ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Research on original authenticity construction and tourist perception in cultural heritage tourism

Jingjing Zhang[1,*]

[1] School of Tourism, Yellow River Conservancy Technical University, Kaifeng 475004, Henan, China

SUMMARY: *Under the circumstances where cultural tourism integration continues to deepen and digital technology is widely involved in cultural dissemination, the construction of original authenticity in cultural heritage tourism and the mechanism of tourists' perception have received increasing attention. This paper focuses on the cultural heritage tourism scenario and constructs a multi-source heterogeneous data system integrating text, image, behavior and structured evaluation. It proposes a feature system of original authenticity, a modeling method for tourists' perception correlation, and a five-layer perception analysis model based on multimodal feature fusion, and designs an intelligent computing framework for the relationship between original authenticity construction and tourists' perception. Experimental results show that the model in the original authenticity feature recognition task achieves an Accuracy of 0.918 and a F1 value of 0.908, which are 0.039 and 0.040 higher than those of earlier splicing and fusion models respectively; in the ancient architectural site scenario, the Accuracy reaches 0.923, and the cross-scenario test Accuracy remains at 0.894, demonstrating good stability and generalization ability. The research can provide technical support and method references for the optimization of cultural heritage tourism display, the improvement of tourists' experience and digital governance.*
Povzetek: This paper focuses on the construction of authenticity in cultural heritage tourism and the issue of tourists' perception. By integrating multi-source heterogeneous data such as text, images, behaviors, and structured evaluation, a system of authenticity characteristics, a multimodal perception analysis model, and a relationship calculation framework are constructed. The experimental results show that the model's Accuracy is 0.918, F1 is 0.908, the cross-scenario Accuracy is 0.894, and it has good stability and generalization ability, which can provide references for the optimization of heritage display and digital governance.

KEYWORDS: *Cultural heritage tourism; Construction of original authenticity; Tourist perception; Multimodal integration*

1 Introduction

In the context of the continuous advancement of cultural tourism integration and the continuous embedding of digital technology in tourism scenarios, cultural heritage tourism is shifting from traditional sightseeing activities to comprehensive practices that involve knowledge acquisition, emotional experience, and cultural identification. Compared to general tourism resources, cultural heritage not only carries historical memories and regional culture, but also has distinct public attributes and social value. Therefore, what tourists pay attention to during the visit often

*bjttzjj@126.com
https://doi.org/10.65102/is2026287

Published: 30 April, 2026

goes beyond the appearance of the landscape, service facilities, or convenience of the tour, and extends to whether the displayed content is close to the historical context, whether the space creation has credibility, and whether the narrative of communication can accurately convey the connotation of the heritage. The cognition and judgment formed around authenticity will directly affect the quality of the tourists' experience, emotional investment, satisfaction evaluation, and subsequent communication behaviors, and also relate to whether there can be a relatively coordinated relationship between cultural heritage protection and tourism development.

The original authenticity issue has always been an important topic in the study of cultural heritage. In the context of cultural heritage tourism, original authenticity is not a static single attribute, but is gradually constructed through the joint action of multiple factors such as the preservation of the heritage itself, scene display, cultural interpretation, tourism services, and media dissemination. In recent years, more and more cultural heritage sites have introduced digital navigation, immersive performances, virtual reconstruction, intelligent interpretation, and social media dissemination mechanisms in their display methods, and the channels for tourists to obtain information have significantly increased, and the participation methods have become more diverse. Such changes have enriched the means of heritage dissemination and expanded the boundaries of tourism experiences, but also make the authenticity judgment faced by tourists no longer solely depend on the site, building, objects, or historical scenes themselves, but are also influenced by platform evaluations, image reproduction, space packaging, interactive design, and online public opinion. Thus, the construction of original authenticity in cultural heritage tourism presents a more complex structural feature, and tourists' perception also shows obvious dynamics, differences, and context dependence.

From the existing research, the academic community has conducted a large number of discussions on themes such as authenticity in cultural heritage tourism, local identity, tourists' experience, and destination image, and has formed relatively rich theoretical achievements. However, from the perspective of research methods, the related analyses still mainly rely on questionnaire surveys, interview inductions, and case descriptions, and the attention to multi-source data integration, fine-grained feature extraction, cross-modal knowledge identification, and relationship modeling is still insufficient. Especially in the context of the continuous development of digital tourism, the formation process of tourists' perception is no longer limited to offline contact, but is jointly shaped by graphic evaluations, short-video dissemination, online recommendation information, on-site behavior trajectories, and environmental feedback. A single method is no longer able to fully reveal the complex correlation between the construction of original authenticity in cultural heritage tourism and tourists' perception. With the development of text mining, image recognition, multimodal fusion, and intelligent analysis methods, integrating cultural heritage tourism research with computer technology already has a strong practical basis and methodological support.

Based on this, this paper conducts research on the construction of original authenticity in cultural heritage tourism and tourists' perception, attempting to start from multiple heterogeneous data sources, systematically extract the factors of authenticity representation in cultural heritage scenes, and combine tourists' comment texts, image information, and behavioral characteristics to construct a tourists' perception analysis model and a correlation calculation framework. The research hopes to, while maintaining the core issue awareness of cultural heritage research, promote the related analysis to extend from empirical explanations to intelligent identification based on data support, thereby providing more targeted research ideas for the optimization of cultural heritage tourism display, the improvement of tourists' experience, and digital governance.

2 Related Work

Regarding the original authenticity and tourists' perception in cultural heritage tourism, existing research has been carried out from multiple directions such as sensory experience, landscape interference, spatial cognition, consumption behavior, and digital interaction. Xie et al. (2025) linked audio-visual information with cultural perception in daily heritage scenarios, indicating that the simultaneous presentation of auditory and visual cues significantly influences tourists' understanding of cultural atmosphere and heritage significance [1]. Wei et al. (2025) constructed a visual landscape interference model for heritage sites based on the AHP-fuzzy comprehensive evaluation method, explaining that external construction projects would weaken the completeness and authenticity of the heritage environment through landscape invasion [2]. Xiang and Qin (2024) discovered in the study of urban historical and cultural districts that there is tension between commercial packaging, space renewal, and historical expression in the global tourism context, and tourists' judgment of authenticity has obvious contextual dependence characteristics [3]. From these studies, we can clearly see that the original authenticity is no longer regarded as a single attribute of heritage, but a comprehensive perception result closely related to sensory cues, environmental intervention, and spatial expression methods.

In terms of the formation mechanism of tourist experience, Wani et al. (2024) started from local food consumption and believed that the degree of participation in local cuisine would affect tourists' perception of the authenticity of the destination culture and further influence their attitudes towards sustainable tourism [4]. Fan et al. (2025) analyzed AR-enhanced heritage landscapes using eye movement and heart rate variability indicators, finding that digital overlay technology can enhance attention investment and spatial perception, but its effect is constrained by content matching and display rhythm [5]. Wu et al. (2025) started from the Silk Road heritage tourism, pointing out that the risk of cultural distortion would be mediated through perceived value and satisfaction to influence tourists' loyalty, indicating that the original authenticity is not merely a cognitive judgment but also further transformed into behavioral results [6]. We believe that such research has pushed the discussion of original authenticity from the static identification of "whether it is real" to the dynamic analysis of "how it affects experience and behavior", and also indicates that there is a strong transmission relationship between tourists' perception and subsequent feedback.

