Home-based Immersive Web Rehabilitation Gaming with Audiovisual Sensors

Hayette Hadjar
Faculty of Mathematics and Computer Science, Hagen University, Germany
hayette.hadjar@fernuni-hagen.de

Paul Mc Kevitt
Ulster University (Magee), Derry/Londonderry, Northern Ireland
p.mckevitt@ulster.ac.uk

Matthias Hemmje
Research Institute for Telecommunication and Cooperation (FTK), Dortmund, Germany
mhemmje@ftk.de

ABSTRACT

Early, intensive, and repetitive physical rehabilitation is critical, but it can be difficult to keep patients motivated and engaged. The use of games in immersive reality makes physiotherapy more enjoyable and engaging. During play, many unpleasant emotions can occur as a result of interacting with the joyful activities of the game. This difficulty can be overcome by adjusting game features to the players' emotions and body gestures. This work contributes to the Sensor Enabled Affective Computing for Enhancing Medical Care (SenseCare) project for remote home healthcare applications, and its related SenseCare KM-EP (Knowledge Management-Ecosystem Portal) Affective Computing (AC) platform. In this paper, we propose web-based Augmented Reality (AR) and Virtual Reality (VR) games for home rehabilitation deploying audiovisual sensors (e.g., Camera, Microphone, Kinect), to monitor the patient's emotional well-being and calculate body estimation gestures during the gameplay session, with the aim of giving a quick return to the therapist. In parallel with SenseCare's methods for audiovisual monitoring of gamers, methods for using VR and AR web-based games using the WebXR API are also introduced, and customer satisfaction was determined through an online survey, showing an 81.8% satisfaction rate.

CCS CONCEPTS

• **Applied computing**;; • **Life and medical sciences**,; • **Health care information systems**;

KEYWORDS

WebXR game, emotion recognition, body pose perception

ACM Reference Format:
Hayette Hadjar, Paul Mc Kevitt, and Matthias Hemmje. 2022. Home-based Immersive Web Rehabilitation Gaming with Audiovisual Sensors. In *33rd European Conference on Cognitive Ergonomics (ECCE 2022), October 04–07, 2022, Kaiserslautern, Germany.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3552327.3552347

ECCE 2022, October 04–07, 2022, Kaiserslautern, Germany

ACM ISBN 978-1-4503-9808-4/22/10…$15.00
https://doi.org/10.1145/3552327.3552347

1 INTRODUCTION

Diseases that have their beginnings in neurological disorders can result in a loss of movement in certain regions of the body. In these situations, therapy comprises a variety of rehabilitation treatments delivered over a prolonged period in order to improve the patient's quality of life. Several research studies focus on the benefits of supplementing home rehabilitation for regular hospital treatment [6] [10], particularly with the Covid-19 Pandemic [27] [5], such as the MERLIN system [11].

Emotion Artificial Intelligence (AI) is another term for Affective Computing (AC). Emotion AI is a branch of cognitive computing concerned with the research and development of systems and devices that can recognize, simulate, interpret, and process human affect by collecting data from faces, voices, and body language to measure human emotions, as well as the emotional state of a user, sensed via sensors, microphones, cameras, and/or software logic. This research began with the SenseCare (Sensor Enabled Affective Computing for Enhancing Medical Care) project [9]. The SenseCare Platform was developed as a model AC R&D platform that provides software services for the treatment of patients with a variety of support needs in a mental healthcare setting. In AC, we've initiated work on facial expression analysis as a method of emotion recognition, using Convolutional Neural Networks (CNNs) with TensorFlow [12], and contactless vital signs using remote-PhotoPlethysmoGraphy (rPPG) [13]. However, this previous work only supports emotion detection from images, not from audio-video and other sensors. Figure 1 shows the conceptual architecture of the SenseCare KM-EP (Knowledge Management-Ecosystem Portal) AC platform based on audiovisual emotion recognition.

SenseCare's Sensor Infrastructure uses three different types of sensors: Cameras, Microphones, and Kinect. Cloud computing and storage are also utilized here for data processing and storage. At the next level, Sensor Applied Processing is conducted on the sensor data. At higher levels in the architecture, the AC Fusion Services are concerned with fusing the collected AC data characteristics. SenseCare's user interface layer sits on top of this data fusion layer and produces high-end services for sensor data reporting.

