International Journal of Multimedia Data Engineering and Management
Volume 13 • Issue 1

Multimodal Data Integration and User Interaction for Avatar Simulation in Augmented Reality

Anchen Sun, University of Miami, USA
Yudong Tao, University of Miami, USA*
Mei-Ling Shyu, University of Miami, USA
https://orcid.org/0000-0003-0902-0844
Angela Blizzard, University of Miami, USA
William Andrew Rothenberg, University of Miami, USA
Dainelys Garcia, University of Miami, USA
Jason F. Jent, University of Miami, USA

ABSTRACT

Augmented reality (AR) allows users to interact with the virtual world in the real-world environment. This paper proposes a child avatar simulation framework using multimodal data integration and user interaction in the AR environment. This framework generates the child avatar that can interact with the user and respond to his/her behaviors and consists of three subsystems: (1) The avatar interaction system scrutinizes user behaviors based on the user data. (2) The avatar action control system generates naturalistic avatar activities (actions and voices) according to the avatar internal status. (3) The avatar display system renders the avatar through the AR interface. In addition, a child tantrum management training application has been built based on the proposed framework. And a light machine learning model has been integrated to enable efficient and effective speech emotion recognition. A sufficiently realistic child tantrum management training based on the evaluation of clinical child psychologists is enabled, which helps users get familiar with child tantrum management.

KEYWORDS

Augmented Reality (AR), Behavior Analysis, Child Tantrum, Data Integration, Gaze Tracking, Human-Machine Interaction, Immersive Media, Machine Learning, Multimodality, Speech Emotion Recognition

INTRODUCTION

Simulating avatars in an Augmented Reality (AR) environment becomes an important step for the users to attend to the virtual world. A realistic, vivid, and friendly avatar helps users quickly integrate

DOI: 10.4018/IJMDEM.304391
*Corresponding Author

into the virtual world (Feng et al., 2014). Existing algorithms have already simulated the voice and movement of an avatar and made it appear as a real person in the virtual world (Yang & Bhanu, 2011). Combined with AR technology, avatar simulation has the potential to evolve traditional human-computer interaction in 2D space to 3D mixed reality interaction in realistic environments (Pouyanfar, Yang, et al., 2018). In addition, services such as intelligent voice assistants and intelligent butlers are integrated to further enhance avatar-based interaction in the AR environment. With the development of avatar simulation technology in the AR environment and deep learning, more and more industries begin to take notice of this technology and apply it in the fields of entertainment, education, and healthcare (Pouyanfar, Sadiq, et al., 2018).

Childcare has always been a crucial part of the field of healthcare. Taking care of the children in the early childhood is extremely challenging since they are not yet able to express their needs with words, and their emotions can be intense (Österman & Björkqvist, 2010). Teachers, caregivers, parents, and other groups who have long-term contact with children need the training to manage children's tantrums. Good child tantrum management skills can also help improve teacher satisfaction (Schaack et al., 2020). Teachers in special education schools need to develop their skills to manage children's emotions and attend to tantrums when dealing with children with disabilities (Konst et al., 2013). Due to the environment and time, it is difficult for people to receive targeted professional training in the real world. Therefore, the introduction of avatar simulation combined with AR technology into pediatrics is one of the possible ways to address such childcare challenges.

Confronting the tantrum behaviors that result from early childhood children in real-life situations is the best training for teachers, caregivers, parents, and others who come into direct contact with children (Eisbach et al., 2014). However, it is impossible to create such a training environment in the reality nor provide a tantrum child to train teachers, caregivers, and parents (Belden et al., 2008). In the past, they have gradually learned from their experience in practice and trained newcomers through textual abstraction. However, introducing virtual avatar simulation in the real-world environment to simulate the child's behavior and provide interactive training for teachers, caregivers, and parents through AR can fundamentally revolutionize the existing childcare training methods, significantly improving the training effect.

In recent years, although virtual avatars have become widely available in education (Davis, 2019), healthcare (Foronda et al., 2020), and entertainment (Rusdorf et al., 2007), existing applications and frameworks have apparent limitations. Many nursing and educational applications simply place a fixed-process avatar in a virtual environment that cannot interact with the user in real-time (Kidd et al., 2012). Even for applications that interact with users in real-time, they mostly use controllers instead of realistic human interaction such as eyes, body movements, and voices (Albright et al., 2012).

Therefore, this paper proposes an avatar simulation framework that runs in an AR environment, aggregates multimodal data, and interacts with users. The framework combines the environment detected by the AR device, the avatar's actions, voice, and the user's behavioral data to naturally perform actions and voices on the avatar simulated in the AR environment to interact with the user. A child tantrum management training application was developed in collaboration with a team of clinical child psychologists to demonstrate the proposed framework and meet the team's expectations. The application simulates a child having a tantrum in an AR environment and interacts with the user's gaze and speech. Using it in different scenarios (home, school, hospital, etc.) creates a training environment that links to the real world and allows the user to experience more realistic child tantrums in each scenario. The application analyzes the user's behaviors of different modalities and simulates the child's behavior, helping the user learn how to manage child tantrum properly.

The rest of the paper is organized as follows. The paper discusses the related work in the second section and our proposed framework in the third section. The fourth section presents the experiments and analyses the proposed framework based on the child tantrum management training application as the use case. Conclusions and future work are given in the fifth section.