

Construction of Anhui Music Culture Immersive Experience System Supported by Virtual Reality Technology

Xiaobing Zhang, Meng Ning

Music Conservatory, Huainan Normal University, Huainan 232038, Anhui, China

Corresponding author: Xiaobing Zhang (e-mail: bingbing12290@163.com)

ABSTRACT To address decontextualization and insufficient participation in the dissemination of Anhui music culture, this study constructs a virtual reality-based immersive experience system. First, a multimodal resource library integrating music, movement, and spatial scenes is established through field investigation and high-fidelity sampling. Second, ancient opera stages and representative cultural spaces are reconstructed in a virtual environment, and spatial audio is implemented in the Unity engine to achieve consistency between sound and scene. Third, gesture recognition is introduced to design a three-level progressive interaction mode of “travel-viewing-learning to sing,” enabling users to move from passive observation to embodied participation. Finally, adaptive rendering and cloud streaming technologies are adopted to support cross-platform dissemination. Experimental results show that the VR group has higher θ/α wave ratio (0.85) and skin conductance response ($0.42 \mu\text{S}$) than the video group, while memory retention reaches 86.0%. The system frame rate reaches 89 fps, indicating good immersion, learning effectiveness, and technical usability.

INDEX TERMS virtual reality, anhui music culture, immersive experience, spatial audio, digital cultural protection

I. INTRODUCTION

ANHUI music culture, as an important part of the Huizhou style and charm, encompasses national intangible cultural heritage such as Huizhou opera and Chaohu folk songs. This cultural landscape carries profound historical memories and regional sentiments [1,2]. However, this type of intangible cultural heritage music currently faces many difficulties in its inheritance and dissemination: on the one hand, traditional recording methods are mainly based on two-dimensional audio and video, and music is artificially separated from its original cultural space (such as ancient stages, ancestral halls, and village environments), resulting in the loss of cultural context; on the other hand, the audience's acceptance mode is singular, mostly passive viewing, lacking a sense of embodied participation and presence, making it difficult to evoke deep emotional resonance [3]. The above-mentioned problems of “decontextualization” and “lack of participation” seriously restrict the effective dissemination and living inheritance of Anhui music culture, and it is urgent to explore new technical means and cultural presentation methods.

In recent years, digital technology has provided new possibilities for the protection of intangible cultural heritage. Hariri used literature research and content analysis methods to explore the important significance of the digitization of

personal cultural heritage archives for the legal protection of archive intellectual property rights, so as to promote the preservation and maintenance of national cultural heritage [4]. Zhao proposed a digital protection method for intangible cultural heritage based on customized three-dimensional convolutional neural networks, which is used to accurately identify, classify and archive various forms of intangible cultural heritage [5]. Wei analyzed the integration and development of network media and intangible cultural heritage from the perspective of inheritance and dissemination, in order to address the challenges faced by the protection and inheritance of intangible cultural heritage in the context of globalization [6]. Adewumi reviewed the protection of intangible cultural heritage under the impact of globalization and social transformation, and took Nigerian traditional crafts as an example to explore the significance of its maintenance and inheritance for national identity and future development [7]. Xu pointed out that the folk songs of the She ethnic group in Ningguo, Anhui, have formed a musical form of integration and continuous reconstruction in more than 100 years of development, reflecting the influence of the common beliefs and cultural identity of the ethnic group [8]. Tong and Hee-Gyun explored the optimization path of the development and protection of intangible cultural heritage from the perspective of digitalization and

augmented reality, with the help of big data and advanced technology, so as to provide a reference for related practices [9]. However, existing methods mostly remain at the level of planar integration of audiovisual information, and have failed to truly build an interactive, sound-scene integrated cultural space. Users are still in the role of “observers” in the experience and cannot obtain an immersive and interactive experience.

