International Journal of Natural Resources and Environmental Studies
ISSN: 3006-2012 (Print), ISSN: 3006-0834 (Online) | Volume 7, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijnres.v7n3.13
Journal homepage: https://wepub.org/index.php/IJNRES/index

A Study on the Guilin Tourism Visual Corpus Based on Multi modal Geospatial Information Fusion

Yuhe Sun, Jingjing Wu

Guilin University of Technology, Guilin, 541004, China

ABSTRACT

The global tourism industry is undergoing a rapid transition from traditional service-based models to those driven by multimodal data. As a key tourist destination situated in the heart of China's karst region, Guilin is rich in natural landscapes, cultural heritage, and multilingual tourism content. However, these resources are dispersed and fragmented, which hinders the integration required for effective smart tourism management. Despite substantial domestic research in areas such as geospatial multimodal fusion, tourism corpora, and geovisualization, these efforts are often siloed within individual disciplines and struggle to establish a cohesive system that links multimodal data integration, corpus construction, and visualization. This paper addresses these issues by reviewing the evolving trends in these three core fields and emphasizing the critical role of multimodal data integration, the function of tourism corpora as data storage and semantic linkage tools, and the transformative potential of visualization in making multimodal data more accessible and understandable. The study develops a framework for collaborative multimodal fusion, corpus support, and visualization output, focusing on Guilin as a case study. By integrating diverse datasets, such as terrain remote sensing data, images of scenic areas, tourist reviews, and spatio-temporal paths, this work provides valuable data support for Guilin's tourism strategy and fills gaps in domestic research on the integration of multimodal geospatial data, tourism corpora, and visualization. Furthermore, it proposes a transferable model for similar international tourist destinations, advancing interdisciplinary research at the intersection of geospatial information science and tourism management.

KEYWORDS

Multimodal Geospatial Data; Tourism Corpus; Geovisualization; Smart Tourism; Guilin.

1. INTRODUCTION

The global tourism industry is currently experiencing significant restructuring (World Economic Forum, 2025). With the increasing integration of artificial intelligence (AI), big data, and geographic information technologies, the industry is shifting from a traditional service-oriented model to one driven by multimodal data. This transition is not only a technological advancement but also represents a fundamental rethinking of how tourism supply and demand are matched. The widespread adoption of technologies such as high-resolution remote sensing satellites (RSS), unmanned aerial vehicles (UAVs), and social media has introduced a new layer of complexity to tourism information, which now includes visual images, text reviews, spatio-temporal paths, geospatial data, and other related forms (Fitzpatrick et al., 2024; Yang & Chen, 2025). Traditional decentralized models of information management, which rely on single databases and static text, can no longer meet the demands of smart tourism that require personalized services for tourists and dynamic decision-making for managers (Rogowski, 2020; Juan et al., 2024). Furthermore, multimodal information fusion technologies have overcome the limitations of single-modal data processing, enabling deeper semantic connections

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

through cross-modal attention mechanisms. Notably, systems like the multimodal machine learning classification system developed by Baltrušaitis et al. (2017) and the Vision-and-Language Transformer (ViLT) model proposed by Kim et al. (2021) have provided a framework for the integration of Guilin's tourism information, covering images, text reviews, and geospatial data. Additionally, visualization technologies, which serve as key carriers linking multimodal data and user cognition, are reshaping the transmission and interaction of tourism information (Qin & Pan, 2023). Their intuitive and interactive nature effectively reduces the cognitive load on users, enhancing decision-making efficiency.

Currently, both domestic and international studies in geospatial multimodal fusion, tourism corpora, and geovisualization have made significant strides (Yang & Chen, 2025, on landscape planning models; Juan et al., 2024, on the Self-AM-V Trans algorithm; Šoltésová et al., 2025, on GIS-based tourism resource evaluation). However, existing research is mostly confined to specific technical dimensions and lacks a comprehensive, cross-disciplinary approach. As a key international tourist city in China's karst region, Guilin is home to various forms of tourism data, including natural landscapes, cultural heritage, and multilingual tourist reviews (Lan et al., 2022; Cao et al., 2023). Despite these abundant resources, the scattered nature of tourism information poses challenges for implementing an integrated tourism strategy. Issues such as the lack of connection between ecological data from the Lijiang River basin and tourist data from nearby scenic spots, and the separation of historical information about the Jingjiang Prince's Mansion from its physical context, create difficulties in both regional management coordination and efficient access to information for tourists. Additionally, domestic research has predominantly focused on isolated technological applications, which has led to challenges in interdisciplinary collaboration, platform integration, and the bridging of theory and practice. This has resulted in the failure to establish a complete chain from data fusion to corpus construction and visualization (Han et al., 2025).

This study addresses these gaps by analyzing Guilin's tourism data, integrating multimodal geospatial information, and employing geovisualization technologies to construct a visualized tourism information corpus. This corpus combines multimodal data (e.g., geospatial coordinates, landscape images, and tourist reviews) and enables an intuitive visual presentation of the tourism landscape in Guilin. Its core value lies not only in supporting the city's tourism strategy but also in facilitating the sustainable development and intelligent management of the tourism industry. Moreover, the study advances domestic research in multimodal geospatial data fusion and tourism visualization technologies and offers a practical, transferable model for similar international tourist cities.

2. DEVELOPMENT BACKGROUND OF MULTIMODAL INFORMATION FUSION IN GEOSPATIAL INFORMATION SCIENCE

Since the 1990s, the emergence of Geographic Information Systems (GIS) has enabled researchers to investigate spatial information using computerized tools such as spreadsheets, databases, and graphic software. Maps have evolved from static representations into interactive geospatial data interfaces that facilitate data exploration and information acquisition (Kraak, 2003). At the same time, rapid advances in information science have transformed data storage, analysis, decision-making, and visualization. The explosive growth of complex multimodal data, together with the increasing demand for dynamic interaction, has accelerated the development of new information technologies such as virtual reality, machine learning, cloud computing, big data, and the Internet of Things (Lee & Kang, 2015). The fixed functions of traditional GIS are no longer adequate for handling the complexity of multimodal data (Shaw et al., 2008; Han et al., 2025). This has encouraged the application and optimization of cross-modal information fusion analysis, which has become a key approach to enhancing the understanding of complex geographical phenomena and supporting decision-making. The theoretical foundation of multimodal fusion lies in the interdisciplinary evolution of information fusion theory. Lahat et al. (2015) described multimodality as a framework that enhances usability more effectively than single-modality systems. Originally applied to multi-sensor data integration in the military domain, this theory has since expanded into geospatial applications. Within the context of information systems and human-computer interaction, Ramachandram and Taylor (2017) and Baltrušaitis et al. (2017) further refined the theory, defining multimodal information fusion as an interdisciplinary system that integrates information from visual, textual, and geospatial modalities to achieve semantic association and synergy. Its essence lies in enabling heterogeneous data to collaborate in ways that generate greater analytical and decision-making value. This theoretical evolution provides a framework for integrating the heterogeneous data of Guilin's tourism sector, encompassing natural landscape imagery, cultural heritage texts, and the spatio-temporal paths of tourists.