With the continuous deepening of digital technology intervention, research has begun to focus on the construction of authenticity in virtual dissemination and mixed experiences. Shim et al. (2025) discussed the multi-terminal user experience in the social heritage metaverse, arguing that virtual heritage dissemination expands the boundary of cultural reach and also reshapes tourists' expectations of real scenes [7]. Kühne et al. (2025) reflected on the common binding relationship between authenticity, landscape, and tourism from a theoretical perspective, reminding researchers to avoid simply equating authenticity with physical preservation [8]. Literature [9] examined the quality of real experiences and tourists' behaviors in the integrated heritage scenarios of physical and digital entities through experimental design, proving that immersive interaction significantly changes tourists' perception paths of reality [9]. Overall, the existing results provide an important foundation for this study and also make us realize that the issue of original authenticity in cultural heritage tourism is extending from the protection of physical heritage to more complex research scenarios such as digital display, media dissemination, and mixed experiences. However, most literature still focuses on single-scenario explanations or single-class data analysis, and the joint modeling of multi-source information such as text, images, and behaviors is still insufficient. There is still room for further exploration of the authenticity construction mechanism in cultural heritage tourism and its dynamic impact on tourists' perception. The review of related studies is shown in Table 1.

Table 1: Review of Related Studies

References	Research Focus	Methodological Features	Main Implications
[1]	Audiovisual information and cultural perception	Coupled analysis of sensory information	Multisensory cues can influence tourists' perception of heritage culture
[2]	Landscape interference in buffer zones	AHP–fuzzy comprehensive evaluation	External visual intrusion can weaken perceived authenticity
[3]	Evaluation of authenticity in historic districts	Field investigation and evaluation of street-block scenes	Tension exists between commercialization and historical expression
[4]	Local cuisine and tourism perception	Analysis of tourist perception relationships	Local consumption experiences can strengthen cultural authenticity perception
[5]	Perception of AR heritage landscapes	Eye-tracking and HRV experiments	Digital augmentation can alter tourists' attention and experience
[6]	Risk of cultural distortion and loyalty	Mediation effect analysis	Authenticity affects satisfaction and behavioral outcomes
[7]	Heritage metaverse communication	Multi-terminal user experience research	Virtual communication is reshaping the understanding of authenticity
[8]	Theoretical reflection on authenticity	Conceptual analysis	Authenticity should not be simplified
[9]	Physical–digital integrated heritage experience	Experimental design	Tourists' authentic experience pathways become more complex in integrated scenarios

3 Research Methods

3.1 Collection and Preprocessing of Multi-source Heterogeneous Data for Cultural Heritage Tourism

To avoid the narrowness of information caused by relying solely on questionnaires or single comment texts, this paper expands the research object of original authenticity in cultural heritage tourism scenarios to a heterogeneous data set composed of text, images, behaviors, and structured evaluations. The data mainly comes from the official websites and digital museum pages of cultural heritage sites, online travel platform comment areas, social media image content, interaction logs of scenic area guide systems, and supplementary questionnaires on-site. The text data is used to extract the semantic expressions of tourists regarding historical atmosphere, display methods, and scene credibility, the image data is used to identify the spatial form, display environment, and visual symbol characteristics of the heritage space, the behavior logs are used to depict the duration of stay, click frequency, path switching, and navigation preferences, and the structured questionnaire is used to provide perception labels and verification basis. To enhance data comparability, three types of constraints - source credibility, content completeness, and time proximity - are introduced during the collection stage to screen the original samples, and the sample quality score is defined as:

$$\mathrm{Q_i} = \alpha \mathrm{C_i} + \beta \mathrm{I_i} + \gamma e^{-|\Delta t_i|/T}, \quad \alpha + \beta + \gamma = 1 \tag{1}$$

where, C_i is the source credibility, I_i is the content completeness, Δt_i is the time difference between the sample time and the on-site sampling window, and T is the time decay constant.

Only samples with $Q_i \geq 0.72$ are retained for the subsequent processing.

After the screening is completed, this paper conducts targeted preprocessing for different modalities of data. The text part deletes advertisement phrases, repetitive comments, and invalid symbols, and combines the cultural heritage context to construct an original authenticity vocabulary list, mapping the high-frequency semantics such as history, ritual, craftsmanship, space, and restoration in a standardized manner. The image part uniformly resizes to 224×224, eliminates overexposed, severely occluded, and missing sample images, and then retains the core display units of the heritage through target area cropping. The behavior log part is segmented by access sessions, and indicators such as duration of stay, page jump count, and trigger frequency of explanations are standardized. To achieve unified expression of multi-modal data at the sample level, a heterogeneous feature concatenation vector is constructed:

$$\mathbf{H}_i = [\eta_t \mathbf{T}_i \parallel \eta_v \mathbf{V}_i \parallel \eta_b \mathbf{B}_i \parallel \eta_s \mathbf{S}_i] \tag{2}$$

where, T_i, V_i, B_i, S_i represent the features of text, image, behavior, and structured evaluation, η_t, η_v, η_b, η_s are modal weight coefficients, and // represents vector concatenation. Through this processing, the originally scattered heterogeneous information is transformed into a unified sample representation that can be input into subsequent models, laying the data foundation for the extraction of original authenticity features and the modeling of tourists' perception.

3.2 Construction of the Original Authenticity Feature System for Cultural Heritage Scenarios

In the context of cultural heritage tourism, original authenticity is not a single indicator that can fully describe the attribute. It is reflected not only in the preservation status of the heritage itself and the accuracy of historical information transmission, but also in the real experiences formed in the display environment, narrative methods, and interactions with tourists. To avoid simply understanding original authenticity as a static preservation level, this paper combines the actual characteristics of cultural heritage tourism scenarios and decomposes it into four primary dimensions: heritage authenticity of the entity, spatial environment authenticity, cultural narrative authenticity, and experience interaction authenticity. Further, it is refined to form a multi-level indicator system. Among them, the authenticity of the heritage entity mainly reflects the coordination of architectural form, material texture, retention degree of components, and repair traces; spatial environment authenticity is used to describe the consistency of the surrounding landscape, visual interference, commercial packaging intensity, and historical atmosphere of the heritage; cultural narrative authenticity emphasizes the degree of alignment between the explanation content, display text, and guide information and local historical culture; experience interaction authenticity focuses on the sense of credibility and immersion formed by tourists during visits, viewing, triggering of guided tours, and immersive experiences. Through hierarchical construction, the abstract judgment of original authenticity can be transformed into a structured feature set that is recognizable and calculable.

Considering that the scales of different indicators are different and the representation intensity of original authenticity varies, this paper first performs interval normalization processing on each feature:

$$z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j) + \varepsilon} \tag{3}$$

where, x_{ij} represents the original value of the i-th sample in the j-th indicator, z_{ij} is the normalized result, and ε is a small constant introduced to prevent the denominator from being zero. Subsequently, based on expert scoring and the degree of data dispersion, the comprehensive weights of each indicator are determined, and the original authenticity representation value of the sample is calculated:

$$A_i = \sum_{m=1}^{M} w_m z_{im} + \lambda \cdot \frac{1}{K}\sum_{k=1}^{K} r_{ik} \tag{4}$$

where, A_i is the comprehensive score of the original authenticity of the i-th sample, w_m is the weight of the m-th indicator, r_{ik} is the correction feedback of the i-th scene by the tourists in the k-th perception item, and λ is the perception correction coefficient. This formula links the objective features with the tourists' immediate cognition, reducing the deviation caused by solely relying on static scene indicators.

Based on the above ideas, this paper finally forms a feature system for the original authenticity of cultural heritage tourism scenarios. This system retains the attention to historical continuity and material integrity in heritage research, and also absorbs new perception factors in digital guides, interactive displays, and tourism communication environments, providing a unified input for subsequent tourists' perception calculation and correlation modeling. The original authenticity feature system is shown in Table 2.

