Artificial Intelligence and Human-Computer Interaction
Y. Ye and P. Siarry (Eds.)

doi:10.3233/FAIA240162

Intelligent Control of Character for AR-Based Remote Psychological Counseling System

Yihua BAO [a], Zhen SONG [b,1], Dongdong WENG [a], Yue LIU [a], Mo SU [a] and Junyu CHEN [b]
[a] *Beijing Engineering Research Center of Mixed Reality and Advanced Display, Beijing Institute of Technology, China*
[b] *Advanced Research Center for Digitalization of the Traditional Drama, The Central Academy of Drama, China*

Abstract. As living standards escalate, the need to prioritize psychological and mental health has become increasingly clear. Traditional psychological counseling, constrained by fixed locations and schedules, cannot adequately meet the growing demand for informatization and intelligence. To overcome these constraints, we developed an Intelligent Control Auxiliary System (ICAS) for use in an Augmented Reality (AR)-based remote psychological counseling setting. A comparative experiment was conducted to assess the efficacy of ICAS against the system powered by motion capture technology, with participation from eight licensed psychological counselors who carried out system operations and contributed their professional insights. Most counselors recognized the superior convenience of the intelligent control mode when operating the remote psychological counseling system. Demonstrating similar effectiveness to traditional motion capture control systems, ICAS suggests a promising new pathway for remote counseling methodologies.

Keywords. ICAS, augmented reality, remote psychological counseling, motion capture system

1. Introduction

Remote psychological counseling is a new form of psychological health service that has emerged in recent years. Global rates of depression and anxiety have continued to rise since the new coronavirus disease pandemic [1], which has increased demand for counselors as well as for remote psychological counseling. Although certain remote counseling sessions can be supported by internet and video communications, these mediums are insufficient for the transfer of information such as body language in interpersonal encounters, which can cause some disruptions to counselors' work. Virtual human technology is becoming widely used as a method to treat mental disorders such as anxiety and depression, thanks to the advancement of virtual reality technology. The use of virtual human technology to aid in the early detection and management of

[1]Corresponding Author: Zhen Song, Advanced Research Center for Digitalization of the Traditional Drama, The Central Academy of Drama, China. E-mail: songzhen@zhongxi.cn.

depression and anxiety is receiving increasing interest from researchers [2-5]. Studies have shown clients prefer virtual humans to anonymous surveys or face-to-face conversations when discussing mental health difficulties [4]. The current focus of research is on improving the human-like look and expressive capabilities of virtual humans, but it pays less attention to the influence of the practicality and intelligence of the technology powering the virtual human on the work of counselors.

The present study introduces an innovative Intelligent Control Auxiliary System (ICAS) that utilizes Augmented Reality (AR) technology. Through simple button controls, this system enables therapists to control virtual agents that interact with clients and speak with them while demonstrating care and purpose. Due to its simplicity of use, ICAS stands in sharp contrast to the conventional approach to controlling virtual agents through motion capture. While ensuring the integrity of information transfer in remote counseling, this simplicity can significantly reduce the strain for counselors.

In comparative experiments and interviews with eight counselors, ICAS effectively replaced motion capture, with the benefit of being simpler. In conclusion, ICAS based on Augmented Reality provides a user-friendly interface that not only reduces counselor workload but enhances convenience while maintaining information integrity during remote interactions. Further exploration of this promising approach to improve remote psychological counseling is warranted.

Key contributions include:

- Introducing ICAS, a novel system reducing counselor strain and boosting effectiveness.
- A prototype incorporating ICAS and standard motion capture.
- Counselor study results validating ICAS' potential to replace motion capture through simplified operation, reduced workload, and improved efficiency.

2. Related work

2.1. Remote psychological counseling technology

With the development of information technologies, new remote counseling technologies based on video and audio have emerged, greatly expanding the reach of counseling services. Existing studies have shown the advantages of remote counseling in improving accessibility and protecting privacy [6]. However, these technologies also have limitations like information loss and restricted interaction [7]. In traditional face-to-face counseling, nonverbal cues are very important for understanding clients' emotions and thoughts. Zeren's study found 73% of counselors encountered difficulties of mutual understanding in online counseling [8], and Haberstroh et al.'s study suggested the lack of nonverbal cues restricted online counselors' ability to convey empathy and build rapport [9]. Nevertheless, with the increase of available technologies, more and more mental health services are provided virtually [10,11]. Persky's study found immersive technologies have more advantages over traditional therapies, including easier experimental control and higher sense of reality [12].

