Non-Visual AR Physiotherapy

Ananth Nayak*
ananth_nayak@berkeley.edu
University of California, Berkeley
California, USA

Grace Thompson*
grace_thomp@berkeley.edu
University of California, Berkeley
California, USA

Shikha Shah*
shikha.shah@berkeley.edu
University of California, Berkeley
California, USA

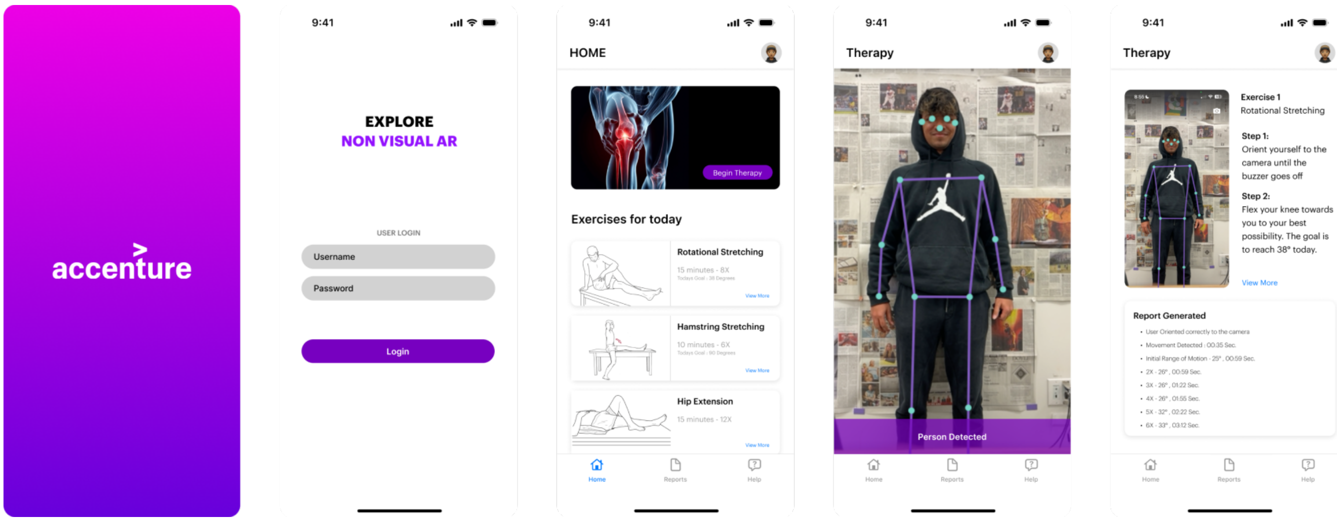

Figure 1. Navigation of app's interface and user flow.

Abstract

Total knee arthroplasty (TKA) is one of the most common surgeries performed each year. In addition to post-operative rehabilitation, patients often require personal aid post surgery to expedite getting back to their full range of motion. While current analog physiotherapy methods aid patient recovery, we wanted to explore means of enhancing personal therapy using non-visual augmented reality (AR). The focus of this research is to address the questions: "How can non-visual AR aid in physical therapy task guidance for post-surgery recovery?" and "How can non-visual AR support physiotherapists and carers in their roles?". Our method presents a unique use of non-visual AR in TKA recovery physiotherapy. Our goal is to increase the efficacy and efficiency of rehabilitation by substituting conventional instruments with augmented reality (AR)-based solutions that make use of auditory input, as well as AI and large language models (LLM) for activities such as filing reports and taking session notes. This technological integration aims to lessen the administrative load on therapists, enabling patients to be more independent in conducting home therapy sessions, in addition to improving exercise precision and reducing recovery downtime. By combining AR technology with user-centered healthcare, our project demonstrates a novel opportunity for non-visual AR to enhance patient recovery.

*All authors contributed equally to this research.