In order to solve the above problems, this paper proposes to build an Anhui music culture immersive experience system based on virtual reality technology that integrates the technical heritage of smart weaving and high-performance fiber artistry. Compared with the previous researches, the digital original cultural space of the music is sampled besides the sampling of music, not only, but also. The music is integrated into virtual acoustic environment by spatial audio algorithm, forming the auditory field. At the same time, gesture recognition and head-mounted display are introduced, and three-level interactive module is designed to let the audience “travel–watch–learn to sing” in the virtual space, and turn from passive observer into active participant. The core of the method innovation is to present the music in the complete cultural environment, destroy the one-way communication mode of traditional media through the embodied interaction, and solve the problems of context fragmentation and participation absence in the transmission of the intangible cultural heritage music.

The main contributions of this paper are as follows: First, it constructs an immersive experience system that includes an Anhui music resource database, a virtual cultural space, and a natural interaction mechanism, providing a new paradigm for the digital protection of regional intangible cultural heritage music; Second, through multimodal physiological measurement and behavioral analysis experiments, it systematically verifies the advantages of this method in creating immersion, maintaining learning effects, and system usability; Third, it forms a reproducible technical route, providing a reference for the immersive dissemination of music culture in other regions.

II. DIGITAL EXTRACTION OF CULTURAL GENES

In order to make the Anhui music culture reflected in the virtual experience system true and accurate, we carried out fieldwork in Huizhou, Anqing and Chaohu Lake core cultural ecological areas. We collected representative works such as Huizhou opera “The Flood of the Seven Armies”, Qingyang opera “The Brocade Weaving Story” and Chaohu folk song “The Sisters-in-Law’s Flower-Sharing Song”. We also recorded the traditional musical instruments used in the above performances (Huizhou fiddle and tanggu drum), costume and prop, as well as the ancient stage style.

In the process of audio and video acquisition, we used Blackmagic URSA Mini Pro 12K high-definition camera to record the performance scene from the panoramic view of the 360 degree view. In addition, we used omnidirectional condenser microphones to record the details of actors’

singing and the band accompaniment. In order to retain the sound field information of music in its original space, we collected the impulse response data of the ancient stage and ancestral hall at the same time as the ancient stage and ancestral hall at the time of recording, which will be used for convolutional reverberation processing of spatial audio after. For motion capture, we used NOKOV optical motion capture system to record the inheritor’s movement, gesture and facial micro-expression at a sampling rate of 120Hz. A total of 40 marker points were attached to the head, torso and extremities of the key parts of the body. In addition, we focused on capturing stylized movements such as “sleeve shaking” and “posing” in Huizhou opera and rhythmic movement in Chaohu folk song performance. This ensures that the movement of virtual characters is natural, smooth and conforms to the norms of opera.

III. CONSTRUCTION OF A VIRTUAL REALM INTEGRATING SOUND AND SCENERY

According to the previously collected 3D point cloud and oblique photogrammetry data, the typical cultural places in Anhui Province were reconstructed in the Unity 3D engine. To achieve a balance between geometric precision and visual realism, we employed a hybrid rendering approach: the structural framework was modeled using AutoCAD and 3ds Max based on point cloud dimensions, while the high-definition 12K panoramic video data was processed into high-resolution PBR (Physically Based Rendering) textures and skyboxes to wrap the 3D assets, ensuring that the synthetic environment maintains the photorealistic aesthetic of the original fieldwork.

The configuration of background ambient sound is carefully layered to preserve acoustic integrity. While primary cultural sounds are processed via convolutional reverberation from site-specific impulse responses to maintain ‘acoustic truth,’ secondary environmental cues such as wind chimes or distant animal sounds are limited to the ‘exploration layer’ only. These are implemented as separate spatial audio objects with attenuation curves that do not interfere with the pure impulse response data of the musical performances, thus providing environmental context without compromising the historical acoustic profile of the ancestral hall. The synchronization of vision and hearing space and time is the basis of the integration of soundscape. This research adopts timecode synchronization technology to synchronize the collected motion capture data, audio waveform and 3D scene animation at the frame level, and strictly matches the virtual character’s lip, body movement, vocal technique and drum beat. Scene lighting system simulates the change of sunlight angle and color temperature in southern Anhui based on the real geographical data, and introduces dynamic weather system (morning fog, twilight halo, etc.) to endow the virtual space with the vitality that spreads over time. Finally, all resources are packaged and managed through Unity’s Addressable Assets system, which can load on demand and schedule efficiently.

