Title	バーチャルリアリティ作業空間の展開：設計基準の策定と低ジッタータイピングシステムによる効率性の向上
Author(s)	XU, TIANSHU
Citation	
Issue Date	2025-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/19927
Rights	
Description	supervisor: 長谷川 忍, 先端科学技術研究科, 博士

Abstract

Open-Plan Workspace (OPWS) is an office style that allows many employees to work simultaneously in a wall-less, partition-less environment. OPWS is characterized by a high sense of openness, low cost, promotion of cooperation, and enhancement of the collective wisdom of the team. Although OPWS has already proven its value, many shortcomings still exist. OPWS is not only full of auditory and visual interference, but the low level of privacy protection also causes psychological stress to employees and reduces work efficiency. Although many researchers have been working on solving these problems, they still cannot declare that the issues have been entirely resolved. One prominent limitation is that most proposals suggest creating additional workspaces, which incur extra costs. Virtual Reality Workspace (VRWS), a virtual personal space independent of OPWS, can alleviate the psychological pressure caused by the lack of privacy protection in public office environments. It can also reduce auditory and visual interference in the workspace. Thus, we hypothesize that VRWS technology has great potential to address the challenges in OPWS. However, no studies have shown how Virtual Reality (VR) environments can be designed to maintain or improve work efficiency. In traditional workspaces, some studies suggest that the office environment has a significant influence on work efficiency. Therefore, we believe that a properly designed VRWS can improve users' work efficiency.

Previous research has shown that a pleasant office environment should be a cozy space free from visual and auditory interference, with good lighting, controlled sound, and plenty of natural light. Although some studies have proposed solutions to improve the shortcomings of OPWS, it remains unclear whether these solutions can be applied to VRWS. This research aims to create a VRWS with the favorable characteristics of OPWS to maintain or improve work efficiency.

We adopted Semantic Differential (SD) analysis to compare the emotional responses of participants in both OPWS and VRWS, identifying differences between the two environments and exploring factors unique to VRWS that influence work efficiency.

Due to the difficulty of finding a typical noisy OPWS in the author’s region, we decided to use the Cave Automatic Virtual Environment (CAVE) system to simulate a typical noisy OPWS. The CAVE system is a projection-based virtual reality system, which consists of several projection screens surrounding the participants and can produce a completely immersive virtual environment. At the same time, mini speakers were arranged around the CAVE system to restore the simulated OPWS sound environment as much as possible. For the simulated OPWS content in the CAVE system, we selected

NASA's mission center, where one of the frequent activities involves information exchange among employees.

We hypothesized that a VRWS with excellent OPWS characteristics—namely, an environment free from visual and auditory interference, with good lighting, sufficient natural light, and privacy protection—would maintain or improve work efficiency. To meet these requirements, we implemented the following measures: A combination of Head Mounted Display (HMD) and noise-canceling headphones was used to eliminate visual and auditory interference. To create a pleasant lighting environment, we increased the brightness of the virtual model and used natural light sources instead of ordinary light sources, ensuring the entire virtual space was well-lit. To provide ample natural light, we designed large floor-to-ceiling windows to replace the walls on either side of the VRWS and positioned the virtual desk near the windows, allowing users to enjoy the scenic views outside. For privacy protection, we designed VRWS to be single use, ensuring users experience a personal office environment.

Despite participants' overall satisfaction with working in VR, some expressed concerns regarding efficiency. A significant factor was the perceived challenge of prolonged and consistent touch typing while wearing an HMD. Consequently, participants indicated the need to frequently remove and wear the HMD if needed to type something, which negatively impacted work efficiency and user enthusiasm.

An excellent VRWS should not only provide a well-designed virtual office environment but also feature robust typing assistance functionality. To address the challenges of typing in VR, we summarized the drawbacks of not supporting a physical keyboard and relying on additional auxiliary devices. To comprehensively solve these issues, we decided to use only the HMD-mounted camera to capture typing actions and reproduce them in real time in VR. This approach eliminates the need for auxiliary devices while allowing users to verify their typing hand positions in VR. However, typing in VR presents unique challenges. When using the HMD's camera to capture typing hand positions, the fingers are often obscured by the palm, making it difficult to obtain a complete hand contour. Therefore, a dataset of "obscured typing hands" was required to train the hand-tracking model.

We conducted a data collection experiment to create this dataset. Each participant engaged in a one-hour typing session, collecting a total of 21,900 images. Using OpenCV for data augmentation, we expanded the dataset to approximately 438,000 images, 20 times larger than the original dataset. Manual annotation was performed by human annotators with the assistance of MediaPipe to extract meaningful images of typing hands.

Additionally, Motion History Image (MHI), a computer vision technique for capturing temporal motion patterns, was applied to the data to extract motion-related features.

To construct the neural network model, we prioritized minimizing VR latency. We employed a 2-stream (2S) ResNet18 to process typing images and MHI data, followed by LSTM for further processing. The model outputs the coordinates of 42 key points, which are transmitted in real time to the VR controller, enabling accurate reconstruction of typing hand positions in VR. Kalman Filtering (KF) was applied to reduce jitter. An ablation study was conducted to evaluate the model's effectiveness.

To identify the optimal network framework for VR typing tasks, we compared several models, including some recently proposed ones, focusing on metrics such as latency, accuracy, and jitter. Our analysis confirmed that the 2S-LSTM model performed among the best. To further evaluate the typing assistance solution, we conducted comparative experiments. Participants performed typing tasks under normal typing conditions, using the 2S-LSTM model, and with two other VR typing solutions, followed by a questionnaire. Statistical analysis of the experimental and questionnaire results demonstrated that the 2S-LSTM solution effectively maintained typing efficiency.

Finally, we hypothesized that the better the typing support system, the smaller the impact on users (with normal typing having no impact). To test this hypothesis, we analyzed participants' typing behavior under different conditions. Results showed that the 2S-LSTM solution had the least impact on users.

This research addresses two major challenges in virtual office systems: designing an effective VRWS and overcoming text input difficulties in VR, laying a solid foundation for the future development of VRWS systems. By referencing the findings and conclusions of this study, future researchers and developers can build upon this work to create improved VR office assistance solutions for a broader range of users.

In terms of VRWS design, this research establishes a foundational comprehensive design standard by integrating principles from OPWS studies to create an environment with minimal visual and auditory interference, enhanced lighting, and strong privacy protection. These features contribute to a more comfortable and efficient VR office experience, supported by empirical evidence from user evaluations.

For typing in VR, a key novelty of this research lies in the integration of the Two-Stream (2S) architecture and the Kalman Filter (KF) to develop a low-jitter hand tracking system. This innovative combination enables the system to accurately track and replicate typing hand positions in real time, addressing a critical challenge in VRWS. The proposed 2S-LSTM-KF solution not only improves typing efficiency but also reduces latency and jitter, maintaining users' natural typing habits and enhancing overall productivity.

Through comprehensive evaluations comparing this model with other recent solutions, including state-of-the-art models, the 2S-LSTM-KF system demonstrated the best overall performance across latency, accuracy, jitter, and deployment feasibility.

Moreover, this study opens new possibilities for applying VR technology in fields such as remote work, education, and virtual collaboration. By establishing a design standard for VRWS and addressing usability challenges in VR text entry with a robust and scalable solution, this research paves the way for the widespread adoption of VRWS. The findings offer valuable insights for researchers and developers aiming to design more efficient and user-friendly virtual environments, contributing to the theoretical and practical advancement of VR technology in virtual office settings.

Keywords: virtual reality; workspace; work efficiency; low jitter; typing support; hand tracking.