# Research on the Structure and Key Algorithms of Smart Gloves Oriented to Middle School Experimental Scene Perception

Hongyue Wang[1,2], Xin Meng[1,2], and Zhiquan Feng[1,2](✉)

[1] College of Information Science and Engineering, University of Jinan, Jinan 250022, China
1536748981@qq.com

[2] Shandong Provincial Key Laboratory of Network Environment Intelligent Computing Technology, University of Jinan, Jinan 250022, China

**Abstract.** The existing virtual experiment platform mainly uses virtual reality technology or animation technology to assist students in experimental teaching, but it lacks the standardized supervision of users' experimental behaviors. To address the above problems, this paper designs a prototype smart glove application for middle school experimental scenarios and proposes a scene perception algorithm based on the smart glove, so as to obtain the user's experimental behavior more accurately. Based on the perception of the experimental scene, this paper also proposes a multimodal fusion of intelligent navigation interaction paradigm to obtain the user's experimental intention, thus allowing students to conduct exploratory experiments on a virtual experimental platform with targeted guidance and monitoring of user behavior. Experiments show that the smart glove designed in this paper can sense the relative relationship between experimental equipment and objects in the scene in real time. Based on the user's experimental behavior, the smart glove can also infer the operator's experimental intent and provide timely feedback and guidance on the user's experimental behavior.

**Keywords:** Virtual experiment platform · Smart gloves · Scene perception · Intention inference

## 1 Introduction

Experimental teaching is an important part of secondary school teaching, but in practice there are problems such as the existence of dangerous experimental supplies for operation and irregularities in experimental operations. With the development of Internet technology, the virtual experiment platform has solved these problems. However, there are still some defects in the existing virtual experiment platform. On the one hand, the virtual experiment platform tends to use animation to present the experimental process. On the other hand, most virtual experiment platforms lack the monitoring of users' experimental behaviors.

In order to solve the above problems, this paper designs a smart glove that can be used for experimental teaching in secondary schools and also has the ability of field


Y. Sun et al. (Eds.): ChineseCSCW 2021, CCIS 1491, pp. 409–423, 2022.
https://doi.org/10.1007/978-981-19-4546-5_32

perception. The smart glove can sense the information of objects in the experimental scene and their corresponding position relationship in real time. Students can operate the real experimental apparatus on the virtual experimental platform to conduct experiments. Based on the field perception, this paper also presents an interactive example of intelligent navigation, which can synthesize the information of speech, vision and sensors used to infer the user's experimental intention and provide feedback and guidance to the user.

## 2 Related Work

The development of information technology has provided a broad platform for virtual labs. Experimental teaching through the use of virtual laboratories has the following advantages: 1. Virtual laboratories can save the cost of experiments and allow students to conduct more types of experiments for schools with insufficient experimental funds; 2. The use of virtual laboratories can better ensure the safety of students; 3. Traditional laboratories require teachers to spend time on equipment setup and maintenance.

In recent years, with the continuous development of Internet technology, virtual laboratories have become a research hotspot for scholars at home and abroad. Sotomayor-Moriano J [1] enable students to conduct experiments in a real environment through a Web browser, and can design and practice controls in an interactive virtual laboratory. Khulood Aljuhani [2] have developed a virtual laboratory platform based on the Web. Users can use the mouse to perform experiments in the virtual laboratory individually and perform multiple experiments as needed. It can not only enhances the flexibility of teaching technology, but also deepens students' understanding of experimental phenomena. Web-based platforms can only immerse students in the learning experience to a limited extent. In order to improve immersion, Ioannis Doumanis [3] uses game mechanics and game design thinking to design a multi-mode immersive teaching environment. Experiments show that the platform can better improve students' academic performance and subjective experience. Dongfeng Liu [4] also used computer simulation technology to design a low-cost virtual laboratory. Compared with the above methods, this virtual laboratory allows students to assemble instruments and use their own assembled instruments to perform physical experiments, so it can help students understand the scientific process better. Although the above virtual laboratory design methods have achieved good results, the attractiveness of the virtual learning environment is relatively weak due to the operation of the mouse and keyboard.

In response to the problems of insufficient experimental funds and lack of experimental equipment in some middle schools, Francisco Torres [5] built a virtual laboratory on the Unity platform, allowing students to interact with experimental equipment on the virtual platform and design experiments based on learning topics. Diana Bogusevschi [6] uses virtual reality technology to teach water cycle experiments in order to save experimental costs. Students can conduct experiments in accordance with the guidelines of the program, which greatly enhances students' interest in learning and deepens students' understanding of water cycle experiments. Augmented reality technology (AR) can increase the user's immersion, so many researches also focus on augmented reality technology. For example, Joanne Yip [7] applied AR video to students' learning process to help students understand space-related knowledge. Mustafa Fidan [8] integrated

AR technology into the teaching process of physics and invited 91 students to conduct experiments. The experimental results show that with the assistance of AR technology, students' learning attitudes are more active and their academic performance is improved.

