RESEARCH ARTICLE

ADVANCED MATERIALS TECHNOLOGIES
www.advmattechnol.de

Multimodal Information Perception and Understanding: Application of Smart Glove in Virtual-Reality Fusion Chemistry Experiment Platform

Hongyue Wang, Zhiquan Feng, and Xin Meng*

Virtual reality (VR) experimental platforms are frequently utilized in the teaching of chemical experiments because of its high level of interest, life-like effect, and safe and dependable advantages. However, users are unable to operate genuine experimental equipment on the present VR experimental platform, and these systems are unable to precisely assess and manage users' operation intentions. In light of the aforementioned concerns, this paper develops a smart glove with cognitive capabilities, which is composed of a variety of simple commercial sensors and binocular cameras. The smart glove can be applied to the virtual-reality fusion chemical experiment platform established in this paper. Additionally, a navigation interaction algorithm based on multimodal intention understanding is also proposed in this study (hereinafter referred to as NIAMIU algorithm). The smart glove can accurately and efficiently obtain the user's operation intention through the algorithm. The experiment showcased that the smart glove detailed in this paper is capable of precisely determining the position of users' hands in space and guiding and rectifying the user's interactive behavior. Meanwhile, users may converse with the smart glove in an unsupervised setting and conclude the experiment by following the directions supplied by the smart glove. In comparison to a conventional data glove, the smart glove designed in this paper considerably augments the accuracy and efficiency of human-computer interaction.

1. Introduction

In the elementary and secondary school classroom setting, experimental teaching may not only develop the students' technical skills but also their understanding of experimental class concepts. The current experimental teaching, on the other hand, has a lot of weaknesses. To begin with, several experimental materials are potentially hazardous. If students operate in an unsafe manner, their safety will be jeopardized. Second, since teachers cannot monitor all of their students' operations in an experimental class, some students are more likely to engage in nonstandard experimental behavior during the experiment. As a precaution, the introduction of a virtual experiment platform reduces the risk of an experiment accident. The latest virtual experiment platform,[1–6] however, has three deficiencies. On the one hand, most virtual experimental platforms use animation or simulation software to visualize the experimental process, and users do not have access to real experimental equipment; on the other hand, most traditional virtual experimental platforms employ single-mode interaction. For example, if users only use their gestures as input information, it is difficult for the experimental platform to appropriately ascertain the user's intent. The third argument is that when utilizing a virtual experiment platform, users are required to memorize numerous sets of rules, which places them under stress and found it challenging to observe the experimental process and outcomes.

The following original contributions are produced in this paper to address the challenges mentioned above.

1) This paper designs a smart glove with multiple sensors for secondary school experimental teaching. It eliminates the difficulties of unnecessary wiring and the exorbitant prices of installing multiple sensors on chemical instruments in conventional virtual labs. Furthermore, smart glove is more cognitively capable than traditional data glove.[7–13]
2) Based on multimodal intent understanding,[14–26] this paper establishes a navigation interaction algorithm. By integrating information from the user's voice channel and behavior channel, the algorithm can ascertain the user's experimental intent and then guide or rectify them based on their intent.
3) This paper creates a virtual-reality fusion chemistry laboratory. Users wearing the smart glove can operate real experimental equipment in the laboratory, compensating for the fact that users in conventional virtual experiment education

H. Wang, Z. Feng, X. Meng
School of Information Science and Engineering
University of Jinan
Jinan, Shandong 250022, China
E-mail: ise_fengzq@ujn.edu.cn
H. Wang, Z. Feng, X. Meng
Shandong Provincial Key Laboratory of Network Based Intelligent Computing
University of Jinan
Jinan, Shandong 250022, China

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/admt.202200549.

DOI: 10.1002/admt.202200549

are unable to operate real experimental objects. Furthermore, the laboratory can track and guide the user's behavior in real-time, improving interaction efficiency and reducing stress from the process of memorization.

2. Results and Discussion

2.1. Intelligent Cognitive Module

Cognitive ability is the skill of the human brain to process, store, and extract information. The smart glove's cognitive ability incorporates not only the perception and recognition of the user's hand tactile information, gestures, experimental object information, and voice commands in the context of middle school experimental teaching, but also the ability to com prehend the user's operation intent. As a result, this paper develops a NIAMIU algorithm for smart glove based on multimodal intent understanding on the one hand and a module for multimodal information perception and recognition on the other. The overall framework of the smart glove cognitive module is depicted in **Figure 1**. The smart glove's data acquisition layer captures multimodal data, which is then processed and sent to the cognitive layer. The smart glove can perceive and recognize experimental objects and users' hand behavior and other information in the cognitive layer; on the other hand, they can collect and comprehend the user's intent under various interaction conditions to differentiate the user's operation behavior. The cognitive layer's results are finally presented in the interaction layer, where the virtual-reality fusion experiment platform reacts to the user's actions in real-time and guides the user through the experimental procedure.

This paper collects data from the user behavior and voice channels through smart glove device, and fills the information of the two channels into the task slot. For the behavior channel, smart glove uses a variety of sensor data to obtain the user's action information, and utilizes the scene perception function to determine the object that the user is operating; For the voice channel, the smart glove system has the capability of voice recognition. This paper uses Baidu speech recognition and part of speech analysis to process the voice instructions input by users. Finally, this paper uses the NINMIU algorithm to analyze the user's final experimental intention, generate corresponding experimental phenomena and guide the user. The schematic figure of smart glove is shown in **Figure 2**.

2.2. Behavior Channel

2.2.1. Smart Glove Sensor Design

An attitude sensor, flex sensors, pressure sensors, infrared rangefinder, camera, vibration module, and RFID reader are among the smart glove's hardware components. The hand's 3D data can be obtained by the attitude sensor, which is then applied to display the hand's posture. The flexure degree of fingers can be determined using flexible sensors in a smart glove. The purpose of pressure sensors is to detect the pressure exerted by the user's fingers, which is used to define whether the user has successfully grasped the object. The goal of an infrared rangefinder is to calculate the distance between the target object and the user in the real world. The vibration module's purpose is to provide vibration feedback to the user when he or she grabs a real or virtual object. The glove's dialogue with the user is accomplished by the voice input and output device, which allows the glove to interact with the user more effectively.