At the technical level, Furgale et al. (2013) jointly estimated the temporal offset and relative spatial displacement between measurements from different sensors. Serving as a bridge between theory and practice, their research offered technical guidance for synchronizing and calibrating multi-sensor data in tourism scenarios. Over the following decade, the use of electronic devices for tourism information acquisition became increasingly prevalent (Amadu et al., 2025), alongside significant advances in sensing technologies such as high-resolution remote sensing satellites, UAV-borne multispectral sensors, and IoT-based devices. For example, Arora et al. (2023) employed sensor technologies to monitor tourist flow within scenic areas, extending their application scope within tourism contexts. The smart tourism sensor network established by Tang et al. (2022) and the sensor research on trail tourist flow conducted by Madden et al. (2021) have further refined the use of sensing technologies in tourism information collection from system-level and scenario-specific perspectives. Rogowski (2020) combined infrared pyroelectric sensors with digital imaging and transmission technologies, overcoming data acquisition challenges in mountainous tourism areas and providing technical references for similar work in the karst mountain regions surrounding Guilin.

Meanwhile, non-sensor-based methods of information acquisition have also gained traction. Global Positioning System (GPS) technology can accurately record tourists' outdoor movement paths, compensating for the limitations of sensors in capturing dynamic trajectory data. For instance, Bielański et al. (2018) examined outdoor leisure activities in protected areas using GPS tracking, while Jurado Rota et al. (2019) employed Volunteered Geographic Information (VGI) to monitor tourist movements in similar environments. Both studies offer valuable methodological references for acquiring tourist trajectory data in tourism research. Motion-sensing cameras have also been widely used to capture interactive imagery between tourists and landscapes, adding dynamic visual data that complement sensor-based information (Miller et al., 2017; Sitarz et al., 2018). The integrated use of tracking devices and emerging technologies has advanced tourism information acquisition toward multidimensional collaboration. Gao and Schmöcker (2021) recorded tourists' movement patterns using Wi-Fi data packet sensors, demonstrating a novel integration of static sensing data and dynamic path analysis. Similarly, research on IoT sensors by Sehrawat and Gill (2020), along with studies on smart wearable devices by Jovanov (2019), has contributed to the collection of behavioral data from tourists, broadening the range of device-linked information acquisition in tourism contexts. Together, these advancements have endowed geographic information science with robust multimodal data acquisition capabilities. They have led to the development of comprehensive systems characterized by multi-device, multi-dimensional, and multi-scenario data collection, thereby laying a solid foundation for subsequent multimodal data fusion and analysis.

Traditional GIS has long been centered on geometric space, while multimodal fusion extends its analytical scope to include semantic and cognitive spaces. Tatane et al. (2024) and Chantrapornchai et al. (2018), for example, integrated textual emotion data with the spatial distribution of Points of Interest (POIs) to examine tourists' perceptions of urban landscapes, thereby demonstrating the feasibility of fusing geometric and semantic spaces. Yang et al. (2024) conducted a case study of Xuanwu Lake in Nanjing, proposing a method for mining and visualizing cultural imagery related to tourist attractions. By combining cognitive dimensions, such as semantic network analysis of high-frequency terms, with emotional dimensions represented through pie charts and word clouds, they designed a mobile interactive cultural tourism map that enhanced communication within scenic sites. Recent advances in artificial intelligence, particularly deep learning and cross-modal attention mechanisms like Vision-Language Transformers, have provided crucial technical support for this line of research. Building on multimodal information fusion theory, Baltrušaitis et al. (2017) developed a multimodal machine learning classification framework that clarified the technical pathways for cross-modal data processing, including feature extraction, alignment, and fusion. This laid the methodological foundation for multimodal semantic association in geospatial analysis. Similarly, the multimodal convolutional neural network proposed by Srivastava et al. (2019), which integrates map imagery and Google Street View features, achieved semantic consistency across modalities and expanded the framework established by Baltrušaitis and colleagues. Du et al. (2021) applied a multi-scale semantic segmentation network and object-based methods to delineate detailed urban functional zones from remote-sensing imagery, extending multimodal semantic analysis to large-scale spatial planning. Song et al. (2025) combined mobile sensing technology with machine learning to perform semantic segmentation on panoramic images of Suzhou's classical gardens, enabling the evaluation of tourism space quality. This work demonstrates the practical adaptability of multimodal techniques and presents a transferable analytical paradigm for cultural and tourism contexts such as Guilin's karst landscapes and the Jingjiang Prince's Mansion.

In practice, multimodal information fusion has been applied across various fields, including smart cities, public health, and smart tourism. Zhou et al. (2020) proposed a classification model for urban functional zones that confirmed the effectiveness of integrating frequency statistics with convolutional neural networks for multimodal data analysis. Seghieri (2023) designed an optimization framework for medical resource allocation, integrating geospatial, mobility, and healthcare data. This framework also offers a useful reference for coordinating tourism resources, visitor flows, and service infrastructure in Guilin. Yang and Chen (2025) further developed a rural tourism landscape planning model that incorporates geospatial data, social media imagery and text, and online tourist reviews. The resulting three-dimensional evaluation framework, emphasizing ecology, culture, and commerce, aligns well with Guilin’s multi-dimensional development needs, particularly in relation to its karst ecology, cultural heritage, and tourism economy.