IV. EMBODIED INTERACTIVE MECHANISM DESIGN FOR TRADITIONAL OPERA

In order to help users switch from passive observers to active participants, the system uses the HTC VIVE Pro Eye head-mounted display and Leap Motion gesture recognition module to design an embodied interactive method based on body movements. The interaction logic adopts a three-level progressive model of “exploration—viewing—learning to sing”. Users can gradually experience Anhui’s musical culture in the virtual space. In the exploration layer, users can walk freely in the virtual village. By using six degrees of freedom tracking function of the head-mounted display and joystick control, users can explore different scenes in the cobble stone road. When users approach different scenes (such as the stage or the entrance to an ancestral hall), the system pops up a cultural prompt in the collision detection area and briefly describes the musical genre and its historical background. When users stand in the courtyard and look around, they will hear faint opera sounds coming from somewhere in the distance—“hear the sound before seeing the performer”. In the viewing layer, users can enjoy a multi-angle line-of-sight. When users enter the stage scene, the system automatically selects the best viewing angle according to the standing position. Users can also switch view points actively. By using the controller to select virtual seats (such as the main hall, side rooms or platform), they can switch to different viewing angles. When the performance begins, users can use gesture recognition to grab virtual props (such as folding fans or gongs) and view their names and uses; they can also grab a virtual instrument in the virtual orchestra area and zoom in to view the playing process and timbre. To enhance users’ immersion, the system places visual guide points in key parts of singing, prompting users to focus on the body movement and facial expression of the actor.

In the singing practice layer, it is the innovation of the interactive method. The system labels the movement and singing waveform data of inheritors in advance and creates a “singing style - body movement” library. When users stand in the reserved area on the stage, the system projects a semi-transparent virtual character outline in the air and prompts users to imitate it. Leap Motion captures the user’s hand skeletal movement in real time and matches it with the standard motion library by dynamic time-warping matching. When the similarity reaches the threshold, the virtual character gives applause or cheers. In the vocal imitation module, the system extracts the fundamental frequency from the user’s vocal recording collected through the microphone and compares it with the original voice recording; it then shows the pitch deviation as a visual waveform to help users adjust their vocalization. In addition, there is a rhythm interaction component. When the drumbeat sounds, users can tap a virtual drum in the air with gestures; the system shows real-time color light effects according to the error between the tapping time and the standard beat. The error is small and the light is golden; the error is large and the

light is blue and makes ripples.

V. ADAPTIVE RENDERING AND CROSS-PLATFORM DISTRIBUTION

The system also adopts an intelligent priority-based rendering strategy. When the frame rate drops below 90fps, the dynamic resolution scaling technology selectively reduces the resolution of peripheral background elements and non-essential shadows. Crucial ‘cultural detail’ assets, such as the intricate textures of the opera stage and the facial expressions of virtual performers, are locked at high-precision rendering priority to prevent visual inconsistency and ensure that the core cultural heritage remains legible even under system stress. For texture and lighting, the system adopts progressive loading strategy based on streaming texture. When user enter the scene, diffuse map and basic lighting will be loaded first and after viewpoint be stabilized, the normal map and specular map will be gradually loaded. The dynamic lighting baking adopts lightmap, precomputed global illumination for static object, real-time shadow with limited rendering distance for dynamic character.