Keywords: Augmented Reality (AR), Virtual Reality (VR), Physiotherapy, Total Knee Arthroplasty (TKA), Rehabilitation, Non-Visual Interaction, Audio Task Guidance, Large Language Models (LLM), Artificial Intelligence (AI), User-Centered Healthcare, Accessibility in Physiotherapy, Multi-Tasking Challenges, Pseudo-Disability, Computer Vision, Pose Estimation

1 Introduction

More than 600,000 total knee arthroplasty (TKA) procedures are performed each year, with an expected increase in the number of procedures by 673% by the year 2030 (1). Recovery from surgery can take up to 1 year, during which the patient is recommended to complete postoperative physical therapy to increase range of motion, decrease pain, and expedite overall recovery. Despite this necessity, rehabilitation programs vary greatly in treatment, frequency, and cost (2). The average patient completes 33.2 physical therapy visits on average post-procedure, accounting for the majority of their outpatient costs (3). The aim of this project is to create an intuitive and affordable improvement upon the essential postoperative physical therapy experience, specifically regarding the issues of cost, and lack of treatment consistency. We have developed a solution that demonstrates how non-visual AR in the form of audio task guidance could improve and therefore expedite post-TKA surgery-recovery for patients. This project's conceptual framework is based on the integration of three main technological components: a large language model (LLM) with Retrieval Augmented Generation (RAG) capabilities to provide audio responses; a cloud-based architecture that leverages Firebase/Firestore; and a computer vision model for precise pose estimation. These components merge within iOS applications that can be installed on iPhones running iOS 13.0 or later. When the user opens the app, the user's medical history and a customized treatment plan—either recommended by a physician or based on medical necessity—are downloaded from the cloud to the device. This data informs the treatment plan that is communicated to the patient via audio. Using the iPhone's camera and computing power, a Mediapipe model integrates advanced computer vision to determine the user's range of motion. Simultaneously, an LLM triggers an audio feedback system, synthesizing the collected data to provide contextually relevant and personalized recommendations regarding the user's form and progress. Using these technologies, we leverage non-visual AR to improve TKA patients' ability to maintain proper rehabilitation forms.

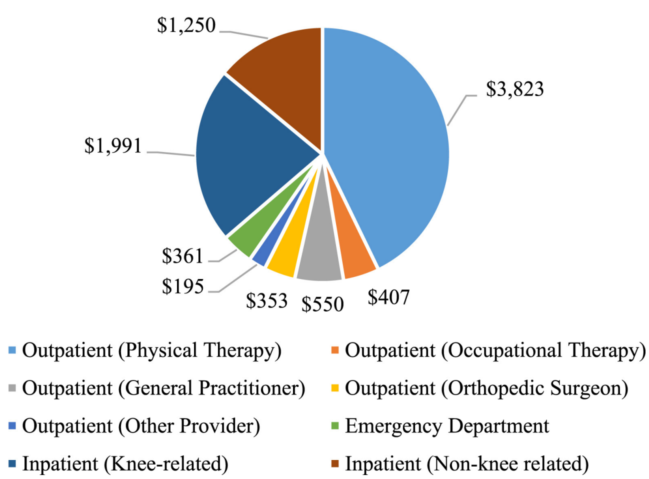

Figure 2. Cost drivers in the year after knee arthroplasty. [28]

2 Related Work

2.1 The Rise of AR and VR in Training and Education

Augmented Reality (AR) and Virtual Reality (VR) have been increasingly acknowledged for their revolutionary impact on training and education (23). AR glasses, for instance, offer real-time guidance for complex tasks, providing step-by-step instructions and contextual feedback, which proves invaluable in hands-on learning such as car maintenance or medical procedures (24). In the realm of physical therapy and similar fields where visual distractions might be a hindrance, non-visual augmented reality, particularly audio coaching, has shown potential. This method enables users to concentrate on their tasks, aided by auditory cues and instructions, potentially leading to more efficient and effective learning experiences. This approach's effectiveness was corroborated by Accenture's research, which highlighted the utility of auditory cues in maintaining task flow during audio and visual AI coaching sessions for training baristas (25).

2.2 Technological Evolution in Physiotherapy

With the increasing prevalence of total knee implant surgeries each year (4), there is a growing emphasis on utilizing technology to improve rehabilitation outcomes, including Total Knee Arthroplasty (TKA). Clinicians face the challenge of delivering positive results within limited time and resources, compounded by a rise in patient numbers and falling reimbursement rates (5). Traditional tools like the Goniometer, used for measuring a patient's range of motion, are being replaced by more advanced technologies due to their cumbersomeness and potential inaccuracies (6). Innovations such as sensor-based devices by Yo-Ping Huang et al. (7) and flexible wearable sensors by Enrica Papi et al. (8) have improved range of motion measurements, yet challenges remain in their practical home application.

