

vrCAPTCHA: Exploring CAPTCHA Designs in Virtual Reality

Xiang Li

Exertion Games Lab, Department of Human-Centred
Computing, Monash University
Melbourne, Australia

Xi'an Jiaotong-Liverpool University
Suzhou, China

xiang.li18@student.xjtlu.edu.cn

Rakesh Patibanda

Exertion Games Lab, Department of Human-Centred
Computing, Monash University
Melbourne, Australia

rakesh@exertiongameslab.org

Yuzheng Chen

Exertion Games Lab, Department of Human-Centred
Computing, Monash University
Melbourne, Australia

Xi'an Jiaotong-Liverpool University
Suzhou, China

yuzheng.chen18@student.xjtlu.edu.cn

Florian 'Floyd' Mueller*

Exertion Games Lab, Department of Human-Centred
Computing, Monash University
Melbourne, Australia

floyd@exertiongameslab.org

ABSTRACT

With the popularity of online access in virtual reality (VR) devices, it will become important to investigate exclusive and interactive CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) designs for VR devices. In this paper, we first present four traditional two-dimensional (2D) CAPTCHAs (i.e., text-based, image-rotated, image-puzzled, and image-selected CAPTCHAs) in VR. Then, based on the three-dimensional (3D) interaction characteristics of VR devices, we propose two vrCAPTCHA design prototypes (i.e., task-driven and bodily motion-based CAPTCHAs). We conducted a user study with six participants for exploring the feasibility of our two vrCAPTCHAs and traditional CAPTCHAs in VR. We believe that our two vrCAPTCHAs can be an inspiration for the further design of CAPTCHAs in VR.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**; *Visualization techniques*.

KEYWORDS

vrCAPTCHA, virtual reality, CAPTCHA

ACM Reference Format:

Xiang Li, Yuzheng Chen, Rakesh Patibanda, and Florian 'Floyd' Mueller. 2021. vrCAPTCHA: Exploring CAPTCHA Designs in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '21 Extended Abstracts)*, May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3411763.3451985>

*Corresponding author.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI '21 Extended Abstracts, May 8–13, 2021, Yokohama, Japan

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8095-9/21/05.

<https://doi.org/10.1145/3411763.3451985>

1 INTRODUCTION

CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) are used to protect the security of data [15]. These authentication protocols can allow users to access websites while preventing automatic bots. The most common type of CAPTCHA is a series of graphical transformations on a two-dimensional (2D) screen, which makes it is nearly unsolvable by to an automatic bot. Recently, most mobile phones and computers support various CAPTCHAs to ensure user's access to websites. However, with the development of three-dimensional (3D) computing technology such as Virtual Reality (VR) headsets, user's pursuit of 2D space has moved to 3D environments, where we use mid-air [8] gestures [7] to interact in 3D space and explore more immersive experiences.

In this paper, we present six CAPTCHA designs in VR, including four traditional CAPTCHAs (i.e., text-based, image-rotated, image-puzzled, and image-selected CAPTCHAs) and two interactive vrCAPTCHAs (i.e., task-driven and bodily motion-based CAPTCHAs). The aim is to explore the future CAPTCHA designs in a VR environment through leveraging the user's full physical interactions. We argue that the use of CAPTCHAs with entertaining interactivity features will allow users to immerse themselves in 3D. Besides, we believe that research on new virtual reality CAPTCHAs will become more important due to the proliferation of virtual reality devices, and we propose that our work can serve as a useful springboard for such future investigations.

2 RELATED WORK

2.1 Traditional CAPTCHAs

Constrained by the structure of 2D displays, traditional CAPTCHAs are almost always based on a 2D planar design. According to Feng et al. [5], traditional 2D CAPTCHAs can be divided into two types: (1) text-based and (2) image-based CAPTCHAs.

2.1.1 Text-based CAPTCHAs. It is generally believed that the first practical text-based CAPTCHA was invented by Lillibridge [10] in 1997. Then, von Ahn et al. [14] proposed reCAPTCHA, an authentication protocol that requires users to transcribe scanned text. As this is increasingly easier for bots, more recent research has tried



Figure 1: Four traditional CAPTCHAs: (1) Text-based CAPTCHA; (2) Image-rotated CAPTCHA; (3) Image-puzzled CAPTCHA; and (4) Image-selected CAPTCHA.

to make OCR more difficult so that the bots cannot recognize the texts. For example, Baird et al. [2] proposed ScatterType CAPTCHA, which prevents the computer from splitting characters from a word. However, the feasibility of text-input CAPTCHA in VR environments has not been fully discussed. Based on previous research on inputting text in VR [19], it may be a misconception to still opt for text-based CAPTCHAs in VR.