Research on multimodal fusion in geospatial information science not only enhances analytical precision but also contributes to holistic social development. It helps resolve practical challenges in fields like public welfare, urban management, and resource utilization, delivering both scientific and societal value. As a vital driver of regional economic growth, the tourism industry features strong spatial attributes, diverse data types, and flexible application scenarios, all of which align closely with the core capabilities of geospatial information science. This synergy not only provides a critical testing ground for multimodal fusion technologies but also stimulates their continuous refinement, forging a mutually reinforcing cycle between technological innovation and application practice.

Even with these advancements, multimodal data fusion still faces several challenges. Dalla Mura et al. (2015) noted that fusing irrelevant data can distort analytical outcomes, highlighting the need to avoid misassociations between non-tourism social media data and tourism-related geospatial data in the context of Guilin. Another issue concerns multimodal data heterogeneity, as discussed by Baltrušaitis et al. (2017), which arises from differences in resolution, source, and consistency. In Guilin's tourism data, this problem is reflected in the disparity between the high resolution of remote-sensing imagery and the unstructured nature of tourist reviews. Han et al. (2025) further emphasized the quality issues present in user-generated content, such as advertisements or malicious reviews, that need to be filtered out during data processing. These challenges represent key directions for improving the construction of Guilin's tourism information corpus, which this study aims to address.

3. RESEARCH BACKGROUND OF THE TOURISM INFORMATION CORPUS

Several factors, including urbanization and the upgrading of tourism consumption, have jointly driven the digital transformation of the tourism industry (Vujko et al., 2025). Within this context, tourism information has gradually come to exhibit characteristics of multi-sourcing and heterogeneity (Cerba et al., 2015; Wang et al., 2022), which has rendered traditional management models insufficient to meet the current demands of smart tourism. In particular, tourists now seek personalized and immersive experiences (Juan et al., 2024), while management departments must monitor tourist flows and consumption preferences (Rogowski, 2020). For example, Hocevar et al. (2021) emphasized that tourism research should focus on regional dimensions like competitiveness and spatial structure, providing a theoretical basis for the regional orientation and focus of tourism corpora. Furthermore, Rogowski's (2020) work on tourist monitoring, along with Juan's (2024) research on route recommendation, introduces demands from both the management and user perspectives. These dual demands have contributed to the evolution of tourism information corpora from a focus on single-text to a multi-modal integration approach. A corpus is defined as a collection of text fragments that reflect specific language variants or usage contexts and address targeted research questions (Clancy & Vaughan, 2023). Corpus linguistics theory forms the foundation for the systematic management of tourism big data (Qin et al., 2019). In the early stages of development, tourism corpora were primarily text-based. For example, Kirilenko et al. (2017) conducted sentiment analysis on tourism reviews, and Xu & Li (2016) performed text mining to assess hotel satisfaction. While these studies successfully applied text data, they lacked the integration of multi-modal and geospatial information. Recently, with the advancement of multimodal technologies, tourism corpora have started incorporating diverse data types, such as images and POIs coordinates. For instance, Maeda et al. (2018) utilized geotagged Twitter and Foursquare data to analyze variations in tourist preferences, establishing a connection between text and geospatial information. Wang et al. (2022) analyzed tourist distribution and sentiment through social media data, further clarifying the value of integrating multi-source data. Mattei (2024) quantified differences in tourism visual communication by classifying corpus data, supplementing the integration of text and visual information. These studies collectively contributed to the development of a multimodal tourism corpus, integrating text, visual information, and geospatial data, providing valuable insights for designing the multimodal data architecture of the Guilin tourism information corpus.

As an essential part of digital technologies, a high-quality tourism information corpus can offer data support and underpin the digital transformation of the tourism industry, propelling the sector from basic informatization to digitalization and intelligentization (Yu, Min, et al., 2024). In tourism informatization, such corpora facilitate the identification of geographic entities and the extraction of relevant information. The 2023 European Common Tourism Data Space document issued by the European Commission highlights the potential of data sharing to break down data silos (European Commission, 2023), providing a policy basis for data integration in tourism corpora. Additionally, relevant text-based collections can support the targeted delivery of tourism products and AI-driven personalized recommendations, generating strategic commercial value (Lan et al., 2022; Shrestha et al., 2024). For example, the Self-AM-V Trans algorithm proposed by Juan et al. (2024) employs Vision Transformer (ViT) to extract image features and Long Short-Term Memory (LSTM) for sequential data encoding, addressing the challenge of linking visual and sequential information. The studies mentioned above have established a comprehensive pathway for the Guilin tourism information corpus to integrate multi-source data and support personalized recommendations. For instance, by integrating images of Guilin's scenic spots, multilingual reviews, and POIs data, the algorithm can generate personalized travel itineraries, such as a combination of a Li River cruise and a visit to Elephant Trunk Hill.

In geographic semantic analysis, multi-modal fusion emphasizes deeper semantic-level associations (Baltrušaitis et al., 2017). Both Chantrapornchai et al. (2018) and Tatane et al. (2024) have noted that ontology-based knowledge representation methods can effectively integrate semantic relationships within tourism corpora, enhancing the accuracy of semantic analysis and refining the processing of tourism texts. These methods can create tight semantic links between elements such as the Guilin landscape, karst landforms, Li River cruise, and Elephant Trunk Hill, offering a new solution to address semantic disconnection among natural landscapes, cultural heritage, and tourist behavior in the Guilin tourism context. This approach will improve the accuracy of the corpus's semantic analysis. For example, Malamatidou (2024) analyzed a multilingual tourism corpus and found that English texts focus on natural experiences, while French texts emphasize cultural atmosphere. This difference provides a basis for developing cross-cultural tourism destinations. In the context of Guilin's positioning as an international tourist city, such corpus analysis can identify preference differences between English-speaking tourists, who focus on karst landscapes, and Japanese-speaking tourists, who are more interested in folk culture, thereby enabling targeted optimization of cross-cultural tourism services.