2.3 Non-Intrusive Methods and AI in Rehabilitation

The field of computer vision and AI, particularly pose estimation, is being explored to provide non-intrusive methods of measuring range of motion, detecting and tracking the position and orientation of body parts in images or videos (9). Significant advancements have been made in this domain, although challenges persist in hardware, camera calibration, feature extraction, multi-view matching, and 3D reconstruction (15). Research by Milind Shah et al. using OpenCV and Mediapipe libraries has highlighted the accuracy of these tools in human behavior analysis (10), and Jan Stenum and colleagues have emphasized their transformative impact in remotely assessing human movement (11). Samkit Jain's development of a real-time virtual trainer for physiotherapy exercises showcases the potential of on-device pose estimation for post-TKR rehabilitation (15).

2.4 Integration of AI and LLMs in Physiotherapy

The role of AI, particularly Large Language Models (LLMs), is evolving in the realm of physiotherapy. These technologies offer potential for developing more personalized and responsive patient interaction systems (17,18,19). Esposito et al. (2021) discussed the integration of biosignal-based human-machine interfaces in rehabilitation, which could lead to more individualized rehabilitation protocols based on biosignal feedback (20). However, there remains a need for incorporating comprehensive patient data, including comorbidities and physiotherapist inputs, to tailor rehabilitation exercises effectively.

2.5 Emerging Applications of Non-Visual AR in Physiotherapy

Non-visual AR technologies, particularly advanced audio feedback systems, offer real-time, context-sensitive guidance. When combined with AI and machine learning, these technologies create personalized and responsive learning environments, enhancing the effectiveness of training programs. An exemplary application is the "Vi" app, an AI-powered personal running coach that uses bio-sensing earphones for real-time auditory feedback and coaching (27). This technology exemplifies how continuous engagement with tasks can be facilitated without the need for visual cues, underscoring the potential of non-visual AR in various training and educational fields.

3 System Description and implementation

Figure 3 depicts the intricate design of a Non-Visual AR Physiotherapy Application, which is envisioned to provide a dynamic and interactive rehabilitation platform through the utilization of augmented reality (AR) and audio feedback mechanisms. At the forefront of the system's architecture is the User Database, powered by Google Firebase. This database acts as a secure vault that houses essential user data, including authentication credentials via Google Auth, an archive of recent and historical health records, and intricately detailed rehabilitation or therapy schedules tailored to each user. The Computer Vision provides the real-time capture and interpretation of the user's movements. Leveraging advanced pose estimation technologies such as PoseNet or Mediapipe, the module meticulously detects and analyses the user's posture and movements. Central to the application's intelligence would be the Intelligent Agent module, which integrates a Large Language Model (LLM) API. This intelligent agent would be responsible for digesting the rich data stream from the Computer Vision module and interpreting it to derive meaningful insights. It would be further augmented by a RAG Pipeline, complemented by a dual Text-to-Speech (TTS) and Speech-to-Text (STT) framework, which establishes a bidirectional auditory communication channel with the user. The audio output is made spatial to direct the user towards the correct orientation without them looking at the screen for feedback all the time. In this project, we explored various computer vision models to facilitate human pose detection through pose estimation techniques. The process typically involves inputting a processed camera image into a pose estimation model, which then outputs information about key anatomical points. These points are identified by a component ID and are accompanied by a confidence score, ranging from 0.0 to 1.0, indicating the likelihood of each point's accurate detection at a specific location.

