Online-Journals.org

iJIM International Journal of
Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 8 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i08.61245

PAPER

An ML-Optimized Mobile Virtual Tourism Assistant with Adaptive Interaction and Cross-Cultural Performance Evaluation

Xiaozhou Peng[1], Yuqin Huang[1](✉), Lin Zhao[2]

[1]Hunan First Normal University, Changsha, China

[2]Changsha Intelligent Driving Institute, Changsha, China

hyqlyx@hnfnu.edu.cn

ABSTRACT

The deep integration of mobile Internet and artificial intelligence (AI) is reshaping tourism services, with mobile virtual tourism assistants enhancing travel experiences. However, existing systems have limitations in interaction paradigms, contextual awareness, and cross-cultural adaptability while struggling to balance real-time and lightweight performance and service depth on mobile platforms. To address these challenges, a mobile virtual tourism assistant optimized by machine learning (ML) was proposed, focusing on adaptive interaction and cross-cultural performance optimization. A three-layer end–to–edge collaborative architecture—comprising data perception, intelligent processing, and application presentation—was constructed to accommodate multimodal inputs and the resource constraints inherent to mobile devices. Three core technical innovations were introduced. First, an end-edge collaborative multimodal adaptive interaction mechanism was developed, transitioning from passive question-answering to proactive service delivery through lightweight hybrid dialogue management and contextual prediction algorithms. Second, a device behavior–driven cross-cultural ML model was established, with quantifiable cultural feature vectors and adaptive interface generation logic constructed, supporting dynamic cross-cultural service adaptation. Third, a joint optimization model integrated with mobile edge computing (MEC) was designed, incorporating a real-time replanning algorithm to balance itinerary personalization with mobile resource efficiency. This study bridges the gap between deep personalized interaction on mobile platforms and systematic cross-cultural evaluation, providing technical foundations and theoretical insights for the global deployment of intelligent mobile tourism services.

KEYWORDS

mobile virtual tourism assistant, machine learning (ML), adaptive interaction, cross-cultural human–computer interaction, mobile edge computing (MEC), contextual awareness

Peng, X., Huang, Y., Zhao, L. (2026). An ML-Optimized Mobile Virtual Tourism Assistant with Adaptive Interaction and Cross-Cultural Performance Evaluation. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(8), pp. 118–132. https://doi.org/10.3991/ijim.v20i08.61245

Article submitted 2025-12-03. Revision uploaded 2026-01-27. Final acceptance 2026-02-04.

1 INTRODUCTION

The deep integration of mobile Internet technologies and artificial intelligence (AI) continues to reshape both the form and boundaries of cultural and tourism services [1, 2]. With the widespread adoption of mobile devices and ongoing improvements in portability and computational capability, mobile virtual tourism assistants have been increasingly positioned as core enablers for optimizing travel experiences [3, 4]. Although existing mobile virtual tourism assistants have achieved basic consultation and itinerary planning functions, effective adaptation to the dynamic characteristics and heterogeneous demands of mobile usage scenarios remains limited, and several critical challenges have become increasingly prominent [5]. First, rigid interaction paradigms constrain efficient responses to multimodal inputs, including speech, gestures, and augmented reality (AR) imagery [6]. Second, contextual awareness mechanisms remain lagging behind the real-time update cycles of mobile data streams, such as user location, environmental conditions, and device states [7]. Third, cross-cultural adaptation strategies are largely grounded in static theoretical frameworks [8, 9], with limited capacity for dynamic adjustment based on user-specific device behavior. To address these limitations, a machine learning (ML)–empowered mobile virtual tourism assistant is developed in this study, with emphasis placed on the adaptive interaction mechanism and cross-cultural performance optimization. The proposed approach is designed to directly address the real-time responsiveness, lightweight deployment, and multi-scenario adaptability required by mobile environments, thereby supporting the deep integration of intelligent tourism services on mobile platforms.

Existing research has produced preliminary advances across three related domains: intelligent tourism planning systems, tourism dialogue systems, and cross-cultural human–computer interaction. Nevertheless, substantial gaps persist with respect to mobile scenario adaptation. Intelligent tourism planning systems have predominantly emphasized knowledge graph construction and reinforcement learning application, while insufficient attention has been given to model architectures optimized for mobile computational and energy constraints [10, 11]. Tourism dialogue systems have evolved from rule-based approaches toward architectures driven by large language models (LLMs); however, their multimodal interaction capabilities remain insufficiently aligned with the input characteristics of mobile devices [12, 13]. Research on cross-cultural human–computer interaction has introduced theories of culturally adaptive user interfaces (UIs), yet the systematic exploitation of mobile users' digital behavioral traces remains limited. As a result, dynamic service adaptation remains difficult to achieve in practice [14]. Moreover, existing studies have not effectively integrated lightweight adaptive interaction mechanisms, device behavior–driven cross-cultural optimization, and collaborative scheduling with mobile edge computing (MEC) [15, 16]. The research gap provides the primary motivation for the present study.

The core contributions of this study are reflected in three key technical advances. First, a mobile-oriented multimodal adaptive interaction framework is proposed, enabling real-time multimodal fusion and lightweight semantic understanding, thereby facilitating a transition from passive question–answering to proactive contextual anticipation. Second, a device behavior–driven cross-cultural ML model and quantitative evaluation framework are established, overcoming the subjective limitations of traditional cultural adaptation approaches and enabling computationally tractable and dynamically optimizable cross-cultural services. Third, a personalized planning mechanism jointly coordinated between mobile devices and MEC infrastructure is designed, in which resource allocation and itinerary planning are tightly coupled to address the low-latency and low-energy requirements of resource-intensive mobile services.