Table 2: Feature System of Original Authenticity for Cultural Heritage Scenarios

Primary Dimension	Secondary Indicator	Indicator Meaning	Data Source	Weight
Authenticity of Heritage Entity	Preservation of architectural form	Degree of preservation of the main structure and original layout	Images, field records	0.14
Authenticity of Heritage Entity	Consistency of material texture	Degree of match between restoration materials and the original appearance	Images, expert annotations	0.11
Authenticity of Heritage Entity	Integrity of component continuity	Condition of component loss, replacement, and preservation	Images, archival materials	0.09
Authenticity of Spatial Environment	Degree of landscape coordination	Overall consistency between the surrounding environment and the heritage space	Images, GIS annotations	0.12
Authenticity of Spatial Environment	Intensity of commercial interference	Degree to which shop signs and modern facilities intrude into the scene	Images, text	0.08
Authenticity of Spatial Environment	Degree of restoration of historical atmosphere	Tourists' ability to identify the sense of historical period and local context in the scene	Review texts, questionnaires	0.10
Authenticity of Cultural Narrative	Accuracy of interpretive content	Degree of alignment between guiding information and historical facts	Guide texts, expert verification	0.13
Authenticity of Cultural Narrative	Relevance to local culture	Degree of coupling between displayed content and local traditional culture	Text, questionnaires	0.09
Authenticity of Experiential Interaction	Perceived credibility of interactive experience	Whether interactive devices and digital content weaken the sense of authenticity	Behavioral logs, questionnaires	0.08
Authenticity of Experiential Interaction	Adaptability of immersive participation	Whether tourists' participation modes are coordinated with the heritage scene	Behavioral logs, review texts	0.06

3.3 Description of the Problem of Association Modeling between Visitor Perception Computing and Authenticity

In the context of cultural heritage tourism, visitor perception is not a simple mechanical reception of the objective environment; rather, it is a comprehensive judgment formed through the combined effects of information exposure, spatial stay, emotional response, and individual experience. The reason why different evaluation results can arise in the same heritage scene often does not lie in the fundamental changes of the heritage itself, but rather in the differences in the focus of tourists' attention to scene cues, understanding paths, and experience rhythms. Therefore, if the original authenticity research merely stops at the level of extracting objective features, it is difficult to fully explain why there are varying degrees of strength of the sense of authenticity, different levels of satisfaction, and differentiated subsequent communication intentions among tourists. Based on this understanding, this paper defines visitor perception computing as a potential representation learning process driven by multi-modal input, and places it within a unified association modeling framework with the original authenticity feature system to reveal the structural relationship between the two.

Specifically, let the observation set of the i-th cultural heritage tourism sample be:

$$\mathcal{O}_i = \{t_i, v_i, b_i, c_i\} \tag{5}$$

Among them, ti represents the tourist comments and guide text information, v_i represents the image information of the heritage scene, bi represents the behavioral characteristics such as staying, clicking, and path switching, and c_i represents the contextual variables such as the time period, visit density, and scene type that the tourist is in. The goal of the tourist perception calculation is to extract low-dimensional latent vectors that can represent the realism, immersion, credibility, and cultural identification tendency from the observation set. This paper defines the tourist perception representation as:

$$\mathbf{p}_i = \phi(\mathbf{W}_t\mathbf{t}_i + \mathbf{W}_v\mathbf{v}_i + \mathbf{W}_b\mathbf{b}_i + \mathbf{W}_c\mathbf{c}_i + \mathbf{b}_0) \tag{6}$$

Here, p_i is the latent vector of tourist perception, $\phi(\cdot)$ is a nonlinear mapping function, Wt, Wv, Wb, Wc are parameter matrices corresponding to different modalities, and b0 is the bias term. The significance of this representation lies in that it no longer simply understands tourist perception as a single rating, but rather regards it as a comprehensive representation jointly projected by multiple types of information.

Based on this, this paper defines the original authenticity feature vector constructed in Section 3.2 as ai, and further characterizes the coupling relationship between tourist perception and the original authenticity. Considering that the two are not simply linearly corresponding, this paper uses a bilinear correlation function to describe the matching degree between them:

$$\mathrm{r_i} = \mathrm{p_i^T M a_i} \tag{7}$$

Among them, r_i represents the association strength of the i-th sample, and M is the mapping matrix between the perceptual space and the authenticity space. If r_i is large, it indicates a high consistency between the tourists' perception and the original authenticity features; conversely, it suggests that although the scene has a certain basis of authenticity, it has not effectively transformed into a real tourist-perceptible experience, or there are problems such as display deviation, information interference, and experience mismatch.

To enable the model to not only reflect the perception results but also constrain the association direction, this paper further assigns the comprehensive evaluation labels of tourists as y_i, and defines the predicted value $\hat{y}_i$ as:

$$\hat{y}_i = \psi(r_i + \omega d_i + \xi s_i) \tag{8}$$

where d_i represents the intensity of digital display intervention, s_i represents the level of scene service support, ω and ξ are adjustment coefficients, and ψ(·) is the output mapping function. This equation indicates that the final perception result formed by tourists is not only affected by the matching degree between the original authenticity and the perceptual representation, but also corrected by the digital display method and the service environment. To ensure the stability of model training, this paper adopts the following objective function:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{g}_i)^2 + \lambda_1 \|M\|_F^2 + \lambda_2 \sum_{i=1}^{N} \| p_i - a_i\|_2^2 \tag{9}$$

where the first term is used to constrain the error between the predicted results and the true labels, the second term is used to suppress overfitting of the mapping matrix, and the third term is used to enhance the structural alignment ability between the tourists' perceptual space and the original authenticity space.

Through the above definitions, this paper transforms the tourists' perception problem in cultural heritage tourism into a representation learning and association prediction problem driven by multi-modal features, and also advances the original authenticity research from static description to a computable, comparable, and optimizable analysis level. This problem description provides a clear variable basis and mathematical expression for the subsequent design of tourists' perception analysis models and the implementation of intelligent computing frameworks.

3.4 Design of Tourist Perception Analysis Model Based on Multi-modal Feature Fusion

To more accurately depict the differences in tourists' perception of the original authenticity in cultural heritage tourism scenarios, based on the aforementioned data processing and feature construction, this paper further designs a tourist perception analysis model based on multi-modal feature fusion. Different from the traditional analysis methods based on single text ratings or questionnaire statistics, this model takes tourist comment texts, heritage scene images, and behavior sequences as parallel inputs. Through hierarchical encoding, cross-modal interaction and fusion representation, it realizes the joint identification of tourists' authenticity perception, immersion perception, and cultural identity perception. The model adopts a five-layer structure, including multi-source input layer, modality encoding layer, cross-modal interaction layer, fusion representation layer, and perception output layer, as shown in Figure 1.