As an international tourist city within the core karst region, Guilin has complex and diverse tourism information. However, its decentralized management model has limitations in effectively managing this information. For example, Lan et al. (2022) studied tourist reviews of high-star hotels in Guilin and found that multilingual reviews were not linked to geospatial data, hindering the precise analysis of tourist preferences. Additionally, unstructured data, such as videos of folk activities and texts about landscape legends, lacks systematic integration (Cao et al., 2023), making it difficult to design all-inclusive tourism products. Therefore, it is crucial to use corpus technology as the core, integrating the geospatial-text fusion logic proposed by Maeda et al. (2018) with the semantic association method developed by Tatane et al. (2024) to consolidate Guilin's multi-source tourism information. This integration will enable the deep coupling of multi-modal data, geospatial information, and semantic analysis, driving the high-quality development of Guilin's tourism industry.

4. APPLICATION OF VISUALIZATION TECHNOLOGY IN THE INTERSECTION OF GEOGRAPHY AND TOURISM

Visualization technology serves as a powerful tool for extracting critical information and interpreting multimodal spatio-temporal data. It enables the real-time display of dynamic data, supports cross-platform collaborative analysis, and plays a crucial role in both analytical and communicative aspects of geovisualization and scientific visualization.. Visualization technology, defined by three core features including cross-platform compatibility, interactivity, and real-time performance, has significantly advanced the way geospatial data is analyzed (Mwalongo et al., 2016). The Earth's inherent attributes pertaining to three-dimensional geometric features and four-dimensional temporal characteristics provide the foundational logic for the development of geovisualization and geovisual analytics (Fitzpatrick & Hedley, 2023). Historically, Andrienko & Andrienko (2007) emphasized the integration of computational analysis within visualization technology to overcome the limitations of traditional methods. This shift has driven the evolution of geovisualization from static displays to dynamic analytics, a trend that persists to this day (Carbonell-Carrera & Hess-Medler, 2019). With the advancement of massive data collection and computing capabilities, visualization technology has integrated a range of disciplines, including cartography, scientific visualization, data analysis, and geospatial information science, reinforcing its central role in these fields (Fitzpatrick & Hedley, 2023).

Over the past seventy years, the development of geovisualization and modeling technologies has been substantial. The introduction of GIS in the 1990s laid the groundwork for dynamic spatial interaction of geospatial data (Kraak, 2003), and the extensive use of Synthetic Aperture Radar (SAR) over the last two decades has significantly improved data acquisition precision. Traditional cartographic models, often challenged by accuracy issues and difficulties in conveying complex geospatial information, have benefited from the integration of GIS and SAR technologies, which has driven the widespread adoption of three-dimensional geovisualization (Juřík et al., 2020). Earth scientists have expanded data sources via various technical methods, broadening the application scope of geovisualization. High-resolution topographic and geomorphic data, acquired through photogrammetry and Light Detection and Ranging (LiDAR) technologies have provided high-precision and multi-type data for geovisual analytics. This field has focused on the development of interactive spatial analysis methods to address the challenges of analyzing three- and four-dimensional geospatial data (Robinson, 2017). This expanded application of geovisualization has been demonstrated through case studies. For example, Maio et al. (2013) used geospatial visualization techniques to reconstruct the historical landscape of the American Revolutionary War battlefield. Tyner et al. (2018) employed spatial heat maps to analyze the geographical impacts of the Cambodian genocide. These studies validate geovisualization's ability to restore spatial contexts and illustrate distributions, providing a replicable framework for practices in Guilin, such as visualizing historical scenes at the Jingjiang Prince's Mansion and spatially mapping tourist flows within scenic areas.

As the field of geographic visualization has developed, researchers have gained valuable insights from geographic visualization software and emerging hardware devices (Rhyne et al., 2006). High-resolution datasets and faster computing speeds have further advanced these technologies (Mitasova et al., 2012). Big data analytics platforms, such as Hadoop and Spark, allow for rapid cleansing and correlation of multi-source data, which, when paired with visualization tools such as ArcGIS Pro and QGIS, creates a robust data processing and visualization workflow. The advancements in sensing technologies like LiDAR, photogrammetry, and others have generated massive datasets, driving the need for enhanced processing capabilities in 3D and 4D environments (Havenith et al., 2019). The evolution of sensor technologies, including satellite remote sensing, UAV-based aerial photography, and IoT terminals, has broadened the scope of geospatial information, incorporating imagery, text, and real-time data (Fitzpatrick & Hedley, 2023). However, as Afzal et al. (2019) noted, large-scale data processing and the selection of suitable visualization tools remain significant challenges. These constraints have led to the development of technology-enabled platforms that integrate multiple tools, enabling seamless presentation of multi-modal data across various scales (Lovelace et al., 2016). Multi-window dashboard-based visualization architectures have been developed, leveraging different visualization technologies to present complex information synchronously across multiple dimensions (Hollberg et al., 2021). This approach holds great potential in fields such as geological exploration and tourist guidance.

In tourism, visualization technologies can support the development of intelligent management systems by processing multi-source operational data. These systems can facilitate knowledge discovery and provide targeted support for both tourists and managers (Qin & Pan, 2023). Case studies have shown the feasibility of such systems across different regions. For example, in Pati Regency, Indonesia, the local government utilized QGIS software to digitize and classify spatial data, which was then imported into a WebGIS system and used to create an interactive tourism map. This system provided effective decision-making support for developers, government agencies, and tourism professionals (Saputro et al., 2024). Similarly, Marieta Šoltésová et al. (2025) used GIS-based visualization, integrating data acquisition and spatial analysis technologies, to analyze tourism patterns and dynamics. This work has proven the practical value of visualization technologies in tourism planning and sustainable destination management. Furthermore, Amadu et al. (2025) advanced the field by integrating mobile navigation systems, Mixed Augmented Reality (MAR), and animation, creating a comprehensive application that allows tourists to interact with the system in real-time without relying on physical maps or needing directions. The technical framework described above is particularly suitable for tourism destinations like Guilin, which possess unique ecological and cultural features. By integrating spatial, visual, and audio modalities, such as real-time navigation of bamboo rafting on the Yulong River and audio guides for Zhuang ethnic cultural fairs, Guilin can enhance the visitor experience and provide managers with real-time insights into visitor distribution and ecological conservation needs. These insights will contribute to the balanced development of tourism, culture, and ecology in Guilin.