2.1.2 Image-based CAPTCHAs. Bongo [6] was one of the first visual pattern recognition problems for CAPTCHAs. Chew et al. [3] were among the first to work on tagging image-based CAPTCHAs. After that, many algorithms based on image recognition were proposed and developed. Asirra [4] challenged the user in selecting images with cats among twelve images of cats and dogs. Matthews et al. [11] presented the scene marker test, where users must recognize relationships between irregularly shaped objects embedded in a background image. However, the flat image-based CAPTCHA mainly considers devices with a 2D display plane as the carrier, and the interactive nature of 3D virtual objects unique to VR has not been fully investigated.

2.2 Mid-air Interaction

Koutsabasis and Vogiatzidakis [8] pointed out that mid-air interactions should have the following three characteristics: (1) touchless interaction, (2) real-time sensors that can track some of the user's physical activity and movement, and (3) body movements, postures, and gestures need to be recognized and matched to specific user intents, goals and commands [16]. However, traditional CAPTCHAs are difficult to create such mid-air interactions environment in VR. Since most traditional CAPTCHAs rely on horizontal and vertical planes of interaction using mouse clicks and touch screen taps. We conjecture that these types of interaction are not applicable to the user experience of interacting with CAPTCHAs in VR. Therefore, we selected four traditional text- and image-based CAPTCHAs and implemented them in a virtual reality environment as prototypes for a preliminary study to examine the associated user experience.

3 DESIGN AND DEMONSTRATION

3.1 Four Traditional CAPTCHAs in VR

To guide our design, we firstly considered four traditional CAPTCHAs (i.e., text-based, image-rotated, image-puzzled, and image-selected CAPTCHAs).

3.1.1 Text-based CAPTCHA. In Figure 1.1, the participants were asked to use the trigger button of the right controller to click on the blank input field for text input to bring up the virtual keyboard. Participants then used the VR controller to click the corresponding letters or numbers and finally clicked the verify button.

3.1.2 Image-rotated CAPTCHA. In Figure 1.2, the participants were provided with a tilted image. Then, the participant was asked to drag the slider on the progress bar by holding the trigger button on the controller. The movement of this slider would change the rotation of the image. Once the participant thought that the rotation angle of the image had reached the correct position, they should release the trigger button to verify it.

3.1.3 Image-puzzled CAPTCHA. In Figure 1.3, the participants were asked to drag the slider by holding the trigger button on the right controller. Unlike the previous technique, the movement of the handle changed the horizontal position of the puzzle piece. Once the participant believed that the puzzle piece had been stitched to the original image, they should release the trigger button.

3.1.4 Image-selected CAPTCHA. In Figure 1.4, the participants were asked to follow the text to select all of the nine images that meet the requirements. Participants selected images by moving the slider along the controllers and pressing the trigger button when they finished their selections. When all images were selected, they were asked to press the verify button.



Figure 2: Take-driven CAPTCHA.

3.2 Early Exploration of vrCAPTCHAs

3.2.1 Task-driven CAPTCHA. We first designed a CAPTCHA that requires the user to perform a simple action to authenticate. This CAPTCHA requires the user to lift a virtual object (e.g., apple, wine bottle or silver knife and fork) by holding the trigger button and moving it to a specified location (e.g., purple bowl (Fig. 2.1), trash

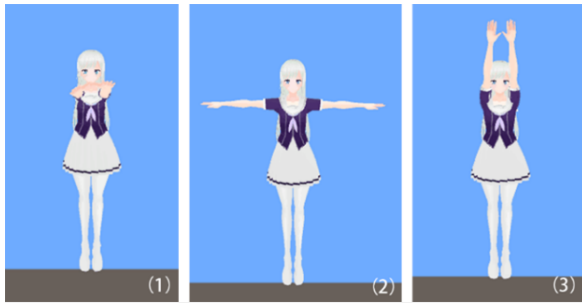


Figure 3: Motion-based CAPTCHA.

can (Fig. 2.2), or dinner plate (Fig. 2.3)). This task-driven CAPTCHA is based on the participant's feedback. "I think you can design a real-life scenario to enhance the interaction of CAPTCHA." (P5).