Virtual Reality (VR), Augmented Reality (AR), and gaming applications have been shown to help in the treatment of neurodegenerative diseases in a growing body of research [19] [22]. Emotions also have a crucial role in player-game interaction. Designers are faced with the challenge of telling compelling stories utilizing new and more popular technology inside Extended Reality (XR) environments [8]. Here, we discuss: (1) How VR and AR games are hosted on web browsers?, (2) How can the SenseCare platform for emotion

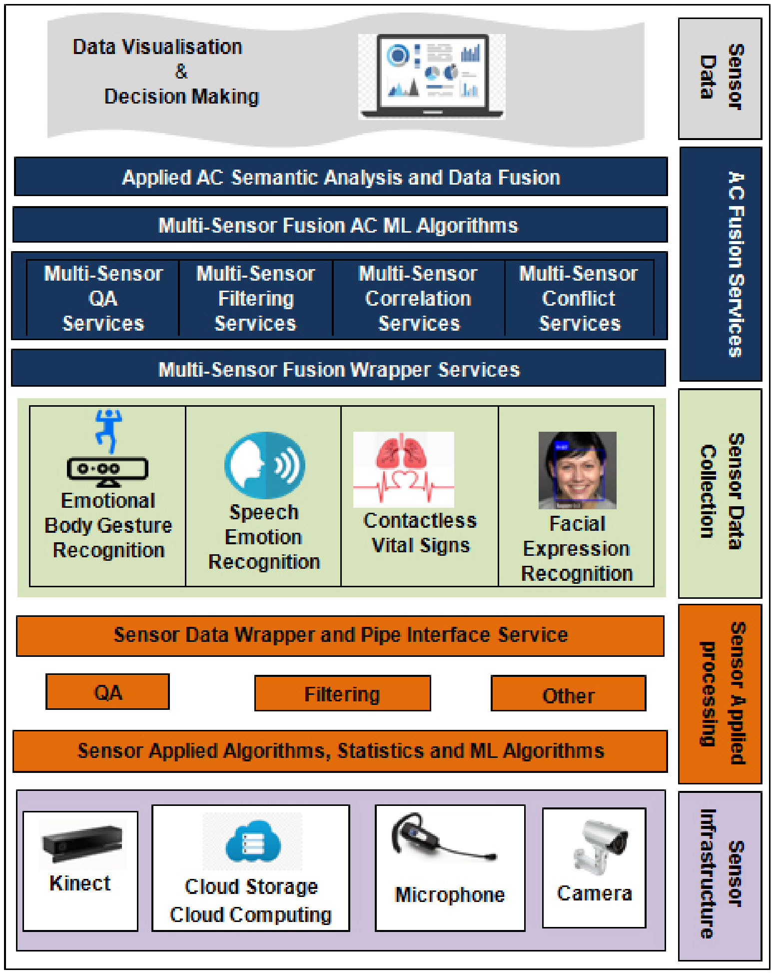

Figure 1: SenseCare's Affective Computing Platform's Conceptual Architecture [13].

recognition detect all of the patient's emotions during a game? , and (3) How are human gestures recognized?

This work encompasses VR and AR games rehabilitation and physical activity for varying-intensity exercises to reestablish motor function (e.g., hands, feet), with the motions being adapted to the patient's physical needs. SenseCare audiovisual monitoring involves: (1) emotion monitoring (e.g., Facial Emotion Expression Recognition (FER), Speech Emotion Recognition (SER)), (2) non-contact vital signs (e.g., heart rate), and (3) measurement of body gestures. Using web-based immersive games for motor rehabilitation has a number of advantages:

1. Support for a wide variety of devices and operating systems.
2. Full accessibility for users globally, allowing for collaboration between patient and therapist, and including patients with patients.
3. Interaction and communication interfaces in games are changing rapidly.

Here are some of the proposed system's benefits:

1. It uses audiovisual sensors for player emotion and gesture monitoring without physical touch.
2. Games are not impacted by the sensors; monitoring happens outside the VR environment, so there will be no changes or interruptions for the player.
3. It allows rapid recognition of emotions in pre-recorded speech for each session.
4. It reduces the need for a human therapist.

The remainder of this paper is structured as follows: Section 2 explores related work on emotion recognition studies to monitor players, as well as the use of XR (VR, AR, Mixed Reality (MR)) in the treatment of neurodegenerative illnesses. The conceptual design for the proposed system is detailed in Section 3. Section 4 discusses implementation and demonstration, and Section 5, evaluation. Finally, Section 6 concludes and discusses future work.

