Mathematical Modeling and Algorithm Application
ISSN: 3006-0842 | Vol. 4, No. 3, 2025

Interdisciplinary Perspectives on Intangible Cultural Heritage Digitization: A Review of Vision-Based Human Pose Estimation Research

Zixuan Lu, Mingyuan Guo*

College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China
* **Corresponding author**: Mingyuan Guo

Abstract: The digitization of intangible cultural heritage (ICH) has emerged as a pivotal direction for cultural preservation and transmission. Among enabling technologies, human pose estimation offers critical support for digitally documenting and disseminating performance-based ICH, such as Tai Chi and traditional dances. This paper systematically reviews the applications, technical challenges, and future trends of computer vision-based human pose estimation in ICH digitization. First, we analyze the strategic significance of ICH digitization and the unique requirements for motion capture in performance-based ICH. Next, we trace the technological evolution from traditional sensor-based methods to deep learning-driven approaches, with a focus on the applicability of core techniques, including pose estimation algorithms and temporal modeling in ICH contexts. Subsequently, we identify persistent bottlenecks, such as occlusion by intricate costumes, low-quality archival footage processing, and cross-cultural semantic understanding of movements. Finally, we prospect future research directions, including lightweight models, immersive virtual-augmented teaching systems, and digital twins, while proposing a synergistic "technology-culture-industry" innovation framework to bolster the dynamic preservation and global dissemination of ICH. By bridging computer vision and cultural heritage studies, this work aims to advance interdisciplinary research and foster sustainable digital safeguarding of traditional culture.

Keywords: Intangible cultural heritage(ICH); Human Pose Estimation(HPE)

1. Introduction

As a living embodiment of human civilization, intangible cultural heritage (ICH) carries the historical memory and cultural wisdom of nations. The Convention for the Safeguarding of the Intangible Cultural Heritage, adopted by UNESCO in 2003, particularly emphasizes the critical role of digital technologies in ICH preservation, urging member states to leverage modern technological means to ensure its sustainable transmission[1]. China's 14th Five-Year Plan for Cultural Development further identifies ICH digitization as a national strategy, explicitly calling for "strengthening the digital preservation and utilization of cultural heritage to facilitate the creative transformation and innovative development of traditional culture"[2]. This strategic framework provides policy support and developmental opportunities for the digital preservation of traditional performing arts such as Tai Chi.

However, the living transmission of performance-based ICH faces unprecedented challenges. Take Tai Chi as an example: despite its inscription on UNESCO's Representative List of the Intangible Cultural Heritage of Humanity, its preservation has long relied on master-apprentice oral transmission, a model constrained not only by geographical and temporal limitations but also by pressing issues such as the aging population of practitioners and declining interest among younger generations[3]. Traditional teaching methods struggle to achieve scalable dissemination, and many nuanced movements and intrinsic rhythms are gradually disappearing as elder practitioners pass away[4]. Similar challenges persist in other performance-based ICH, such as Peking Opera and ethnic dances, art forms rich in cultural meaning and aesthetic values embedded in bodily movements. Conventional video recordings and textual documentation often fail to comprehensively capture their dynamic characteristics and subtle expressive qualities[5].

Against this backdrop, computer vision-based human pose estimation technology offers new possibilities for the digital preservation of ICH. This technique can automatically extract spatial positions and motion trajectories of human joints from video footage, transforming unstructured performance data into computable and analyzable structured representations[6][7]. Compared to traditional motion capture systems[8], vision-based solutions are more cost-effective and scalable, making them better suited for grassroots ICH preservation efforts[9]. Recent advances in deep learning have significantly improved the accuracy and robustness of pose estimation, creating new opportunities for digitizing and analyzing complex ICH movements[10].