The system also adopts dynamic resolution scaling technology, when the frame rate drop to lower than 90fps, the system will automatically reduce the rendering resolution to ensure the stability of frame rate and avoid user suffer from motion sickness. For different device adaptation, the system adopts Unity general rendering pipeline for multi-platform development and on PC, it supports HTC VIVE and Oculus Rift headset for SteamVR output, resolution is 2448×1224 and refresh rate is 90Hz, and it will use highest level of texture filtering and anti-aliasing. On mobile device, for all-in-one device, the system will use hardware performance detection to automatically disable real-time shadow, use simple ambient occlusion instead of screen space reflection, and adopt 3rd order decoding format instead of 7th order Ambisonics for audio decoding, to ensure power consumption in a reasonable range while keeping immersion.

In order to reduce the hardware cost for users even more, the system innovatively introduces cloud streaming technology for lightweight rendering, namely, to implement the core rendering computation on the cloud server cluster, encode the rendered video footage in real time through the NVIDIA CloudXR protocol into H.264/HEVC video streams via 5G or gigabit network and push to user terminals, users only need to install lightweight client on the local end or access the system through browser based on WebRTC protocol, the head-tracking data of headset and controllers uploaded by users are processed on the cloud and then returned as rendered video footage, end-to-end latency is controlled within 20ms. This measurement accounts for the Motion-to-Photon (MTP) delay, including 5G air interface transmission (approx. 10ms) and cloud rendering/encoding (approx. 8ms). Under a stable gigabit environment with an edge computing node located within 150km of the user, empirical tests showed a network jitter of <2ms, ensuring the stability required to suppress vestibulocochlear-

induced motion sickness. With this technology, ordinary smartphones can access a high-fidelity visual experience. However, it should be noted that while cloud streaming provides PC-level graphics to mobile users, hardware such as Cardboard is restricted to the ‘viewing’ and ‘exploration’ layers due to its lack of six degrees of freedom (6DoF). For the ‘singing practice’ layer, which requires precise 6DoF tracking and Leap Motion gesture recognition, users must utilize compatible standalone or PC-VR headsets to access the full range of embodied interactions.

VI. PERFORMANCE VERIFICATION OF IMMERSIVE EXPERIENCE

A. EXPERIMENTAL PREPARATION AND SETUP

To comprehensively evaluate the system’s performance in three dimensions—immersive experience, learning effectiveness, and usability—this study designed three independent experiments, each using different participant samples to avoid the learning effects and fatigue associated with repeated measurements. The experimental equipment included the self-developed “Anhui Music Culture Immersive Experience System,” deployed on a high-performance workstation (CPU: Intel i9-13900K, GPU: NVIDIA RTX 4090, Memory: 64GB). The output device was an HTC VIVE Pro Eye headset (resolution 2448×1224, refresh rate 90Hz, integrated eye-tracking module). Physiological signal acquisition devices included an Emotiv EPOC X 14-channel wireless EEG helmet (sampling rate 128Hz) and a Shimmer3 GSR+ electrodermal response meter (sampling rate 128Hz). Eye-tracking data was acquired by the headset’s built-in Tobii module (sampling rate 120Hz). The system automatically records interaction logs (dwell time, scene switching, gesture trigger count, viewpoint coordinates, rendering frame rate, interaction latency, etc.) in the background. It also includes a built-in interactive knowledge testing module to record accuracy and reaction time during learning effect experiments.

The experiment was conducted in a soundproof, temperature-controlled ($24\pm1^{\circ}\text{C}$), and dimmed (50 lux) virtual reality laboratory. Safety boundaries were set within the room, and an experimental assistant was present to monitor the process throughout. All equipment was calibrated daily before the experiment. The Lab Streaming Layer protocol was used to achieve real-time synchronization of multimodal data, and MATLAB R2023b and Python 3.9 were used for data processing and analysis.