The above research schemes have proved that virtual experimental teaching is feasible in teaching work. It can not only enhance students' interest in learning, help students understand knowledge, but also make up for the shortcomings in traditional experimental teaching. Lowell M [9] has further confirmed this point by comparing traditional teaching methods and virtual experimental teaching methods. They selected 50 students to use traditional multimedia presentation methods and virtual assisted teaching methods using VR equipment for chemistry teaching, and tested two groups of users. The final score shows that the performance of students who use virtual assisted teaching is significantly higher than that of students who use traditional methods. Also using two different teaching methods, Sarah Sullivan [10] asked students to conduct pulley experiments. The results proved that using virtual experiments can be more beneficial to deepen students' understanding of scientific concepts.

Whether using a web browser or using virtual reality technology, most studies tend to use a single channel to interact with users, and the process of human-computer interaction is often accompanied by the collaborative work of multiple channels. For example, Sidenmark [11] proposed in the study of the user's gaze that the user's gaze is often accompanied by the coordinated operation of the eyes, head and body. Therefore, it is very important to treat the line of sight as a coordinated operation of multiple modalities. Using multiple channels as input in the human-computer interaction process not only improves the efficiency of human-computer interaction, but also enables more flexible and free communication. In the design of the car driving user interface, Jingun Jung [12] combines touch and voice to solve the problem of low control efficiency caused by only using voice for input in some aspects, and enhance the user experience.In the AR environment, Ismail A W [13] allows users to interact with virtual objects in two ways: voice commands and gestures, making user operations more natural. The advantages of using multi-channel interaction are more prominent in robot application scenarios. Deng Yongda [14] proposed the use of gestures and voice commands to control the robot to make the interaction process more natural. It also integrates the data of gestures and voice channels. TingHan [15] proved that the result of multi-channel fusion is better than the result of using only a single channel.

In summary, although the virtual experiment platform has shown excellent results in teaching results, there are still shortcomings. In actual experimental teaching, the standardization of experimental procedures is also very important, which puts forward new requirements for the virtual experimental platform, that is, the virtual experimental platform should have the ability to perceive and understand the user's experimental scene.To this end, this paper designs a smart glove with the ability to perceive the scene, which can perceive the object information and the corresponding position relationship in the experimental scene in real time. On the basis of scene perception, this paper also proposes an intelligent navigational interaction paradigm based on multi-modal fusion to infer the user's experimental intention, so as to give the user corresponding feedback and guidance during the experiment. Finally, the user can complete the experimental operation under the guidance of smart gloves.

# 3 Scene Perception Algorithm Based on Smart Gloves

Scene perception technology allows smart gloves to understand the environment like a human. When smart gloves perform scene perception, they can use the camera to detect experimental objects and thus understand the experimental scene. In secondary school experiment scenarios, when users operate experimental equipment with smart gloves, they often cannot obtain a complete image containing the experimental object, so it is difficult to identify and locate the experimental equipment. To address the above problems, this paper proposes an algorithm for perceiving and locating experimental objects in specific scenes based on smart gloves, so as to obtain information about the experimental objects operated by users and their location relationships with other objects in real time.

## 3.1 Smart Glove Structure Design

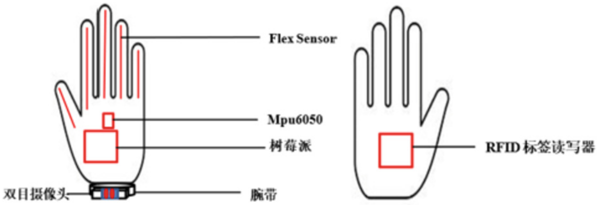


**Fig. 1.** Hardware structure design of the palm of smart glove

The hardware structure design of the smart glove is shown in Fig. 1. The structure of the smart glove is divided into several parts: 1. We use the curvature sensor to obtain the change of the user's finger curvature; 2. The Mpu6050 sensor is small in size and can directly obtain the three-axis acceleration, angular velocity, and angle of the user during movement. Therefore, it is used to restore the posture of the user's hand; 3. As the functional module of the smart glove, the Raspberry Pi is mainly responsible for processing various sensor data; 4. In the palm of the smart glove, we place a small RFID tag to read and write The smart glove can obtain real-time information about the object currently being operated by the user's hand; The data transmission mode of the smart glove designed in this paper adopts wireless transmission, so that users can conduct experiments more conveniently.