2 OVERALL SYSTEM ARCHITECTURE DESIGN

The system is designed with lightweight deployment, real-time responsiveness, and low energy consumption on mobile devices as its core principles. A three-layer end–edge collaborative architecture—comprising data perception, intelligent processing, and application presentation—is constructed, in which mobile devices and network characteristics are deeply coupled to enable efficient multimodal data flow and a closed-loop intelligent service pipeline. At the data perception layer, a mobile-oriented input framework is established through real-time data acquisition and end–to–edge collaborative fusion. Multimodal preprocessing pipelines are optimized to accommodate mobile computational constraints. For speech input, a compact feature extraction scheme based on Mel-frequency cepstral coefficients (MFCCs) is adopted. Gesture input is processed using a feature dimensionality reduction algorithm designed for accelerometer sensor data. For AR image input, object detection is performed on-device using a streamlined version of the You Only Look Once (YOLO) model, substantially reducing data transmission overhead. Contextual data acquisition is governed by an adaptive strategy, in which the sampling frequencies of data sources such as Global Positioning System (GPS) and weather are dynamically adjusted according to battery level and network types. High-priority contextual information is collected at higher frequencies, whereas secondary context signals are activated on demand, achieving a balance between contextual awareness and energy efficiency.

Tourism-related knowledge is accessed through an end–to–edge collaborative mechanism, in which a lightweight knowledge graph retaining core relational structures is extracted from the data backbone and deployed on MEC nodes to ensure rapid on-device query responses. The intelligent processing layer constitutes the technical core of the system. Within this layer, the adaptive interaction engine, utilizing a lightweight LLM quantized by 8-bit integer (INT8) and a compact model based on Bidirectional Encoder Representations from Transformers (BERT), constructs a two-stage semantic processing mechanism that supports rapid on-device recognition and deeper edge-side understanding. A dialogue state incremental update model based on Gated Recurrent Units (GRUs) is integrated, in which a state vector incorporating user location trajectories and device interaction logs is constructed to avoid full recomputation, thereby enabling efficient adaptation to the computational constraints of mobile devices. The cross-cultural adaptation module constructs multidimensional cultural feature vectors based on users' digital footprints collected from mobile devices. Hofstede's cultural dimensions theory is integrated. A random forest algorithm is employed to enable rapid on-device classification of cultural labels and dynamic iteration of profiles. By establishing a mapping repository between cultural dimensions and interaction rules, automated adaptation of interface content and interaction logic is achieved through a component-based development framework.

The personalized recommendation and planning engine operates on a lightweight knowledge graph and employs an improved reinforcement learning algorithm for itinerary planning. A reward function incorporating mobile energy consumption and network latency weights is designed as $R = \omega_1 R_u + \omega_2 R_e + \omega_3 R_l$, where R_u denotes user satisfaction, R_e represents the energy efficiency coefficient, and R_l corresponds to the latency optimization coefficient. The weighting factors ω_1, ω_2, and ω_3 satisfy $\omega_1 + \omega_2 + \omega_3 = 1$. A collaborative strategy is adopted in which preliminary planning is performed on the mobile device, followed by dynamic optimization at MEC nodes. Lightweight replanning is triggered when route deviation or dwell-time thresholds—defined according to user behavior—are exceeded, thereby avoiding excessive computation and unnecessary resource consumption.

The application and presentation layer is designed with a primary focus on optimizing the mobile user experience. For AR interaction, a spatial interaction paradigm combining gesture-based triggering and voice-assisted control is adopted. Instance-based rendering techniques are applied to improve on-device AR model rendering efficiency, thereby significantly reducing mobile GPU workload. Interface presentation is implemented through a component-based dynamic adaptation strategy. Based on the user's cultural profile, interface elements—including icon size, text density, and navigation logic—are automatically adjusted to support stable presentation of core functionalities and culturally adaptive interfaces under offline conditions, ensuring service continuity across diverse usage scenarios. The technology stack is selected to align closely with end–edge collaboration and mobile adaptation requirements. On-device lightweight models are deployed using TensorFlow Lite, while AR functionalities are developed using ARKit and ARCore. Cross-platform interface development is supported through Corona SDK. On the edge side, MEC platforms—such as Huawei CloudEdge and Alibaba Cloud edge nodes—are utilized, with a streamlined version of Neo4j deployed for edge knowledge graph storage. At the coordination layer, the Message Queuing Telemetry Transport (MQTT) lightweight communication protocol is adopted to ensure low-latency data transmission between mobile devices and edge nodes. On-device dialogue management is supported through Rasa Lite, forming an integrated technical support framework that enables efficient module coordination, performance optimization, and the practical deployment of the proposed system's core innovations.

3 KEY TECHNICAL MODULE DESIGN

3.1 Mobile multimodal and context-aware adaptive interaction mechanism

A lightweight hybrid dialogue management architecture is designed through the collaboration of an on-device rule-based engine and an edge-side LLM, achieving a balance between low-latency responses in high-frequency scenarios and flexible handling of open-domain dialogue. Using a finite-state machine, the on-device rule engine constructs core conversational flows, including attraction inquiries and route navigation, which account for approximately 80% of high-frequency usage scenarios. When processed independently on the device, response latency is maintained below 200 ms. The edge-side LLM is dedicated to handling open-ended interactions, such as culturally grounded social conversation and personalized demand articulation. Seamless switching between on-device and edge-side dialogue processing is enabled through a customized end-to-edge collaboration protocol. An incremental data transmission strategy is adopted within this protocol, whereby only changes in dialogue state are synchronized rather than full dialogue contexts. This design significantly reduces network transmission overhead on mobile devices and constrains end-to-edge dialogue state synchronization latency to within 300 ms, thereby accommodating network variability commonly observed in mobile environments.