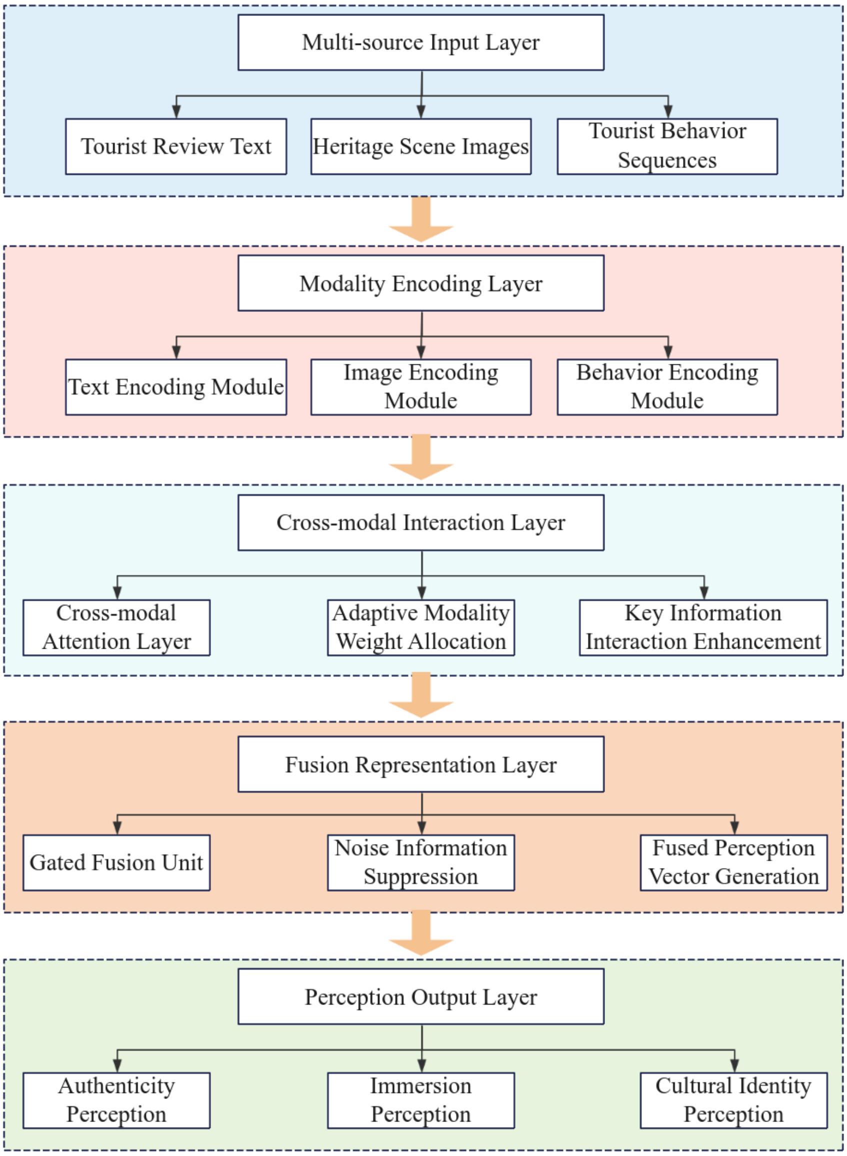

Figure 1: Hierarchical structure diagram of the tourist perception analysis model based on multimodal feature fusion

In the first layer, the multi-source input layer, the model receives three types of core data: one is tourist comment text, used to extract cultural expression, authenticity judgment, and emotional tendency; one is heritage scene images, used to identify spatial form, material texture, and visual interference information; and the other is tourist behavior sequence, used to reflect interaction features such as stay, jump, guide triggering, and browsing path. After entering the second layer of modal encoding, the text, images, and behavior data are respectively mapped to a unified feature space. Let the input samples of the three types be $X_i^{(t)}, X_i^{(v)}, X_i^{(b)}$, then the encoding result is expressed as:

$$E_i^{(m)} = \delta\left(P_m X_i^{(m)} + a_m\right), \qquad m \in \{t, v, b\} \tag{10}$$

Here, P_m is the projection matrix of the mth modal type, am is the bias vector, and δ(·) is the nonlinear activation function. The purpose of this formula is to compress data of different dimensions and distributions into a comparable latent representation space.

In the third layer of cross-modal interaction, the model does not directly concatenate the three types of features, but first measures the consistency between each modal, to determine whether the textual expression, visual scene, and behavioral response are clustered towards the same perceptual direction. For this purpose, this paper introduces the modal synergy coefficient:

$$C_i = \frac{1}{3}\left[\begin{array}{c}\cos\left(E_i^{(t)}, E_i^{(v)}\right) + \cos\left(E_i^{(t)}, E_i^{(b)}\right) + \\ \cos\left(E_i^{(v)}, E_i^{(b)}\right)\end{array}\right] \tag{11}$$

Among them, C_i is used to measure the degree of synergy among the i-th sample in the three modalities. When C_i is high, it indicates that the textual expression, image features, and behavioral responses of the tourists have a strong consistency; conversely, it suggests that the tourists' perception may be interfered by noise information or local stimuli.

In the fourth fusion representation layer, the model maps the encoding results and the synergy information into a unified perception representation vector. The fusion process is defined as:

$$\mathrm{F_i} = \mathrm{LN}\left(\mathrm{W_f}\left[\mu_t \mathrm{E_i^{(t)}}; \mu_v \mathrm{E_i^{(v)}}; \mu_b \mathrm{E_i^{(b)}}\right] + \mathrm{r}\,\mathrm{C_i} + \mathrm{b_f}\right) \tag{12}$$

where, W_f is the fusion matrix, μ_t, μ_v, μ_b respectively represent the contribution coefficients of the three modalities, r is the synergy enhancement factor, bf is the bias term, and LN(·) represents the layer normalization operation. Compared with the simple concatenation method, this design can retain the consistency information between modalities during the feature combination process, thereby improving the stability and interpretability of the fusion representation.

In the fifth perception output layer, the model identifies the tourists' perception based on the fusion vector F_i and outputs results such as authenticity perception, immersion perception, and cultural identity perception. To balance the continuous intensity and category division, this paper adopts a piecewise regression method to generate the final score:

$$\mathrm{G_i} = \theta_1 \mathrm{g_i^{(a)}} + \theta_2 \mathrm{g_i^{(im)}} + \theta_3 \mathrm{g_i^{(c)}} \tag{13}$$

where, $\mathrm{g_i^{(a)}}, \mathrm{g_i^{(im)}}, \mathrm{g_i^{(c)}}$ respectively represent the output values of the i-th sample in authenticity perception, immersion perception, and cultural identity perception, and θ_1, θ_2, θ_3 are normalization weights. Through this structure, the model not only can provide the tourists' perception results, but also can provide hierarchical explanatory basis for the subsequent original authenticity association analysis. The functions of each layer of the model and the main tasks are shown in Table 3.

Table 3: Hierarchical structure and function description of the tourist perception analysis model

Level	Component module	Input content	Output result	Main function
Level 1	Multi-source input layer	Comment text, scene image, behavior sequence	Original heterogeneous data	Builds the data foundation for tourist perception analysis
Level 2	Modal encoding layer	Three types of original input	Text features, image features, behavior features	Completes the unified representation of heterogeneous data
Level 3	Intermodal interaction layer	Encoded features	Modal synergy information	Strengthenes the association expression between different modalities
Level 4	Fusion representation layer	Encoded features and synergy coefficients	Fusion perception vector	Forms a unified tourist perception representation
Level 5	Perception output layer	Fusion perception vector	Authenticity perception, immersion perception, cultural identity perception	Output tourist perception analysis results

3.5 Intelligent computing framework and algorithm flow for original authenticity construction and tourist perception relationship

After extracting the original authenticity features and representing the tourists' perception, how to place the two types of information in the same analysis system for linked calculation becomes the key of the method design in this paper. Considering the characteristics of cultural heritage tourism scenarios such as multiple scene elements, fast perception feedback, and phased relationship changes, this paper constructs an intelligent computing framework for original authenticity construction and tourist perception relationship. This framework takes the heritage scene unit as the basic node, uses the tourists' perception feedback as the dynamic response signal, and through the three links of relationship matching, state update, and result output, forms a cyclic iterative analysis link. The operation flow of the framework is shown in Figure 2.