3.2.2 Motion-based CAPTCHA. We also considered using bodily motions to implement authentication interaction, and we named it motion-based CAPTCHA. With this CAPTCHA, the users are asked to follow the movements of a virtual character that appears in front of them (e.g., body movement like front flat raise (Fig. 3.1), side flat raise (Fig. 3.2) and up raise (Fig. 3.3)). This motion-based CAPTCHA is inspired by previous exergame research studies [1, 9, 13, 17, 18] which demonstrate the immersion that upper body limb movements can bring in VR. It is engaging for users to follow the virtual character and move their bodies.

4 PRE-STUDY AND FEEDBACK

We recruited six participants to finish the pre-user study for these four traditional CAPTCHAs and our two new vrCAPTCHAs. According to our subjective ranking feedback from participants, all participants ranked the task-driven CAPTCHA as their favourite. "I think it is the most interesting CAPTCHA!" (P1). "It is very interesting; I enjoy using this CAPTCHA in VR." (P6). Four participants thought that the motion-based CAPTCHA is the second best. "I like this (motion-based) CAPTCHA, it makes me feel that I am playing a VR game." (P2, P4). All participants thought that the two vrCAPTCHAs could greatly improve their motivation to experience CAPTCHA in VR. Four participants ranked the text-based CAPTCHA last in their preferences. For example, P1 complained that the "text-based CAPTCHA was too time consuming and cumbersome in VR."

5 LIMITATIONS AND DISCUSSION

We acknowledge the limitations of our two vrCAPTCHAs. For task-driven CAPTCHAs, we only considered simple movements of objects. In fact, for virtual 3D spaces, we can design richer interactive environments, such as simulating daily activities like playing basketball or pouring tea. For motion-based CAPTCHAs, we considered only the top half of the user's body. This can be done by introducing other sensors to enable full body interaction and thus provide fuller immersion, for example, using a Kinect camera.

We also acknowledge that the security of our vrCAPTCHAs is insufficiently validated. Given that, we here provide a discussion about interactive vrCAPTCHAs. Unlike the traditional CAPTCHAs,

previous CAPTCHAs are verified by computational processing of images or text that cannot be recognized by computers. Our CAPTCHA requires the user to follow the CAPTCHA to perform the relevant physical actions to complete the submission of the CAPTCHA. Besides, according to Mueller et al. [12], our subsequent work in the future will focus on proposing fuzziness validation tests based on the characteristics exhibited by the user as a human interaction, thus distinguishing the concept of absolute accuracy achieved by the computer.

6 CONCLUSION

In this paper, we propose two interactive vrCAPTCHAs (i.e., task-driven and bodily motion-driven CAPTCHAs), inspired by the previous 2D traditional CAPTCHA designs (i.e., text-based, image-rotated, image-puzzled, and image-selected CAPTCHAs) and combined with the 3D interactivity of VR. We conducted a user study with six participants to evaluate the feasibility of traditional CAPTCHAs and vrCAPTCHAs in VR. We believe that our two vrCAPTCHAs can be an inspiration for the future design of CAPTCHAs in VR.

ACKNOWLEDGMENTS

We are grateful to the six volunteers who were willing to come and participate in our pre-user study. Xiang Li would like to thank his internship advisors Prof. Florian 'Floyd' Mueller and Rakesh Patibanda, and other members at the Exertion Games Lab of Monash University for their valuable feedback during the prototyping of this work.