To support proactive services driven by contextual awareness, a dual-dimensional prediction model integrating contextual features and user behaviors is introduced. On-device real-time extraction is performed for contextual attributes—including GPS location, time, weather conditions, and device battery level—as well as behavioral attributes such as historical dwell time, click preferences, and search records. These features are aggregated into a 12-dimensional raw feature vector, which is subsequently reduced to a 6-dimensional feature vector $X \in R^6$ using principal component analysis (PCA) to lower computational complexity. A mobile-adapted lightweight Long Short-Term Memory (LSTM) network is employed. The network architecture is optimized by constraining the number of hidden-layer neurons to 128, while an attention mechanism is incorporated to amplify the weights of critical features, including user location and dwell time. Under this configuration, the target prediction accuracy is maintained at no less than 85%. In addition, a multi-level proactive service triggering mechanism is designed. A prediction confidence threshold of 0.7 is specified, and proactive service delivery is activated only when the device battery level exceeds 20% and network conditions are assessed as stable, thereby ensuring service proactivity while minimizing unnecessary user interruptions and avoiding excessive consumption of resources.

Multimodal interaction fusion techniques are designed with a focus on mobile input characteristics and user experience optimization, enabling the construction of an efficient and accurate multimodal processing framework. To address the high randomness inherent in mobile hand-drawn inputs, a dual recognition algorithm integrating contour feature and keypoint matching is developed. On-device Canny edge detection is employed to extract hand-drawn contour features, which are subsequently matched against a landmark contour feature repository pre-stored on MEC nodes. Dynamic time warping (DTW) is applied to achieve rapid contour matching, with recognition latency constrained within 500 ms and accuracy maintained above 80%, thereby satisfying the real-time input requirements of mobile scenarios. An emotion-driven multimodal feedback mechanism is further introduced to enable personalized interaction adjustments through the fusion of multidimensional affective information. On-device lightweight convolutional neural networks are utilized to extract facial expression features, while MFCCs combined with an emotion classifier are employed to analyze vocal prosody. These signals are synchronously integrated with text-based sentiment analysis outputs to construct a three-dimensional affective vector, $E = (v, a, d)$, where v denotes valence, a represents arousal, and d corresponds to dominance. Driven by this affective vector, the AR-based virtual assistant adaptively adjusts facial expressions and speech prosody. When negative emotional states are detected, speech rate is automatically reduced and calming cultural experience content is recommended, enabling a transition from passive response mechanisms to emotion-aware adaptive interaction. Notably, all affective feature extraction and vector construction processes are executed entirely on-device, thereby mitigating the risk of sensitive user data transmission and preserving user privacy. Figure 1 illustrates the three-layer architecture comprising data perception, intelligent processing, and application presentation, providing an intuitive visualization of the functional allocation between on-device lightweight processing and edge-side intensive computation.

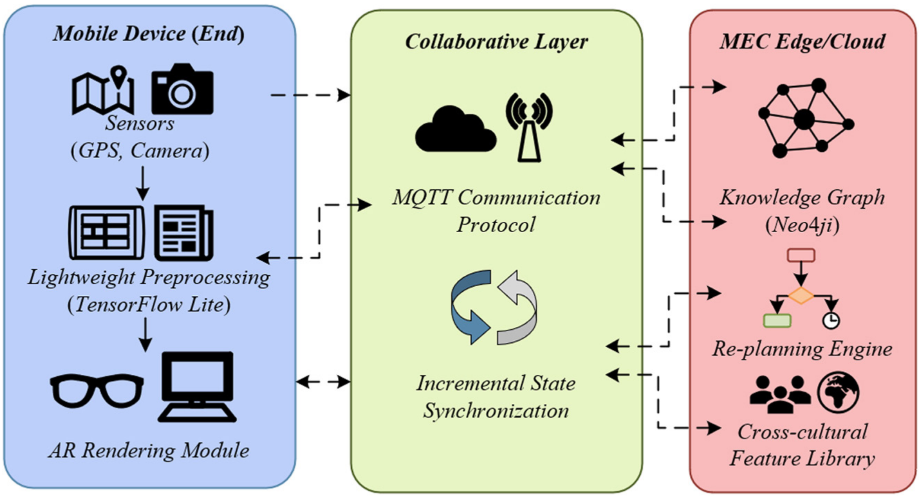


Fig. 1. Overall architecture of the end-edge collaborative mobile virtual tourism assistant

3.2 Mobile ML model for cross-cultural performance optimization

A device behavior–driven cultural feature modeling framework is developed to overcome the limitations of traditional culture adaptation approaches that rely predominantly on subjective theoretical assumptions. Through this framework, cross-cultural characteristics are rendered quantifiable and suitable for efficient on-device processing. From mobile user behavioral data, six categories comprising 18 quantifiable features are extracted to construct a cultural feature vector. These categories encompass language settings, regional interface preferences, interaction behaviors, content preferences, social interaction patterns, and device usage characteristics, thereby comprehensively capturing cultural tendencies embedded within users' digital footprints. Feature weighting is determined through a combined strategy integrating the analytic hierarchy process (AHP) and the entropy weight method. AHP is employed to establish a structured subjective weighting framework, while entropy weighting is applied to correct potential bias and enhance objectivity. Among the core features, language type and confirmation dialog interaction frequency are assigned weights of 0.15 and 0.12, respectively, enabling precise alignment with key indicators defined in Hofstede's cultural dimension theory. Cultural profile classification is conducted using an improved K-means algorithm. By optimizing the initialization strategy for cluster centroids, convergence to local optima is effectively mitigated, and users are categorized into seven representative cultural groups. For on-device deployment, the model is subjected to quantization-based compression, resulting in a 60% reduction in parameter scale. Under this configuration, classification latency is strictly constrained to within 300 ms, satisfying the real-time processing requirements of mobile environments. Based on the feature system, a collaborative filtering algorithm integrating cultural distance and user interest (CD-CF) is proposed. A cultural distance matrix is derived from Hofstede's six dimensions, and the related equation is as follows:

$$CD_{ij} = \sqrt{\sum_{k=1}^{6} w_k (C_{ik} - C_{jk})^2} \tag{1}$$

where, w_k denotes the weight of the cultural dimension, and C_{ik} represents user i's score on dimension k. To optimize the traditional collaborative filtering objective function, a cultural distance penalty term is introduced, yielding the following formulation:

$$arg\,\min \sum_{u,i} (r_{ui} - \hat{r}_{ui})^2 + \lambda \times CD_{u,c} \tag{2}$$

where, $\hat{r}_{ui}$ denotes the predicted rating, $CD_{u,c}$ represents the cultural distance between the user and the recommended content, and the regularization coefficient λ is set within the range of 0.1–0.3 and dynamically adjusted based on user feedback. An end-edge partitioned computation strategy is adopted. The user similarity matrix and cultural distance matrix are stored on MEC nodes, while the mobile device transmits only the user interest vector and location information. Following edge-side computation, Top-*N* results are returned to the device, substantially reducing on-device processing load.

To address cross-cultural adaptation requirements in mobile interfaces, a component-based culturally adaptive UI rendering engine is designed to enable dynamic optimization of interface appearance and interaction logic. The mobile interface is decomposed into more than 20 atomic components, encompassing core elements such as icons, buttons, text fields, and navigation bars. For each component, three to five culturally adaptive style variants are predefined, enabling differentiated presentation based on the user's cultural profile. In cultural contexts characterized by high power distance, larger button dimensions and stronger color contrast are applied, whereas flatter interface layouts with reduced visual hierarchy are adopted for low power-distance contexts. Dynamic adaptation logic is constructed using a decision tree model, in which the user's cultural profile serves as the input for determining the adaptive style of each component. In parallel, component priority rules are defined, ensuring that navigation components are assigned higher adaptation priority than interaction and presentation components, thereby preserving the cultural adaptation effectiveness of core functionalities. To mitigate the impact of network instability in mobile environments, offline adaptation capabilities are incorporated into the engine. Interface configuration files corresponding to frequently encountered cultural scenarios are pre-stored on the device and automatically activated in offline mode. When network connectivity is available, adaptive rules are synchronized and updated via MEC nodes, ensuring a consistent user experience across heterogeneous network conditions. Collectively, the proposed model establishes a closed-loop pipeline encompassing feature modeling, recommendation optimization, and interface adaptation. All core computational modules are optimized for lightweight on-device execution, enabling high-precision cross-cultural adaptation while simultaneously satisfying the low-latency and low-energy consumption requirements inherent to mobile platforms.

3.3 MEC-integrated mobile resource management and itinerary planning optimization

An end-edge collaborative bi-objective joint optimization model is developed to achieve a dynamic balance between itinerary experience and resource efficiency, thereby accommodating the resource-intensive requirements of mobile services such as AR navigation and high-definition content streaming. The model is formulated with the objective of maximizing overall system performance, and the objective objection is defined as $maxF = \alpha \times UE + (1 - \alpha) \times RE$, where UE denotes the user experience score, incorporating key indicators such as itinerary rationality and AR smoothness, and RE represents the resource efficiency coefficient, jointly reflecting bandwidth utilization and mobile energy consumption. The weighting parameter α is constrained within the range of 0.6–0.8 and is dynamically adjusted according to mobile network conditions. Under stable network connectivity, higher values of α are assigned to prioritize user experience, whereas under constrained network conditions, α is moderately reduced to favor efficient resource utilization. Four core constraints are imposed on the model. First, the bandwidth allocated to AR navigation is constrained to be no less than 2 Mbps. Second, the energy consumption associated with a single itinerary is limited to within 10% of the device's battery capacity. Third, deviation in total itinerary duration is restricted to no more than 10%. Fourth, the computational load of MEC nodes is maintained below 80%. Collectively, these constraints establish an optimization boundary that jointly balances service quality and resource constraints.

To address dynamic variations during itinerary execution, an event-triggered on-device rapid replanning algorithm (ET-RPA) is introduced to overcome the resource overhead associated with conventional full replanning strategies. Triggering events are categorized into three classes: location deviation, temporal delay, and resource anomalies. For each class, quantifiable thresholds that can be directly detected on mobile devices are defined. In particular, a resource anomaly is triggered when AR stuttering persists for no less than 1 s, thereby ensuring both detection accuracy and real-time responsiveness. Event states are continuously monitored on the device. Once a predefined threshold is satisfied, a lightweight replanning procedure is initiated. A hybrid strategy combining local adjustment with global adaptation is adopted, whereby optimization is applied only to the affected itinerary segments. Pre-stored alternative plans residing on the device are accessed, and real-time resource states from MEC nodes are incorporated to rapidly generate substitute routes. Through this design, the computational overhead and latency associated with full replanning are effectively avoided. The core modules of the algorithm are implemented in C to enhance on-device execution efficiency. Under this configuration, single replanning latency is constrained within 800 ms, while itinerary rationality is preserved at a rate of no less than 90%. Compared with scenarios without replanning, AR smoothness is improved by more than 30%. The proposed algorithm operates in a complementary manner with the end-edge collaborative joint optimization model, forming a full-process resource management framework spanning initial planning and dynamic adjustment. Through this integration, itinerary personalization is maintained while the fundamental trade-off between mobile resource constraints and low-latency service requirements is effectively resolved. Figure 2 illustrates the event-triggered end-to-edge collaborative replanning process, demonstrating how the mobile device monitors the three classes of triggering events and executes the corresponding route update procedure.