In comparison to a traditional virtual chemistry experiment, the advantage of the smart glove in this paper is that it adopts

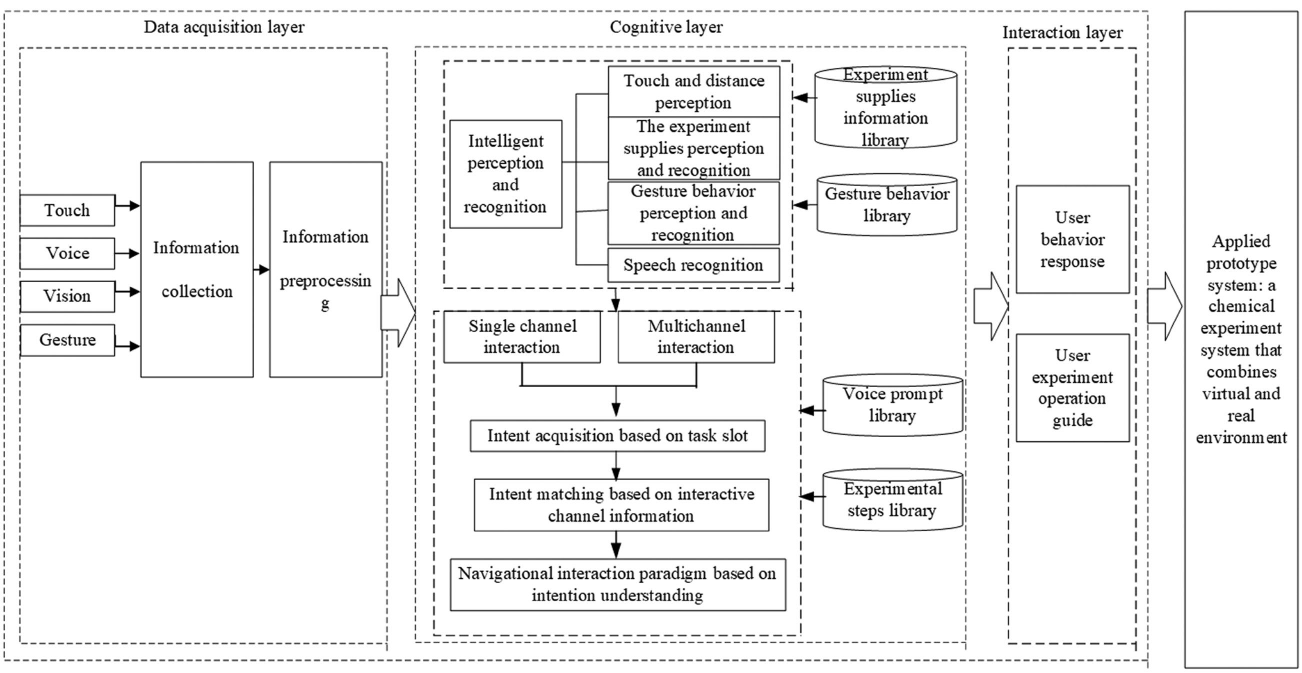

Figure 1. Overall framework diagram of cognitive module of smart glove.

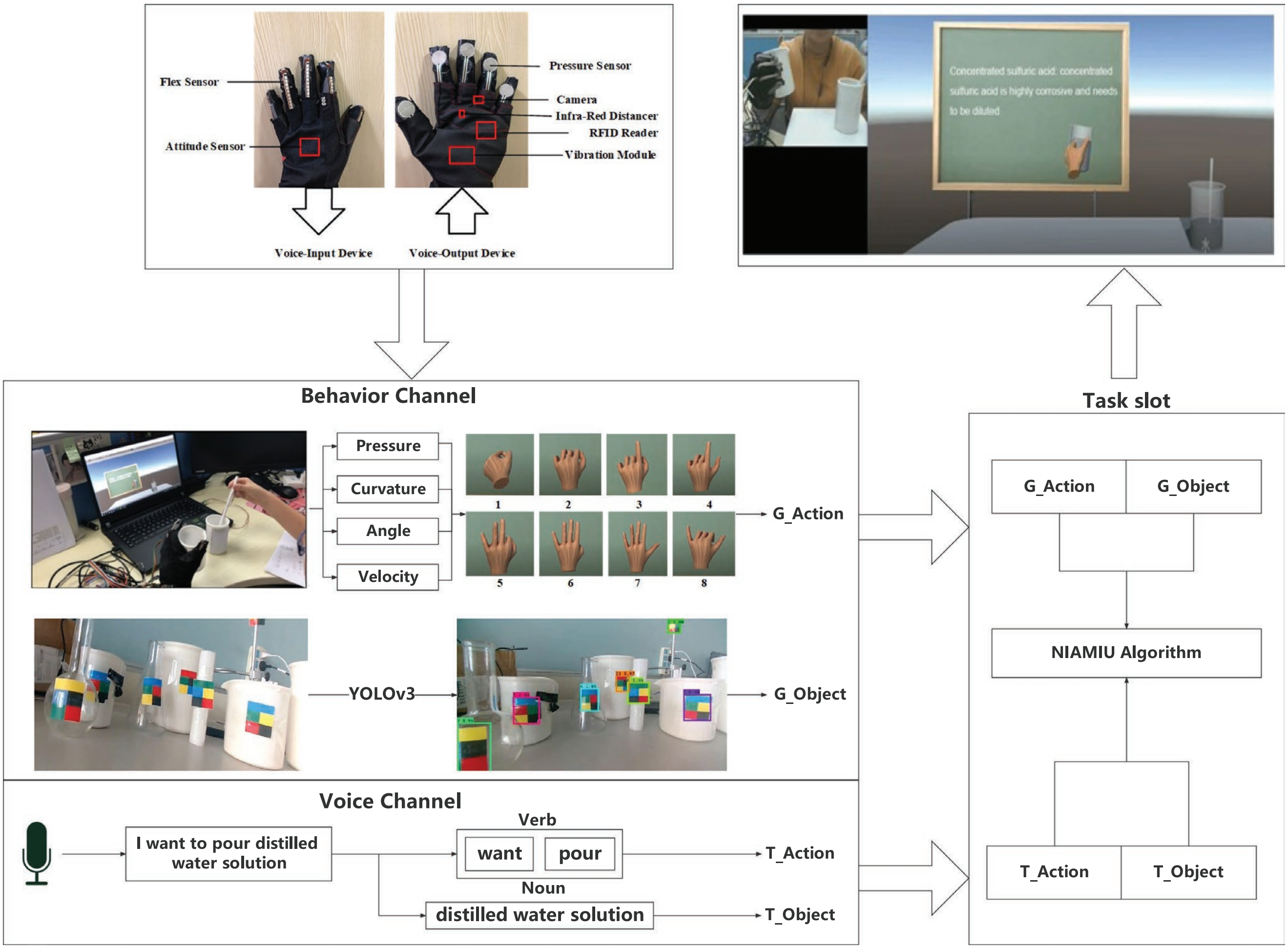

Figure 2. Smart glove schematic figure.

virtual reality technology to augment the user's sensation of activity. Unlike traditional data glove, which only possess one interactive channel, the smart glove integrates multimodal data from users' voices and behavior, enabling it to truthfully interpret and offer feedback on the users' operation intent. In addition, the traditional data glove has an excessive number of cables and is troublesome to operate. The smart glove, on the other hand, incorporates the sensor into its components, enhancing the degree of operational freedom while also ensuring the safety of the experiment.