Given the ubiquity and accessibility of mobile phones, our focus shifted toward implementing pose estimation models on iOS devices. Initially, we utilized the PoseNet model, available through Apple's developer platform and compatible with their native CoreML framework. The PoseNet model enabled us to identify 17 distinct body parts, including major joints like the shoulders, hips, elbows, knees, wrists, and ankles, in images captured by the camera. This model identifies key points, and for our purpose, we concentrated on the hip, knee, and ankle joints. We were able to calculate vectors for the femur and tibia by representing these joints as coordinates:

$$(xhip, yhip), (xknee, yknee), and(xankle, yankle)$$

$$vecA = xknee - xhip, yknee - yhip$$

$$vecB = xankle - xknee, yankle - yknee$$

Using these vectors, we applied the dot product formula to calculate the angle between them. However, we soon recognized limitations in the PoseNet model, particularly in producing low-resolution outputs, which constrained the precision of keypoint estimations. This limitation was critical for our application, which demands high accuracy in pose estimation for medical assessments. This realization led us to explore alternative methods, and we subsequently discovered the MediaPipe framework developed by Google. MediaPipe offered a more accurate model for detecting keypoints in the human body, allowing us to perform more precise vector calculations and angle determinations between the femur and tibia. We integrated this model with a conversational interface, transforming the device into a personal trainer that interacts with users, accesses their medical records, and incorporates inputs from their physiotherapists. By implementing text-to-speech and speech-to-text functionalities, the system engages in dialogues with users, tracking their progress and providing feedback based on thresholds set by their physiotherapists

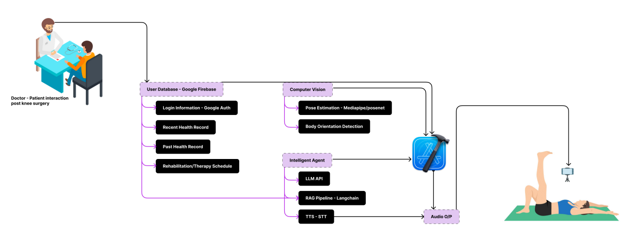

Figure 3. System Architecture for a native ios app to aid Non-Visual Augmented Reality in Postoperative Knee Rehabilitation.

4 Evaluation

4.1 Methodology

The evaluation of our Non-Visual AR Physiotherapy system was conducted using a cursory mixed-methods approach. We used quantitative measures to assess the accuracy of the system in facilitating the knee-flexion exercise used by post-TKA surgery patients. Additionally, qualitative methods were used to gauge user experience and satisfaction.

4.2 Quantitative Evaluation

The quantitative evaluation focused on accuracy of pose estimation: We measured the precision of the pose estimation technology in tracking and analyzing patient movements during exercises. This involved comparing the angles and positions detected by the system with those measured by a goniometer to assess accuracy.

4.3 Quanlitative Evaluation

We collected qualitative feedback regarding the usability, convenience, and overall satisfaction with the Non-Visual AR system. This included their perception of the audio feedback, the ease of following exercises, and the system's impact on their motivation and engagement.

Figure 4. We tested the Non-Visual AR system on 5 participants to evaluate angle detection accuracy and gather qualitative feedback on the user experience.

4.4 Data Analysis

Data from both quantitative and qualitative sources were analyzed to identify common themes and areas for improvement. All participants reported enjoying completing the exercises without having to mentally track progress or refer to a visual aid for form feedback. Two key themes for improvement included the implementation of more pleasant-sounding audio beeping in the orientation process, and more varied and descriptive audio instruction. We also observed that the joint-angle accuracy was significantly hindered when participants wore loose-fitting clothing, such as a long skirt.

4.5 Limitations and Future Directions

Our evaluation acknowledges certain limitations, including the sample size of participants and robustness of data collection. Future research will aim to include a broader population to assess the system's effectiveness across different age groups and physical conditions. Long-term studies will be conducted to evaluate the sustained impact of the Non-Visual AR system on rehabilitation outcomes. Finally, A/B testing will be used to directly compare the Non-Visual AR system with traditional visual-based physiotherapy methods to determine user preference, exercise quality, and adherence to the rehabilitation regimen.

Figure 5. Spatial Audio to augment the interaction and provide spatial cues to orient users in front of the camera.

