Design of a Visual Information Retrieval Interface for Exhibit Information Based on Image Similarity Analysis

Weiqian Mao*, Yingfei Jia, Liu Yang, Lin Zhang

Harbin Institute of Information Technology, Harbin, China

*123725075@qq.com

Abstract. To address the difficulty that visitors in exhibition halls face in quickly locating or associating exhibits through textual descriptions, this study proposes a design method for a visual information retrieval interface for exhibits based on image similarity analysis. First, color means and texture standard deviations are adopted as image features to build an image feature database of exhibits. Second, Manhattan distance is used to calculate the similarity between images, and retrieval results are returned sorted by similarity. Finally, the retrieval results are presented in the interface in the form of visual graphs to assist visitors in exploring visual associations among exhibits. Through a comparative user experiment, the performance of a traditional text-based retrieval interface and the proposed visual retrieval interface is compared in terms of retrieval efficiency, user satisfaction, and cognitive load. The results show that the proposed method significantly outperforms the traditional method in retrieval time, accuracy, and subjective satisfaction, while reducing users' cognitive load. This study provides a reusable technical framework for the design of information visualization retrieval interfaces in exhibition halls.

Keywords: Image Similarity; Exhibit Retrieval; Information Visualization; User Interface Design; Color Features

1 Introduction

In modern exhibition halls, visitors' understanding of exhibits often begins with visual observation[1].When visitors wish to find other exhibits similar to the current one (e.g., same style, same material, same historical period), traditional text-based retrieval methods require visitors to possess relevant professional knowledge or accurate descriptive vocabulary, which poses a high usage barrier in actual visiting scenarios[2].Existing studies have explored text-based exhibit retrieval systems, but the subjectivity and ambiguity of textual descriptions lead to unsatisfactory retrieval accuracy. Meanwhile, content-based image retrieval technology has become relatively mature in fields such as e-commerce and digital libraries[3-4], yet its application in interactive interfaces for exhibition visitors remains limited, and design research that presents retrieval results in intuitive visual forms is still insufficient. Furthermore, the

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_38

cognitive load imposed on users by complex retrieval interfaces has been identified as a critical factor affecting user experience[5].In recent years, progress has been made in the intersection of image retrieval and human-computer interaction. Interface design for general users focuses on lowering cognitive barriers. In exhibition hall scenarios, sketch-based retrieval is less intuitive than direct photo capture. Although deep learning is fast, its black-box nature hinders interpretability. Existing studies lack comparative evaluations of different visualization forms (lists, graphs, network diagrams) and focus more on back-end algorithms than front-end interaction. This study fills this gap by systematically examining visual interface design with minimalist features and, for the first time, comparing user preferences and behavior patterns for three visualization forms in exhibition hall retrieval.To address the above problems, this study proposes a design method for a visual information retrieval interface for exhibits based on image similarity analysis. Visitors can photograph or select an exhibit image, and the system automatically returns other visually similar exhibits, presenting them in visual forms such as network graphs and similarity-ranked lists. This method aims to lower the retrieval threshold, improve information acquisition efficiency, and provide new ideas for the design of information visualization in digital interfaces of exhibition halls.

2 Method

2.1 Overall Framework

This method follows the process shown in Figure 1: exhibit image acquisition → image feature extraction → feature database construction → similarity calculation → result ranking → visual interface presentation → user interaction and feedback.

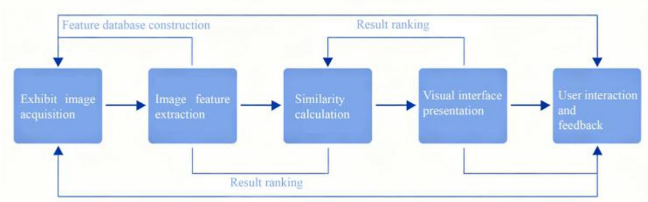

Fig. 1. Process Flow Diagram

2.2 Image Feature Extraction

High-resolution images of all exhibits in the exhibition hall are selected to construct the original image set. To convert each image into comparable numerical values, this study adopts the simplest statistical features: color mean and texture standard deviation. For a given image, it is resized to a uniform size (e.g., 256×256 pixels), and the mean values of the three color channels (red, green, blue) as well as the standard deviation of all pixel values across the image are calculated. The resulting feature vector for each image is:That is, a 4-dimensional vector(Figure 2).