2.2.2. Gesture Recognition Strategy Based on Sensor Data

The gesture behavior database of this paper currently consists of three types of gestures: grasping, releasing, and dumping. The smart glove judges the user's behavior based on the captured data while recognizing the user's gesture behavior. The gesture behavior library will expand in tandem with the expansion of experimental types.

The value obtained by the pressure sensor at the tip of the smart glove finger when the user operates on the experimental instruments is $\beta\{\beta_0,\beta_1,\beta_2,\beta_3,\beta_4\}$, If the value of β_i is greater than the operation threshold ω(ω is used to determine whether the user's hand interacts with the object), set the pressure mark p to true, otherwise set p to false. The smart glove then performs the following operation on the bending module data $f_t\{f_0,f_1,f_2,f_3,f_4\}$ at time $t$and data $f_{t+1}\{f_0,f_1,f_2,f_3,f_4\}$ at time $t+1$

$$\alpha_i = \sum_{i=0}^{i=4} \left(f_{t+1}\right)_i + \left(f_t\right)_i \tag{1}$$

$$\phi = \begin{cases} \alpha_i > 0 \rightarrow 1 \\ \alpha_i < 0 \rightarrow 0 \end{cases} \tag{2}$$

where α_i is the result of subtracting each value in f_t and f_{t+1}, i is the finger label of the smart glove, and $\varphi\{\varphi_0,\varphi_1,\varphi_2,\varphi_3,\varphi_4\}$ is the processing result of the bending module data at the front and back moments. If each value in φ is 1, then σ is true, and if each value in φ is 0, then σ is false. The discrimination process of the user's hand behavior

$$H\left(\sigma, p\right) \tag{3}$$

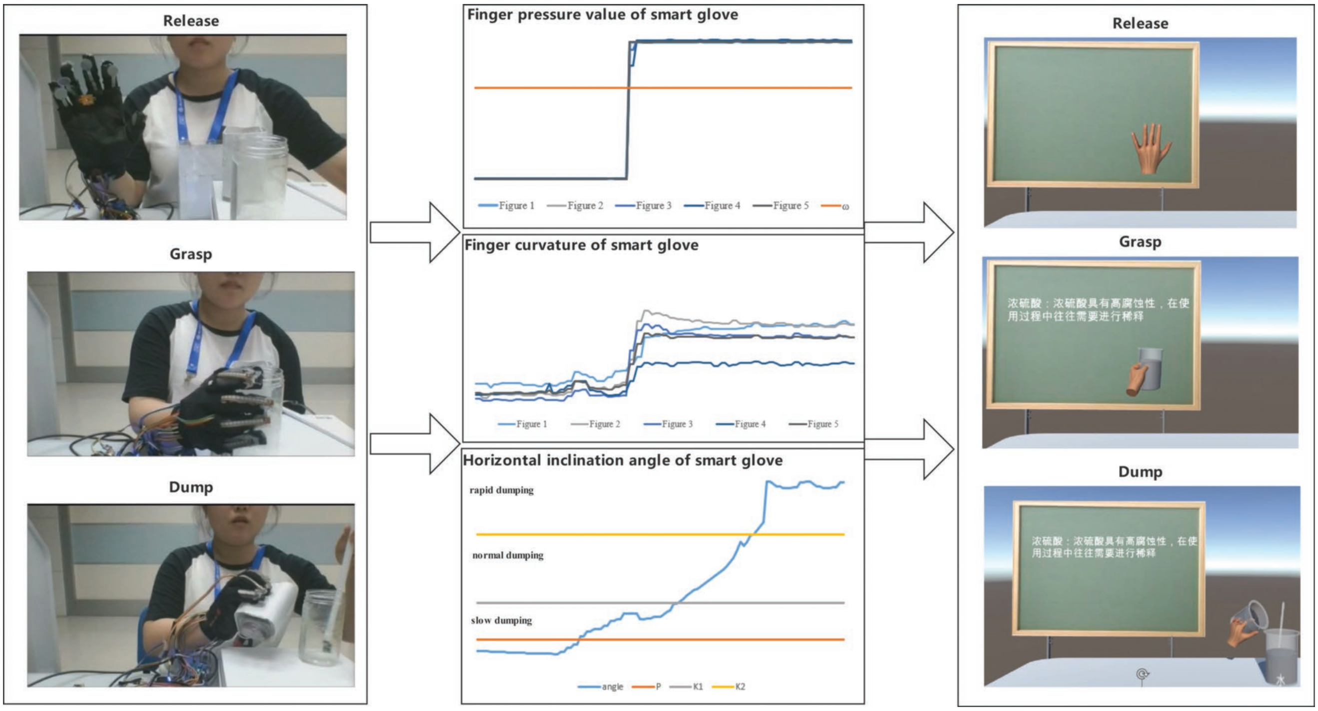

Figure 3. Data changes of grasping, releasing, and dumping of smart glove.

If both σ and p are true, the current user gesture is judged to be a grasping action; if both σ and p are false, the user gesture is judged to be a releasing action.

For the judgment of dumping action, this paper relies on the horizontal tilt angle θ of the attitude sensor to make the judgment. The judgment process of dumping action is as follow:

$$H(\sigma, p, \theta) \tag{4}$$

When the following conditions are met, user behavior can be identified as dumping.

$$\begin{cases} \sigma = \text{true} \\ p = \text{true} \\ \theta \geq P \end{cases} \tag{5}$$

where P is the threshold value of θ when determining the user's dumping behavior, k_1 and k_2 is the threshold for judging slow dumping, normal dumping and rapid dumping. If $P < \theta < k_1$, the behavior in motion is judged as slow dumping; If $k_1 \leq \theta \leq k_2$, the behavior in motion is judged to be normal speed of dumping; If $\theta > k_2$, the behavior in motion is judged to be rapid dumping. **Figure 3** shows the sensor data and virtual scene in which the user operates the grasping action, release action and dumping action respectively.

2.2.3. Scene Perception Function of Smart Glove

Two techniques for obtaining information from the experimental objects are presented in this paper. One method is to use the smart glove wrist camera to capture image label information on the instrument and then detect it using the YOLOv3[27] recognition algorithm. The other option is to use RFID reader to detect the instrument's RFID tag.

Due to space constraints, the smart glove is frequently incapable of capturing an image of the entire experimental instrument throughout the experimentation procedure. As a result, the experimental instrument in this paper is labeled with a variety of colors. The target detection algorithm must ensure a high recognition rate while maintaining real-time performance because the smart glove designed in this paper will be used in middle school experiments. Users can choose different recognition methods (YOLOv3 or RFID) under different experimental scenes or experimental steps.