2 RELATED WORK

A large body of evidence suggests that XR (VR/AR/MR) and gaming applications can assist in the treatment of neurodegenerative

Table 1: Emotion-based Adaptive XR Serious Game Steps

Step N°	Step	Description
Step 1	Creation of XR serious game scenarios.	The developer creates XR serious games scenarios in collaboration with a medical expert.
Step 2	Audiovisual emotion monitoring of patient during immersive session.	The SenseCare audiovisual emotion recognition platform for classification and visualization of emotions collected from audiovisual sensors during therapy sessions.
Step 3	Visualization and evaluation of emotions detected by sensors.	Caregiver explores (positive/negative) emotion results. If SenseCare results show a higher rate of negative vs. positive emotions in the XR room, the caregiver sends a notification to developer to update the 3D virtual scenario.
Step 4	Adopt another scene for XR serious games.	The developer collaborates with a medical specialist to upgrade the XR serious game, where SenseCare has detected the XR room's negative emotions rate is higher than its positive emotions rate.

diseases [7] [1]. In [15], VR was found to be feasible and beneficial to people's arm function with multiple sclerosis when used in a serious gaming environment.

A variety of methodologies are employed in emotion recognition studies to monitor players' emotions during gameplay. These techniques can be divided into three groups: (1) emotion recognition based on self-reported data, (2) emotion recognition based on physiological responses, and (3) gameplay-based emotion recognition. For example, [23] proposes a new generic method/algorithm for emotion recognition to assess timely changes in the player's emotional state during gameplay, and practical methodology to build the affective game module for any serious game. [19] proposes a non-adaptive fuzzy logic model of emotion to estimate player emotion. AR serious games and wearable sensor networks are employed in [22] to boost patient engagement during physical rehabilitation. The proposed system enables health status assessments and the collection of vital data for clinical professionals to analyze. The use of VR serious games has shown efficacy in improving upper limb TeleRehabilitation (TR) following stroke [2]. The results reported in [18] support the use of serious games in the therapy of hemiplegia patients, and highlights the advantages of playing serious games in post-stroke movement therapy. However, further studies are required, due to the absence of Randomized Controlled Trials (RCTs), a limited number of participants, and heterogeneous samples. [17] discusses home-based rehabilitation with virtual coaching where the revolutionary virtual coaching system addresses these problems by merging recent technical advancements with therapeutic pathways.

Pose estimation is the most important task in any AR rehabilitation system. ExerCam is a TR tool discussed in [25]. The application developed includes a task mode for calculating Range Of Motion (ROM) and a game mode to encourage patients to improve their performance during therapy via the web. MediaPipe is an open source framework proposed in [26] for real-time human pose detection and recognition. It enables accurate and quick customized machine learning applications. The machine learning classification techniques employed in [25] are random forest, linear regression, ridge classifier, and gradient boosting classifier. MediaPipe is employed in our work to detect the user's pose, face, and hand landmarks in real time during gameplay.

3 CONCEPTUAL ARCHITECTURE

Our proposed conceptual architecture is shown in Figure 2. Steps 1, 2, 3, & 4 detailed in Table 1 are indicated in Figure 2.

The SenseCare monitors patients during gameplay. It employs Machine Learning (ML) algorithms to recognize audiovisual emotions (e.g., depthwise separable convolutions, densely connected blocks) via sensors. Kinect detects the emotional body language of the player. A Camera is employed for facial expression recognition and contactless vital signs such as heart rate and respiration rate. A Microphone is employed for the player's speech emotion recognition.

4 IMPLEMENTATION

The tools and activities used in play therapy sessions should be comfortable and enjoyable for patients, enabling them to express a range of emotions without fear of underperformance. Serious games include the following tasks for the purpose of neurological or cognitive rehabilitation:

1. Brain game activities to improve mental functioning and prevent brain ageing.
2. Gesture activities: adapts play and exercises to the physical needs of patients.