As mentioned earlier, geographic visualization has evolved significantly, going beyond simple visual communication or the use of digital technologies. Empirical research shows that geographic visualization can improve spatial thinking and decision-making (Carbonell-Carrera & Hess-Medler, 2019). Practitioners can tailor designs to specific geoscientific phenomena and dynamic user needs, contributing to more effective tourism marketing (Qin & Pan, 2023). These advancements not only support data communication and geovisual analysis but also facilitate data exchange among key stakeholders, such as government officials, the general public, and experts (Jacquinod & Julia Bonccorsi, 2019).

Despite these advancements, challenges remain, especially in tourism cities like Guilin, where complex terrain and cultural diversity create operational hurdles. Fitzpatrick & Hedley (2023) noted that large-scale data processing and 3D modeling remain significant challenges. Traditional technologies often struggle with the volume of tourism data, and the irregular topography of karst landscapes complicates 3D modeling. However, innovations including Neural Radiance Fields (NeRF) and Non-Photorealistic Rendering (NPR) have unlocked new possibilities for modeling dynamic landscapes. NeRF can accurately reconstruct the 3D features of karst landforms, while NPR enhances the aesthetic qualities of Guilin's scenery. Together, these technologies provide an effective solution for 3D modeling and dynamic landscape visualization, addressing the unique challenges faced by Guilin.

5. SYNERGISTIC LOGIC AND PRACTICAL CHALLENGES OF MULTIMODALITY, CORPUS, AND VISUALIZATION

Data serves as a core resource for all organizations (Stienmetz, 2024). In the tourism domain, it is further integral as the foundational pillar of smart tourism (Gretzel et al., 2015). The synergy among multimodal technologies, corpora, and visualization plays a pivotal role in unlocking the value of tourism data. Specifically, multimodal technologies have addressed the challenge of integrating different modes of geo-tourism data (Baltrušaitis et al., 2017; Kim et al., 2021). Corpora function as foundational carriers for storing multimodal data and enabling semantic association (Juan et al., 2024; Tatane et al., 2024). Visualization technologies provide intuitive presentations of integrated data, supporting cognitive transformation (Qin & Pan, 2023; Šoltésová et al., 2025). Together, these components form a cohesive system for data integration, storage association, and visualization output, propelling geo-tourism research beyond traditional qualitative descriptions toward quantitative visualization and interactive experiences. Han et al. (2025) noted that multi-modal spatio-temporal data visualization technology is integrated throughout the full lifecycle of landscape projects, spanning from investigation and analysis to planning and design, as well as operation and management. This application logic provides a framework for the Guilin tourism information corpus, covering the entire process from data collection to visualization deployment. In 2025, the University of Central Florida (UCF) introduced the GAEA model (Geolocation Aware Conversational Model), which for the first time integrates precise geolocation with conversational capabilities. By integrating geospatial functionality and multimodal dialogue, this model offers valuable insights for enhancing the visualization interaction of the Guilin tourism information corpus. For instance, when tourists inquire about Guilin's karst landscapes through conversational interfaces, the corpus automatically associates remote sensing imagery with POIs, presenting them in a visualized format.

At present, most domestic studies have applied multimodal data fusion technology to specific aspects of urban tourism. However, a systematic and integrated visualized tourism information corpus grounded in multimodal geospatial information fusion remains underdeveloped. As a key international tourist city and a renowned historical and cultural city in China, Guilin boasts abundant tourism resources, diverse scenic areas, and favorable geographical conditions. The circular tourism zone centered on Guilin extends outward to surrounding regions, covering over 20,000 square kilometers (Cao et al., 2023), offering a unique scenario for multimodal geospatial information fusion. The visualized Guilin tourism information corpus developed in this study addresses this research gap, establishing a reusable fusion paradigm defined by geospatial and cultural-tourism multimodality. It aims to overcome the limitations of traditional visualization, which only presents a single data dimension. Specifically, it integrates Guilin's multi-modal information, including topographic remote sensing data, 360-degree VR panoramas of scenic areas, real-time positioning data for Li River cruise ships, Gui opera performance footage, and tourist review texts, with geospatial coordinates as the core anchoring reference. This integration will result in a corpus system that creates dynamic linkages between spatial location, multimodal content, and user needs, offering foundational support for the precise implementation of smart tourism. It facilitates the transformation of tourism research from a single data-driven paradigm to one driven by multi-modal cognition, while simultaneously opening new avenues for interdisciplinary inquiry, such as human geography and data visualization.

However, the field still faces multiple core challenges that require targeted solutions tailored to the tourism context of Guilin. These challenges include multi-dimensional hurdles in corpus construction, where tourism information corpora must address issues related to multi-source heterogeneous data integration, semantic consistency maintenance, and visualization readability (Han et al., 2025). The approach proposed by Różycki et al. (2024), which involves constructing geospatial vector layers using QGIS and ArcGIS Pro, can be applied to integrate heterogeneous data, such as historical and current remote sensing imagery and scenic area POIs. Additionally, it is essential to apply the ontological knowledge representation method developed by Tatane et al. (2024) to ensure semantic consistency across Guilin's landscapes, karst landforms, and folk customs, thus mitigating discrepancies in data association. In terms of accurately matching unstructured data with geospatial information, unstructured data such as landscape legend texts and folk activity videos are difficult to directly anchor to geospatial coordinates. The SOS-Match framework proposed by Thomas et al. (2024), which extracts target masks through a zero-shot segmentation model and achieves frame alignment based on geometric consistency, can be adapted for precise matching between folk activity videos (e.g., the March 3rd Song Festival) and their respective locations (e.g., Yangshuo West Street), addressing the spatial association challenge for unstructured data. Furthermore, in addressing the balance between privacy protection and visualization, for Guilin tourism's Volunteered Geographic Information (VGI), including location-tagged photos uploaded by tourists, the HyperLogLog (HLL) data abstraction format proposed by Dunkel et al. (2020) can preserve data utility while safeguarding privacy, making it suitable for visualization processing of Guilin's specific VGI. Regarding the need for visualization design to balance information richness and simplicity, the geometric simplification and layered rendering strategy developed by Dübel et al. (2017) can be applied to the design of Guilin's comprehensive tourism map. For example, scenic area distribution can be shown on the bottom layer, tourist flow on the middle layer, and review sentiment on the top layer. This strategy enhances information density while preventing visual clutter.