However, applying pose estimation to ICH preservation presents unique challenges[11]. First, costume occlusion, such as the flowing sleeves in Peking Opera or the loose garments worn in Tai Chi, often obscures key body parts, leading to detection errors in appearance-based algorithms. Studies show that traditional algorithms exhibit significantly larger joint localization errors when dealing with voluminous clothing[12]. Second, movement characteristics pose technical hurdles: the slow motions of Tai Chi may cause motion blur in videos, compromising temporal modeling accuracy, while rapid theatrical movements demand higher real-time processing capabilities[13]. Furthermore, ICH movements often carry profound cultural semantics. For instance, poses like "White Crane Spreads Its Wings" in Tai Chi or "striking a stage pose " in Peking Opera embody specific cultural imagery, requiring digitization technologies to capture not only external kinematics but also the intrinsic

linkages between movements and their cultural contexts[14].

Currently, commercial motion capture systems (e.g., Vicon, OptiTrack) achieve millimeter-level precision but remain prohibitively expensive and operationally complex for widespread ICH applications[15]. Consumer-grade devices like Kinect[16] improve accessibility but suffer from limited accuracy and field-of-view, often failing to meet the demands of intricate ICH movements [17]. This gap between technological capabilities and preservation needs motivates our research toward algorithmic optimization.

As the first systematic review of pose estimation technologies for ICH, this study establishes a comprehensive technical framework and research paradigm for this interdisciplinary field. Our contributions are threefold:

Comparative Analysis: We systematically evaluate the applicability of traditional sensors, 2D pose estimation, and 3D reconstruction in ICH scenarios, proposing multimodal fusion (e.g. RGB videos, inertial sensors) as a future direction.

Evaluation Framework: We innovatively construct a three-dimensional assessment matrix (accuracy–real-time performance–cost) to guide technology selection for diverse ICH projects.

Interdisciplinary Methodology: We propose a "technology–culture" dual-driven framework, emphasizing that algorithm development must prioritize both technical metrics and cultural fidelity to ensure digitized outcomes authentically preserve ICH's essence.

Beyond advancing computer vision applications in cultural heritage, this research explores a development pathway for deep integration of technology and culture. By creating digital archives of ICH movements, we not only safeguard endangered practices but also lay a data foundation for education, creative dissemination, and cultural industries. From a broader perspective, this work operationalizes UNESCO's principle of "living heritage", offering China's technological solutions for global ICH preservation.

2. Related Work

2.1. Evolution of Human Pose Estimation Technologies

The development of human pose estimation technology reflects significant advancements in the field of computer vision. Early motion capture techniques primarily relied on optical marker-based systems, which utilized high-reflectivity markers attached to key body parts and synchronized multiple infrared cameras to achieve 3D motion reconstruction. Commercial systems such as Vicon achieved sub-millimeter-level precision and demonstrated remarkable success in film special effects and sports analysis[18]. However, these systems impose stringent environmental requirements, including specialized capture spaces and complex calibration procedures, while their exorbitant costs excluded most ICH preservation institutions from adoption.

The advent of depth-sensing technologies introduced new possibilities through markerless systems, exemplified by Microsoft Kinect[19]. Kinect v2 employed time-of-flight (ToF) ranging principles, reducing costs to consumer-grade levels. Nevertheless, its limited field of view and increasing depth measurement errors with distance hindered its effectiveness in capturing complex ICH movements [20].

Breakthroughs in deep learning have opened new avenues for human pose estimation. In 2D pose estimation, the OpenPose framework pioneered real-time motion capture in multi-person scenarios, with its innovative Part Affinity Fields (PAFs) mechanism providing a technical foundation for processing group-based ICH performances[21]. However, OpenPose's limb association algorithms frequently misjudge movements when confronted with traditional loose-sleeved attire. HRNet achieved notable improvements on standard benchmarks by maintaining high-resolution feature representations throughout processing, yet its sensitivity to lighting conditions limited its applicability in real-world ICH settings[22]. LiteHRNet adopted an ingeniously lightweight design, significantly improving computational efficiency while maintaining high accuracy, thereby enabling mobile-based ICH applications[23]. While these advancements expand possibilities for ICH digitization, they also reveal adaptability challenges in culturally specific scenarios.