4.1 Game web-based VR and AR implementation

Our serious games include tasks for the purpose of neurological or cognitive rehabilitation of the activities of upper limb gestures for the first time, adapting their exercises to the physical needs of the

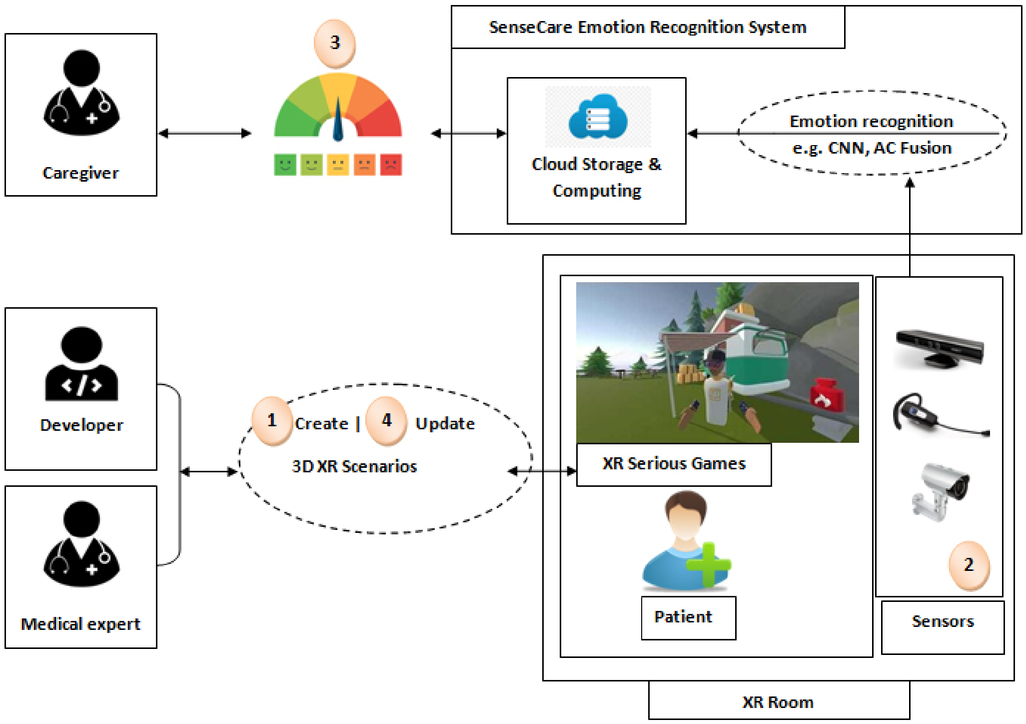

Figure 2: Conceptual design of WebXR-Based Immersive Serious Game for Rehabilitation.

patient. The platform is hosted by a Ubuntu Linux server operating system and Node.js web server. It has a list of serious games developed using the WebXR API [3] [29]. WebXR was selected due to the following attributes:

1. The API accesses input (headset and controller pose information) and output (hardware display) capabilities usually associated with VR and AR devices. It allows users to create and host web-based VR and AR experiences [30].
2. The WebXR Device API delivers the interfaces needed for developers to create appealing, practical, and secure immersive web apps on a range of hardware solutions.
3. The API enables the development, and hosting, of VR and AR.
4. Immersive web environments allow interaction, collaboration, and communication between patient and therapist, and also amongst patients without physical contact.
5. Users may access their experience without leaving their website because no App Stores or large downloads are required.

The WebXR device API is dependent on graphics APIs such as WebGL, WebGL2, and A-Frame [31].

4.1.1 WebVR game experimentation. Example games shown in Figure 3 (Game A, Game B) use A-Frame, and apply real-world physics to entities in a scene using the attributes of an A-Frame physics system. The use of 3D A-Frame models generates impressions of virtual controllers that allow a user to interact with a scene [3].

Game A: This game enables natural and easy interactions with the connected controller. The objective of this game is to rehabilitate upper limbs. The patient uses VR controllers to grab, stretch, and drag-drop virtual cubes inside a color-changing box to change their color. URL link: https://studev4.fernuni-hagen.de:20283/vrhand2/.

Game B: A chess game enables users to interact with and move chess pieces in a 3D world using VR controllers. To utilize it, a web browser (Chrome) is required. URL link: https://studev4.fernuni-hagen.de:20283/vrchess/.

4.1.2 WebAR game experimentation. Figure 4 shows the architecture and procedure of a WebAR rehabilitation game. A screenshot from a Smartphone of output from AR rehabilitation of an upper limb (Game C) is shown in Figure 5.

Game C: An example of Marker-based AR, uses ar.js [3] with Marker Tracking (e.g., Hiro, Kanji, custom pattern) for upper limb gestures. The user uses tangible objects to move virtual cylinders from one position to another.

A demonstration of web-based AR is available at this URL link: https://studev4.fernuni-hagen.de:20283/arhand.