At the same time, to better recognize the experimental scene and monitor the user's experimental behavior, the smart glove must obtain the user's hand movement trajectory during the user's experiment and map it in real-time to the virtual-reality fusion experiment platform built in Unity. This paper uses the open-source ORB-SLAM2 technology to facilitate the smart glove to perceive the user's movement trajectory.

This paper uses the formula (6) to process the coordinate information according to the coordinate mapping relationship between the virtual scene and the camera position after obtaining the user's hand movement trajectory

$$\begin{bmatrix} \mathrm{Pos}_x \\ \mathrm{Pos}_y \\ \mathrm{Pos}_z \end{bmatrix} = k \times \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix} \tag{6}$$

Among them, $(\mathrm{Pos}_x\ \mathrm{Pos}_y\ \mathrm{Pos}_z)$ is the position coordinate of the smart glove obtained by the ORBSLAM2 system. $(U_x\, U_y\, U_z)$

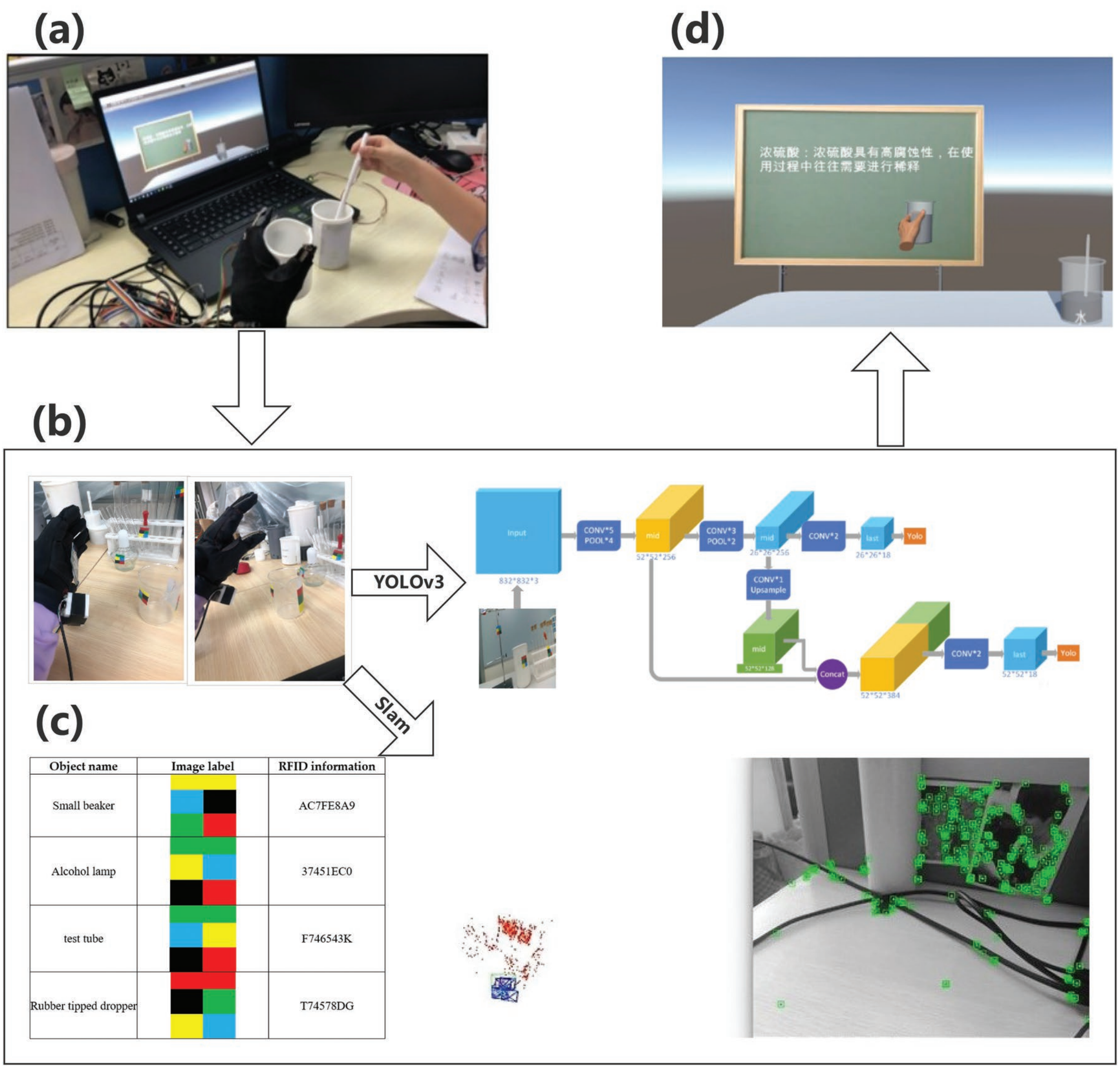

Figure 4. Schematic diagram of scene perception of smart glove. (a) User wears a smart glove to conduct the concentrated sulfuric acid dilution experiment. (b) The scene perception module of smart glove. (c) Image tags and RFID information corresponding to different experimental objects. (d) User operations are mapped to Unity virtual scene in real time.

is the 3D position of the smart glove mapped to the virtual scene. k is the scale factor for coordinate conversion. The 3D position of the smart glove is dispatched to the Unity platform via socket communication in this paper.

When the user performs chemical experiment operations on the virtual-reality fusion experimental platform, the user's operations can be mapped to the virtual scene created by Unity in real time, allowing the user to watch the experimental phenomena produced by the reaction instantaneously. **Figure 4** depicts the scene perception module of smart glove.

2.2.4. Behavior Channel Information Processing

In the process of the user completing the intelligent experiment, the smart glove obtains the user's current gesture information G_action by analyzing the sensor data, obtains the user's current experimental instrument information G_object by analyzing the RFID data or the label recognition result of the camera, and passes it into algorithm 1 (Supporting Information) for splicing operation to obtain the user's current possible behavior intent G. When G successfully matches the experimental steps in ES, it is evidenced that G denotes the user's intent in the behavior channel. The following is the specific implementation procedure.

2.3. Voice Channel

Users can enter voice information at any time during the intelligent experiment. At the same time, the smart glove system features a voice recognition function to offer a more comprehensive human interaction. The voice instructions input by users are processed using Baidu speech recognition and part of speech analysis in this paper. The smart glove system divides the user's voice instruction S into a verb set S_v and a noun set S_n via part of speech analysis. The smart glove system then calculates the Cartesian product of the elements in the two sets to obtain the intent set V that may be included in the user's voice instruction, and then matches V with the experimental step set ES corresponding to the current experiment to obtain the current voice instruction's intent T. Algorithm 2 in the Supporting Information depicts the specific implementation process.