2.2. Virtual characters in psychological counseling

In recent years, virtual characters have been increasingly applied in psychological counseling and achieved promising results [13-16]. With the development of artificial intelligence and computer graphics, more advanced virtual characters are created and utilized to enhance counseling practices. Studies have shown virtual characters as assistive tools can encourage self-disclosure and improve counseling satisfaction [13]. Meanwhile, with the fast development of avatar technologies, increasing research has implemented avatars in VR environments for psychological therapies and achieved results. Ilona Halim et al. proposed personalized VR therapies [17]. Lin et al. developed an online chatbot testing platform to relieve student stress [18]. Currently, virtual characters are mainly controlled by motion capture [19], which is complex to operate. Joy O. Egede's study demonstrated virtual human mediators improved users' attitude toward tasks [20]. Although applications of virtual characters have shown some effects, their control methods like motion capture are complex and hard to scale up. Therefore, a more intelligent and convenient control method is needed.
First paragraph.

In current evaluation scenarios, while existing remote counseling technologies and virtual character applications provide considerable value, they often do not completely meet the growing demands for improved informatization and intelligence. To fill these noted gaps, we propose the novel ICAS system. Our goal with ICAS is to surmount specific limitations and to offer an optimized and intelligent pathway for psychological counseling. In the following sections, we discuss the design and empirical results of this innovative system.

3. System design

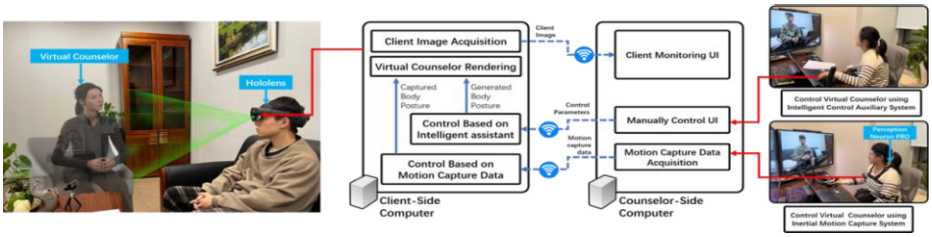

Figure 1. Remote users interact with a virtual counselor by wearing an AR headset, and the local counselor controls the virtual counselor using two methods: ICAS and motion capture system.

We propose an augmented reality system grounded in remote psychological counseling, providing local users, i.e., counselors, the ability to manipulate virtual agents via two distinct approaches, thereby interacting with remote users, or clients. The effectiveness of the ICAS has been verified through comparative analyses involving traditional motion capture systems, the fundamental structure of which is depicted in Figure 1. In demonstrating the viability of the ICAS, we will employ preliminary user research.

3.1. Early user research

To gain an in-depth understanding of the challenges faced by counselors in remote psychological counseling, we conducted preliminary user research involving structured

interviews and demand surveys with four senior counselors from fields like adolescent health. Prior to the interviews, we first explained key technologies to be used in our research, including augmented reality head-mounted displays, remote proxy operating systems, and motion capture devices. The counselors then discussed issues encountered in remote counseling and provided suggestions on potential technical assistance programs. The majority preferred virtual agents with realistic human images, while one was open to animal or cartoon images. They unanimously recommended that a counselor take on simulated client scenarios to validate the system's effectiveness. One counselor contributed an anonymized real case as a template for testing, with patient consent. Based on counselor feedback, we categorized clients into venting, defensive, and super-rational types. Accordingly, virtual agent reactions like nodding, smiling, leaning forward, hand gestures, and arm movements were customized to suit counselor needs. These insights provide important references for our subsequent experimental designs.

3.2. Avatar generation and animation

Studies show clients prefer virtual agents over real humans, with higher realism increasing trust [21]. To achieve realism, we used a light stage system to obtain neutral and multi-expression facial images of people. High quality multi-view images allowed reconstructing intricate 3D head models with rich expressions. We further converted these into detailed 3D face models with blendshapes imported into Unreal Engine 5 to precisely drive highly realistic virtual agents. Abundant facial and subtle motions enhanced realism for accurate control. We captured 40 typical expressions to build the animation library. In constructing the virtual counselor, we ensured suitable attire without distraction. Referring to animation techniques like keyframing, motion capture, and simulation [22], we used Noitom motion capture to acquire counselor body animations and redirect them to the virtual agent in MotionBuilder. iPhone 13 Pro's ARKit captured facial animations to drive the virtual counselor's face, building a nonverbal animation database.

Figure 2. Three-level control model for ICAS.