B. PHYSIOLOGICAL RESPONSE EXPERIMENT OF PRESENCE AND SPATIAL COGNITION

To explore the advantages of this system compared to traditional 2D video in creating a sense of spatial presence, a multimodal physiological measurement method was used to compare the neurological and physiological response characteristics of users under the two media. Forty participants were recruited and randomly divided into a VR

group and a 2D video group, experiencing Anhui music and culture scenes respectively. EEG, electrodermal data, and eye-tracking data were collected simultaneously to quantify the differences in immersive experience using objective indicators. This study was approved by the Ethics Committee of Huainan Normal University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

TABLE 1. Comparison of Physiological Indicators Between the Two Groups (M $\pm$ SD)

Measurement Indicator	VR Group (n=20)	2D Video Group (n=20)	t-value	p-value
θ/α ratio	0.85 ± 0.12	0.52 ± 0.09	9.86	< 0.001
Skin conductance response (μS)	0.42 ± 0.15	0.18 ± 0.08	6.23	< 0.001
Mean pupil diameter (mm)	3.86 ± 0.31	3.21 ± 0.24	7.41	< 0.001
Gaze hotspot concentration (%)	67.3 ± 8.2	45.6 ± 7.5	8.75	< 0.001
Head rotation amplitude ($^{\circ}/\text{s}$)	24.6 ± 5.3	3.1 ± 1.2	23.52	< 0.001

In Table 1, the theta-alpha wave ratio in the VR group (0.85 ± 0.12) was significantly higher than that in the video group (0.52 ± 0.09), indicating that the VR experience induced a deeper state of immersion. The change in skin conductance in the VR group ($0.42\pm0.15\mu\text{S}$) was 2.3 times that of the video group ($0.18\pm0.08\mu\text{S}$), reflecting a stronger level of emotional arousal. Regarding eye-tracking data, the average pupil diameter in the VR group increased by 23.7%, and the focus of fixation increased by 47.6%, indicating that users invested more cognitive resources in paying attention to cultural details in the virtual space. Head rotation data ($24.6^{\circ}/\text{s}$) further confirmed that VR users constructed spatial cognition through active exploration. These physiological indicators consistently demonstrate that the immersive experience constructed by this system is significantly superior to traditional video media in creating a sense of presence and guiding spatial cognition, providing effective technical support for the contextual reproduction of Anhui music culture.

C. EXPERIMENT ON THE RETENTION OF MUSIC AND CULTURE LEARNING EFFECTS

In this study, in order to analyze the long-term memory retention effect of the system on Anhui music and culture knowledge, 2×2 mixed experimental design was adopted. Thirty subjects with no music knowledge were randomly allocated into VR group and video group. The VR group learned the same contents of Anhui opera through the immersive system, and video group learned the same contents of Anhui opera through 2D video. The immediate test was carried out after learning, and delayed test was carried out one week later. The accuracy rate of the two groups were recorded, and the retention rate was calculated.

In Figure 1, the accuracy rate of the VR group in the immediate test was $87.3\pm5.3\%$, significantly higher than that of the 2D video group ($72.1\pm7.1\%$) ($t=6.85$, $p<0.001$). In the delayed test one week later, the accuracy rate of the