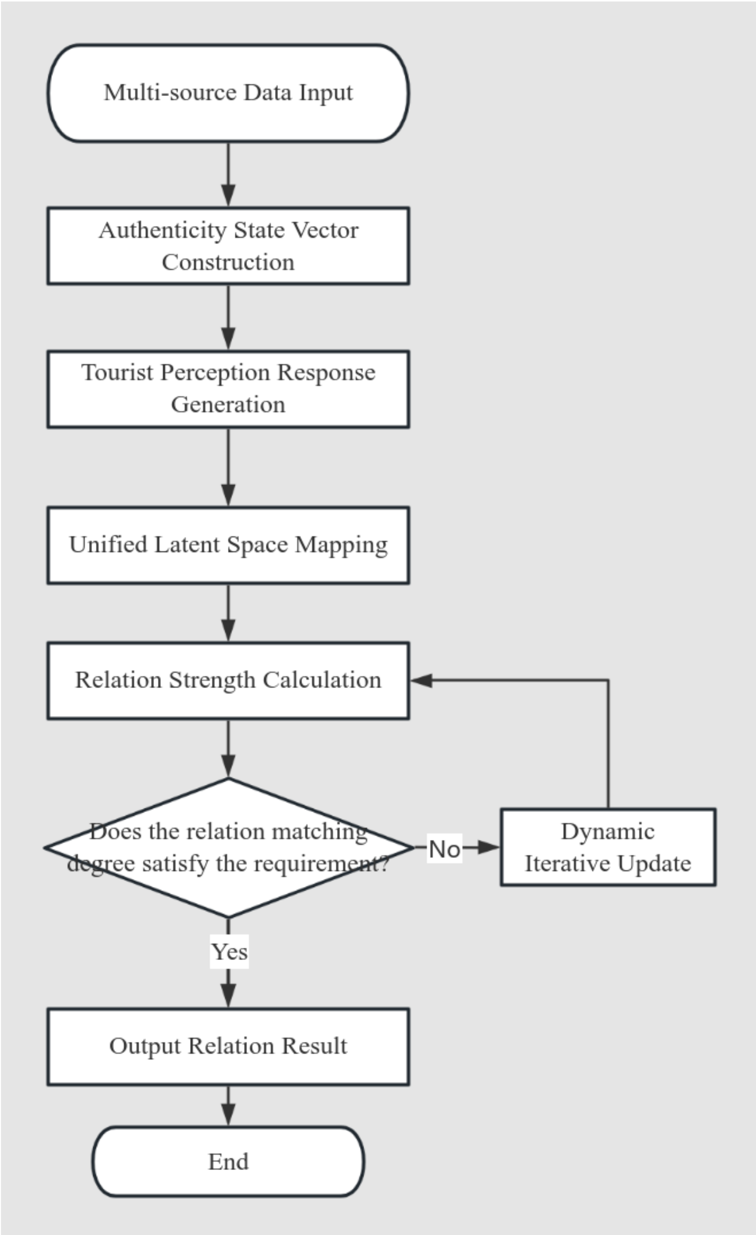

Figure 2: Intelligent computing flowchart of the relationship between authenticity construction and tourists' perception

Firstly, for the i-th heritage scene sample, this paper represents its original authenticity status as:

$$o_i = [r_i, e_i, n_i] \tag{14}$$

Among them, "r_i" represents the characteristic of the heritage itself, "e_i" represents the characteristics of the environment and display, and "ni" represents the characteristics of narration and interaction. On the side of tourists' perception, it is recorded as:

$$u_i = [s_i, m_i, c_i] \tag{15}$$

Among them, s_i represents the authenticity perception response, m_i represents the immersion perception response, and c_i represents the cultural identity response. After entering the relationship calculation module, both are first mapped to the same latent space through a transformation matrix to reduce the influence of dimensional differences and modal deviations.

To depict the driving intensity of the original authenticity construction on tourists' perception, this paper defines the relationship response function as:

$$\zeta_i = \sigma(o_i W_r u_i^T + b_r) \tag{16}$$

where, ζ_i represents the authenticity-perception relationship strength of the i-th scene sample, W_r is the relationship mapping parameter, b_r is the bias term, and $\sigma(\cdot)$ is the compression function. If ζ_i is high, it indicates that the constructed authenticity factors in the scene can be effectively transformed into perceptible and recognizable perception results by tourists; if ζ_i is low, it suggests that although tourists have come into contact with the heritage scene, the formation of perception has experienced attenuation, deviation or mismatch.

Considering that cultural heritage tourism experiences have dynamic update characteristics, this paper further introduces an iterative correction mechanism to update the scene state in a phased manner:

$$o_i^{(t+1)} = o_i^{(t)} + \lambda(u_i^{(t)} - \zeta_i^{(t)} o_i^{(t)}) \tag{17}$$

where, λ is the update step size. The meaning of this formula is that if the tourists' perception feedback is consistent with the current scene authenticity expression, the system maintains the original state; if the gap between the two is large, the scene representation is corrected through the feedback term to make the subsequent analysis closer to the real tourism situation. To evaluate the overall relationship modeling effect, this paper adopts a comprehensive optimization objective:

$$J = \frac{1}{N}\sum_{i=1}^{N}[\,\omega_1(1-\zeta_i) + \omega_2\|o_i - u_i\|_2 + \omega_3\Pi_i] \tag{18}$$

where, Π_i represents the deviation penalty term for the scene affected by external interference, ω_1, ω_2, ω_3 are weight coefficients. The smaller the optimization objective, the more stable the coupling between the original authenticity construction and tourists' perception.

Based on the above design, this paper forms an intelligent computing process consisting of data input, relationship mapping, dynamic correction and result output, which not only can identify the matching degree between the original authenticity construction and tourists' perception, but also can locate the specific links where authenticity expression is insufficient or perception response is weak, providing a computational basis for subsequent experimental analysis and scene optimization.

4 Experimental Results and Analysis

4.1 Experimental Settings and Evaluation Metrics

To verify the applicability of the model in cultural heritage tourism scenarios, the experiment takes the samples after cross-modal alignment as the analysis objects, and divides the training set, validation set and test set according to the scene units. Considering that the identification of tourists' perception and the analysis of original authenticity correlation analysis have both classification and prediction characteristics, this paper conducts comparative tests in a unified experimental environment. The hardware platform is an Intel Xeon processor, 32 GB memory and RTX 4060 graphics card. The batch size for model training is set to 32, the initial learning rate is set to 0.001, and the number of iteration rounds is set to 80. To reduce the interference caused by accidental errors, all experiments are repeated 5 times and the average value is taken. The evaluation metrics selected to measure the model's recognition performance are Accuracy, Precision, Recall and F1 value. The data set division situation is shown in Table 4.

Table 4: Data Set Division Situation

Dataset	Simple Size
Training Set	6048
Validation Set	864
Test Set	1728
Total	8640

4.2 Comparison of Original Authenticity Feature Recognition Results and Model Performance

After completing data preprocessing, the construction of the original authenticity feature system, and multimodal perception modeling, this paper further validates the original authenticity feature recognition effect and conducts a comparison with common baseline methods. Considering that tourists' judgments of authenticity in cultural heritage tourism scenarios are often influenced by text expression, visual information, and behavioral feedback simultaneously, single-modal models are prone to problems such as feature omission， and result fluctuations. Therefore， in this section， the text feature model， image feature model， behavior sequence model， early stitching fusion model， and the multi-modal fusion model proposed in this paper are selected for comparison to test the comprehensive performance of the established method in original authenticity feature recognition. The evaluation indicators are Accuracy， Precision， Recall， and F1 value. The relevant results are shown in Table 5.

Table 5: Comparison of Original Authenticity Feature Recognition Results of Different Models

Model	Accuracy	Precision	Recall	F1
Text feature model	0.831	0.824	0.817	0.820
Image feature model	0.846	0.839	0.832	0.835
Behavior sequence model	0.812	0.805	0.798	0.801
Early stitching fusion model	0.879	0.871	0.866	0.868
This paper model	0.918	0.911	0.906	0.908

From Table 5, it can be seen that the model in this paper achieves the best results in all four indicators, indicating that the multimodal feature fusion mechanism can more fully integrate text, image and behavioral information, and shift the original authenticity feature recognition from single-point judgment to comprehensive representation. Especially in complex cultural heritage scenarios, the model can not only identify the explicit clues such as the heritage itself and the environmental features, but also capture the implicit features such as the historical atmosphere, narrative credibility and experience coordination in tourist comments, thereby improving the stability and completeness of the recognition. Overall, the Accuracy of the model in this paper reaches 0.918, which is 0.039 higher than the earlier splicing and fusion model and 0.072 higher than the image feature model; the F1 value reaches 0.908, which is 0.088 higher than the text feature model, indicating that this method has better comprehensive performance and application value in the original authenticity feature recognition task.