3D pose estimation presents even greater challenges. VideoPose3D[24] employed temporal convolutional networks to reconstruct 3D motion from 2D keypoint sequences, demonstrating robust performance on standard datasets. The METRO model[25] innovatively utilized graph convolutional networks to directly model relationships between human mesh vertices, achieving relatively accurate 3D reconstruction even without depth information. These methods provide technical support for stereoscopic documentation of ICH movements, though their accuracy remains insufficient for ICH practices involving extensive self-occlusion, such as Tai Chi. Multi-view fusion techniques[26] are emerging as a viable solution, where multiple ordinary cameras combined with Structure-from-Motion (SfM) technology enhance capture robustness while maintaining cost efficiency. A multi-view motion analysis system developed by Tsinghua University demonstrated that a six-camera configuration could improve recognition accuracy by over 15% for complex movements, making this approach particularly suitable for fixed installations like ICH training centers.

2.2. Current Research on ICH Digitization and Technological Adaptability

International research on ICH digitization demonstrates diversified development trends. The National Museum of Ethnology in Japan adopted a hybrid acquisition strategy for its Noh theater digitization project, first using professional motion capture equipment to record representative movements and then employing computer vision techniques to expand motion variations. Notably, the project emphasized the digital representation of cultural semantics by establishing detailed cultural background annotations for each movement[27]. Similarly, research on Bharatanatyam dance at the Indian Institute of Technology Madras focused on addressing technical challenges posed by traditional attire. The team developed specialized semantic segmentation algorithms to mitigate interference from dancers' ankle bells (salangai) in movement analysis, with successful applications in cross-cultural dance teaching systems[28]. These international cases highlight the importance of integrating technical methods with cultural context.

Domestic research on ICH digitization has also achieved significant progress. The "Peking Opera Digital Project" led by the National Academy of Chinese Theatre Arts established a systematic movement coding framework, decomposing the hand-eye-body-step techniques of Peking Opera into basic

motion units, each associated with detailed biomechanical features and cultural interpretations[29]. The ethnic dance preservation project at Minzu University of China developed a Long Short-Term Memory (LSTM)-based movement quality assessment model capable of automatically detecting discrepancies between trainees' performances and standard templates, providing intelligent auxiliary tools for ICH transmission[30]. Notably, Xia et al. proposed a Tai Chi training and scoring assistance method based on motion capture technology and Dynamic Time Warping (DTW) algorithm, enabling learners to correct erroneous movements in real-time and improve accuracy and fluency—demonstrating the innovative value of interdisciplinary research [31]. These practices have not only accumulated valuable technical experience but also provided critical references for understanding the unique requirements of ICH digitization.

The diversity of ICH categories necessitates differentiated technological solutions. Performance-based ICH (e.g., dance and theater) prioritizes movement continuity and expressiveness, requiring high-frame-rate temporal analysis; craft-based ICH (e.g., traditional craftsmanship) relies more on precise hand motion capture; while language-based ICH, though primarily audio-driven, benefits from synchronized facial expression and body language recording to enhance communicative effects[32]. Taking Tai Chi as an example, its slow and fluid movements present specific challenges for video analysis. When the video frame rate falls below 60 fps, the continuity of Tai Chi motion analysis is compromised; moreover, loose practice garments cause limb occlusion, reducing recognition accuracy to below 60% for certain movements in traditional algorithms. These findings underscore the importance of algorithm optimization tailored to specific ICH types and reveal current technological limitations in practical scenarios.

Critical issues remain unresolved, including the cost disparity between commercial motion capture systems and the limited budgets of ICH preservation institutions, as well as the lack of unified data standards and long-term preservation solutions. Future research must emphasize deeper integration of technology and culture, enhancing algorithmic performance while strengthening the understanding and expression of ICH's intrinsic values.