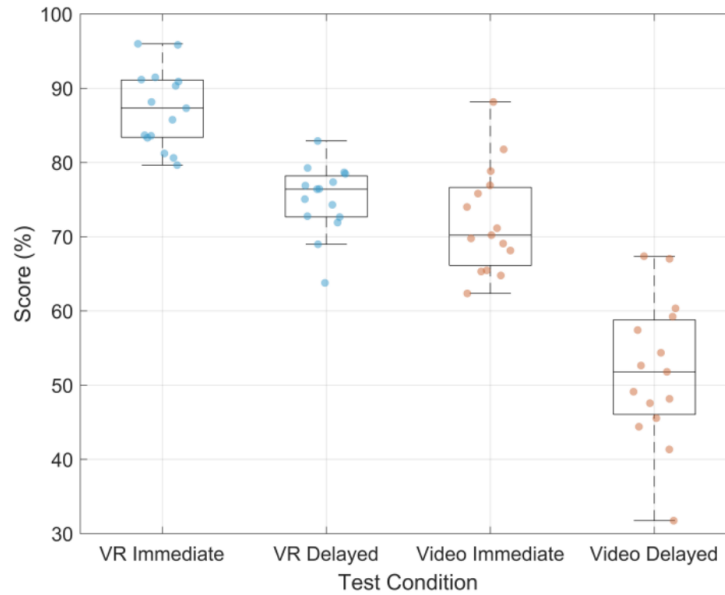


FIGURE 1. Comparison of Music and Culture Learning Effect Retention

VR group remained at $75.1 \pm 4.6\%$, while that of the video group dropped to $51.9 \pm 9.6\%$, further widening the difference between the two groups ($t=10.21$, $p<0.001$). In terms of memory retention rate, the VR group reached 86.0% , significantly better than the video group's 72.0% ($t=6.58$, $p<0.001$). The results indicate that immersive virtual reality learning not only facilitated the initial acquisition of Anhui music and culture knowledge but also significantly improved long-term memory retention, with a knowledge decay rate far lower than traditional video learning methods.

D. SYSTEM USABILITY AND INTERACTIVE BEHAVIOR ANALYSIS EXPERIMENT

To subjectively assess the interaction response quality of the system and the continuity of user experience, we recruited 20 participants (with different VR experience, 10 novices and 10 experts) to finish the predefined task flow: enter the scene, find target stage, trigger performance, do gesture imitation interaction on the performance, switch scene. The whole interaction process was recorded by the system backend, including the rendering frame rate, interaction latency, time cost of finishing the task, number of operational errors and number of abnormal viewpoint jumps.

TABLE 2. Comparison of System Usability and Interaction Behavior Indicators between Two User Groups ($M \pm SD$)

Measurement Indicator	Novice Group (n=10)	Experienced Group (n=10)	t-value	p-value
Average frame rate (fps)	88.6 ± 2.3	89.1 ± 2.1	0.51	0.617
Interaction response latency (ms)	24.3 ± 3.2	23.8 ± 2.9	0.37	0.715
Total task completion time (s)	126.4 ± 15.7	98.2 ± 12.3	4.46	< 0.001
Number of operation errors	4.7 ± 1.5	2.1 ± 0.9	4.69	< 0.001
Number of abnormal gaze jumps	3.2 ± 1.8	1.8 ± 1.1	2.11	0.049
Frequency of scene switching	5.8 ± 2.1	7.3 ± 1.9	1.69	0.108

Table 2 shows that the system performed stably in terms of technical performance. Both groups of users perceived an average rendering frame rate close to 90fps, with an interaction response latency of approximately 24ms, meeting the standards for a comfortable VR experience, and there were no significant differences between the groups ($p>0.05$). In terms of behavior, the experienced group's task completion time ($98.2 \pm 12.3s$) was significantly shorter than the novice group's ($126.4 \pm 15.7s$), and the number of operational errors (2.1 ± 0.9 times) was also significantly less than the novice group's (4.7 ± 1.5 times), indicating that the system's interaction logic conforms to conventional VR operating habits, and novices can complete the core tasks after a short period of adaptation. The number of abnormal viewpoint jumps was slightly higher in the novice group (3.2 ± 1.8 times vs 1.8 ± 1.1 times, $p=0.049$), suggesting the need for further optimization of visual guidance during scene transitions to reduce the possibility of dizziness. Overall, the system performed well in terms of usability and was able to support users in completing the immersive exploration task of Anhui music culture.

