C2016-0-02395-1
6. Rosenholtz R. Capabilities and limitations of peripheral vision. Annual Review of Vision Science, 2016, 2: 437-457. https://doi.org/10.1146/annurev-vision-082114-035733