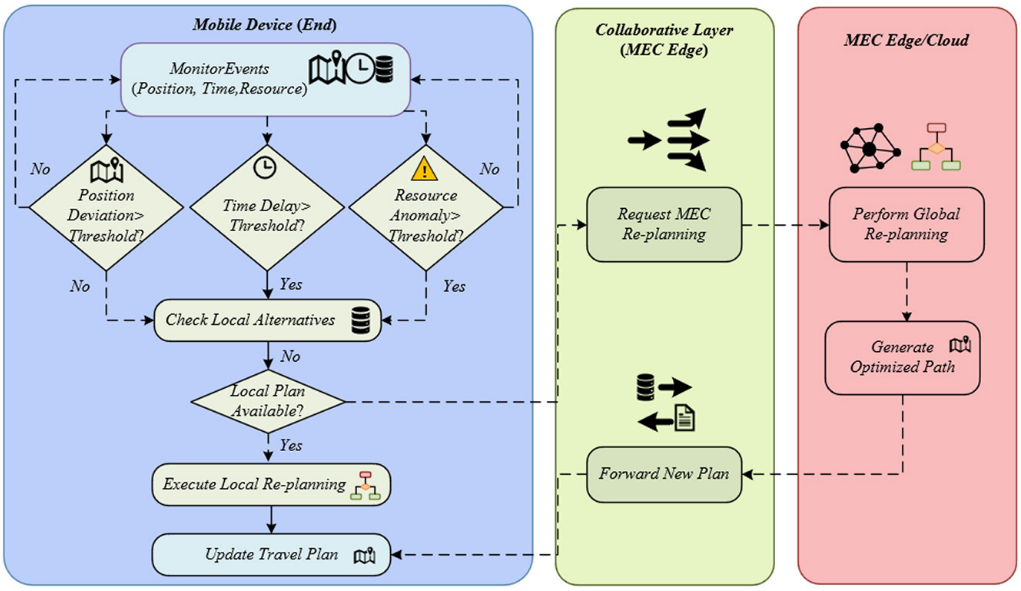


Fig. 2. Event-triggered end-edge collaborative replanning algorithm

4 SIMULATION EXPERIMENTS AND CROSS-CULTURAL PERFORMANCE EVALUATION

The experimental evaluation was conducted across four primary dimensions: functional performance, cross-cultural adaptability, module effectiveness, and representative scenario adaptation. Experiments encompassed three categories of mobile devices and four cultural groups. Comparative testing against three baseline systems—Baseline 1 (traditional rule-driven), Baseline 2 (LLM-based adaptive interaction without cross-cultural optimization), and Baseline 3 (cross-cultural adaptation without MEC collaboration and lightweight optimization)—was combined with ablation studies to comprehensively assess the technical advantages and innovative contributions of the proposed system. Key experimental results and analyses are presented below.

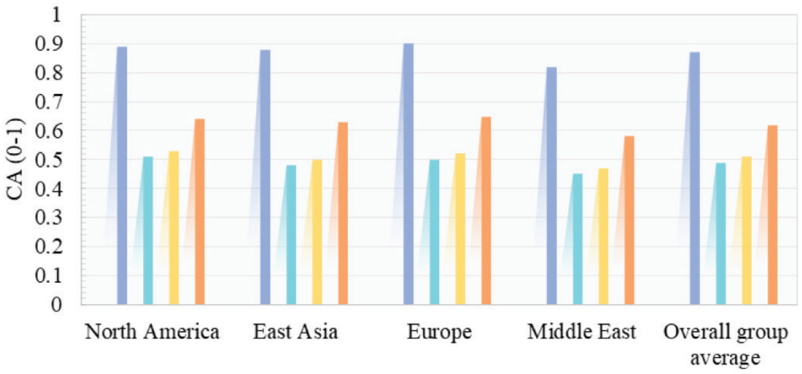
Functional performance evaluation focused on core mobile requirements, including real-time responsiveness, accuracy, and resource efficiency. Comparative experiments were conducted across flagship, mid-range, and entry-level devices. The results are summarized in Table 1.

Table 1. Cross-device comparison of functional performance metrics

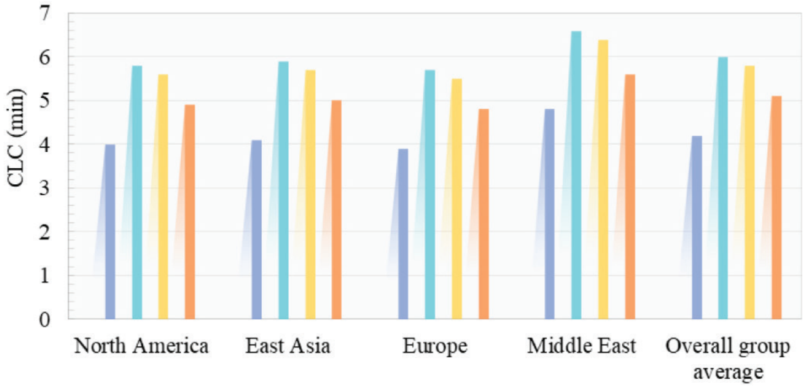
Performance Metric	Device Type	Proposed System	Baseline 1	Baseline 2	Baseline 3	Improvement vs. Baseline 1 (%)	Improvement vs. Baseline 3 (%)
Interaction response latency (ms)	Flagship	175	510	240	400	65.7	56.3
	Mid-range	260	570	310	450	54.4	42.2
	Entry-level	370	640	440	520	42.2	28.8
Multimodal input recognition accuracy (%)	Flagship	91.5	64.2	86.8	80.7	42.5	13.4
	Mid-range	89.2	62.5	83.7	78.3	42.7	13.9
	Entry-level	86.7	60.8	80.5	75.9	42.6	14.2
Replanning latency (ms)	All devices (avg.)	710	1820	1080	1500	61.0	52.7
Replanning rationality retention (%)	All devices (avg.)	92.8	68.1	84.3	81.0	36.3	14.6
Single-itinerary energy reduction (%)	All devices (avg.)	29.3	–	12.1	8.3	–	253.0
AR navigation smoothness (FPS)	All devices (avg.)	29.8	15.1	22.5	18.2	97.3	63.7
Mobile bandwidth utilization (%)	All devices (avg.)	77.1	42.5	61.2	52.9	81.4	45.7