3.3. Intelligent control auxiliary system (ICAS)

We developed the Intelligent Control Auxiliary System (ICAS) to aid remote psychological counseling. Early counselor interviews identified three typical client communication styles: cathartic, defensive, and hyper-rational. As clients may switch between these, counselors need to dynamically adjust.

As shown in Figure 2, ICAS has three activation modes-high(level 1), medium(level 2), and low(level 3). The high mode represents the virtual agent's highest intelligent behavior level, followed by decreasing levels. For instance, in high mode, the virtual counselor persistently exhibits attentive body language and expressions for engagement. Specifically, the virtual counselor maintains an open frontal posture, continually expressing care through micro-expressions, leaning forward, while staying calm and rational to build trust. The virtual counselor's head and gaze focus steadily on the client.

In medium mode, the counselor uses more moderate expressions, smiling, and questions to show friendly care. In low mode, the counselor reduces engagement, occasionally nodding and adjusting posture to relax the client. Counselors can easily switch expression based on the client's changing states to efficiently utilize time and record information.

The ICAS system intelligently retrieves nonverbal behaviors for the three modes from an animation database for natural, coherent facial and body motion without additional motion capture. Counselors only make selections through our system, altering activeness levels to cater to different client needs, providing intelligent and efficient assistance.

3.4. Motion capture system

Traditional motion capture-based control of virtual agents requires professional motion capture equipment and smartphone cameras. Specifically, counselors need to wear Noitom Perception Neuron PRO inertial motion capture devices and set up iPhone 13 Pro phone cameras for facial expression capture. The motion capture device needs to be calibrated, and the phone camera parameters like ARKit IP settings need to be configured. Currently, motion capture techniques can be divided into three types: optical, inertial, and optical-inertial hybrid capture. Considering factors such as cost, operational convenience, and capture accuracy, this study chose the solution combining Noitom inertial motion capture devices and iPhone cameras. Through this system, counselors can drive virtual agents using real-time capture of facial and body motions, enabling natural interaction with clients.

4. Experiment

Figure 3. (a) Perspective of the remote user. (b) The local space. (c) The remote space.

4.1. Experiment environment

As shown in Figure 3, we established an experimental environment to validate and compare two systems. Based on early user research, we simulated remote psychological

counseling using different systems for comparison. An experienced counselor portraying a scripted client did not limit dialogue but primarily presented cathartic, defensive, and hyper-rational phases based on the character's background, aiming to elicit counselling techniques and nonverbal information from counselors. Equipped with augmented reality, the client directly observed and interacted with the virtual counselor. Counselors facilitated interaction through two discrete systems controlling virtual agents, while observing the client's state via camera.

Due to spatial constraints in counseling activities, we used two 3m x 2m spaces: a remote home-like space for the client, ensuring open space to display the counselor's virtual agent, and a local space with basics like a chair, computer, and motion capture devices for operation. The spaces were isolated by a partition, preventing the client from viewing the local space while the counselor monitored the client.

4.2. Supported unit

Regarding computing needs, both spaces were furnished with high-performance machines (Intel Core i9-7900X processor, NVIDIA GeForce RTX 3080 graphics card, 32GB RAM, 1TB SSD) to ensure seamless system operation. For an immersive virtual experience, users wore HoloLens 2 augmented reality headsets (monocular resolution 2560 * 1440, the field of view 52 °, and refresh rate 60 fps). Remote C920 cameras were deployed to enable counselors to discern subtle emotional shifts on the clients' faces. High-fidelity external microphones captured voice data. All devices were interconnected via a high-speed LAN, ensuring low-latency, stable system communication.

The Noitom Perception Neuron PRO inertial capture device collected the counselor's limb movements, controlling the movements of virtual character. Facial expressions and lip movements were relayed in real time using the live link face software on the iPhone 13 Pro.

4.3. Participants

Eight professional psychological counselors (four men, four women, with ages ranging from 27 to 58 and an average age of 42) were invited to participate in our study. All were certified as national second-level psychological counselors with a minimum of three years' professional experience and substantial practice in counseling. They were recruited online and compensated RMB 100 each for their involvement. We surveyed participants' prior experience with virtual reality and motion capture. Five participants had no knowledge of VR, three had VR experience, three had used motion capture devices before, and five had no knowledge of motion capture.

4.4. Experimental procedure

The local psychological counselors, acting as study participants, were required to conduct a counseling session using both the ICAS system and the motion capture system. The order in which the systems were used was randomized to avoid sequence effects. All sessions lasted a minimum of 15 minutes to ensure a thorough evaluation of the system's functionality.