4.2 SenseCare KM-EP Audiovisual Monitoring

SenseCare focuses on improving medical care by creating a variety of input interfaces for various sensory devices, such as cameras and wearable sensors from the Internet of Things. The SenseCare web platform already has a number of analysis techniques for emotion recognition [14] [13]. The SenseCare KM-EP provides player monitoring using audiovisual sensors as discussed below.

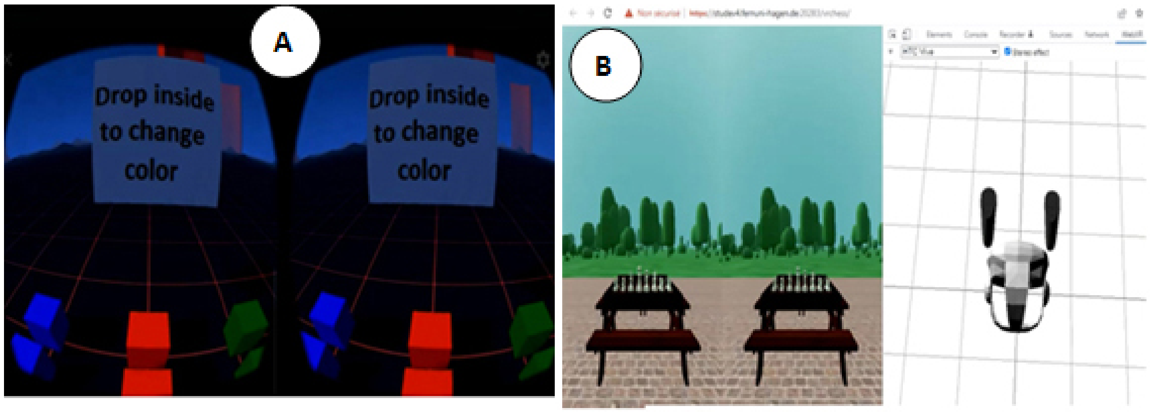

Figure 3: Conceptual Screenshots of example WebVR games (A: Dropping; B: Chess).

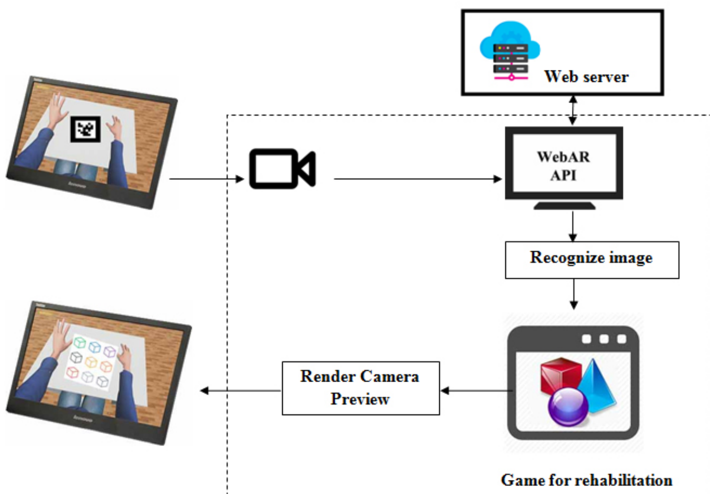

Figure 4: Architecture and Procedure of a WebAR Rehabilitation Game.

Figure 5: Screenshot of WebAR game.

4.2.1 Facial Emotion Recognition (FER).. The SenseCare KM-EP employs CNNs and TensorFlow for FER. This methodology, detailed previously in [13], can identify automatically facial emotion expressions from the camera whilst playing the game. The API classifies FER into 7 categories in real time for every 500 milliseconds (happy, sad, angry, disgusted, fearful, neutral, and surprised). We separate them into 3 categories (neutral, positive, negative). From time (t = 0) through time (t = n), positive and negative emotions are calculated as follows:

$$\sum_{t=n}^{t=0}(positive) = \sum_{t=n}^{t=0}(happy) + \sum_{t=n}^{t=0}(surprised)/2 \qquad (1)$$

$$\sum_{t=n}^{t=0}(negative) = \sum_{t=n}^{t=0}(angry) + \sum_{t=n}^{t=0}(fearful) + \sum_{t=n}^{t=0}(disgusted) + \sum_{t=n}^{t=0}(sad) + \sum_{t=n}^{t=0}(surprised)/2 \qquad (2)$$

The following formula determines the emotion recognition accuracy from facial expressions and speech:

$$Accuracy = \frac{\text{Accuracy SER} + \text{Accuracy FER}}{2}$$

4.2.2 Speech Emotion Recognition (SER).. The MFCCs (Mel-Frequency Cepstral Coefficients) technique is the most extensively used method for extracting characteristics from audio sources [16]. Hence, it is used here to recognize the emotion in the player's recorded voice during gameplay. The Librosa Python package [20] was employed for digital representation of the results with Matplotlib and TensorFlow.