As indicated in Table 1, functional performance across all device categories is consistently superior to that of the three baseline systems, while effective adaptation to heterogeneous computational capability levels is maintained. For mid-range devices, which represent the dominant usage segment, interaction response latency is constrained to 260 ms, satisfying real-time mobile interaction requirements and corresponding to improvements of 54.4% and 42.2% relative to Baseline 1 and Baseline 3, respectively. Even on entry-level devices with limited computational resources, response latency remains below 400 ms, underscoring the effectiveness of the lightweight model design and end-to-edge collaborative architecture in accommodating mobile hardware constraints. Multimodal input recognition accuracy reaches an average of 89.1% across all devices, with 89.2% achieved on mid-range devices. Improvements exceeding 13% relative to the baseline systems are observed, demonstrating the precision of the multimodal fusion algorithm under complex mobile input conditions.

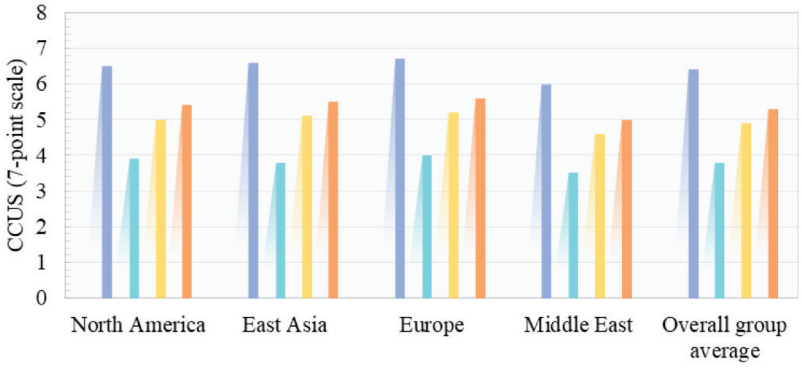
Resource efficiency and dynamic planning performance are also notably enhanced. Replanning latency is reduced to 710 ms, representing a reduction of more than 50% relative to the baseline systems, while itinerary rationality retention remains at 92.8%, ensuring the reliability of itinerary adjustments. Average per-itinerary energy consumption is reduced by 29.3%, corresponding to an improvement of 253% relative to Baseline 3. In parallel, AR navigation smoothness is maintained at 29.8 FPS, and mobile bandwidth utilization reaches 77.1%, both of which substantially exceed baseline performance. These results confirm that the MEC-enabled collaborative resource management mechanism effectively balances service quality with mobile energy consumption and bandwidth constraints, particularly in resource-intensive scenarios such as AR navigation.



a) Cultural adaptation accuracy (CA)



b) Cultural learning curve (CLC)



c) Cross-cultural user satisfaction (CCUS)

Fig. 3. Cross-cultural performance comparison across multiple user groups

Cross-cultural performance evaluation encompassed four cultural groups—North America, East Asia, Europe, and the Middle East—and focused on adaptation accuracy, usability, neutrality, and user acceptance. The corresponding results are presented in Figure 3.

Overall cross-cultural performance is demonstrated to be strong. The average CA across all user groups reaches 0.87, corresponding to improvements of 77.6% relative to Baseline 1 and 40.3% relative to Baseline 3, thereby meeting the predefined performance targets. Among the individual groups, the European cohort exhibits the highest CA value of 0.90. Although the Middle Eastern cohort records the lowest CA value at 0.82, performance remains substantially improved relative to the baseline systems, indicating that the device behavior–driven cultural feature modeling and adaptive interface generation mechanisms provide generalizable adaptation across diverse cultural contexts. The CLC averages 4.2 min across all user groups, representing a 30% reduction relative to Baseline 1 and reflecting a lowered adoption threshold for users from different cultural backgrounds. Notably, the European cohort achieves the shortest CLC at 3.9 min, highlighting a pronounced usability advantage.

Cultural neutrality (CN) is maintained at 0.16, substantially lower than the values exceeding 0.32 observed in the baseline systems. This result indicates the absence of pronounced cultural bias and confirms that equitable adaptation is achieved across cultural groups. CCUS attains an overall average score of 6.4. Scores of 6.7 and 6.6 are observed for the European and East Asian cohorts, respectively, while a score of 6.0 is maintained even for the Middle Eastern cohort, which exhibits greater cultural divergence. These results collectively validate the effectiveness of cross-cultural adaptation in interface presentation, content recommendation, and interaction logic, and demonstrate broad user acceptance across cultural groups.

To verify the necessity of each proposed innovation, ablation experiments were conducted by selectively removing single core modules and examining the resulting performance variations. The corresponding results are reported in Table 2.