3. Method

3.1. Data Collection Program Design

Data acquisition in the digital preservation of non-heritage faces unique challenges and needs. Multimodal data fusion technology provides an effective way to solve this problem, and by combining the advantages of RGB video, inertial sensors and depth cameras, a more comprehensive and reliable database of NRL movements can be constructed.RGB video, as the most basic acquisition method, is characterized by low cost and easy access, and it is able to retain the visual appearance information of the movements, but it is prone to produce the following problems when dealing with rapid movements or complex costumes motion blurring and occlusion problems. Inertial sensors, although capable of providing high-frequency motion data without the influence of visual occlusion, have the limitations of cumulative error and lack of absolute position information. Depth cameras, on the other hand, are able to directly acquire 3D spatial information and provide geometric constraints for movement reconstruction. It has been shown that in taijiquan motion capture, the tri-modal fusion scheme can improve the completeness of motion capture by more than 35% compared to a single modality.

Cultural sensitivity processing is an important feature that distinguishes non-heritage digitization from conventional motion capture. Traditional costumes such as the water sleeves of Peking Opera and Taiji practice uniforms often cause serious limb occlusion, and specialized semantic segmentation algorithms need to be developed to differentiate between costumes and human body regions. The interference exclusion of props is equally important, for example, the metal reflective properties of Taiji swords can affect vision-based tracking algorithms, which need to be differentiated by material recognition and motion characterization. The improved Mask R-CNN model developed by a team from the University of Science and Technology Beijing[33] achieves a segmentation accuracy of 89.7% in traditional costume scenes, providing cleaner input data for subsequent pose estimation. In addition, the acquisition process needs to consider cultural taboos and the comfort level of the inheritors, such as certain religious dances that may restrict the shooting of specific angles, which requires the technical solution to be more adaptable and flexible.

Data labeling also needs to incorporate cultural understanding. Simple joint labeling is difficult to fully express the cultural connotation of non-heritage movements, and it is necessary to add semantic labels of movements, mechanical characterization and other diversified information. When the China Academy of Art established the folk dance database, it not only labeled the basic joint coordinates, but also recorded the movement's emotional expression, cultural implication and other rich attributes, laying a solid foundation for subsequent in-depth analysis and application. This culturally-enhanced data representation is becoming a new trend in the digitization research of NRL.

3.2. Algorithm Optimization Path

For the special needs of NRL digitization, algorithm optimization needs to be innovative from multiple dimensions. Lightweight model design is crucial to promote the application of the technology in grassroots NRL organizations. Traditional pose estimation algorithms often require powerful computational resources, which is contradictory to the limited hardware conditions in most NRM sites. lightweight networks such as LitePose compress the model size to half of its original size while maintaining high accuracy through deep separable convolution and feature distillation techniques, making it possible to run in real time on mobile devices such as smartphones. The Edge Pose framework proposed by Tsinghua University further optimizes the memory access pattern and achieves a real-time performance of 25 FPS on edge devices such as Raspberry Pi, which provides a convenient tool for field investigation and on-site recording of NRM[34].

Spatio-temporal feature enhancement is another key path to improve the performance of the algorithm. Non-heritage movements often have specific rhythms and rhymes, which require algorithms to have strong temporal modeling capabilities. Spatio-temporal convolutional network [35] can effectively model the trajectory of slow continuous movements such as taijiquan by capturing the long-distance dependency through dilated convolution, while the Transformer architecture[36] utilizes the self-attention

mechanism to automatically focus on the key frames in the sequence of movements, which is excellent in recognizing transient gestures such as the "debut" of Peking Opera. For the occlusion problem caused by the wide robe and big sleeves, a related team utilizes the attention mechanism to reduce the joint positioning error in occlusion scenarios by combining the clothing region segmentation information and the human body structure a priori.