5 Discussion

The integration of computer vision in physiotherapy, particularly in the context of TKA post-surgery rehabilitation, presents several ethical challenges. These challenges are primarily regarding data privacy, potential biases in the technology, and the necessity for transparency. Data privacy and security are paramount, given the sensitive nature of the patient data involved, including physical movements and possibly biometric data. Ensuring confidentiality and security of this data, in compliance with HIPAA regulations and through robust encryption methods, is essential. Considering most TKA patients are older adults who may have caregivers, the design of non-visual AR systems should cater to both adults and their caregivers. For older adults who may have varying levels of comfort with technology, it is crucial to have simple and clear instructions. This can be extended to include a caregiver mode, providing guidance on assisting the patient with exercises. Additionally, given the potential age-related impairments such as reduced hearing, vision, or cognitive abilities, the system should be designed for accessibility. The use of spatial audio for camera alignment in physiotherapy is innovative and could be enhanced by a system that detects the positions of both the patient and caregiver, adjusting the audio accordingly. Such a dual-position system would ensure effective delivery of instructions, tailored to who needs to act on them. For instance, the caregiver might receive detailed exercise instructions, while the patient gets simpler cues or encouragement. However, developing a system that accurately detects and differentiates between two individuals' positions in real-time is technologically challenging and requires advanced sensors and algorithms. This aligns with the trends in AI and IoT towards creating more adaptive and responsive environments. Moreover, this approach raises additional privacy considerations, necessitating clear consent protocols and stringent data security measures.

6 Conclusion and Future Works

Non-Visual AR Physiotherapy has the potential to significantly enhance the physiotherapy experience for TKA surgery patients by integrating non-visual AR, computer vision, and large language models for intuitive audio feedback. Looking forward, we aim to refine this system with a computer vision model for optimal patient positioning and direct health data integration from physiotherapists, enhancing personalization and effectiveness of rehabilitation. The reports and data from physiotherapy would be used as an initial input for the system. This data will not only assist in training the model but also in tailoring daily goals for the user, potentially accelerating their recovery process. Expanding on accessibility and usability, considering the demographic of knee replacement patients, will be a significant area of focus. The goal is to create interfaces and interactions that are straightforward and easy to use, catering especially to older adults who may not be as technologically adept. Finally, conducting long-term impact studies will be essential to assess the real-world efficacy of the non-visual AR system in physiotherapy. These studies will provide valuable insights, enabling data-driven improvements to the system. The integration of Large Language Models is anticipated to offer a more intuitive and natural experience during therapy sessions, further augmenting the system's utility in rehabilitative care.

References

[1] N. Yohe, A. Funk, M. Ciminero, O. Erez, and A. Saleh, "Complications and Readmissions After Total Knee Replacement in Octogenarians and Nonagenarians," Geriatr Orthop Surg Rehabil, vol. 9, p. 215145931880411.

[2] H. Jacobs, GH. Seeber, K. Allers, F. Hoffmann, "Utilisation of outpatient physiotherapy in patients following total knee arthroplasty - a systematic review. BMC Musculoskelet Disord," 2021 Aug 18;22(1):711. doi: 10.1186/s12891-021-04600-2. PMID: 34407785; PMCID: PMC8375073.

[3] E. Papi, Y. N. Bo, and A. H. McGregor, "A flexible wearable sensor for knee flexion assessment during gait," Gait & Posture, vol. 62, pp. 480–483, May 2018.

[4] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," Computer Vision and Image Understanding, vol. 81, no. 3, pp. 231–268, 2001, doi: https://doi.org/10.1006/cviu.2000.0897.

[5] A. Begum, G. Geerthana, S. Sakthiya, S. Banu M, and P.Anand, "Application of SRF-PI Current Control in the Design of a Single-Phase Asymmetrical Inverter for Use in Weak Grid Environments," vol. 5, pp. 24–40, Oct. 2023.

[6] J. Stenum, K. M. Cherry-Allen, C. O. Pyles, R. D. Reetzke, M. F. Vignos, and R. T. Roemmich, "Applications of Pose Estimation in Human Health and Performance across the Lifespan," Sensors, vol. 21, no. 21, p. 7315, Nov. 2021, doi: 10.3390/s21217315.

[7] D. Esposito, J. Centracchio, E. Andreozzi, G. D. Gargiulo, G. R. Naik, and P. Bifulco, "Biosignal-Based Human–Machine Interfaces for Assistance and Rehabilitation: A Survey," Sensors (Basel), vol. 21, no. 20, p. 6863.

[8] A. Pabst, "CARE: A platform for upskilling workers, at scale.," Labs Notebook.

[9] A. Pabst, "CARE: A platform for upskilling workers, at scale.," Labs Notebook.