4.2.3 Contactless vital signs. Pulse rate monitoring from the camera using the rPPG technique [13] in the browser provides contactless vital signs. The Viola-Jones algorithm [28] is employed to define the Region of Interest on the face for signal extraction processing, and Haar-like features [21] are employed for face detection. URL demo link: https://studev4.fernuni-hagen.de:20286/.

4.2.4 Body monitoring gestures. For body monitoring gestures, we use MediaPipe BlazePose [4] for pose estimation. It provides real-time human body pose perception in the browser, works up to 4 meters from the camera and predicts the location of 33 pose keypoints (e.g., hands, feet). A demonstration using WebGL and TensorFlow.js is available at this URL link: https://studev4.fernuni-hagen.de:20284/body. Position data can be used to visually model 3D objects in space. These elements can be combined with a real-world scene captured with the camera. The pose estimate score is utilized to evaluate a patient's performance and provide real-time feedback to the therapist.

The following requirements are necessary for the system to perform successfully: (1) the HTTPS (Hypertext Transfer Protocol Secure) is needed on the web server so that the camera can be started in browsers on various devices (e.g., Android, Windows) both for playing AR games and for visual monitoring of the player, and (2) for accurate visual monitoring, the playroom environment should have high quality interior lighting. The Oculus Quest hand tracking API is now supported by WebXR in the Oculus Browser. Using hands without controllers is possible according to Quest 2's new Hand Tracking 2.0 mode (Oculus). This approach facilitates hand movement, range of motion, and muscle development.

The following steps are required for the for the treatment system to function successfully: (1) develop a training plan for the patient's recovery and utilization of the VR Room on a regular basis, (2) show patients how to use the VR Room's equipment for brain rehabilitation and physical exercises, and (3) notify a person (family member or others) of the training plan for assisting the patient and prepare to commence the treatment session within the set time frame.

5 EVALUATION

Effective healthcare information system evaluations are necessary to ensure that these systems meet the needs and expectations of users and healthcare organizations. Early in the design cycle, the cognitive walkthrough technique [24] conducts theory-based assessments of user interface designs. Cognitive walkthroughs enable systems to be evaluated for usability from the perspective of system developers, as well as providing direct input to help with the design process. Currently, the proposed VR/AR games have not tested directly with patients. An online survey (Google Forms) was conducted with a URL[1] link to the VR game (Game A) on the web in order to determine satisfaction ratings. The survey URL link was emailed to 21 persons who volunteered to participate and to 2 medical experts. Initial results of 11 responses show an 81.8% satisfaction rate.

6 CONCLUSION AND FUTURE WORK

To recover from a stroke, rehabilitation is essential. Here, we propose techniques for using VR and AR games on the web applied to motor rehabilitation of upper limbs. We have discussed SenseCare software platform methods for player audiovisual monitoring with a camera, Kinect, and microphone sensors, i.e., emotions (FER, SER), contactless vital signs (heart rate), and pose body estimation. An online survey was hosted on the web to measure user satisfaction with Game A (Dropping box to change color). There was an 81.8 % satisfaction rate in the first round of 11 responses. Future work includes: (1) enhancing AR-related features (e.g., therapist Avatar), (2) calculation of a range of motions, comparing results with the demo therapist score, (3) more interactive games could be developed including, e.g., audio, multi-users, and (4) evaluation of the system with actual patients.

REFERENCES

[1] Hamdi Ben Abdessalem, Yan Ai, K S Marulasidda Swamy, and Claude Frasson. 2021. Virtual Reality Zoo Therapy for Alzheimer's Disease Using Real-Time Gesture Recognition. Adv. Exp. Med. Biol. 1338, (2021), 97–105. DOI:https://doi.org/10.1007/978-3-030-78775-2_12

[2] Paula Amorim, Beatriz Sousa Santos, Paulo Dias, Samuel Silva, and Henrique Martins. 2020. Serious games for stroke telerehabilitation of upper limb-a review for future research. Int. J. telerehabilitation 12, 2 (2020), 65.