Cross-modal knowledge migration provides innovative ideas for solving the problem of scarcity of annotated data in the digitization of non-legacy. Aiming at the limited sample size of non-legacy items, the researchers explored an effective way to utilize rich action data from other domains for migration learning. Knowledge migration from data-sufficient domains (e.g., sports) to the NRL domain can be achieved through an adversarial training strategy, and this small-sample learning method significantly reduces the dependence on labeled data in the target domain, providing a feasible solution for the digital preservation of endangered NRL items. It is also found that the feature alignment method combined with the attention mechanism can further enhance the cross-domain migration effect and better adapt to the data distribution characteristics of the target domain while maintaining the knowledge of the source domain. These technological breakthroughs provide an important reference for cross-domain knowledge migration in the digitization of NRM. This small-sample learning method has special value for the protection of endangered NRL projects.

The interpretability of the model should not be neglected as well. Unlike conventional action analysis, NRH digitization requires algorithmic decision-making processes that can be understood and verified by cultural researchers. Methods such as visual attention maps and key feature interpretation can help inheritors understand the basis of the algorithm's judgment and build trust in the digitization results. This transparent dialogue between technology and culture is an important guarantee to ensure that the digitized results truly reflect the connotation of the NRL.

3.3. Evaluation System Construction

The establishment of a scientific assessment system is an important foundation for promoting the healthy development of digitization technology of NRH. In terms of quantitative indexes, in addition to the gesture estimation indexes such as PCKh@0.5 and MPJPE, it is also necessary to design specialized assessment standards for the characteristics of NRLs. The movement coherence index can quantify the integrity performance of smooth movements such as taijiquan, while the cultural semantic retention measures the ability of the digitization results to convey the cultural connotations of the movements. The evaluation framework proposed by some scholars[37], which considers both static gesture accuracy and dynamic movement fluency, is more suitable for the evaluation needs of NRL movements.

The evaluation of cultural fidelity requires the in-depth participation of inheritors and cultural researchers. The traditional technical assessment is often limited to numerical indicators and ignores the most important cultural value of NRL. The establishment of a hybrid assessment mechanism involving experts, combining quantitative analysis with qualitative evaluation, is the only way to fully grasp the cultural appropriateness of digitized results. The practice of the China Intangible Cultural Heritage Protection Center shows that organizing inheritors to certify the cultural authenticity of the digitized results can effectively avoid the distortion of cultural information in the process of technical processing. This assessment concept, which respects cultural subjectivity, is gaining wider and wider recognition.

Long-term effect tracking is also an important component of the assessment system. Digitization of NRLs is not a one-time data collection, but a continuous cultural recording process. Establishing a mechanism for regular review of digitization results, tracking the impact of technological aging on data quality, and evaluating the long-term applicability of different technological routes are crucial to the formation of sustainable NRH preservation programs. The "10-20-50" long-term preservation framework proposed by the European Committee for Digital Cultural Heritage, which emphasizes the readability and usability of digital resources after 10, 20 and 50 years, provides a reference for long-term planning for the digitization of NRH.

3.4. Typical application scenarios

Digital teaching system is the most directly valuable application direction of non-heritage digital technology. Traditional non-heritage teaching mainly relies on face-to-face teaching by teachers and apprentices, which is limited by time and space constraints and difficult to meet the learning needs of modern society. Intelligent teaching systems based on posture estimation can capture students' movements in real time and compare them with standard templates to provide instant feedback and targeted guidance. The Taiji auxiliary training system developed by Beijing Sport University helps students accurately master the essentials of complex movements such as the "knee-straddling step" by analyzing the angle deviation of the joints and visualizing the movement trajectory, so as to increase the learning efficiency by more than 40%. This kind of system is especially suitable for promotion in remote areas and school education, effectively expanding the coverage of non-genetic inheritance.