4.3 Analysis of Tourist Perception Difference Recognition Results and Key Influencing Factors

Based on the original authenticity feature recognition, this paper further examines the recognition effect of tourist perception differences and analyzes the main factors influencing the formation of perception. To enhance the interpretability of the results, this paper divides tourist perception into three dimensions: authenticity perception, immersion perception and cultural identity perception, and makes a comprehensive judgment based on the model output results and feature contribution degrees. Overall, there are certain differences in the recognition results of different dimensions, among which the authenticity perception identification effect is the most stable, immersion perception is second, and cultural identity perception is relatively more affected by individual experience and scene context differences. The relevant recognition results and key influencing factors are shown in Table 6.

Table 6: Analysis of Tourist Perception Difference Recognition Results and Key Influencing Factors

Perception Dimension	Accuracy	F1	Key Influencing Factors	Impact Weight
Authenticity Perception	0.912	0.904	Conservation Degree of Heritage Elements	0.31
Immersion Perception	0.893	0.887	Spatial Environment Coordination	0.28
Cultural Identity Perception	0.876	0.869	Cultural Narrative Relevance	0.34
Overall Perception	0.901	0.893	Interactivity Experience Credibility	0.27

It can be seen from Table 6 that the Accuracy of authenticity perception reaches 0.912, indicating that tourists' judgments on the form, material style and restoration coordination of the heritage are relatively concentrated, and the model can more easily capture the stable features. The recognition result of immersion perception is slightly lower, indicating that tourists' feelings about the spatial atmosphere, visual continuity and interaction rhythm during the visit are more easily affected by scene changes. The Accuracy of cultural identity perception is 0.876, which is the lowest among the three dimensions. This means that whether tourists have a deeper cultural understanding depends not only on the display content itself, but also on their personal knowledge background and the narrative style on the spot. Overall, the three core factors influencing tourist perception differences are the conservation degree of the heritage elements, spatial environment coordination and cultural narrative relevance, among which the impact weight of cultural narrative relevance is the highest, reaching 0.34. This indicates that in cultural heritage tourism, whether tourists truly establish a strong perception identity largely depends on whether the scene expression can clearly and naturally present historical information.

4.4 Comparative Experiment Analysis of Different Cultural Heritage Types and Tourism Scenarios

To test the applicability of the model in different cultural heritage types and tourism scenarios, this paper selected four typical scenarios: historical and cultural districts, ancient architectural ruins, museum exhibitions, and intangible cultural heritage live display, and conducted comparative experiments, with Accuracy and F1 value as the main evaluation indicators. The model recognition results in different scenarios are shown in Figure 3.

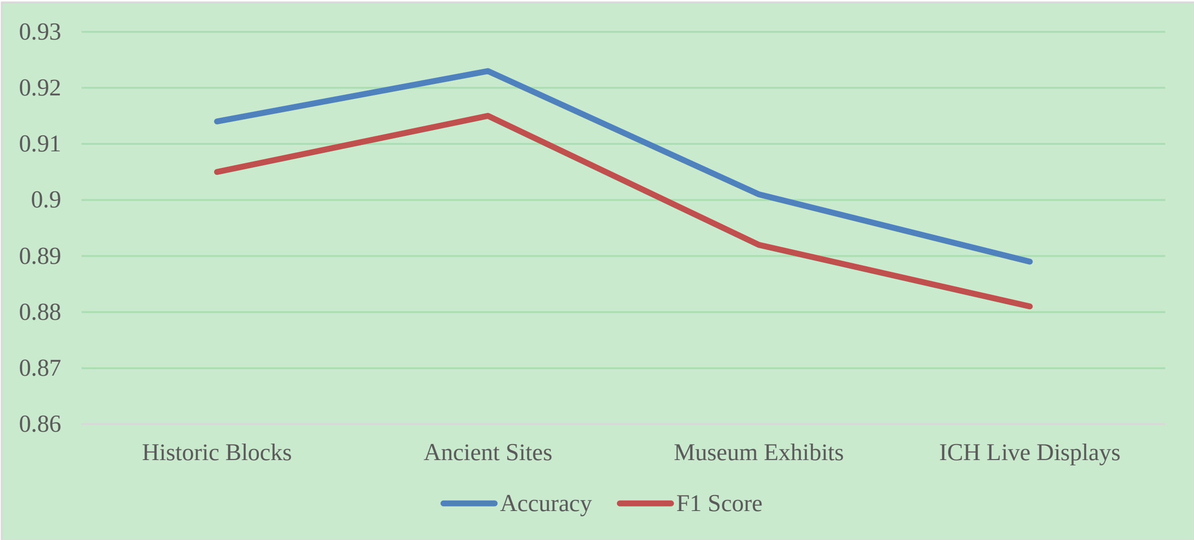

Figure 3: Line graph of model recognition results for different types of cultural heritage and in tourism scenarios

Overall, the model in this paper maintains a high recognition level in all four scenarios, but due to differences in scene stability, information density, and interaction methods, there are still certain fluctuations between different scenarios. Among them, the recognition effect of the ancient architectural site scenario is the best, with an Accuracy of 0.923 and an F1 value of 0.915, indicating that the authenticity clues of this type of scenario are relatively concentrated, and the model is more likely to complete stable recognition. The historical and cultural district scenario ranks second, with Accuracy and F1 reaching 0.914 and 0.905 respectively. In the museum exhibition scenario, tourists' perception is more dependent on the content of the explanation and the organization of the exhibition, and the model indicators slightly decrease. The recognition results of the dynamic display of intangible cultural heritage scenarios are relatively the lowest, with an Accuracy of 0.889 and an F1 of 0.881, indicating that after enhancing the interactivity, the differences in tourists' perception are more obvious. In summary, the mean Accuracy of the four scenarios reaches 0.907, and the mean F1 value is 0.898, with the maximum and minimum Accuracy differing by only 0.034, indicating that the model still has good adaptability and stability in different cultural heritage tourism scenarios.

4.5 Model Robustness Analysis

To test the stability of the model under perturbation conditions and its cross-scenario transfer ability, this paper sets up noise injection experiments and cross-scenario testing experiments. The noise proportion is gradually increased from 0 to 0.20 to observe the performance changes of the model under input disturbances; at the same time, some scene data is used as an external test set to evaluate the model's generalization ability. The curve of model robustness changes under noise perturbation is shown in Figure 4.