At the onset of the experiment, the participants completed a demographic questionnaire. The experimenter explained the purpose and tasks of each session before beginning. Each session was preceded by a 10-minute briefing on the respective system.

Once participants were ready, timing began for the simulated counseling. After the first system trial, participants answered a questionnaire, rested for 5-10 minutes as the room and system were reset, then began the second system trial.

The experiment concluded with participant interviews about the functionality and convenience of the two systems. On average, the entire experiment lasted 45 minutes. The motion capture system trial often lasted an hour due to the additional time required to equip and calibrate the motion. In the motion capture system, it is important to ensure the counselors have correctly worn the devices and powered on each tracker, and for facial capture the counselors face needs to stay fully and completely visible to the front camera of the iPhone.

4.5. Measurements

To evaluate the performance of the two systems, we used the following metrics:

System Usability Scale. We evaluated the usability of the system for each condition using the System Usability Scale (SUS) [23]. This metric employs a 5-point scale, where 1 signifies "strongly disagree", and 5 indicates "strongly agree". Users assigned an appropriate score for each question, reflecting their subjective perceptions. For scoring consistency, we deducted 1 from all odd-numbered responses and subtracted the score obtained from 5 for even-numbered responses. The final usability score was the sum of all these adjusted scores, multiplied by 2.5.

Nasa-Task Load Index (NASA-TLX). To assess the users' workload for each condition, we utilized the NASA-TLX [24]. This scale employs a 7-point Likert scale, where 1 indicates "strongly disagree, " and 7 denotes "strongly agree. " The final scores were compartmentalized into 21 distinct ratings.

4.6. Result

Utilizing IBM SPSS statistical tool, we subjected all pairwise comparisons of the experimental result data to Bonferroni correction. The Shapiro-Wilk normality test verified the conformance of the score data to a normal distribution. Subsequently, a two-sample t-test facilitated an analysis of the SUS scores under two system conditions, observing that the difference in the overall scores of the two systems is not statistically significant ($p > 0.05$). This suggests a near-equivalence in user experience from a usability standpoint for both systems.

The workload assessment results did not display significant variances between the two system conditions. Interestingly, participants utilizing the ICAS reported notably lower workload scores in terms of mental workload and physiological needs along with a faster task completion time. This denotes a comparative advantage of the ICAS in diminishing the workload of counselors. However, participants expressed greater satisfaction with the performance of the motion capture system.

5. Discussion and future work

Our study represents one of the earliest investigations comparing AR-based intelligent remote agent systems to augment counselor efficiency in China. We hypothesized the ICAS could supersede traditional remote agent control systems due to superior convenience and acceptance. Results partially confirmed no significant usability

difference between ICAS and motion capture, supporting our hypothesis. However, ICAS reduced mental and temporal demand versus traditional systems, demonstrating potential to address high counselor workload. But ICAS' performance falls short of motion capture, suggesting traditional systems may provide superior real-time control, based on discussions. Future studies could verify counseling effectiveness by collecting data from remote clients.

Compared to motion capture, ICAS offers simplified substitute features, making it particularly beneficial for remote counseling aids and lightweight modalities. We initially performed a single camera comparison study but its occlusion susceptibility and two-dimensional information limitations made it suboptimal for high-fidelity digital human reproduction. Thus, we chose an inertial six-degree-of-freedom motion capture system with wider usability as our control group. Currently, ICAS primarily conveys nonverbal cues like expressions and postures. As these are finite, no large model reasoning algorithms were used. Counselors still drive content and progression, now with enhanced nonverbal communication. Future iterations could incorporate reasoning algorithms into remote aids for assistance, requiring an agent-counselor collaboration paradigm.

Preliminary user studies validated reliability of ICAS' remote agent control modes with varying proactivity degrees. Counselors generally accepted distinct control modes derived from different client typologies. Interviews revealed comparable expressiveness between systems, but ICAS' wider movement ranges fostered greater coexistence and validation, warranting further investigation. Counselors acknowledged the virtual agent model, desiring future customizations like attire, hairstyle, and gender which could influence outcomes.

Limitations include HoloLens 2's limited field of view and resolution, and rainbow light effects. ICAS also has limitations, requiring additional algorithms if more functions like intent inference are needed.