[10] M. Karabacak and K. Margetis, "Embracing Large Language Models for Medical Applications: Opportunities and Challenges," Cureus, vol. 15, no. 5, p. e39305.

[11] V. N. Gandbhir and B. Cunha, "Goniometer," in StatPearls, Treasure Island (FL): StatPearls Publishing, 2023.

[12] J. G. Owens et al., "How New Technology Is Improving Physical Therapy," Curr Rev Musculoskelet Med, vol. 13, no. 2, pp. 200–211, Apr. 2020, doi: 10.1007/s12178-020-09610-6.

[13] D. Spilka, "How VR And AR Are Revolutionizing eLearning For Learners Of All Ages," eLearning Industry.

[14] Odemakinde and Elisha, "Human Pose Estimation with Deep Learning – Ultimate Overview in 2024."

[15] C. Brandebusemeyer, A. R. Luther, S. U. König, P. König, and S. M. Kärcher, "Impact of a Vibrotactile Belt on Emotionally Challenging Everyday Situations of the Blind," Sensors (Basel), vol. 21, no. 21, p. 7384.

[16] S. Borgeaud et al., "Improving language models by retrieving from trillions of tokens." arXiv, Feb. 07, 2022. doi: 10.48550/arXiv.2112.04426.

[17] Q. Ai et al., "Information Retrieval meets Large Language Models: A strategic report from Chinese IR community," AI Open, vol. 4, pp. 80–90.

[18] A. Alfieri, R. Wolter, and S. Hashemi, "Intent Disambiguation for Task-oriented Dialogue Systems," Oct. 2022, pp. 5079–5080. doi: 10.1145/3511808.3557516.

[19] Y.-P. Huang, Y.-Y. Liu, W.-H. Hsu, L.-J. Lai, and M. S. Lee, "Monitoring and Assessment of Rehabilitation Progress on Range of Motion After Total Knee Replacement by Sensor-Based System," Sensors, vol. 20, no. 6, p. 1703, Mar. 2020, doi: 10.3390/s20061703.

[20] “Motion Tracking Exercise Platform for Physio and Fitness,” Kemtai. Accessed: Nov. 14, 2023. [Online]. Available: https://kemtai.com/

[21] A. Hung et al., “Ninety-day and one-year healthcare utilization and costs after knee arthroplasty,” Osteoarthritis and Cartilage, vol. 27, no. 10, pp. 1462–1469, Oct. 2019, doi: 10.1016/j.joca.2019.05.019.

[22] “Physio Pose: A virtual physiotherapy assistant | by Samkit Jain | Medium.” Accessed: Nov. 14, 2023. [Online]. Available: https://medium.com/@_samkitjain/physio-pose-a-virtualphysiotherapy- assistant-7d1c17db3159

[23] S. Bubeck et al., “Sparks of Artificial General Intelligence: Early experiments with GPT-4.” arXiv, Apr. 13, 2023. doi: 10.48550/arXiv.2303.12712.

[24] L. Mündermann, S. Corazza, and T. P. Andriacchi, “The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications,” J NeuroEngineering Rehabil, vol. 3, no. 1, p. 6, Dec. 2006, doi: 10.1186/1743-0003-3-6.

[25] C. Bregler and J. Malik, “Tracking people with twists and exponential maps,” in Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231), Jun. 1998, pp. 8–15. doi: 10.1109/CVPR.1998.698581.

[26] H. Jacobs, G. H. Seeber, K. Allers, and F. Hoffmann, “Utilisation of outpatient physiotherapy in patients following total knee arthroplasty – a systematic review,” BMC Musculoskelet Disord, vol. 22, no. 1, p. 711, Dec. 2021, doi: 10.1186/s12891-021-04600-2.

[27] B. Rose, “Vi review: a personal assistant for your workouts, but it needs some help,” The Verge.

[28] A. Hung et al., “Ninety-day and one-year healthcare utilization and costs after knee arthroplasty,” Osteoarthritis and Cartilage, vol. 27, no. 10, pp. 1462–1469, 2019, doi: https://doi.org/10.1016/j.joca.2019.05.019.

A Acknowledgement

Mentored by “Accenture Labs”.

B Disclosure

I used ChatGPT for copy-editing the first draft of this section; and to brainstorm topic titles given my set of references.