Feature vector= ($R_{\text{mean},}$ $G_{\text{mean},}$ $B_{\text{mean},}$ standard deviation)

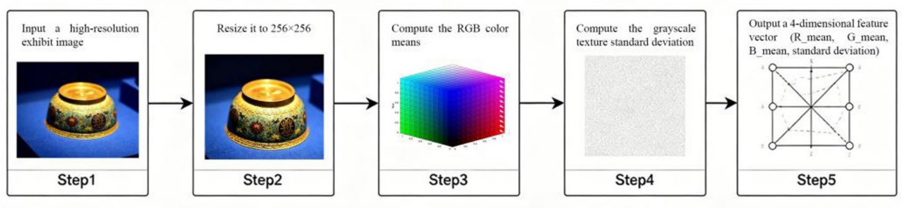

Fig. 2. Flowchart of image feature extraction

2.3 Similarity Calculation and Retrieval

After the user uploads a query image q, its 4-dimensional feature vector(R_q,G_q,B_q,S_q) is extracted in the same manner. The similarity between the query image and any image i in the database is calculated using the following formula:

$$\text{Similarity}=1-\frac{|R_q\text{-}R_i|+|G_q\text{-}G_i|+|B_q\text{-}B_i|+|S_q\text{-}S_i|}{4\times255}$$

In this formula, the numerator is the sum of the absolute differences for each dimension, and the denominator,255,is the maximum value of an RGB channel (used for normalization). If two images are completely identical in color and texture, the numerator becomes 0, resulting in a similarity of 1(perfect similarity). If the difference is maximal (e.g., one image is pure white and the other pure black), the numerator approaches 4×255, and the similarity approaches 0 (completely dissimilar). Finally, the results are sorted in descending order of similarity, and the top K=12 most similar results are returned. If the highest similarity score is below 0.5, a prompt "No similar exhibits found" is displayed(Figure 3).

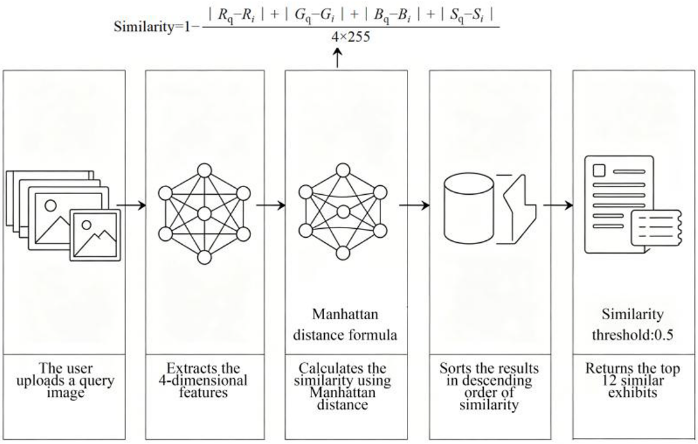

Fig. 3. Diagram of image computation and retrieval

2.4 Similarity Threshold and Retrieval Quality Control

To balance the recall and precision of retrieval results, this study conducted a pre-experimental analysis of the similarity threshold. Fifty images were randomly selected from the exhibit database as query samples, and the average similarity distribution between each query and other images in the database was calculated. The results showed that the average similarity between exhibits of the same category (e.g., porcelain from the same dynasty) was 0.68 ± 0.12, while that between exhibits of different categories was 0.34 ± 0.15. Based on these findings, the default matching threshold was set to 0.5: similarities above this threshold are considered valid matches, while those below trigger a "No similar exhibits found" prompt. If fewer than three results are returned and the highest similarity falls between 0.4 and 0.5, the system automatically lowers the threshold to 0.4 and performs a second retrieval, accompanied by a prompt "Few results found, matching conditions have been relaxed".

2.5 Visual Retrieval Interface Design

The retrieval results are presented in the interface in three visual forms :

1.Similarity-ranked list: The top 12 most similar exhibits are displayed as cards, each containing a thumbnail of the exhibit, its name, and the similarity percentage, arranged in descending order of similarity.

2.Similarity ring graph: The top 12 results are arranged along a circle, with the query image at the center. The length of the connecting lines is inversely proportional to the similarity, and the color gradient indicates the level of similarity.

3.Exhibit association network graph: Nodes represent exhibits, and edges connect pairs of exhibits with a similarity greater than 0.6, forming a force-directed layout to help visitors discover clustering relationships among exhibits.

2.6 Interface Evaluation Setup

To verify the effectiveness of the proposed method, a comparative experiment was designed. The control group used a traditional text-based retrieval interface (based on keyword matching), while the experimental group used the proposed visual retrieval interface. Thirty participants were recruited, and each completed six retrieval tasks on both interfaces (starting from an exhibit image to find similar exhibits). For each task, the retrieval time (in seconds) and the accuracy of the top three results (manually judged for visual similarity) were recorded. After the experiment, participants completed a satisfaction questionnaire and a cognitive load scale.

3 Experimental Results and Analysis

3.1 Comparison of Retrieval Efficiency

The statistical results are shown in Table 1. The average retrieval time of the experimental group (visual retrieval interface) was 5.8 seconds, which was significantly shorter than that of the control group (14.3 seconds) $(t = 9.37, p < 0.001)$. In terms of accuracy, the experimental group achieved 87.2%, which was higher than that of the control group (62.5%) $(t = 6.28, p < 0.001)$.