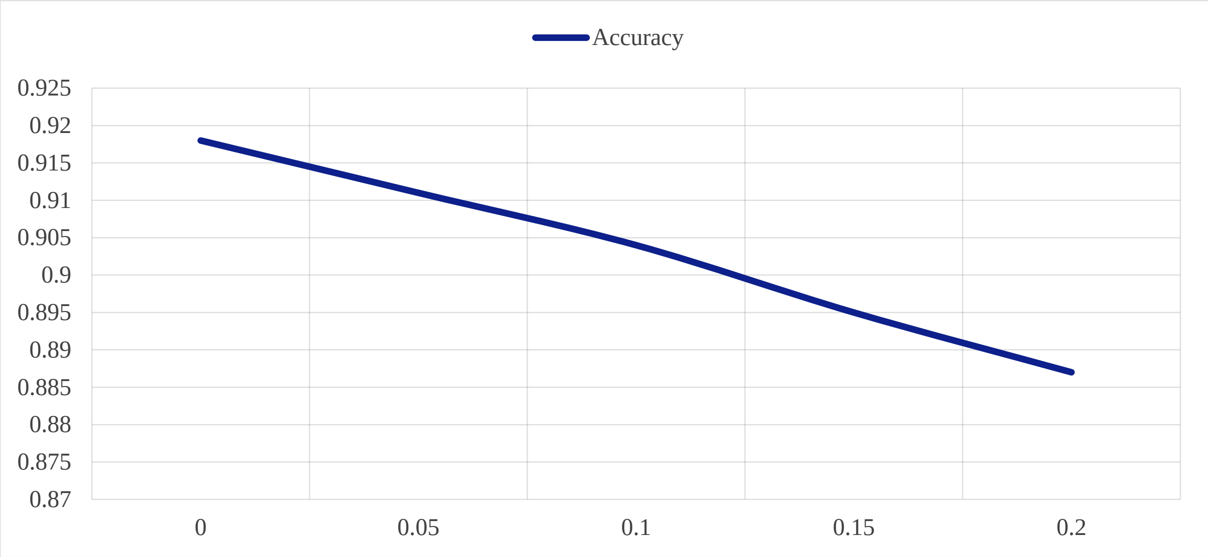

Figure 4: Curve diagram of model robustness change under noise disturbance

As the noise proportion increases, the model's Accuracy shows a slow downward trend, but the overall fluctuation is small; in cross-scenario tests, the model still maintains a high recognition level, indicating that the constructed multimodal fusion mechanism has good stability. In summary, when the noise proportion increases from 0 to 0.20, Accuracy decreases from 0.918 to 0.887, with a decrease of 0.031; the cross-scenario test Accuracy remains at 0.894, indicating that the model in this paper has strong robustness in complex cultural heritage tourism environments.

4.6 Statistical Significance Test and Result Validity Analysis

To determine whether the performance improvement of the model in this paper has statistical significance and to further verify the reliability of the experimental results, this paper conducts a significance test on the recognition results of the model and the text feature model, image feature model, and early splicing fusion model based on the aforementioned 5 repeated experiments. The test indicators are Accuracy and F1, and the paired t-test is used to compare the differences in results between different models. At the same time, the standard deviation and effect size are combined to judge the stability of the model output. The statistical test results are shown in Table 7.

Table 7: Statistical Significance Test Results of Different Models

Comparative Model	Mean Accuracy ± Standard Deviation	Mean F1 ± Standard Deviation	p-value	Effect Size (d)	Result
Text Feature Model	0.831 ± 0.009	0.820 ± 0.011	0.003	1.42	Significant
Image Feature Model	0.846 ± 0.008	0.835 ± 0.010	0.005	1.31	Significant
Early Concatenation Fusion Model	0.879 ± 0.007	0.868 ± 0.009	0.011	0.96	Significant
Proposed Model	0.918 ± 0.005	0.908 ± 0.006	—	—	—

From Table 7, it can be seen that the standard deviation of Accuracy and F1 of this paper model is within 0.006, indicating that the model has a small fluctuation in repeated experiments and a relatively stable output result. Compared with the three types of comparison models, the p-values of this paper model are all less than 0.05, where the p-value when compared with the text feature model is 0.003, and when compared with the early splicing fusion model is 0.011, both reaching significant levels. At the same time, the highest effect size reaches 1.42, and the lowest is 0.96, indicating that the advantage of this paper model is not caused by accidental fluctuations but has a strong practical effect. In summary, this paper model not only has better recognition accuracy but also has stable results and significant differences, and can provide more reliable experimental support for the subsequent reconstruction of original authenticity and the analysis of tourist perception relationship.

4.7 Abandonment Experiment and Key Module Contribution Analysis

To verify the specific contribution of each component of the model to the recognition results, this paper conducts an abandonment experiment based on the complete model, removing the text encoding module, image encoding module, behavior encoding module, and cross-modal interaction layer, and compares the performance changes of different variant models. The abandonment experiment results are shown in Table 8. Overall, after removing each key module, the Accuracy and F1 of the model show a different degree of decline, indicating that the multimodal analysis framework constructed in this paper has strong overall synergy.

Table 8: Abandonment Experiment Results and Key Module Contribution Analysis

Model Variant	Accuracy	F1	Performance Degradation
Full Model	0.918	0.908	—
Without Text Encoding Module	0.891	0.882	0.027
Without Image Encoding Module	0.896	0.887	0.022
Without Behavior Encoding Module	0.902	0.893	0.016
Without Cross-Modal Interaction Layer	0.884	0.875	0.034

As shown in Table 8, the performance of the model decreased most significantly after removing the cross-modal interaction layer. The Accuracy dropped from 0.918 to 0.884, a decrease of 0.034, indicating that this module played a crucial role in coordinating multi-source feature relationships and enhancing joint expression. After removing the text encoding module, the model's Accuracy dropped to 0.891, suggesting that tourist comments and semantic information provide important support for authenticity judgment. The image encoding module and the behavior encoding module also have stable contributions. The removal of the behavior module resulted in a relatively smaller decrease, still reaching 0.016. Overall, the complete model maintained the best performance in Accuracy and F1, indicating that the modules do not simply add up but jointly enhance the visitor's knowledge recognition effect in the cultural heritage tourism scenario.

5 Discussion

The experimental results of this study indicate that the authenticity of cultural heritage tourism is not a one-way concept solely determined by the preservation status of the heritage itself. We tend to view it as a comprehensive outcome resulting from the combined effects of scene expression, the way tourists interact with the environment, and the individual understanding process. From the previous results, the recognition effect in ancient architectural sites and historical cultural districts is relatively high, which indicates that when the form and features of the heritage are clear and the spatial clues are concentrated, tourists are more likely to form a stable perception of authenticity. We also noticed that in scenarios with dynamic display of intangible cultural heritage and interactive elements, the perception results fluctuate more significantly, reflecting that tourists' judgment of authenticity is not solely dependent on the physical heritage itself, but is also influenced by the display method, participation intensity, and the continuous impact of cultural interpretation paths.

From the methodological perspective, after incorporating text, image, and behavioral information into a unified analysis framework, the performance of the model has significantly improved compared to single-modal methods. This indicates that if the research on cultural heritage tourism still remains at the level of interpreting single-source data, it is often difficult to restore the tourists' actual perception process. We believe that, especially in the context of continuous involvement of digital navigation, immersive experiences, and social communication, what tourists are confronted with is no longer a simple heritage space, but a composite scene organized by multiple media. Therefore, we believe that the construction of original authenticity should no longer be simply understood as a passive reproduction of historical states, but should be regarded as an experience result formed through continuous negotiation between the scene, technology, and cognition.

From the application perspective, the results of this study have certain implications for the optimization of cultural heritage tourism displays. We believe that for managers, enhancing tourists' perception of authenticity does not mean simply increasing digital technology or strengthening commercial packaging. Instead, they should pay more attention to whether the heritage itself, the environmental atmosphere, and the cultural narrative are coordinated. If the display design overly pursues visual stimulation while weakening the coherence of historical clues, tourists may temporarily be interested but may not form a deeper cultural identity. On the contrary, the research results of this study suggest that those cultural heritage spaces that can maintain the integrity of scene logic, natural narrative expression, restrained interaction methods, and effective interaction often receive more stable perception feedback.

Of course, this study still has certain limitations. The current research is mainly based on existing samples for modeling analysis. Our discussions on individual background differences of tourists, the influence of long-term memory, and cross-regional cultural understanding biases are still not sufficient. Although some non-structured data have been filtered and cleaned, we also realize that different platforms' expression habits may still bring potential disturbances. This also means that the relationship between original authenticity in cultural heritage tourism and tourists' perception still has room for further deepening. Future research needs to further advance in sample expansion, cross-regional comparison, and dynamic perception tracking.