2.4. Navigation Interaction Algorithm Based on Multimodal Intent Understanding

This paper proposes a navigation interaction algorithm based on multimodal intent understanding to facilitate more natural human interaction with users wearing gloves. To process data from both user behavior and speech channels, the algorithm (algorithm 3, Supporting Information) employs the task slot method. The data from both channels is filled into the task slots designed in this paper using the task slot method.

For the intention from the user behavior channel, this paper inserts the gesture recognition result *action* and the experimental instrument information *object* in the behavior intent G output by algorithm 1 into G_action and G_object respectively to obtain the processed behavior intent $G\{G\_action, G\_object\}$. For the intention from the speech channel, this paper inserts the verb v and noun n of the speech intent T output by algorithm 2 into the T_action and T_object respectively, and obtains the processed speech intention $T\{T\_action, T\_object\}$.

During the experiment, users have two ways to obtain the final intention. One is to experiment with only one channel, and the other is to experiment with two channels simultaneously. Given these two different situations, this paper proposes a multimodal fusion algorithm.

First, this algorithm must determine how many channels users are currently using. Since the system will continuously generate sensor data throughout the user's usage of the smart glove, the G_action information of the behavior channel will be continuously updated. When the user simply enters the voice channel's intent, the smart glove will attempt to get the behavior channel's G_object information within t seconds. If G_object information is not captured, user can continue the experiment; if G_object information is captured, the algorithm will fuse the intentions of the two channels.

When the user simply enters the behavior channel's intent, the smart glove will attempt to get the voice channel's intention information within t seconds. If the voice channel's intention information is not captured, user can continue the experiment; if the intent information is captured, the algorithm will fuse the intentions of the two channels.

Finally, the algorithm will match $G\{G\_action, G\_object\}$ and $T\{T\_action, T\_object\}$ when the intention information from two channels is input at the same time. If the matching is completed, the system will perform corresponding experimental operations; on the contrary, the algorithm will inform the user which channel information needs to be changed, so that the user can operate again. This algorithm can fuse the intentions of two channels and obtain the final intent of users, generate corresponding experimental phenomena and guide users.

3. Experimental Section

3.1. Application of Smart Glove in Virtual Experiment Platform

In this paper, a virtual-reality fusion experimental platform for concentrated sulfuric acid dilution is established. After user sends the command "grabbing the beaker containing concentrated sulfuric acid", **Figure 5**a depicts the situation of grabbing the wrong beaker. The system will ask the user for specific information about the substances in the beaker model. The operation is evidenced in Figure 5b after the user has successfully grabbed the beaker model. When the user takes up the beaker model with concentrated sulfuric acid, the system notifies the user of the description of the concentrated sulfuric acid dilution experiment.

Simultaneously, the virtual-reality fusion experiment platform will display the experimental equipment required for the experiment, as shown in Figure 5c. The dilution operation will be carried out after the user picks up the beaker model, and the virtual hand and virtual beaker in the platform will also move synchronously with the glove. If the user does not meet the experimental guidelines during the dilution process, such as diluting concentrated sulfuric acid too quickly, the glove will prompt the user, "overthrowing the angle and splashing the concentrated sulfuric acid," and displaying the corresponding experimental phenomena on the virtual-reality fusion experimental platform. Figure 5d illustrates how concentrated sulfuric acid splashes and corrodes the desktop when the user makes a mistake. After the user has completed the concentrated sulfuric acid experiment, the smart glove will issue a voice alert to inform users to complete the experiment.

3.2. Comparative Experiment

In order to verify whether the smart glove designed in this paper have acquired the cognitive ability (including sensing the user's hand posture, experimental supplies information, user voice instructions, hand behavior recognition, etc.), and can use the smart glove to discern the intentions of users, we designed a virtual-reality fusion experimental platform in unity and applied the middle school concentrated sulfuric acid dilution experiment as an example. During the experiment, we invited 10 student volunteers (five primary school students and five middle school students), of which they had not received the teaching of concentrated sulfuric acid dilution experiment. The experiments involving the use of the smart glove were approved by the IRB, and all participants in these experiments have provided written informed consent.

3.2.1. Smart Glove Sensing User Hand Posture

In the traditional AR experiment, Xiao[4] uses Kinect to acquire gestures, but it has limitations on the field of vision and misidentification. However, the smart glove can procure the user's hand posture (the rotation angle of the hand and the bending condition of the fingers) in real-time, enabling it to perceive

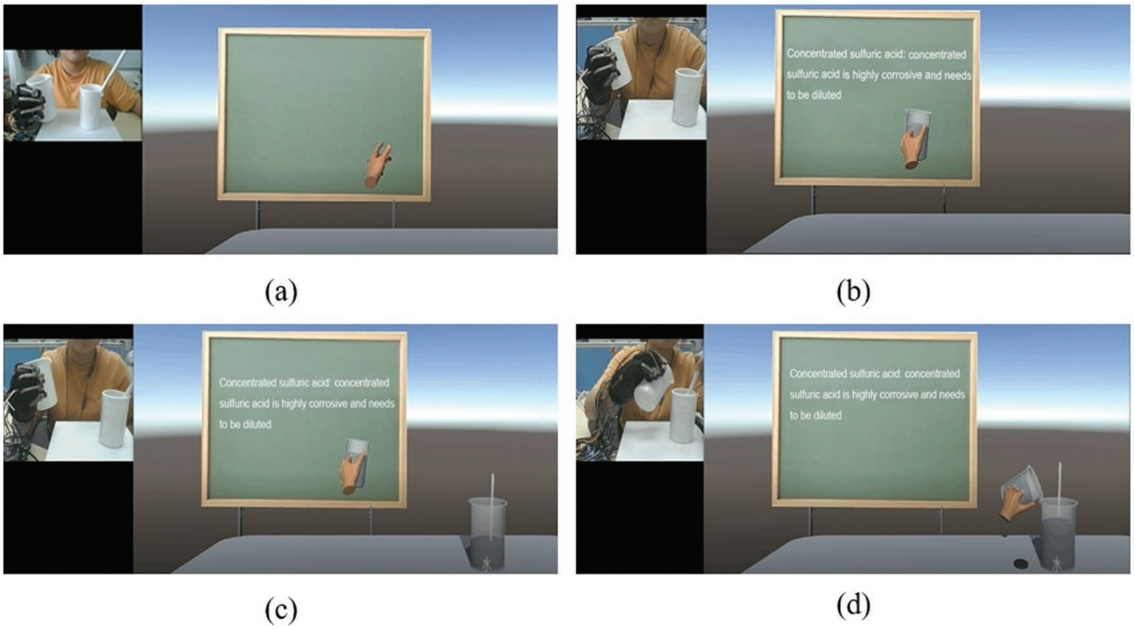

Figure 5. (a) is an example of a user wearing smart gloves to pick up an object that does not match the voice command; (b) is a schematic diagram of a model of a user successfully holding a beaker filled with concentrated sulfuric acid; (c) is a scene diagram of a concentrated sulfuric acid dilution experiment platform; (d) is corruption of desktop caused by user performing wrong operation.

the user's operation intent. To evaluate this function, this paper designed a comparative experiment, in which volunteers were invited to use Kinect and smart glove to complete the eight gestures five times in a short distance and in a long distance, and then record the user's completion to demonstrate that the smart glove can accurately determine hand posture in real-time.

