Information Systems Department
«ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В СУЧАСНОМУ СВІТІ: ДОСЛІДЖЕННЯ МОЛОДИХ ВЧЕНИХ»

UDC 004.5

Syromiatnykova Daria, Suprun Olexander
olexander.suprun@nure.ua, daria.syromiatnykova@nure.ua

Kharkiv National University of Radio Electronics, Kharkiv

TECHNOLOGICAL INNOVATIONS IN AUGMENTED REALITY SYSTEMS FOR ART PERFORMANCES

Artistic performances present unique challenges for AR tracking [1]: unpredictable motion, dynamic lighting changes, audience presence, the necessity for minimal invasiveness and wireless freedom for performers, as well as requirements for sub-millimeter accuracy and minimal latency to preserve the illusion. Traditional marker-based and sensor-based systems often constrain artistic intent. Modern innovations are transitioning towards hybrid, markerless, and context-aware methods based on computer vision (CV), machine learning (ML), and distributed computing.

Recent research focuses on enhancing the robustness of Simultaneous Localization and Mapping (SLAM) systems in complex performance scenarios. Classical SLAM algorithms operate on low-level point features. A key innovation lies in the integration of neural networks for real-time semantic segmentation of the environment. Studies, such as [2], demonstrate how object recognition (stage, actors, props) enables the system to dynamically differentiate between static background and moving targets, improving map stability and reducing drift. This is critical for performances where the stage is constantly changing. Traditional SLAM assumes a static environment. For performances with complex choreography, Dynamic SLAM algorithms [3] are being developed, which explicitly model moving objects, not discarding them as outliers but integrating them into the map context. This is achieved using optical flow combined with motion estimation for trajectory prediction.

For performances in large-scale spaces (museums, historical buildings), research involving pre-loaded 3D maps (prior maps) is highly relevant. Systems similar to [4] utilize pre-scanned point clouds of the location, enabling rapid camera relocalization after occlusions or extreme motion, thereby reducing the real-time computational load. Accurate tracking of an actor's skeleton and facial expressions is key for interactive AR embodiment of avatars or visual effects attached to the body. While RGB cameras are prevalent, the integration of depth sensors is becoming standard for professional AR systems. Research [5] shows that combining RGB-D data with Convolutional Neural Network (ConvNet)-based algorithms provides significantly higher accuracy in body pose estimation, particularly under partial occlusions or non-standard viewpoints typical of stage movement.

For performances responsive to a performer's facial expressions, lightweight neural architectures for facial landmark detection and emotion recognition are advancing. Accurate hand and finger tracking is essential for virtuosic interactive manipulation of virtual objects. Innovations in this field include the use of generative models and 3D mesh networks, which, trained on large synthetic datasets, are capable of reconstructing the full 3D pose of a hand from a single RGB image, compensating for finger occlusions.

Devices such as the Microsoft HoloLens 2 and Magic Leap 2 integrate infrared (IR) sensors for precise gaze tracking. This unlocks new artistic techniques: gaze-based interface control, analysis of audience attention focus in interactive performances, or creating effects that respond to where the performer is looking.

Modern standalone AR headsets utilize inside-out tracking, but the innovation lies in expanding the sensor suite: high-precision gyroscopes, accelerometers, magnetometers, as well as ultrasonic sensors for short-range proximity sensing.

Technological innovations in AR tracking are transforming artistic performances, transitioning them into a hybrid, interactive space. Key achievements in recent years include the shift towards semantic and dynamic SLAM, the integration of AI for enhanced accuracy and prediction, edge-computing-based architectures to combat latency, and the emergence of novel hardware interfaces.

References

1. Чеботарьова, І.Б., Гуріна, К.В., & Чеботарьов, Р.І. (2022). AR-технології в сучасній рекламі. Поліграфічні, мультимедійні та web-технології. Т. 2. (с. 34-36).
2. Suprun, O.O., & Taranenko, S.B. (2021). Development of a Hybrid Visual-Inertial SLAM Algorithm for Dynamic Stage Environments. Radio Electronics, Computer Science, Control. (p. 167-170).
3. Pumarola, A., Vakhitov, A., & Agudo, A., (2021). PL-SLAM: Real-time Monocular Visual SLAM with Points and Lines. International Conference on Robotics and Automation (ICRA). (p. 4503-4510).
4. Saito, S., Huang, Z., Natsume, R., & Morishima. (2023). PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization. /CVF. Conference on Computer Vision and Pattern Recognition (CVPR). (p. 81-90).
5. Liu, S., Wei, Y., Li, Q., & Xu, W. (2023). EdgeAR: An Edge-Assisted Framework for Latency-Sensitive Augmented Reality Applications. Proceedings of the ACM SIGCOMM Conference. (p. 345-358).

Academic supervisor: Cand. Sc. (Technic), Ass. Professor Suprun O.