ACKNOWLEDGMENTS

The project was funded by Guangxi Key Laboratory of Spatial Information and Geomatics (Grant No. 21-238-21-11).

REFERENCES

[1] Afzal, S., Hittawe, M. M., Ghani, S., Jamil, T., Knio, O., Hadwiger, M., & Hoteit, I. (2019). The state of the art in visual analysis approaches for ocean and atmospheric datasets. Computer Graphics Forum, 38(3), 882–907.

[2] Amadu, F. O., Nhamo, L., Benzougagh, B., & Turyasingura, B. (2025). Application of geographic information system in ecotourism: a global bibliometric analysis. Cogent Social Sciences, 11(1), 2460711.

[3] Andrienko, N., & Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. Cartographica, 42(2), 117-138. https://doi.org/10.3138/carto.42.2.117

[4] Arora, S., Pargaien, S., Khan, F., Tewari, I., Nainwal, D., Mer, A., Mittal, A., & Misra, A. (2023). Monitoring tourist footfall at Nainital in Uttarakhand using sensor technology. In Proceedings of the Fourth International Conference on Electronics and Sustainable Communication Systems (ICESC-2023) (pp. 200-204). IEEE.

[5] Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2017). multi-modal machine learning: A survey and taxonomy. arXiv Preprint arXiv:1705.09406v2 [cs.LG].

[6] Bielański, M., Taczanowska, K., Muhar, A., Adamski, P., González, L.-M., & Witkowski, Z. (2018). Application of GPS tracking for monitoring spatially unconstrained outdoor recreational activities in protected areas – A case study of ski touring in the Tatra National Park, Poland. Applied Geography, 96, 51–65.

[7] Carbonell-Carrera, C., & Hess-Medler, S. (2019). 3D Landform Modeling to Enhance Geospatial Thinking. ISPRS International Journal of Geo-Information, 8(2), 65.

[8] Cao, R., Mo, Y., & Ma, J. (2023). Carbon footprint analysis of tourism life cycle: The case of Guilin from 2011 to 2022. Sustainability, 15(9), 7124.

[9] Cerba, O., Jedlicka, K., Janecka, K., Mildorf, T., Fryml, J., Vlach, P., ... & CharvÃt, K. (2015). Integration and Visualization of Tourism Data.

[10] Chantrapornchai, C., Bunlaw, N., & Choksuchat, C. (2018). Semantic image search: Case study for Western Region Tourism in Thailand. Journal of Information Processing Systems, 14(5), 1195–1214.

[11] Clancy, B., & Vaughan, E. (2023). Using corpus linguistics to interpret economic news texts. In [Book Title] (pp. [Chapter Page Range]). Routledge.

[12] Dalla Mura, M., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., & Benediktsson, J. A. (2015). Challenges and opportunities of multi-modality and data fusion in remote sensing. Proc. IEEE, vol. 103, no. 9, pp. 1585–1601, Sep. 2015.

[13] Dunkel, A., Löchner, M., & Burghardt, D. (2020). Privacy-aware visualization of volunteered geographic information (VGI) to analyze spatial activity: A benchmark implementation. ISPRS International Journal of Geo-Information, 9(10), 607.

[14] Dübel, S., Röhlig, M., Tominski, C., & Schumann, H. (2017). Visualizing 3D terrain, geo-spatial data, and uncertainty. Informatics, 4(1), 6.

[15] Du, S., Du, S. H., Liu, B., & Zhang, X. Y. (2021). Mapping large-scale and fine-grained urban functional zones from VHR images using a multi-scale semantic segmentation network and object based approach. Remote Sensing of Environment, 261, 112480.

[16] European Commission. (2023). Towards a common European tourism data space: Boosting data sharing and innovation across the tourism ecosystem (Communication No. 2023/C263/01). Official Journal of the European Union, C263(1), 1-13.

[17] Fitzpatrick, M., & Hedley, N. (2023). Review of the state of practice in geovisualization in the geosciences. Frontiers in Earth Science,

[18] Furgale, P., Rehder, J., & Siegwart, R. (2013). Unified temporal and spatial calibration for multi-sensor systems. In Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1280–1286). IEEE.

[19] Gao, Y., & Schmöcker, J.-D. (2021). Estimation of walking patterns in a touristic area with Wi-Fi packet sensors. Transportation Research Part C: Emerging Technologies, 128, 103219. https://doi.org/10.1016/j.trc.2021.103219

[20] Gill, L., Lange, E., Morgan, E., & Romano, D. (2013). An analysis of usage of different types of visualisation media within a collaborative planning workshop environment. Environment and Planning B: Planning and Design, 40(4), 742-754.

[21] Gretzel, U., Werthner, H., Koo, C., & Lamsfus, C. (2015). Conceptual foundations for understanding smart tourism ecosystems. Computers in Human Behavior, 50(1), 558-563.

[22] Guerin, A., Stock, G. M., Radue, M. J., Jaboyedoff, M., Collins, B. D., Matasci, B., Avdievitch, N., & Derron, M.-H. (2020). Quantifying 40 years of rockfall activity in Yosemite Valley with historical Structure-from-Motion photogrammetry and terrestrial laser scanning. Geomorphology, 356, 107069.

[23] Han, X., Li, Z., Cao, H., & Hou, B. (2025). multi-modal spatio-temporal data visualization technologies for contemporary urban landscape architecture: A review and prospect in the context of smart cities. Land, 14(5), 1069.

[24] Havenith, H.-B., Cerfontaine, P., & Mreyen, A.-S. (2019). How virtual reality can help visualise and assess geohazards. International Journal of Digital Earth, 12(2), 173–189.

[25] Hocevar, M., & Bartol, T. (2021). Mapping urban tourism issues: analysis of research perspectives through the lens of network visualization. International Journal of Tourism Cities, 7(3), 818-844.

[26] Hollberg, A., Kiss, B., Röck, M., Soust-Verdaguer, B., Houlihan Wiberg, A., Lasvaux, S., Galimshina, A., & Habert, G. (2021). Review of visualising LCA results in the design process of buildings. Building and Environment, 190, 107530.