## 3.2 Target Detection System Based on YOLOv3

In the process of experimenting, due to the limitation of experimental space and other factors, smart gloves sometimes cannot capture the complete image of the experimental article. Therefore, this article uses a combination of different colors to mark the experimental article. Considering that the smart gloves designed in this article will ultimately need to be applied to middle school experimental scenes, when performing target detection, the target detection algorithm is required to ensure high recognition rate while pursuing real-time performance. In order to achieve faster and better identification of experimental objects, this article uses the YOLOv3 network [16] to detect objects in the experimental scene. First, we use binocular cameras to collect experimental scene data of experimental item tags from different perspectives and distances. In the process of data collection, in order to ensure that smart gloves can realize real-time scene perception in different experimental situations, we use LabelImg tool to label the collected pictures, and send the labeling information as training samples to the YOLOv3 network model After training in YOLOv3, the experimental item detection model FC based on YOLOv3 was finally obtained. Figure 2 shows the target detection process of the YOLOv3 network.

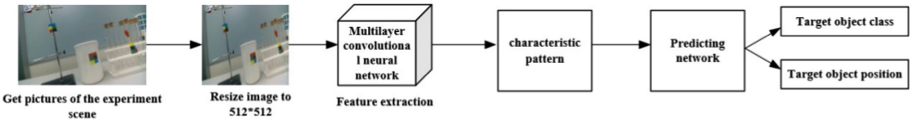


**Fig. 2.** Schematic diagram of YOLOv3 target detection process

### 3.3 Smart Glove Movement Track Acquisition

In the process of the user's experiment, the smart glove needs to obtain the user's hand movement trajectory and map it to the virtual experiment platform built in Unity in real time, so as to lay the foundation for the smart glove to better perceive the experimental scene and track the user's experimental behavior. This paper uses the open source ORB-SLAM2 system [17] to realize the perception of the user's movement trajectory by smart gloves. The schematic diagram of ORB-SLAM2 system obtaining the user's hand movement trajectory is shown in Fig. 3:

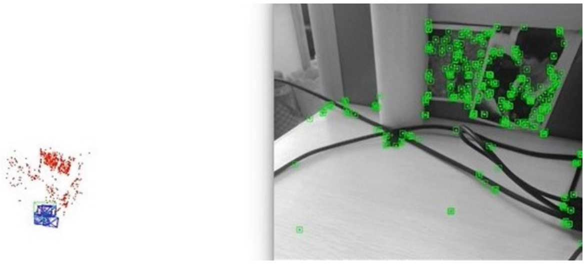

**Fig. 3.** Smart glove's perception of the user's hand trajectory

After obtaining the movement trajectory of the user's hand, we can use the formula (1) to process the coordinate information according to the coordinate mapping relationship between the virtual scene and the camera position.

$$\begin{bmatrix} Pos_x \\ Pos_y \\ Pos_z \end{bmatrix} = k \times \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix} \tag{1}$$

Among them, $(Pos_x, Pos_y, Pos_z)$ is the position coordinate of the smart glove obtained by the ORBSLAM2 system. $(U_x, U_y, U_z)$ is the three-dimensional position of the smart glove mapped to the virtual environment. $k$ is the scale factor for coordinate conversion.

### 3.4 Scene Perception Algorithm Based on Smart Gloves

The scene perception algorithm based on smart gloves considers the following two situations: 1. Target object search, that is, the user uses smart gloves to search for target experimental objects in the experimental scene; 2. Replacement of the target object, that is, the user has an experimental object in his hand and needs to perform an operation on another target object. This paper use the SGBM (Semi-global Block Matching) algorithm to obtain the depth information of the target object. Combining the position information of the target object obtained by target detection, the identification and positioning of the experimental object is realized, and the obtained information is transferred to the Unity platform. During the user's experiment, this article sets the target object set $E$. Each time an experimental operation is performed, the target object $E$ operated by the user will be dynamically updated. The specific steps of the Scene Perception Algorithm Based on Glove (Scene Perception Algorithm Based on Glove, hereinafter referred to as SPABG) are as follows:

Algorithm 1: Scene Perception Algorithm Based on Smart Gloves (SPABG Algorithm)

Input: experimental scene image $pig$ captured by binocular camera, sensor information $O^h$.