The construction of virtual inheritors in the meta-universe scenario represents an innovative form of non-heritage dissemination. Through high-precision motion capture and three-dimensional reconstruction technology, the performance of NGT inheritors can be digitized into a virtual image, which can be displayed and interacted across time and space in the meta-universe space[38]. Dunhuang Research Institute and Tencent developed the "Digital Cave of Scripture" project, through the action redirection technology to make the ancient murals in the dance "live", for the user to bring an immersive cultural experience. This application not only protects endangered non-legacy skills, but also stimulates the cultural interests of the younger generation and injects new vitality into non-legacy.

The cross-cultural movement style migration study, on the other hand, demonstrates the academic value of digitization of non-heritage. Comparing movement characteristics in different cultural contexts through computational analysis methods can reveal deep-seated cultural differences and commonalities. The Ethnic Dance Style Analysis Project conducted by the Central University for Nationalities quantitatively compared the movement parameters of Tibetan, Mongolian, and Uyghur dances, providing new data support for dance anthropology research. Such cross-cultural research not only enhances the understanding of the diversity of human movement, but also provides a theoretical basis for the innovative development of NRM.

With the continuous progress of technology, the digital

application of non-heritage is expanding in a more diversified direction. From the development of non-heritage games based on movement recognition to the production of creative content combined with generative AI, technological innovation is injecting new vitality into traditional culture. It should be noted that while pursuing technological innovation, it is necessary to always adhere to the respect for the authenticity of non-heritage and ensure that digital means serve the core objective of cultural inheritance. This balance and synergy is precisely the most challenging and valuable topic in the study of digitization of non-heritage.

4. Summary

As an important field of integration between traditional culture and modern technology, digital conservation of NRLs has made remarkable progress in the past decade. Through the innovative application of computer vision and artificial intelligence technologies, especially human posture estimation methods, more scientific and systematic digital preservation solutions have been established for performing arts NRLs such as taijiquan, Peking Opera, and folk dances. These technological breakthroughs have not only solved the limitations of traditional protection methods in terms of precision, efficiency and cost, but more importantly opened up a brand-new path for the living inheritance of NRH. The current technical system has been able to realize multi-level digital processing from basic motion capture to advanced cultural semantic analysis, providing solid technical support for the preservation, research and dissemination of non-heritage.

However, NRH digitization technology still faces several key bottlenecks. The small-sample learning problem is particularly prominent in the field of NRL preservation, where many endangered NRL programs may only have a few senior inheritors left, and the action data that can be collected is very limited. Existing deep learning algorithms usually require large-scale training data to achieve the desired results, which contradicts the realistic conditions of NRL protection. Although methods such as data augmentation and transfer learning alleviate this problem to a certain extent, the algorithms still do not perform satisfactorily when dealing with certain unique local NRL items. Insufficient cross-cultural generalization capability is another notable challenge. Most current posture estimation algorithms are trained on common western sports datasets, and when applied to eastern traditional performing arts, it is often difficult to accurately capture their unique movement rhythms and cultural connotations. For example, it is difficult for existing algorithms to fully understand the aesthetic characteristics and spiritual essence of movements that contain specific cultural imagery, such as the “cloud hand” in taichi.

In the future, knowledge distillation technology is expected to be an effective way to reduce computing costs. By distilling knowledge from large models into small models, the demand for computing resources can be significantly reduced while maintaining better performance, which is important for deploying applications in villages and non-heritage protection organizations with limited resources. The combination of digital twin and blockchain may solve the problem of digital authentication and traceability of NRM. The establishment of a complete digital asset chain can ensure that the copyright attribution and dissemination trajectory of the digitalized results are clearly traceable, which not only protects the legitimate rights and interests of the inheritors, but also promotes the rational use of digital resources. The intelligent inheritance mode of human-computer collaboration will be another important development direction. Through the design of more humanized human-computer interaction interface, let the technology really serve the inheritance practice, rather than replacing the traditional master and apprentice inheritance. This model not only retains the essence of interpersonal interaction of non-genetic inheritance, but also incorporates the auxiliary function of modern technology, which may be an ideal path to realize the sustainable development of non-heritage.