6. Conclusions

In this paper, we developed an Intelligent Control Auxiliary System (ICAS) for remote counseling utilizing a virtual agent. The system empowers counselors to communicate non-verbal information via the virtual agent, thereby enhancing the efficacy of remote counseling. A prototype of this AR remote counseling system was formulated and subsequently validated through comparative experimentation with conventional motion capture systems. The findings suggest that the ICAS aligns with the traditional system in terms of usability and efficacy, while offering easier operation. At the same time, the counselors participating in the experiment believed that using virtual agents for remote counseling can better express their posture and emotional states. In summary, ICAS streamlines operational procedures, reduces counselors' workload, and enhances remote counseling efficiency. It offers a promising solution for advancing remote psychological counseling, with significant potential for future applications.

Acknowledgments

This work was supported by Beijing Outstanding Young Scientist Program (BJJWZYJH01201910048035).

References

[1] McKay D, Asmundson G J G. COVID-19 stress and substance use: Current issues and future preparations. Journal of Anxiety Disorders.2020; 74: 102274.

[2] Fitzpatrick K K, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR mental health. 2017;4(2): e7785.

[3] Jaiswal S, Valstar M, Kusumam K, et al. Virtual human questionnaire for analysis of depression, anxiety and personality. Proceedings of the 19th ACM international conference on intelligent virtual agents. 2019; 81-87.

[4] Lucas G M, Rizzo A, Gratch J, et al. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Frontiers in Robotics and AI. 2017;4: 51.

[5] Luerssen M H, Hawke T. Virtual agents as a service: Applications in healthcare. Proceedings of the 18th International Conference on Intelligent Virtual Agents. 2018; 107-112.

[6] Békés V, Aafjes-van Doorn K. Psychotherapists' attitudes toward online therapy during the COVID-19 pandemic. Journal of Psychotherapy Integration.2020; 30(2): 238.

[7] Yellowlees P, Shore J, Roberts L. Practice guidelines for videoconferencing-based telemental health–October 2009. Telemedicine and e-Health. 2010; 16(10): 1074-1089.

[8] Zeren Ş G. Face-to-face and online counseling: Client problems and satisfaction. Education & Science/Egitim ve Bilim. 2015; 40(182).

[9] Haberstroh S, Parr G, Bradley L, et al. Facilitating online counseling: Perspectives from counselors in training. Journal of Counseling & Development. 2008; 86(4): 460-470.

[10] Maulik P K, Kallakuri S, Devarapalli S, et al. Increasing use of mental health services in remote areas using mobile technology: a pre–post evaluation of the SMART Mental Health project in rural India. Journal of global health. 2017; 7(1).

[11] Riemer-Reiss M L. Utilizing distance technology for mental health counseling. Journal of Mental Health Counseling. 2000; 22(3).

[12] Persky S. Employing immersive virtual environments for innovative experiments in health care communication. Patient Education and Counseling. 2011; 82(3): 313-317.

[13] Kang S H, Gratch J. Socially anxious people reveal more personal information with virtual counselors that talk about themselves using intimate human back stories. Annual Review of Cybertherapy and Telemedicine. 2012;181: 202-207.

[14] Lucas G M, Gratch J, King A, et al. It's only a computer: Virtual humans increase willingness to disclose. Computers in Human Behavior. 2014;37: 94-100.

[15] Falconer C J, Rovira A, King J A, et al. Embodying self-compassion within virtual reality and its effects on patients with depression. BJPsych Open. 2016; 2(1): 74-80.

[16] A. Rizzo, B. Lange, J. G. Buckwalter, et al. Simcoach: an intelligent virtual human system for providing healthcare information and support. 2011;277-281.

[17] Halim I, Baghaei N, Stemmet L, et al. Designing and implementing individualized vr for supporting depression. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW): IEEE; 2022. p. 217-221.

[18] Lin APC, Trappey CV, Luan C-C, Trappey AJC, Tu KLK. A test platform for managing school stress using a virtual reality group chatbot counseling system. Applied Sciences. 2021; 11(19):9071.

[19] DeVault D, Artstein R, Benn G, et al. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems. 2014;1061-1068.

[20] Egede J O, Price D, Krishnan D B, et al. Design and evaluation of virtual human mediated tasks for assessment of depression and anxiety. Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents.2021; 52-59.

[21] Kang S H, Gratch J. Virtual humans elicit socially anxious interactants' verbal self-disclosure. Computer Animation and Virtual Worlds. 2010; 21(3-4): 473-482.

[22] O'Brien J F, Hodgins J K. Animating fracture. Communications of the ACM. 2000; 43(7): 68-75.

[23] Brooke J. Sus: A "quick and dirty" usability. Usability Evaluation in Industry. 1996; 189(3): 189-194.

[24] Hart S G, Staveland L E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology. North-Holland; 1988; 52: 139-183.