Output: target object information set $E$ voice prompt V, scene response.
1. Obtain images of experimental scenes $pig$;
2. Use ORBSLAM2 algorithm to obtain current smart glove location information $Oh(x_h, y_h, z_h)$ and output $Oh$ to Unity platform.
3. Use the target detection model FC to detect the experimental objects on $pig$, and store the name and location of the detected target object in $E_n$, $E_p$;
4. Determine whether $E_n$ is empty, if $E_n$ is empty, return to step 1; If $E_n$ is not empty, then obtain the three-dimensional coordinates of $E_p$ and store it in $E_w(x, y, z)$;
5. Pass $E_n$ and $E_w$ into the Unity platform, and use formula (2) to process $E_w$ to obtain the coordinate $E_u$ in the virtual scene;

$$\begin{bmatrix} E_{ux} \\ E_{uy} \\ E_{uz} \end{bmatrix} = r \times \begin{bmatrix} E_{wx} \\ E_{wy} \\ E_{wz} \end{bmatrix} + B \tag{2}$$

Among them, $r$ is the scale factor of coordinate conversion, and $B$ is the position correction matrix from the real environment to the virtual environment.
6. Determine whether the current user operating object set $O^h$ is empty.

(1) If $O^h$ is empty, perform the formula (3) operation on $Oh$ and $E_u$ to obtain the distance $d$ between the two in real time. If the value of $d$ keeps decreasing, output a voice prompt c $V_1$ to the user and return to step 1;

$$d = \sqrt{(x_h - E_{ux})^2 + (y_h - E_{uy})^2 + (z_h - E_{uz})^2} \tag{3}$$

(2) If $O^h$ is not empty and $E$ is empty, then $O^h$ is processed, the target object name $O^n$ corresponding to $O^h$ is queried, and $O^n$ and $Oh$ are stored in the target object set $E$, output $E$ and voice prompt $V_2$ and display the corresponding target object model in the virtual scene, return to step 1;

(3) If $O^h$ is not empty and $E$ is not empty, then:

①Execute formula (3) operation on $O^h$ and $E_u$, and obtain the distance $d$ between the two in real time;
②If $d$ satisfies $d \leq \xi$ ($\xi$ is the threshold for judging whether the target object is maneuverable), the corresponding target object model is displayed in the virtual scene, and the current target object name $E_n$ and target object position $E_u$ are stored in $E$ ,output $E$ and voice prompt $V_3$, and return to step 1.

In the SPABG algorithm, the sensor information $O^h$ is the result obtained by the RFID tag reader placed on the palm of the smart glove; $V_1$ is "currently you are approaching $E_n$"; $V_2$ is "currently you have selected $O^n$"; $V_3$ is "currently you have selected $E_n$ .

# 4 Intelligent Navigational Interaction Paradigm Based on Scene Perception

This chapter will introduce the intelligent navigational interaction paradigm based on scene perception proposed in this article. It processes information from different channels to infer the user's experimental intention and guide the user on the experimental operation. The intention understanding here refers to the fusion of visual data, sensor data and voice data from the smart glove for specific application scenarios and context-based interaction scenarios, so as to infer the user's experimental intent.

## 4.1 Overall Framework

We constructed an overall framework for the navigational interaction paradigm of smart gloves based on multi-modal fusion intention understanding (Fig. 4). The overall process can be divided into three levels: the input layer, the interaction layer, and the presentation layer. In the input layer, smart gloves obtain data from voice, sensors, and visual channels through interactive devices such as microphones and cameras. In the interaction layer, the smart glove first extracts the semantics of the data from the three channels. When processing the data of the voice channel, this article uses Baidu Voice to recognize and analyze the user's voice instructions. This paper proposes a user intention understanding strategy based on intention screening, which integrates information from various channels to obtain the user's experimental intention. According to the different experimental intentions of the user, combined with the set experimental library, the user's experimental behavior is judged, and finally the user's experimental intention, scene response, voice prompt and other information are output. The output result of the interactive layer will finally be presented in the interactive interface in the presentation layer.