From the perspective of social value, technology-enabled rural NHIs are of special significance. Many precious non-heritage items are preserved in remote villages, and the inheritors often face the dilemma of closed information and lack of resources. Through the development of lightweight digital tools suitable for rural conditions, the conservation ability and dissemination effect of these inheritors can be significantly enhanced. For example, a smartphone-based action collection APP and a cloud processing platform will allow rural inheritors to conveniently record and share their skills. Meanwhile, the idea of constructing a global database of NRH movements has also triggered important ethical reflections. In the process of promoting the unification of technical standards and data sharing, how to respect the specificity of different cultures, protect the rights and interests of traditional knowledge, and avoid cultural appropriation and misinterpretation are all issues that need to be treated carefully. International organizations should take the lead in formulating corresponding ethical guidelines and technical norms to ensure that the digitization process truly serves the protection of cultural diversity, rather than leading to new cultural inequalities.

The digital preservation of NRLs is at a critical stage of transition from technological exploration to systematic application. Future development requires the deep integration of technological innovation and cultural understanding, and close cooperation among engineers, cultural scholars and inheritors. Through the establishment of a more open and inclusive research ecology, it is expected to build a digital protection system that meets technical standards and respects cultural authenticity. This will not only change the practice of NRL protection, but also provide useful reference for the digital transformation of other traditional cultural fields.

Conflicts Of Interest

The authors declare that they have no conflict of interest.

References

[1] Directives, U. N. E. S. C. O. "UNESCO Operational Directives for the Implementation of the Convention for the Safeguarding of the Intangible Heritage." 2022.

[2] Wang, Zhuo Yu, Rong Qun Chen, and Jun Yu Zhao. "Research on the development of China’s cultural tourism Industry during the “14th Five-Year Plan” period." SHS Web of Conferences. Vol. 165. EDP Sciences, 2023.

[3] Wang, Xiaobing. "Preservation and transmission of traditional chinese martial arts in the digital era." Of Papers Presented at 2023 4th Asia Sport Science Conference (ASSC). 2023.

[4] Gao, Yang, et al. "Inheritors’ happiness and its relevant factors in intangible cultural heritage." Sustainability 14.21 (2022): 14084.

[5] Hou, Yumeng, et al. "Digitizing intangible cultural heritage embodied: State of the art." Journal on Computing and Cultural Heritage (JOCCH) 15.3 (2022): 1-20.

[6] Kico, Iris, et al. "Digitization and visualization of folk dances in cultural heritage: A review." Inventions 3.4 (2018): 72.

[7] Lan, Gongjin, et al. "Vision-based human pose estimation via deep learning: A survey." IEEE Transactions on Human-Machine Systems 53.1 (2022): 253-268.

[8] Colyer, Steffi L., et al. "A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system." Sports medicine-open 4 (2018): 1-15.

[9] Ghafoor, Mehwish, and Arif Mahmood. "Quantification of occlusion handling capability of a 3D human pose estimation framework." IEEE Transactions on Multimedia 25 (2022): 3311-3318.

[10] Zheng, Ce, et al. "Deep learning-based human pose estimation: A survey." ACM Computing Surveys 56.1 (2023): 1-37.

[11] Stacchio, Lorenzo, et al. "Danxe: An extended artificial intelligence framework to analyze and promote dance heritage." Digital Applications in Archaeology and Cultural Heritage 33 (2024): e00343.

[12] Guigon, Emmanuel, et al. "Experimental and theoretical study of velocity fluctuations during slow movements in humans." Journal of neurophysiology 121.2 (2019): 715-727.

[13] Mei, Xiao, and Yang Xiao. "The “ICH Movement” in China." Music, Communities, Sustainability: Developing Policies and Practices (2022): 101.

[14] Qi, Fengmeng, et al. "Effects of taichi on physical and psychological health of college students: A systematic review." Frontiers in Physiology 13 (2022): 1008604.