[3] Rakesh Baruah. 2021. Physics and User Interaction in A-Frame. In AR and VR Using the WebXR API: Learn to Create Immersive Content with WebGL, Three.js, and A-Frame. Apress, Berkeley, CA, 289–302. DOI:https ://doi.org/10.1007/978-1-4842-6318-1_9

[4] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. 2020. Blazepose: On-device real-time body pose tracking. arXiv Prepr. arXiv2006.10204 (2020).

[5] Sara Bernini, Fabrizio Stasolla, Silvia Panzarasa, Silvana Quaglini, Elena Sinforiani, Giorgio Sandrini, Tomaso Vecchi, Cristina Tassorelli, and Sara Bottiroli. 2021. Cognitive Telerehabilitation for Older Adults With Neurodegenerative Diseases in the COVID-19 Era: A Perspective Study. Front. Neurol. 11, (2021). DOI:https://doi.org/10.3389/fneur.2020.623933

[6] Nunzio Alberto Borghese, Michele Pirovano, Pier Luca Lanzi, Seline Wüest, and Eling D de Bruin. 2013. Computational intelligence and game design for effective at-home stroke rehabilitation. Games Heal. Res. Dev. Clin. Appl. 2, 2 (2013), 81–88.

[1] Survey URL: https://docs.google.com/forms/d/e/1FAIpQLScMk9Bzl2KUxLx5Uo5aBpyF-tC8574HS5LMKuCBqfL6sBtiHA/viewform?vc=0&c=0&w=1&flr=0 .

[7] Tiago Coelho, Cátia Marques, Daniela Moreira, Maria Soares, Paula Portugal, António Marques, Ana Rita Ferreira, Sónia Martins, and Lia Fernandes. 2020. Promoting Reminiscences with Virtual Reality Headsets: A Pilot Study with People with Dementia. Int. J. Environ. Res. Public Health 17, 24 (December 2020). DOI:https://doi.org/10.3390/ijerph17249301

[8] Sanika Doolani, Callen Wessels, Varun Kanal, Christos Sevastopoulos, Ashish Jaiswal, Harish Nambiappan, and Fillia Makedon. 2020. A Review of Extended Reality (XR) Technologies for Manufacturing Training. Technologies 8, 4 (2020). DOI:https://doi.org/10.3390/technologies8040077

[9] Felix Engel, Raymond Bond, Alfie Keary, Maurice Mulvenna, Paul Walsh, Huiru Zheng, Haiying Wang, Ulrich Kowohl, and Matthias Hemmje. 2016. SenseCare: Towards an Experimental Platform for Home-Based, Visualisation of Emotional States of People with Dementia. In Advanced Visual Interfaces. Supporting Big Data Applications, Springer International Publishing, Cham, 63–74.

[10] Cristian Gmez-Portes, Carmen Lacave, Ana I Molina, and David Vallejo. 2020. Home Rehabilitation Based on Gamification and Serious Games for Young People: A Systematic Mapping Study. Appl. Sci. 10, 24 (2020). DOI:https://doi.org/10.3390/app10248849

[11] Silvia Guillén-Climent, Ainara Garzo, María Nieves Muñoz-Alcaraz, Pablo Casado-Adam, Javier Arcas-Ruiz-Ruano, Manuela Mejías-Ruiz, and Fernando Jesús Mayordomo-Riera. 2021. A usability study in patients with stroke using MERLIN, a robotic system based on serious games for upper limb rehabilitation in the home setting. J. Neuroeng. Rehabil. 18, 1 (2021), 41. DOI:https://doi.org/10.1186/s12984-021-00837-z

[12] H Hadjar, J Lange, Binh Vu, Felix C Engel, Gwendolyn Mayer, P McKevitt, and Matthias L Hemmje. 2020. Video-based automated emotional monitoring in mental health care supported by a generic patient data management system. In PSYCHOBIT.

[13] Hadjar Hayette, Vu Binh, Dennis Maier, Gwendolyn Mayer, Paul Mc Kevitt, and Matthias Hemmje. 2021. Video-based emotion detection analyzing facial expressions and contactless vital signs for psychosomatic monitoring. CERC2021 (2021).