Through the comparison between **Figure 6**, it can be seen that Kinect has a problem with misidentification of similar gestures (such as gesture 1 and gesture 2), and that the gesture recognition rate drops dramatically when the user is more than a specific distance from Kinect. However, the smart glove may utilize multisensor fusion technology to track the user's hand changes in real time to display the user's gesture intention, and the virtual hand in the Unity scene can also change synchronously with the user's hand, which is the basis for detecting the user's operation intention.

3.2.2. Verification of Behavior Channel Action Information

To verify the effect of smart glove on user gesture recognition, volunteers were invited to perform "grab the beaker," "release the beaker," and "pour solution" actions on the beaker provided in this paper for 20 times, and the gesture recognition strategy based on sensor data proposed in this paper was used for identification. The experimental result is shown in **Table 1**.

The number of successful recognitions divided by the total number of actions is defined as the recognition rate of a gesture recognition strategy in this paper. Table 1 illustrates that while the recognition rate for picking up the beaker and pouring the solution is relatively high, the recognition rate for releasing the beaker is relatively low, and the overall recognition outcome remains high. We observed a few undetected gestures and false recognitions during the gesture recognition process. When the problems listed above occur, users can use voice commands to prompt the system to operate again.

3.2.3. Verification of Behavior Channel Object Information

Since acquiring the user's gesture intent and the operated object information are the two most essential parts during the experiment, it's necessary to compare two methods (YOLOv3 and RFID) for acquiring the user's operated object and compare the recognition effectiveness of the two methods. At the same time, during the YOLOv3 object recognition process, there is a possibility of occlusion issues between labels on the lab bench. Therefore, volunteers were invited to test the accuracy of two methods of identifying experimental objects in simple and cluttered experimental environment. Each volunteer operates 5 times for different methods.

As it can be seen from **Figure 7**, the recognition effects of the two recognition methods for acquiring and operating objects are great, notwithstanding the chemical basis of the volunteers are different. As a result, the two object recognition systems suggested in this paper have been proven to be effective. We discovered that when users of YOLOv3 recognize the image label, they must control the distance between the glove camera and the label based on the test results. Because YOLOv3 training data is unique, its recognition effect will be influenced by the amount of light and shadow in the scene, as well as the degree of scene confusion, resulting in a small number of false positives or unidentified phenomena. Although the cluttered environment has an impact on object recognition accuracy, the smart glove designed in this paper integrates the camera at the wrist of the glove and can dynamically perceive scene

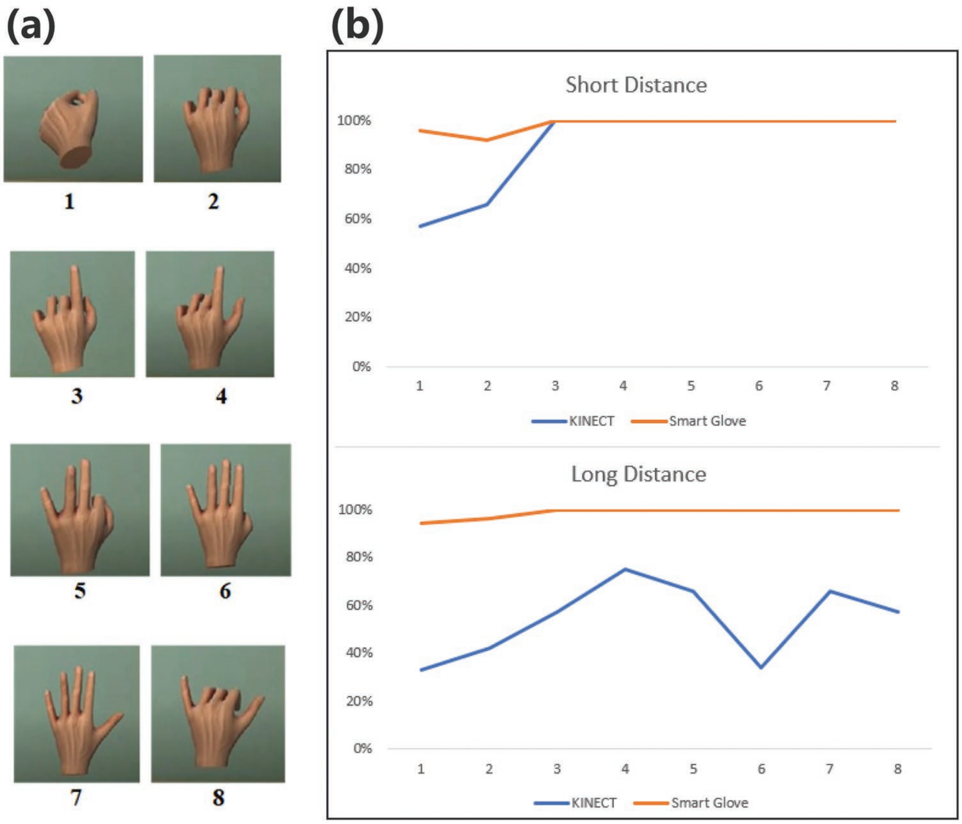

Figure 6. a) Eight gestures that users can perform; b) short and long distance gesture recognition accuracy.

information, so the occlusion problem has little impact on the recognition accuracy of this system. At the same time, users can master the experimental model using natural methods for RFID tag identification, and then can obtain the specific information of experimental supplies in real-time. The results of the experiments also show that RFID technology and YOLOv3 recognition technology improve target recognition efficiency and real-time in traditional human-computer interaction, as well as the naturalness of human-computer interaction.

3.2.4. Verification of NIAMIU algorithm

In this paper, volunteers were asked to participate in two experiments: (1) the AR experiment of intention acquisition using the MMNI algorithm[26] and (2) the smart glove chemistry experiment of intention acquisition using the NIAMIU algorithm.

Table 1. The test result of gesture recognition strategy.

Gesture number	Grab the beaker	Release the beaker	Pour solution
Recognition success	186	175	190
Recognition failure	12	17	4
Other action recognition	2	8	6
Recognition rate	93%	87.5%	95%
Other action recognition rate	1%	4%	3%

Each volunteer must complete concentrated sulfuric acid dilution experiments in the comparison experiment three times, and the intention recognition rate is calculated by recording the number of successful recognitions (successful recognition means that the system prompts the user with the current operation intention by voice, and gives the response results of the operation in real time on the virtual-reality fusion platform). **Table 2** depicts the intention recognition rate for the key steps and all steps.