VII. CONCLUSION

This paper presents an immersive experience system of Anhui music culture that integrates spatial reproduction, soundscape fusion, and the technical heritage of smart weaving and high-performance fiber artistry. Its efficacy is tested in terms of three dimensions physiological response, learning effect and usability. The research indicates that the replenishment of intangible cultural heritage music in its original cultural location and integration with embodied interaction produces a high level of presence and richness of cultural comprehension on the part of the viewers or

observers and the retention of knowledge is also better than traditional media. The system satisfies the requirements set on VR comfort experience in the aspect of technical performance and has excellent usability and generality. The weaknesses of the study are that the type of sample was rather limited, and the information on the subject of the scene should be further enhanced to represent the entire range of the Anhui music culture. Upcoming studies will expand user testing across diverse age and cultural groups while utilizing artificial intelligence to dynamically generate virtual spaces, music content, and the technical heritage of smart weaving and high-performance fiber artistry.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of Huainan Normal University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

AUTHOR CONTRIBUTIONS

Conceptualization – Xiaobing Zhang and Meng Ning; methodology – Xiaobing Zhang and Meng Ning; investigation – Xiaobing Zhang and Meng Ning; writing-original draft preparation – Xiaobing Zhang and Meng Ning. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

The research is supported by: the Philosophy and Social Sciences Planning Project of Anhui Province in 2024, “Research on the Music Culture of Anti-Japanese Base Areas in Anhui Region”, Project No. AHSKYQ2024D066

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] K. Ji and A. Wang, “Reflections on the Integration and Development of Cultural Tourism in Cultural Inheritance and Innovation—Taking Huizhou Culture as an Example,” *Journal of Longdong University*, vol. 33, no. 6, pp. 116-120, 2022, doi: 10.3969/j.issn.1674-1730.2022.06.024.
- [2] C. Zhang, “Exploration of the Cultural Value of Anhui Flower Drum Lantern Music in Intangible Cultural Heritage Inheritance,” *Forestry Science and Technology Information*, vol. 56, no. 1, pp. 209-211, 2024, doi: 10.3969/j.issn.1009-3303.2024.01.063.
- [3] X. Cao, “Exploration of the Education and Inheritance of Anhui Intangible Cultural Heritage Folk Songs in Universities,” *Journal of Chuzhou University*, vol. 25, no. 4, pp. 125-130, 2023, doi: 10.3969/j.issn.1673-1794.2023.04.026.
- [4] A. Hariri, “Digitalization of intangible cultural heritage (ICH) archives as legal protection of intellectual property rights,” *Khizanah al-Hikmah: Jurnal Ilmu Perpustakaan, Informasi, dan Kearsipan*, vol. 9, no. 2, pp. 145-156, 2021, doi: 10.24252/kah.v9i2a7.
- [5] J. Zhao, “Digital protection and inheritance path of intangible cultural heritage based on image processing algorithm,” *Scalable Computing: Practice and Experience*, vol. 25, no. 6, pp. 4720-4728, 2024, doi: 10.12694/scpe.v25i6.3287.
- [6] Y. Wei, “Digital Protection and Inheritance Strategies of Weinan Intangible Cultural Heritage under Network Media,” *Journal of Sociology and Ethnology*, vol. 6, no. 1, pp. 150-157, 2024, doi: 10.23977/jsoce.2024.060120.
- [7] A. Adewumi A, “Protecting intangible cultural heritage in the era of rapid technological advancement,” *International Review of Law, Computers & Technology*, vol. 36, no. 1, pp. 3-16, 2022, doi: 10.1080/13600869.2021.1997084.
- [8] Q. Xu, “Cultural Identity and Morphological Reconstruction of She Ethnic Folk Songs in Ningguo, Anhui,” *Symphony (Journal of Xi'an Conservatory of Music)*, vol. 42, no. 3, pp. 31-36, 2023, doi: 10.3969/j.issn.1003-1499.2023.03.005.
- [9] C. Tong and H. G. Kim, “Intangible cultural heritage based on AR technology,” *Journal of Applied Data Sciences*, vol. 3, no. 3, pp. 118-127, 2022, doi: 10.47738/jads.v3i3.64.