REFERENCES

- [1] Peter Arnold, Rohit Ashok Khot, and Florian 'Floyd' Mueller. 2018. "You Better Eat to Survive". In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, New York, NY, USA, 398–408. <https://doi.org/10.1145/3173225.3173238>
- [2] Henry S. Baird, Michael A. Moll, and Sui Yu Wang. 2005. ScatterType: A legible but hard-to-segment CAPTCHA. In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, Vol. 2005. 935–939. <https://doi.org/10.1109/ICDAR.2005.205>
- [3] Monica Chew and J. D. Tygar. 2004. Image recognition CAPTCHAs. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 3225. 268–279. https://doi.org/10.1007/978-3-540-30144-8_23
- [4] Jeremy Elson, John R. Douceur, Jon Howell, and Jared Saul. 2007. Asirra: A CAPTCHA that exploits interest-aligned manual image categorization. In *Proceedings of the ACM Conference on Computer and Communications Security*. 366–374. <https://doi.org/10.1145/1315245.1315291>
- [5] Yunhe Feng, Qing Cao, Hairong Qi, and Scott Ruoti. 2020. SenCAPTCHA: A Mobile-First CAPTCHA Using Orientation Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (jun 2020), 1–26. <https://doi.org/10.1145/3397312>
- [6] Martin Golubitsky and David G. Schaeffer. 1985. *The Recognition Problem*. Technical Report. 51–116 pages. https://doi.org/10.1007/978-1-4612-5034-0_2
- [7] Karine Jospe, Agnes Flöel, and Michal Lavidor. 2017. The role of embodiment and individual empathy levels in gesture comprehension. *Experimental Psychology* 64, 1 (2017), 56–64. <https://doi.org/10.1027/1618-3169/a000351>
- [8] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human-Computer Interaction* 35, 18 (2019), 1747–1768. <https://doi.org/10.1080/10447318.2019.1572352>
- [9] Xiang Li and Yuzheng Chen. 2020. Auto-Hierarchical Data Algorithm: Focus on Increasing Users' Motivation and Duration in Virtual Reality. In *2020 5th IEEE International Conference on Big Data Analytics, ICBDA 2020*. 150–153. <https://doi.org/10.1109/ICBDA49040.2020.9101254>
- [10] M D Lillibridge, M Abadi, K Bharat, and A Z Broder. 2001. Method for selectively restricting access to computer systems. , 18 pages. <http://www.google.com/patents/US6195698>

- [11] Peter Matthews and Cliff C. Zou. 2010. Scene tagging: Image-based CAPTCHA using image composition and object relationships. In *Proceedings of the 5th International Symposium on Information, Computer and Communications Security, ASIACCS 2010*. 345–350. <https://doi.org/10.1145/1755688.1755736>
- [12] Florian Mueller and Andrea Lockerd. 2001. Cheese: Tracking mouse movement activity on websites, a tool for user modeling. In *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/634067.634233>
- [13] Rakesh Patibanda, Florian 'Floyd' Mueller, Matevz Leskovsek, and Jonathan Duckworth. 2017. Life Tree. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, New York, NY, USA, 19–31. <https://doi.org/10.1145/3116595.3116621>
- [14] Luis Von Ahn, Maunei Blum, and John Langford. 2004. Telling humans and computers apart automatically. *Commun. ACM* 47, 2 (feb 2004), 56–60. <https://doi.org/10.1145/966389.966390>
- [15] Luis Von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum. 2008. reCAPTCHA: Human-based character recognition via web security measures. *Science* 321, 5895 (sep 2008), 1465–1468. <https://doi.org/10.1126/science.1160379>
- [16] Wenge Xu, Hai Ning Liang, Yuzheng Chen, Xiang Li, and Kangyou Yu. 2020. Exploring Visual Techniques for Boundary Awareness During Interaction in Augmented Reality Head-Mounted Displays. In *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020*. 204–211. <https://doi.org/10.1109/VR46266.2020.1581268453415>
- [17] Wenge Xu, Hai-Ning Liang, Qiuyu He, Xiang Li, Kangyou Yu, and Yuzheng Chen. 2020. Results and Guidelines From a Repeated-Measures Design Experiment Comparing Standing and Seated Full-Body Gesture-Based Immersive Virtual Reality Exergames: Within-Subjects Evaluation. *JMIR Serious Games* 8, 3 (jul 2020), e17972. <https://doi.org/10.2196/17972>
- [18] Wenge Xu, Hai Ning Liang, Xiaoyue Ma, and Xiang Li. 2020. VirusBoxing: A HIIT-based VR boxing game. In *CHI PLAY 2020 - Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. 98–102. <https://doi.org/10.1145/3383668.3419958> arXiv:2010.03781
- [19] Shumin Zhai, Per Ola Kristensson, and Barton A. Smith. 2005. In search of effective text input interfaces for off the desktop computing. *Interacting with Computers* 17, 3 (2005), 229–250. <https://doi.org/10.1016/j.intcom.2003.12.007>