According to the comparative experimental data, The NIAMIU algorithm suggested in this paper can assist users in solving the problem of incomplete expression information by merging the intent information of voice and behavior channels. As a result, it has a greater intention recognition rate than the MMNI algorithm. In summary, the NIAMIU algorithm can effectively obtain user's intention, hence improving experimental efficiency and interactivity.

And then, this paper invites volunteers to complete experiments using single voice channels, single behavioral channels, and multimodal fusion, and counts the completion rate of experiments and volunteers' satisfaction with these modalities (out of 10 scores), as shown in **Table 3**, to verify that the algorithm can better recognize user behavioral intentions than single-mode interaction.

As can be seen in Table 3, the experiment completion rate using only the voice channel is 100%, but the user rating is extremely low because there is no process of interacting with real experimental objects; one user behavior during

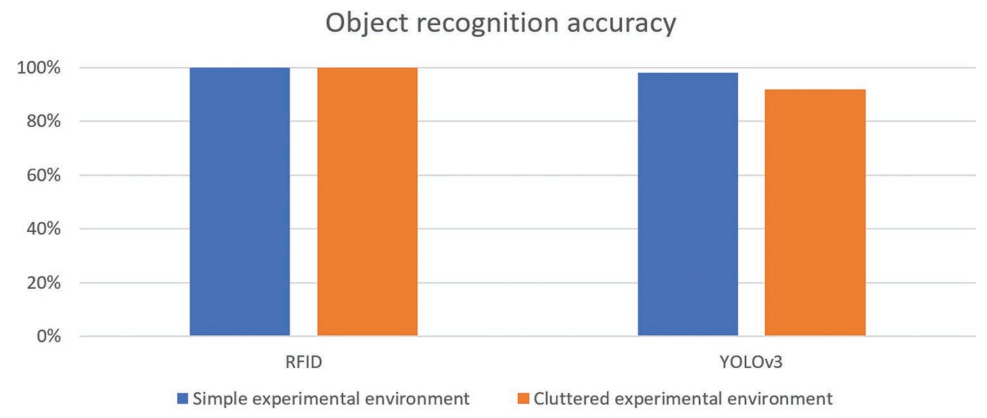

Figure 7. Comparison diagram of object recognition accuracy of behavior channel.

the experiment may correspond to several intentions, so the experiment completion rate using only the behavior channel decreases, but the user rating greatly improves because this approach allows the user to perform experiment. As a result, we can conclude that multimodal fusion of experimental interactions can better meet the needs of users and identify their behavioral intent than single-mode interactions.

In summary, the NIAMIU algorithm proposed in this paper can integrate information from various channels, determine the user's operational intent, implement navigation guidance based on user intent, and reach the objective of intelligent interaction.

3.3. Cognitive Evaluation

This paper invites 10 volunteers to perform an evaluation of a virtual-reality fusion experiment, virtual experiment,[26] NOBOOK[1] experiment, and real experiment based on smart glove using NASA's TLX table.[28] We believe that the lower the final scores of mental demands (MD), physical demands (PD), time demands (TD), effort (E), and frustration (F) in the evaluation of the six indicators, the better the effect. The higher the performance (P) score, the better the effect. The numbers 0~1 indicate a low cognitive burden, 1~2 a small cognitive burden, 2~3 a moderate cognitive burden, 3~4 a large cognitive burden, and 4~5 a very large cognitive burden. The results are shown in **Figure 8**.

According to the users' scores, the users' satisfaction with the smart glove system designed in this paper is the greatest among the overall evaluation results of the six indicators. Because the smart glove system has the lowest score in indicator MD and indicator TD, users would believe that the process of virtual-reality fusion utilizing smart glove is clearer than other experimental platforms. For example, in order to utilize the NOBOOK experimental platform for experiments, users must grasp the platform's many features at first, such as the experimental platform's creation. Because the smart glove system's indicator P score is higher than that of other platforms, this paper conducted interviews with a number of experimenters. The findings demonstrate that when users utilize other experimental platforms, they spend the majority of their time figuring out how to use the platform. When they utilize smart glove for experiments, they will pay greater attention to the results, and the experimenters will be able to better comprehend the experimental phenomena by looking at the screen and listening to the system explain the experimental mechanism. Simultaneously, the intelligent experimental system will rectify nonstandard behavior in the experimental process, allowing them to better comprehend the fundamentals of experimental operation. And then, indicator E and F scored the lowest for the smart glove system, showing that users are more interested in the technology and are less likely to be frustrated. This is because the user cannot control the real experimental object in the virtual experiment, thus the sense of genuine operation is weak; in the real experiment, if there is no instruction, users would struggle to complete the experiment, making it simple to become upset. To summarize, users wearing smart glove can conduct chemical experiments more intelligently and naturally in the intelligent chemical experiment system proposed in this paper. It can boost students' experimental immersion and experimental operating abilities when compared to other experimental platforms.

4. Conclusions

The existing virtual experiment platform has a host of issues. To begin with, the user is unfamiliar with how to operate real experimental equipment. Second, when the user operates on the virtual experimental platform, the memorization burden is

Table 2. Intention statistics.

Intention	MMNI	NIAMIU
Pick up the beaker	93%	95%
Pour distilled water solution	92%	92%
Agitated solution	95%	100%
All intentions	91%	95%

Table 3. Comparison experiments between single-mode and multimodal interactions.

Interaction method	Experiment completion rate	Satisfaction level
Voice channel	100%	3.4
Behavior channel	75%	8
Multimodal fusion	98%	9.5

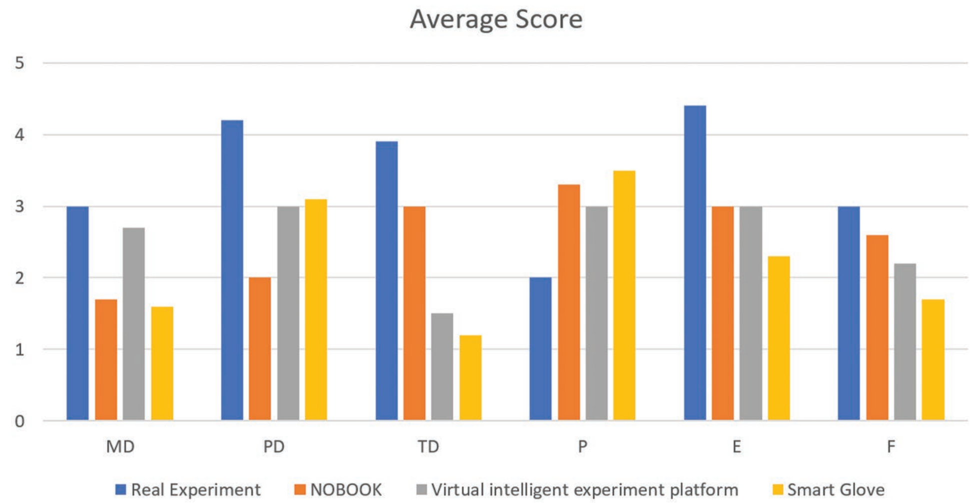

Figure 8. NASA user evaluation.

substantial, and the system is unable to accurately determine the user's operation intent and so on.