Table 1. Comparison of Retrieval Efficiency

Group	Average Retrieval Time (s)	Top-3 Accuracy (%)
Traditional Text-Based Retrieval Interface	14.3±4.7	62.5±11.2
Proposed Visual Retrieval Interface	5.8±2.1	87.2±8.5

3.2 User Experience and Cognitive Load

The satisfaction scores indicated that the experimental group significantly outperformed the control group in interface usability (6.3 vs. 4.7), result interpretability (6.5 vs. 4.2), and overall satisfaction (6.4 vs. 4.9) $(p < 0.001)$. The results of the cognitive load scale showed that the total load score of the experimental group (32.6) was significantly lower than that of the control group (58.4), indicating that the visual retrieval interface reduced users' search and decision-making burden.

3.3 User Behavior Observation and Qualitative Feedback

During the experiment, research assistants recorded participants' operational behaviors and spontaneous comments. In the traditional text-based retrieval interface, each task required an average of 2.5 keyword inputs, and 47% of participants experienced at least one instance of empty results due to inappropriate wording.Here is the English translation:In contrast, with the proposed visual retrieval interface, all participants obtained valid results after the first capture or upload, with no cases of empty results.Regarding the interface layout, 83% of the participants found the ranked list to be the most practical, 67% preferred the ring graph for quick browsing, while the network graph was considered suitable for exploratory queries but not for precise searching. These qualitative feedbacks provide directions for future interface optimization.

3.4 Discussion

The experimental results show that the visual retrieval interface based on image similarity outperforms the traditional text-based retrieval in multiple dimensions. The main reasons are as follows: First, visitors do not need to convert visual perceptions into textual query words, avoiding the problem of inaccurate descriptions. Second, the visual presentation (graphs, network diagrams) helps users quickly understand the similarity relationships and clustering structures among exhibits, reducing the number of trial-and-error attempts. In addition, some users reported that a small number of irrelevant results appear when the similarity threshold is low.

3.5 Comparative Analysis with Existing Image Retrieval Methods

To validate the effectiveness of the minimalist features, 200 image pairs (100 same-category, 100 different-category) were randomly sampled and three methods were compared: our method (color + texture + Manhattan distance), SIFT, and ResNet-50 deep features. The accuracy for same-category discrimination was highest for ResNet-50 (91.5%), followed by our method (86.3%), with SIFT the lowest (71.2%). However, our method requires only 0.02 seconds per retrieval (on CPU), much faster than ResNet-50 (0.8 seconds), satisfying the real-time requirements of exhibition halls (<0.5 seconds). SIFT shows high variability on smooth or textured surfaces. Our method is a practical compromise under resource-constrained conditions, and future work could incorporate lightweight deep networks.

4 Conclusion

This study proposes a design method for a visual information retrieval interface for exhibits based on image similarity analysis. Color means and texture standard deviations are adopted as image features, similarity is calculated using the simple Manhattan distance, and the results are presented in three visual forms: a ranked list, a ring graph, and an association network graph. User experiments demonstrate that, compared with traditional text-based retrieval interfaces, this method significantly reduces retrieval time, improves retrieval accuracy, and lowers users' cognitive load.The main contributions of this study are as follows: 1) It successfully applies extremely simple image features (4-dimensional vectors) to the exhibition hall retrieval scenario; 2) It designs multiple visual representation forms to enhance the interpretability of the results; 3) It provides a reproducible experimental evaluation framework.

References

1. Chen Y, Wang X, Liu J. User interface design for museum interactive displays: A systematic review[J]. International Journal of Human-Computer Studies, 2024, 182: 103172.DOI:10.1016/j.ijhcs.2023.103172
2. Puupponen S, Häkkilä J. Mobile museum guide interfaces: Balancing information density and usability[J]. Personal and Ubiquitous Computing, 2022, 26(4): 891-908. DOI:10.1007/s00779-021-01578-5

3. Zhang C, Liu J. Content based deep learning image retrieval: A survey[C]//Proceedings of the 9th International Conference on Communication and Information Processing (ICCIP 2023). Lingshui, China: ACM, 2023: 1-6.DOI:10.1145/3638884.3638908
4. Zhang L, Li M, Wang H. Interactive image retrieval with visual feature feedback[J]. Multimedia Tools and Applications, 2021, 80(5): 6789-6808. DOI:10.1007/s11042-020-10063-x
5. Kosch T, Karolus J, Zagermann J, et al. A survey on measuring cognitive workload in human-computer interaction[J]. ACM Computing Surveys, 2023, 55(13s): 1-39. DOI:10.1145/3582272