[27] Jacquinod, F., & Bonaccorsi, J. (2019). Studying social uses of 3D geovisualizations: Lessons learned from action-research projects in the field of flood mitigation planning. ISPRS International Journal of Geo-Information, 8(2), 84.

[28] Juan, Z., Zhang, J., & Gao, M. (2024). A multi-modal travel route recommendation system leveraging visual Transformers and self-attention mechanisms. Frontiers in Neurorobotics, 18, 1439195.

[29] Jurado Rota, J., Pérez Albert, M. Y., & Serrano Giné, D. (2019). Visitor monitoring in protected areas: An approach to Natura 2000 sites using Volunteered Geographic Information (VGI). Geografisk Tidsskrift-Danish Journal of Geography,

[30] Juřík, V., Herman, L., Snopková, D., Galang, A. J., Stachoň, Z., Chmelík, J., Kubíček, P., & Šasinka, Č. (2020). The 3D hype: Evaluating the potential of real 3D visualization in geo-related applications. PLOS ONE, 15(5), e0233353.

[31] Jovanov, E. (2019). Wearables meet IoT: Synergistic personal area networks (SPANs). Sensors, 19(19), 4295.

[32] Kim, W., Son, B., & Kim, I. (2021). ViLT: Vision-and-Language Transformer without convolution or region supervision. In Proceedings of the 38th International Conference on Machine Learning (pp. 5583-5594). PMLR.

[33] Kirilenko, Andrei P., Svetlana O. Stepchenkova, Hany Kim, and Xiang (Robert) Li. (2017). "Automated Sentiment Analysis in Tourism: A Comparison of Approaches." Journal of Travel Research 0 (0): 1-14.

[34] Kraak, M.-J. (2003). Geovisualization illustrated. ISPRS Journal of Photogrammetry & Remote Sensing, 57(4), 390–399.

[35] Lahat, D., Adali, T., & Jutten, C. (2015). multi-modal data fusion: An overview of methods, challenges, and prospects. Proceedings of the IEEE, 103(9), 1449–1477.

[36] Lan, Y., Huang, Y., Yao, J., & Lu, Y. (2022). Research on comments preference based on online comments of tourists from high-star hotels in Guilin. In Proceedings of the 2022 International Conference on Intelligent Data and Communication Technologies (ICIDC 2022). EAI.

[37] Lee, J.-G., & Kang, M. (2015). Geospatial big data: Challenges and opportunities. Big Data Research, (00), 1–10.

[38] Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). A review of text corpus-based tourism big data mining. Applied Sciences, 9(16), 3300.

[39] Loidl, M., Wallentin, G., Cyganski, R., Graser, A., Scholz, J., & Haslauer, E. (2016). GIS and transport modeling—Strengthening the spatial perspective. ISPRS International Journal of Geo-Information, 5(6), 84.

[40] Lovelace, R., Birkin, M., Cross, P., & Clarke, M. (2016). From big noise to big data: Toward the verification of large data sets for understanding regional retail flows. Geographical Analysis, 48(1), 59–81. https://doi.org/10.1111/gean.12081

[41] Lu, J. (2022). Personalized Recommendation Algorithm of Smart Tourism Based on Cross-Media Big Data and Neural Network. Computational Intelligence and Neuroscience, 2022, 9566766.

[42] Madden, K., Ramsey, E., Loane, S., & Condell, J. (2021). Trailgazers: A scoping study of footfall sensors to aid tourist trail management in Ireland and other Atlantic areas of Europe. Sensors, 21(6), 2038.

[43] Maeda, T. N., Yoshida, M., Toriumi, F., & Ohashi, H. (2018). Extraction of tourist destinations and comparative analysis of preferences between foreign tourists and domestic tourists on the basis of geotagged social media data. ISPRS International Journal of Geo-Information, 7(3), 99.

[44] Maio, C. V., Tenenbaum, D. E., Brown, C. J., Mastone, V. T., & Gontz, A. M. (2013). Application of geographic information technologies to historical landscape reconstruction and military terrain analysis of an American Revolution Battlefield: Preservation potential of historic lands in urbanized settings, Boston, Massachusetts. Journal of Cultural Heritage, 14, 317–331.

[45] Malamatidou, S. (2024). Translating tourism: Cross-linguistic differences of alternative worldviews. Palgrave Macmillan.

[46] Mattei, E. (2024). Approaching tourism communication with empirical multi-modality: Exploratory analysis of Instagram and website photography through data-driven labeling. Frontiers in Communication, 9, 1355406.

[47] Miller, A. B., Leung, Y.-F., & Kays, R. (2017). Coupling visitor and wildlife monitoring in protected areas using camera traps. Journal of Outdoor Recreation and Tourism, 17, 44–53.

[48] Mitasova, H., Harmon, R. S., Weaver, K. J., Lyons, N. J., & Overton, M. F. (2012). Scientific visualization of landscapes and landforms. Geomorphology, 137, 122-137.

[49] Müller, H., & Würriehausen, F. (2015). Integration of authoritative and volunteered cultural landscape information. In Proceedings of FIG Working Week 2015: From the Wisdom of the Ages to the Challenges of the Modern World (pp. 1-10). Sofia, Bulgaria: International Federation of Surveyors (FIG).

[50] Mwalongo, F., Krone, M., Reina, G., & Ertl, T. (2016). State-of-the-Art Report in Web-based Visualization. Computer Graphics Forum, 35(3), 554-575.

[51] Qin, Z., & Pan, Y. (2023). Design of a smart tourism management system through multisource data visualization-based knowledge discovery. Electronics, 12(3), 642.

[52] Ramachandram, D., & Taylor, G. W. (2017). Deep multi-modal learning: A survey on recent advances and trends. IEEE Signal Processing Magazine, 34(6), 96–108.

[53] Robinson, A. C. (2017). Geovisual analytics. In J. P. Wilson (Ed.), The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2017 Edition).

[54] Romeo, S., Cosentino, A., Giani, F., Mastrantoni, G., & Mazzanti, P. (2021). Combining ground based remote sensing tools for rockfalls assessment and monitoring: The Poggio Baldi landslide natural laboratory. Sensors, 21(8), 2632.

[55] Rogowski, M. (2020). Monitoring System of tourist traffic (MSTT) for tourists monitoring in mid-mountain national park, SW Poland. Journal of Mountain Science, 17(8), 2035-2047.