In order to solve the above problems, this paper has done the following three works:

1) This paper designs a smart glove with cognitive ability. In traditional virtual experiment, Kinect is utilized to gather user's gestures; however, it has the limitations of visual field and similar gesture recognition. Using multisensor fusion technology, the smart glove can accurately gather user gesture information. Furthermore, smart glove can identify the RFID tag and image label information to receive information about the operational object efficiently, steadily and in real time. However, in terms of gesture recognition, the smart glove can currently recognize only a few types of gestures. Meanwhile, some sensor layouts, such as the bending sensor, have flaws in terms of smart glove equipment hardware layout. When the user's hand is small, the smart glove's bending data is limited, which may affect experimental behavior discrimination. In the future, the structure of smart glove will need to be improved.
2) This paper proposes the NIAMIU algorithm, which can integrate the user's voice channel and behavior channel intention information and translates it to the user's final operation intention. Its advantage is that it can minimize user's cognitive pressure and increase their understanding of knowledge in the virtual-reality fusion experiment. However, the major actions in the algorithm's behavior channel are "grab," "release," and "dumping." As a result, the future research will combine gesture actions in more experimental settings to create a rich sensor-based gesture identification library.
3) This paper creates a virtual-reality fusion chemistry laboratory, which solves the problem that users are unable to operate real experimental objects in traditional virtual experiment education. However, in the chemical experiment system based on smart glove, the presentation of experimental phenomena is limited to animation. In the future, we hope to combine virtual animation with real experimental phenomena to improve the authenticity of experimental experience for user.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This paper is supported by the Independent Innovation Team Project of Jinan City (No. 2019GXRC013), and the Shandong Provincial Natural Science Foundation (No. ZR2020LZH004).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

data glove, human-computer interaction, intent understanding, multimodal fusion, virtual-reality fusion experiment

Received: April 5, 2022
Revised: July 1, 2022
Published online:

[1] S. Fang, L. Xue, Y. Liang, et al., in *2020 International Conference on Virtual Reality and Visualization (ICVRV)*, IEEE, **2020**, pp. 346–347, http://school.nobook.com.cn/site.
[2] B. Zeng, Z. Feng, T. Xu, M. Xiao, R. Han, *IEEE Access* **2020**, *8*, 142507.
[3] L. Jiang, *Complexity* **2021**, *2021*, 5553211.
[4] H. Wang, Z. Feng, J. Tian, X. Fan, *Comput. Intell. Neurosci.* **2022**, *2022*, 3545850.

[5] M. H. Elgewely, W. Nadim, A. ElKassed, M. Yehiah, M. A. Talaat, S. Abdennadher, *Open House International* **2021**.
[6] C. L. Huang, Y. F. Luo, S. C. Yang, C. M. Lu, A.-S. Chen, *J. Educ. Comput. Res.* **2020**, *58*, 596.
[7] W. Gao, G. Fang, D. Zhao, Y. Chen, *Pattern Recognit.* **2004**, *37*, 2389.
[8] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, J. Yang, *IEEE Trans. Syst., Man, Cybern., Part A: Syst. Hum.* **2011**, *41*, 1064.
[9] A. B. K, A. G Nair, D. R. K, K. Ananthanarayanan, H. R. Nandi Vardhan, in *2016 Int. Conf.,* Amrita Vishwa Vidyapeetham University, IEEE, **2016**, pp. 1–6.
[10] K. Roy, D. P. Idiwal, A. Agrawal, B. Hazra, in *Proc. 2015 Conf. Advances in Robotics (AIR '15)*, ACM, New York, USA **2015**.
[11] Z. Ma, P. Ben-Tzvi, J. Danoff, *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *24*, 1323.
[12] Noitom, https://www.noitom.com.cn.
[13] Y. Ming, Y. Yang, R. P. Fu, C. Lu, L. Zhao, Y. M. Hu, C. Li, Y. X. Wu, H. Liu, W. Chen, *Adv. Mater. Technol.* **2018**, *3*, 1800257.
[14] M. C. Caschera, F. Ferri, P. Grifoni, *Int. J. Web Grid Serv.* **2007**, *3.1*, 82.
[15] A. Shahroudy, G. Wang, T. T. Ng, in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, IEEE, Athens, Greece **2014**.
[16] K-S. Song, Y.-H. Nho, J.-H. Seo, D.-s. Kwon, in *15th International Conference on Ubiquitous Robots (UR)*, IEEE, Honolulu, Hawaii **2018**.
[17] P. Y. Hui, H. Meng, *IEEE/ACM Trans. Audio, Speech, Lang. Process.* **2014**, *22*, 417.
[18] R. M. Jiang, A. H. Sadka, D. Crookes, *IEEE Trans. Syst., Man, Cybern., Part C: Appl. Rev.* **2010**, *40*, 676.
[19] V. P. Minotto, C. R. Jung, B. Lee, *IEEE Trans. Multimedia* **2015**, *17*, 1694.
[20] M. Ehatisham-Ul-Haq, A. Javed, M. A. Azam, H. M. A. Malik, *IEEE Access* **2019**, *7*, 60736.
[21] M. Ghayoumi, M. Thafar, A. K. Bansal, *Towards Formal Multimodal Analysis of Emotions for Affective Computing[C]//DMS*, **2016**, pp. 48–54.
[22] J. Cacace, A. Finzi, V. Lippiello, in *26th IEEE International Symposium on Robot and Human "Interactive Communication" (RO-MAN)*, IEEE, Lisbon, Portugal **2017**.
[23] D. Yongda, L. Fang, X. Huang, *Comput. Electr. Eng.* **2018**, *72*, 443.
[24] A. W. Ismail, M. Billinghurst, M. S. Sunar, C. S. Yusof, Designing an Augmented Reality Multimodal Interface for 6DOF Manipulation Techniques, in *Intelligent Systems and Applications. IntelliSys 2018. Advances in Intelligent Systems and Computing* (Eds: K. Arai, S. Kapoor, R. Bhatia), Springer, Cham **2018**.
[25] H. Liu, T. Fang, T. Zhou, L. Wang, *IEEE Access* **2018**, *6*, 74762.
[26] M. Xiao, Z. Feng, X. Fan, B. Zeng, J. Li, *IEEE Access* **2020**, *8*, 125982.
[27] L. Zhao, S. Li, *Electronics* **2020**, *9*, 537.
[28] S. G. Hart, L. E. Staveland, *Adv. Psychol.* **1988**, *52*, 139.