6 Conclusion

This paper focuses on the construction of authenticity in cultural heritage tourism and the issue of tourists' perception, proposing an analysis approach driven by multi-source heterogeneous data, and constructing an authenticity feature system, a tourists' perception calculation model, and a relationship intelligent computing framework. It realizes the joint analysis of authenticity expression and tourists' perception feedback. The results show that the multimodal fusion model has a strong recognition ability in complex scenarios, with an Accuracy of 0.918 and a F1 value of 0.908; the authenticity perception dimension has the best recognition effect, with an Accuracy of 0.912; the overall Accuracy of the model in different scenarios reaches 0.907. The ablation experiments show that when the cross-modal interaction layer is removed, the Accuracy drops to 0.884, indicating that this module plays a key role in improving the joint expression ability. Overall, this research promotes the transformation of the authenticity issue in cultural heritage tourism from empirical judgment to computational analysis, and lays a foundation for subsequent more detailed scene optimization and intelligent management research.

Funding

This study was supported by the Project of the 2024 Henan Xing Culture Project: Cultural Research Special Project (2024XWH021): Historical and Cultural Research of the Central Axis in Kaifeng City

References

[1] Xie J, Kang J, Zhao Z,et al.Effects of audiovisual information and its perception on the sense of culture in everyday heritage[J]. Applied Acoustics, 2025, 239. DOI:10.1016/j.apacoust.2025.110795.

[2] Wei W, Zhenrui M A, Shizhen X.Construction of a Linear Engineering Visual Landscape Interference Model for the Buffer Zone of the Libo Heritage Site Based on the AHP-Fuzzy Comprehensive Evaluation Method:A Case Study of the Guiyang-Nanning High-Speed Railway[J].Journal of Landscape Research, 2025, 17(5):12-19.

[3] Xiang Q, Qin W.Evaluation of authentic perception of urban historical and cultural blocks under global tourism[J].Journal of Central South University of Forestry & Technology, 2024, 44(11):165-173.DOI:10.14067/j.cnki.1673-923x.2024.11.017.

[4] Wani M D, Dada Z A, Shah S A.Can consumption of local food contribute to sustainable tourism? Evidence from the perception of domestic tourists[J].The International Journal of Sustainable Development and World Ecology, 2024(1/2):31. DOI:10.1080/13504509.2023.2262950.

[5] Fan W, Li C, Gao S,et al.How AR-Enhanced Cultural Heritage Landscapes Influence Perception in Rural Tourism Spaces: Evidence from Eye Tracking and HRV[J]. Sustainability (2071-1050), 2025, 17(23).DOI:10.3390/su172310575.

[6] Wu K, Su L, Zhang S,et al.Cultural distortion risk and tourist loyalty at silk road heritage: The mediating roles of perceived value and satisfaction[J].PLoS ONE, 2025, 20(11). DOI:10.1371/journal.pone.0335476.

[7] Shim H, Oh K T, Shi C K.Exploring the Social Heritage Metaverse for Virtual Heritage Communication from Multi End-user Centered Experiences[J].ACM Journal on Computing and Cultural Heritage, 2025, 18(3).DOI:10.1145/3736225.

[8] Kühne, Olaf, Vogler R, Berr K.Authenticity, Landscape and Tourism: Criticism of a Common Coalescence[J].Tourism (13327461), 2025, 73(1).DOI:10.37741/t.73.1.6.

[9] ükrü Frat ifti, Izel B.Exploring relations among authentic tourism experience, experience quality, and tourist behaviours in phygital heritage with experimental design[J].Journal of Destination Marketing & Management, 2024, 31(000):18. DOI:10.1016/j.jdmm.2023.100848.

[10] Li L, Li S.Do Tourists Really Care about Authenticity? A Study on Tourists' Perceptions

of Nature and Culture Authenticity[J].Sustainability, 2022, 14.DOI:10.3390/su14052510.

[11] Li X, Wang C.Understanding the relationship between tourists' perceptions of the authenticity of traditional village cultural landscapes and behavioural intentions, mediated by memorable tourism experiences and place attachment[J].Asia Pacific Journal of Tourism Research, 2023, 28:254 - 273.DOI:10.1080/10941665.2023.2217959.

[12] Garau-Vadell J B, Orfila-Sintes F, Batle J.The quest for authenticity and peer-to-peer tourism experiences[J].Journal of Hospitality and Tourism Management, 2021, 47(10): 210-216.DOI:10.1016/j.jhtm.2021.03.011.

[13] Chirieleison C, Montrone A, Scrucca L.Destination labels for historic villages: The impact on perception, experience, and satisfaction[J].Tourism and Hospitality Research, 2021, 22:164 - 179.DOI:10.1177/14673584211020788.

[14] Chen H, Jiao Y, Li X,et al.Family tourism: Interpersonal interaction, existential authenticity and quality of tourist experience:[J].Journal of Vacation Marketing, 2022, 28(1):82-94.DOI:10.1177/13567667211022407.

[15] Wang Y, Wang H.Social Entrepreneurship Promoting the Inheritance of Ceramic Intangible Cultural Heritage: A Value Co-Creation Perspective[J].Journal of Ceramics, 2025, 46(1):195-203.DOI:10.13957/j.cnki.tcxb.2025.01.018.

[16] Wenjie L.Research on Recreational Intention and Experience in Tourism-Take Nantou Ancient City as an Example[J].Psychology Research, 2022, 12(7): 450-479. DOI:10.17265/2159-5542/2022.07.003.

[17] Lan F, Huang Q, Zeng L,et al.Tourism Experience and Construction of Personalized Smart Tourism Program Under Tourist Psychology.[J].Frontiers in psychology, 2021, 12:691183.DOI:10.3389/fpsyg.2021.691183.

[18] SunYingsunyinggufe@outlook.comOuQuanfengSchool of Culture Tourism and Geography, Guangdong University of Finance and Economics, Guangzhou, China. Research on the traditional zoning, evolution, and integrated conservation of village cultural landscapes based on "production-living-ecology spaces"–A case study of villages in Meicheng, Guangdong, China[J]. Open Geosciences, 2021, 13(1):1303-1317. DOI:10.1515/geo-2020-0279.

[19] Salet X.The search for the truest of authenticities: Online travel stories and their depiction of the authentic in the platform economy[J].Annals of Tourism Research, 2021, 88(6):103175.DOI:10.1016/j.annals.2021.103175.

[20] Zhao Y, Zhan Q, Du G,et al.The effects of involvement, authenticity, and destination image on tourist satisfaction in the context of Chinese ancient village tourism[J].Journal of Hospitality and Tourism Management, 2024, 60:51-62. DOI:10.1016/j.jhtm.2024. 06.008.

[21] Zuo Y, Lan T, Liu S,et al.The Post-Effects of the Authenticity of Rural Intangible Cultural Heritage and Tourists' Engagement[J].Behavioral Sciences, 2024, 14(4):17. DOI:10. 3390/bs14040302.

[22] Shen Y, Schwab K, Tsutsumi A,et al.A Comparative Study of VR and 2D Tourism Videos: A Thematic Analysis of Virtual Tourism Experiences Among Generation Z[J].Tourism & Hospitality (2673-5768), 2025, 6(4).DOI:10.3390/tourhosp6040200.

[23] Hua Y, Ding L, Dong H,et al.Influence of User-Generated Content (UGC) in Social Media on the Intangible Cultural Heritage Preservation of Gen Z Tourists in the Digital Economy Era[J].International Journal of Tourism Research, 2024, 26(5).DOI: 10.1002/jtr.2743.