Table 2. Contribution analysis of core innovation modules based on ablation experiments

Experimental Group	Adaptive Interaction Engine	Cross-Cultural Adaptation Module	MEC Collaborative Optimization Module	Overall System Performance (100-Point Scale)	Functional Performance Subscore (50-Point Scale)	Cross-Cultural Performance Subscore (50-Point Scale)	Performance Loss Relative to Full System (%)
Group 1 (full system)	Enabled	Enabled	Enabled	92.6	47.8	44.8	0.0
Group 2 (adaptive interaction engine removed)	Disabled	Enabled	Enabled	59.1	30.2	43.9	36.2
Group 3 (cross-cultural adaptation removed)	Enabled	Disabled	Enabled	64.8	47.1	27.7	30.0
Group 4 (MEC collaborative optimization removed)	Enabled	Enabled	Disabled	69.4	35.7	43.7	25.1
Group 5 (all modules removed; near Baseline 1)	Disabled	Disabled	Disabled	41.3	20.5	20.8	55.4

The ablation results demonstrate that each of the three core innovation modules plays an indispensable role in overall system performance, with their combined operation providing synergistic benefits. When the adaptive interaction engine is removed, overall system performance is reduced by 36.2%, corresponding to an estimated contribution of approximately 35.0%. The functional performance subscore decreases sharply from 47.8 to 30.2, confirming that lightweight hybrid dialogue management and contextual prediction mechanisms provide foundational support for service experience quality. Removal of the cross-cultural adaptation module results in a performance loss of 30.0%, accompanied by a pronounced drop in the cross-cultural performance subscore to 27.7. This outcome indicates that the cross-cultural adaptation module constitutes the primary determinant of culturally adaptive capability and directly governs the feasibility of global service deployment.

When the MEC collaborative optimization module is disabled, a performance loss of 25.1% is observed, with the functional performance subscore declining by 12.1 points. Performance degradation is primarily reflected in energy efficiency, AR smoothness, and bandwidth utilization metrics, demonstrating that MEC collaboration is essential for alleviating mobile resource constraints. When all three modules are disabled, overall system performance declines to 41.3 points, approaching the level of Baseline 1. This result further highlights the synergistic value of the three innovation modules, which together establish the system's core technical advantage and performance barrier.

To analyze differentiated adaptation effects, German users characterized by high uncertainty avoidance and Chinese users characterized by high collectivism were selected as representative case studies. Detailed results are presented in Table 3.

Table 3. Case analysis of representative cultural groups

Analysis Dimension	High Uncertainty Avoidance (German User Group)	High Collectivism (Chinese User Group)
CA	0.91	0.90
Core interface adaptation features	Structured itinerary card layout, high information density, multi-level confirmation dialogs, and a cool color theme	Group itinerary recommendations prioritized quick-access sharing functions, a flat layout, and a warm color theme
Core itinerary recommendation features	Detailed attraction descriptions, precise time-slot planning, explicit alternative plans, and compliance-related information are prioritized	Group-friendly attractions, shared itinerary templates, local cultural experience recommendations, and social check-in point annotations
Core interaction logic features	Stepwise guidance, reversible operations, detailed exception alerts, and minimal proactive interruptions	Voice-based group interaction support, one-click itinerary sharing, multi-user AR co-viewing, and proactive recommendation of popular experiences
Replanning trigger thresholds	Route deviation ≥ 300 m, time delay ≥ 10 min, and resource anomaly ≥ 0.5 s	Route deviation ≥ 500 m, time delay ≥ 20 min, and resource anomaly ≥ 1 s
CCUS	6.8 (7-point scale)	6.7 (7-point scale)
Single-itinerary energy consumption	≤ 8% battery usage	≤ 9% battery usage
AR navigation smoothness (FPS)	31.2	30.5

The case analysis demonstrates that user needs and preferences are accurately accommodated across distinct cultural groups. For German users exhibiting high uncertainty avoidance, structured itineraries and multi-level confirmation logic are automatically generated. High-information-density presentation is applied to reduce decision uncertainty, and stricter replanning trigger thresholds are enforced. As a result, a CA of 0.91 and a CCUS score of 6.8 are achieved, reflecting effective alignment with preferences for rigor. For Chinese users characterized by high collectivism, interface design prioritizes group itineraries and sharing functionalities. Multi-user AR co-viewing and proactive recommendation of socially oriented experiences are supported, enabling effective accommodation of group-based travel behaviors. Correspondingly, a CA of 0.90 and a CCUS score of 6.7 are obtained, indicating strong user acceptance. Across both case studies, energy consumption control and AR navigation smoothness remain at high levels, demonstrating that cross-cultural adaptation and resource efficiency optimization can be jointly achieved.

5 CONCLUSION AND FUTURE WORK

This study addressed the core challenges of adaptive interaction and cross-cultural performance optimization in mobile virtual tourism assistants and established an integrated solution that combines ML with end-edge collaborative technologies. Three principal contributions were achieved. First, a multimodal adaptive interaction framework tailored to mobile environments was developed. Through lightweight hybrid dialogue management and contextual prediction mechanisms, constraints related to mobile computational capacity, energy consumption, and real-time responsiveness were effectively mitigated, enabling an experiential transition from passive question-answering to proactive service delivery. Second, a device behavior–driven cross-cultural ML model was constructed. By leveraging users' mobile digital footprints, a quantifiable cultural feature system was established and integrated with adaptive interface generation techniques, thereby enabling dynamic cross-cultural service adaptation and precision optimization. Third, a collaborative resource management and itinerary planning strategy jointly operated by mobile devices and MEC infrastructure was designed. Through bi-objective optimization and event-triggered replanning mechanisms, performance bottlenecks associated with resource-intensive services—such as AR-based navigation—were effectively alleviated. Simulation experiments and cross-cultural evaluations demonstrated that the proposed system consistently outperformed baseline approaches in functional performance, cross-cultural adaptability, and user experience, addressing a critical gap in existing research by integrating deep mobile personalization with systematic cross-cultural evaluation. At the theoretical level, this study enriches the body of knowledge in cross-cultural human–computer interaction for mobile environments by introducing a novel adaptation paradigm that integrates end–edge collaboration with ML. The proposed quantitative evaluation framework further provides a standardized reference for future research in this domain. At the practical level, the findings offer technical support for the intelligent and globalized development of mobile tourism applications, facilitating deeper integration of MEC, AI, and the cultural tourism industry and advancing the deployment of next-generation intelligent mobile tourism services.