[15] Rallis, Ioannis, et al. "Choreographic pattern analysis from heterogeneous motion capture systems using dynamic time warping." Technologies 7.3 (2019): 56.

[16] Bostanci, Erkan, Nadia Kanwal, and Adrian F. Clark. "Augmented reality applications for cultural heritage using Kinect." Human-centric Computing and Information Sciences 5 (2015): 1-18.

[17] Afyenni, Rita, et al. "Augmented Reality Applications for Cultural Heritage Using Kinect Sensor." International Journal of Advanced Science Computing and Engineering 6.3 (2024): 127-131.

[18] Goldfarb, Nathaniel, et al. "Open source Vicon Toolkit for motion capture and Gait Analysis." Computer methods and programs in biomedicine 212 (2021): 106414.

[19] Schlagenhauf, Franziska, Siddarth Sreeram, and William Singhose. "Comparison of kinect and vicon motion capture of upper-body joint angle tracking." 2018 IEEE 14th international conference on control and automation (ICCA). IEEE, 2018.

[20] Tölgyessy, Michal, et al. "Evaluation of the azure kinect and its comparison to kinect v1 and kinect v2." Sensors 21.2 (2021): 413.

[21] Qiao, Sen, Yilin Wang, and Jian Li. "Real-time human gesture grading based on OpenPose." 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE, 2017.

[22] Sun, Ke, et al. "Deep high-resolution representation learning for human pose estimation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

[23] Yu, Changqian, et al. "Lite-hrnet: A lightweight high-resolution network." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

[24] Wang, Hao, Dingli Luo, and Takeshi Ikenaga. "Image information assistance neural network for videopose3d-based monocular 3d pose estimation." 2021 17th International Conference on Machine Vision and Applications (MVA). IEEE, 2021.

[25] Zhang, Ping, et al. "Model and algorithm of coordinated flow controlling with station-based constraints in a metro system." Transportation Research Part E: Logistics and Transportation Review 148 (2021): 102274.

[26] Zhou, Kangkang, et al. "Efficient hierarchical multi-view fusion transformer for 3D human pose estimation." Proceedings of the 31st ACM International Conference on Multimedia. 2023.

[27] Har, Austin Oting. "Integrating Live Computer Tools into the Creation, Adaptation, and Performance of Japanese Noh Theatre." Audio Engineering Society Convention 155. Audio Engineering Society, 2023.

[28] Shekhawat, Megha Yashwant. Reflecting changes in performing arts over time: Bharatanatyam as a case study. Diss. 2023.

[29] Hou, Jianfeng, et al. "The implementation of a Beijing Opera interactive display system based on motion recognition." 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). IEEE, 2021.

[30] Qin, Wei, and Junying Meng. "The research on dance motion quality evaluation based on spatiotemporal convolutional neural networks." Alexandria Engineering Journal 114 (2025): 46-54.

[31] Xia, Feng, Xin Lu, and Xingwei Si. "Taijiquan Auxiliary Training and Scoring Based on Motion Capture Technology and DTW Algorithm." International Journal of Ambient Computing and Intelligence 14.1 (2023): 1-15.

[32] Shen, Zijun, et al. "Application and Exploration of Computer Technology in Digital Protection of Bashu Intangible Cultural Heritage." 2024 5th International Conference on Information Science, Parallel and Distributed Systems (ISPDS). IEEE, 2024.

[33] He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.

[34] Zhang, Lei, et al. "EdgePose: Real-Time Human Pose Estimation Scheme for Industrial Scenes." IEEE Access (2024).

[35] Hewage, Pradeep, et al. "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station." Soft Computing 24 (2020): 16453-16482.

[36] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).

[37] Ismail-Fawaz, Ali, et al. "Establishing a unified evaluation framework for human motion generation: a comparative analysis of metrics." Computer Vision and Image Understanding 254 (2025): 104337.

[38] Zhang, Xiao, et al. "Metaverse for cultural heritages." Electronics 11.22 (2022): 3730.