[14] Michael Healy, Ryan Donovan, Paul Walsh, and Huiru Zheng. 2018. A Machine Learning Emotion Detection Platform to Support Affective Well Being. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2694–2700. DOI:https://doi.org/10.1109/BIBM.2018.8621562

[15] Johanna Jonsdottir, Rita Bertoni, Michael Lawo, Angelo Montesano, Thomas Bowman, and Silvia Gabrielli. 2018. Serious games for arm rehabilitation of persons with multiple sclerosis. A randomized controlled pilot study. Mult. Scler. Relat. Disord. 19, (2018), 25–29. DOI:https: //doi.org/https://doi.org/10.1016/j.msard.2017.10.010

[16] K V Krishna Kishore and P Krishna Satish. 2013. Emotion recognition in speech using MFCC and wavelet features. In 2013 3rd IEEE International Advance Computing Conference (IACC), 842–847. DOI:https://doi.org/10.1109/IAdCC.2013.6514336

[17] Sofoklis Kyriazakos, Hannes Schlieter, Kai Gand, Massimo Caprino, Massimo Corbo, Peppino Tropea, Elda Judica, Irma Sterpi, Stefan Busnatu, Patrick Philipp, Jordi Rovira, Alvaro Martínez, Marc Lange, Inigo Gabilondo, Rocio Del Pino, Juan Carlos Gomez-Esteban, Lucia Pannese, Morten Bøttcher, and Vibeke Lynggaard. 2020. A Novel Virtual Coaching System Based on Personalized Clinical Pathways for Rehabilitation of Older Adults—Requirements and Implementation Plan of the vCare Project. Front. Digit. Heal. 2, (2020). DOI:https://doi.org/10.3389/fdgth.2020.546562

[18] Minhua Ma and Kamal Bechkoum. 2008. Serious games for movement therapy after stroke. In 2008 IEEE International Conference on Systems, Man and Cybernetics, 1872–1877. DOI:https://doi.org/10.1109/ICSMC.2008.4811562

[19] Minhua Ma, M McNeill, D Charles, and S McDonough. 2006. Virtual Reality Games for Motor Rehabilitation. . University of Wolverhampton.

[20] Brian McFee, Colin Raffel, Dawen Liang, Daniel P Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. 2015. librosa: Audio and music signal analysis in python. In Proceedings of the 14th python in science conference, 18–25.

[21] T Mita, T Kaneko, and O Hori. 2005. Joint Haar-like features for face detection. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 1619-1626 Vol. 2. DOI:https: //doi.org/10.1109/ICCV.2005.129

[22] Joao Monge and Octavian Postolache. 2018. Augmented Reality and Smart Sensors for Physical Rehabilitation. In 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), 1010–1014. DOI:https://doi.org/10.1109/ICEPE.2018.8559935

[23] Belkacem Mostefai, Amar Balla, and Philippe Trigano. 2019. A generic and efficient emotion-driven approach toward personalized assessment and adaptation in serious games. Cogn. Syst. Res. 56, (August 2019), 82–106. DOI:https://doi.org/10.1016/J.COGSYS.2019.03.006

[24] Peter G Polson, Clayton Lewis, John Rieman, and Cathleen Wharton. 1992. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. Int. J. Man. Mach. Stud. 36, 5 (1992), 741–773. DOI:https: //doi.org/https://doi.org/10.1016/0020-7373(92)90039-N

[25] Francisca Rosique, Fernando Losilla, and Pedro J Navarro. 2021. Applying Vision-Based Pose Estimation in a Telerehabilitation Application. Appl. Sci. 11, 19 (2021). DOI:https: //doi.org/10.3390/app11199132

[26] Amritanshu Kumar Singh, Vedant Arvind Kumbhare, and K Arthi. 2022. Real-Time Human Pose Detection and Recognition Using MediaPipe. In Soft Computing and Signal Processing, Springer Singapore, Singapore, 145–154.

[27] John Torous, Keris Jän Myrick, Natali Rauseo-Ricupero, and Joseph Firth. 2020. Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. JMIR Ment. Heal. 7, 3 (March 2020), e18848. DOI:https: //doi.org/10.2196/18848

[28] P Viola and M Jones. 2001. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, I–I. DOI:https://doi.org/10.1109/CVPR.2001.990517

[29] GitHub - immersive-web/webxr: Repository for the WebXR Device API Specification. Retrieved February 23, 2022 from https://github.com/immersive-web/webxr

[30] Immersive Web Developer Home. Retrieved March 10, 2022 from https://immersiveweb.dev/

[31] GitHub - networked-aframe/networked-aframe: A web framework for building multi-user virtual reality experiences. Retrieved March 2, 2022 from https://github.com/networked-aframe/networked-aframe