6 ACKNOWLEDGEMENT

This work was supported by the Social Sciences Project of Hunan Province (No. 24YBA346).

7 REFERENCES

[1] H. Jin, "Integration of mobile interaction technology in the tourism industry and its impact on tourism consumption patterns," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 1, pp. 140–154, 2025. https://doi.org/10.3991/ijim.v19i01.53495

[2] N. M. Al-Hazmi, "Tourism advertising using specialized mobile applications and its impact on increasing the demand for the services of tourism organizations," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 16, pp. 61–69, 2023. https://doi.org/10.3991/ijim.v17i16.42689

[3] P. Tak and M. Gupta, "Examining travel mobile app attributes and its impact on consumer engagement: An application of SOR framework," *Journal of Internet Commerce*, vol. 20, no. 3, pp. 293–318, 2021. https://doi.org/10.1080/15332861.2021.1891517

[4] I. S. Hesjevoll, A. Fyhri, and A. Ciccone, "App-based automatic collection of travel behaviour: A field study comparison with self-reported behaviour," *Transportation Research Interdisciplinary Perspectives*, vol. 12, p. 100501, 2021. https://doi.org/10.1016/j.trip.2021.100501

[5] O. O. Afolabi, A. Ozturen, and M. Ilkan, "Effects of privacy concern, risk, and information control in a smart tourism destination," *Economic Research-Ekonomska Istraživanja*, vol. 34, no. 1, pp. 3119–3138, 2021. https://doi.org/10.1080/1331677X.2020.1867215

[6] G. Srivastava, S. Bag, P. J. Ramudu, S. K. Shrivastav, J. Pueschel, and A. C. Benabdellah, "Theoretical perspectives on the impact of generative AI on the tourism sector: A literature review," *Journal of Global Information Management*, vol. 33, no. 1, pp. 1–28, 2025. https://doi.org/10.4018/JGIM.388177

[7] M. Kenteris, D. Gavalas, and D. Economou, "An innovative mobile electronic tourist guide application," *Personal and Ubiquitous Computing*, vol. 13, no. 2, pp. 103–118, 2009. https://doi.org/10.1007/s00779-007-0191-y

[8] K. Reinecke and A. Bernstein, "Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces," *ACM Transactions on Computer-Human Interaction*, vol. 18, no. 2, p. 8, 2011. https://doi.org/10.1145/1970378.1970382

[9] D. Soldatenko and E. Backer, "A content analysis of cross-cultural motivational studies in tourism relating to nationalities," *Journal of Hospitality and Tourism Management*, vol. 38, pp. 122–139, 2019. https://doi.org/10.1016/j.jhtm.2018.12.004

[10] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7370–7389, 2014. https://doi.org/10.1016/j.eswa.2014.06.007

[11] L. Salau, M. Hamada, R. Prasad, M. Hassan, A. Mahendran, and Y. Watanobe, "State-of-the-art survey on deep learning-based recommender systems for e-learning," *Applied Sciences*, vol. 12, no. 23, p. 11996, 2022. https://doi.org/10.3390/app122311996

[12] J. Arcos-Pumarola, M. S. Almela, and B. R. Maestre, "Volunteer tourism: Reflections from applied ethics," *Tourism Culture & Communication*, vol. 22, no. 3, pp. 219–230, 2022. https://doi.org/10.3727/109830421X16345418234047

[13] K. Schianetz, T. Jones, L. Kavanagh, P. A. Walker, D. Lockington, and D. Wood, "The practicalities of a learning tourism destination: A case study of the Ningaloo Coast," *International Journal of Tourism Research*, vol. 11, no. 6, pp. 567–581, 2009. https://doi.org/10.1002/jtr.729

[14] M. H. Miraz, P. S. Excell, and M. Ali, "Culturally inclusive adaptive user interface (CIAUI) framework: Exploration of plasticity of user interface design," *International Journal of Information Technology & Decision Making*, vol. 20, no. 1, pp. 199–224, 2021. https://doi.org/10.1142/S0219622020500455

[15] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017. https://doi.org/10.1109/MC.2017.9

[16] E. B. Wagiu, C. M. Liu, and Y. Palopak, "Mapping technological trajectories of edge computing: A citation graph analysis," *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 16545–16560, 2024. https://doi.org/10.1109/JIOT.2024.3355056

8 AUTHORS

Xiaozhou Peng holds a Ph.D. in International Tourism and Management from City University of Macau and is a Lecturer of tourism management in Business

School, Hunan First Normal University, specializing in digital tourism and tourism destination management research (E-mail: pxz@hnfnu.edu.cn).

Yuqin Huang holds a Master's Degree in tourism management from Hunan Normal University and is a Lecturer of tourism management in the Business School, Hunan First Normal University, specializing in digital tourism and tourism education research (E-mail: hyqlyx@hnfnu.edu.cn).

Lin Zhao holds a Bachelor's degree in Electronic Information Engineering from Hunan University of Commerce and serves as a Software Engineer in the Connected Transportation Division at CIDI (Changsha Intelligent Driving Institute), specializing in vehicle-infrastructure cooperation and intelligent transportation systems (E-mail: zhao.lin@cidi.ai).