[56] Różycki, S., Michalski, M., & Kobielec, A. (2024). Use of geoinformatics for the digitization and visualization of sensitive space in the urban landscape: A case study of the Gross-Rosen sub-camps systems. Remote Sensing, 16(5), 783.

[57] Rhyne, T.-M., MacEachren, A., & Dykes, J. (2006). Guest editors' introduction: Exploring geovisualization. IEEE Computer Graphics and Applications, 26(4), 20–21.

[58] Saputro, M. F. A., Herlambang, B. A., & Anam, A. K. (2024). Sistem informasi geografis peta interaktif pariwisata kabupaten pati berbasis sistem. Jurnal Ilmiah Research Student, 1(3), 767-776.

[59] Seghieri, M. (2023). Innovative applications of GIS-based methods in healthcare planning and resource allocation. Journal of Remote Sensing & GIS, 12(2), 1000286.

[60] Sehrawat, D., & Gill, N. S. (2020). IoT based human activity recognition system using smart sensors. Advances in Science, Technology and Engineering Systems Journal, 5(4), 516–522.

[61] Shaw, S.-L., Yu, H., & Bombom, L. S. (2008). A space-time GIS approach to exploring large individual-based spatiotemporal datasets. Transactions in GIS, 12(4), 425–441.

[62] Shrestha, D., Wenan, T., Shrestha, D., Rajkarnikar, N., & Jeong, S.-R. (2024). Personalized Tourist Recommender System: A Data-Driven and Machine-Learning Approach. Computation, 12(3), 59.

[63] Sitarz, M., Bukowski, M., Zięba, S., & Zwijacz.,ozica, T. (2018). Monitoring of illegal human activities in the Tatra National Park, Poland. In J. Dehez (Ed.), The 9th International Conference on Monitoring and Management of Visitors in Recreational and Protected Areas (MMV9) Abstract Book (pp. 342). Bordeaux, France: Bordeaux Sciences Agro, Irstea.

[64] Smith, M. (2022). Treatise on geomorphology. Available at: http://www.sciencedirect.com:5070/referencework/9780128182352/treatise-on-geomorphology [book]

[65] Song, H., Chen, J., & Li, P. (2025). Decoding the cultural heritage tourism landscape and visitor crowding behavior from the multidimensional embodied perspective: Insights from Chinese classical gardens. Tourism Management, 110, 105180.

[66] Stienmetz, J. L. (2024). Factors influencing the data sharing practices of European tourism stakeholders: Considerations for a sustainable common European tourism data space. Journal of Smart Tourism, 4(3), 35-43.

[67] Šoltésová, M., Iannaccone, B., Štrba, L'., & Sidor, C. (2025). Application of GIS technologies in tourism planning and sustainable development: A case study of Gelnica. ISPRS International Journal of Geo-Information, 14(3), 120.

[68] Srivastava, S., Vargas-Muñoz, J. E., & Tuia, D. (2019). Understanding urban landuse from the above and ground perspectives: A deep learning, multi-modal solution. Remote Sensing of Environment, 228, 129–143.

[69] Tang, R., Huang, C., Zhao, X., & Tang, Y. (2022). Research on smart tourism oriented sensor network construction and information service mode. Sensors, 22(24), 10008.

[70] Tatane, K., Amalki, A., & Bouzit, A. (2024). Enhancement of semantic analysis based on the ontology reengineering: A case study on Moroccan tourism domain. Journal of Computer Science, 20(8), 872–884.

[71] Thomas, A., Kinnari, J., Lusk, P. C., Kondo, K., & How, J. P. (2024). SOS-Match: Segmentation for open-set robust correspondence search and robot localization in unstructured environments. arXiv preprint arXiv:2401.04791v3.

[72] Tyner, J. A., Curtis, A., Kimsroy, S., & Chhay, C. (2018). The evacuation of Phnom Penh during the Cambodian Genocide: Applying spatial video geonarratives to the study of genocide. Genocide Studies and Prevention: An International Journal, 12(3), 163–176.

[73] Vujko, A., Knežević, M., & Arsić, M. (2025). The Future Is in Sustainable Urban Tourism: Technological Innovations, Emerging Mobility Systems and Their Role in Shaping Smart Cities. Urban Science, 9(5), 169.

[74] Wang, J., Xia, Y., & Wu, Y. (2022). Sensing tourist distributions and their sentiment variations using social media: Evidence from 5A scenic areas in China. ISPRS International Journal of Geo-Information, 11(9), 492.

[75] Westoby, M., Lim, M., Hogg, M., Dunlop, L., Pound, M., Strzelecki, M., & Woodward, J. (2020). Decoding complex erosion responses for the mitigation of coastal rockfall hazards using repeat terrestrial LiDAR. Remote Sensing, 12(16), 2620.

[76] World Economic Forum & Kearney. (2025). Future of Travel and Tourism: Embracing Sustainable and Inclusive Growth [Briefing Paper]. World Economic Forum.

[77] Xu, X., & Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. International Journal of Hospitality Management, 55, 57-69.

[78] Xing, W. (2024). Leveraging GIS for sustainable tourism development: A comprehensive spatial approach. In Proceedings of the 2nd International Conference on Machine Learning and Automation.

[79] Yang, H., Shen, J., Hong, S., & Zhu, F. (2024). Mining and visualization of tourism cultural image based on the information transmission model of tourism cultural map—Taking Nanjing Xuanwu Lake tourist attraction as an example. ISPRS International Journal of Geo-Information, 13(11), 391.

[80] Yang, Z., & Chen, J. (2025). Optimization of tourism scenic area view planning and design based on multi-modal fusion. Comput. Aided Des. Appl., vol. 22, no. S3, pp. 201–214.

[81] Yu, M., Ma, B., Liu, D., & Zhang, A. (2024). Is the digital economy empowering high-quality tourism development? A theoretical and empirical research from China. PLOS ONE, 19(5), e0303087.

[82] Zhou, W., Ming, D., Lü, X., Zhou, K., Bao, H., & Hong, Z. (2020). SO–CNN based urban functional zone fine division with VHR remote sensing image. Remote Sensing of Environment, 236, 111458.