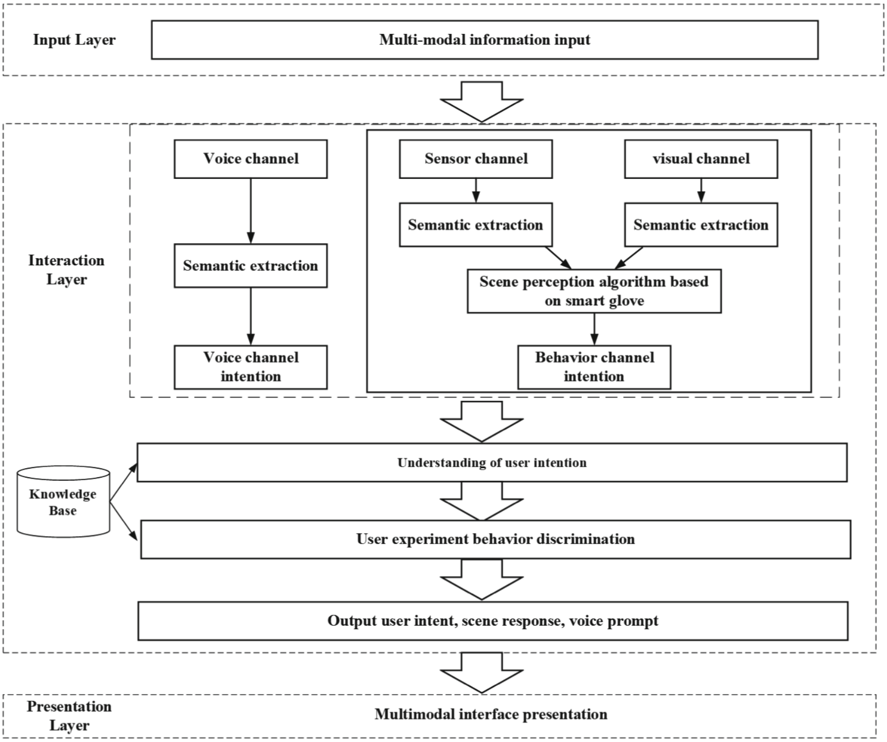


**Fig. 4.** The overall framework of multi-modal intelligent navigation interaction paradigm based on scene perception.

## 4.2 User Intent Understanding

In this paper, the process of inferring the user's most likely intention is defined as the problem of intention screening, and the uncertainty of the user's experimental intention is modeled as the possibility of different categories. We define the intent screening problem as a classification task. In the case of a given experimental operation by the user, the smart glove infers the experimental target from the set of possible experimental targets $I$.

We have pre-established an experiment library for the interaction process between smart gloves and users, which stores the experiment categories and the steps corresponding to each experiment. The set of experiment goals is set to $I = \{I_1, I_2, I_3, ....I_n\}$, and the different experiment categories are set to $I_n$. We use Baidu Voice to recognize and analyze user commands'. For the voice command $T$ acquired by the voice channel, after lexical analysis, the voice channel intention $T_i\left\{T_{s_i}, T_{ni}\right\}$ is obtained, where $T_{s_i}$ and $T_{n_i}$ are the actions and target objects contained in the voice command, respectively.For the data set $Q$ from the sensor channel and the data $pig$ from the vision channel, we integrate the obtained sensor processing results and the target object set $E$ obtained using the SPABG method to obtain the integrated result $G_i\left\{G_{s_i}, G_{ni}\right\}$. Among them, $G_{s_i}$ and $G_{n_i}$ are the actions performed by the user and the names of the target objects operated on respectively. After that, the intentions from the two channels are matched to obtain the user's current operation intention $O_i$, where $i$ is the time label set by the system.

When users conduct experiments, the order of operations is often consistent. Therefore, when we screen user intentions, in addition to using the information in the experiment library, we also need to combine the context information set $\delta\{O_1, O_2 ... O_{i-1}\}$ of the current user operation. Each value in $\delta$ is independent of each other. In the $t$ time period, when the smart glove obtains the user's operation intention $O_i$, it will add $O_i$ to $\delta$. We need to calculate the most likely experimental target H of the current user based on the set $\delta$, namely:

$$P(I_k|\delta) = \max\{P(I_1|\delta), P(I_2|\delta), P(I_j|\delta), ...P(I_n|\delta)\} \tag{4}$$

In the intention screening, we set the offset degree $\vartheta$ for each possible experimental target $I_n$, which is used to calculate the degree of offset of each value in $\delta$ in the experimental step corresponding to $I_n$. Then the deviation $\vartheta_{n_i}$ of the experimental operation $O_i$ in $I_n$ can be calculated by formula (5):

$$\vartheta_{n_i} = i - \beta_n[O_i] \tag{5}$$

where $\beta_n[O_i]$ is the number of the experimental step corresponding to $O_i$ in $I_n$. If $O_i$ is not in the experimental step of $I_n$, set it to blank. The calculation process of $P(I_n|\delta)$ in formula (4) is as follows:

$$P(I_n|\delta) = \frac{\alpha_n - \sum \vartheta_{ni}}{\alpha_n} \times 100\% \tag{6}$$

$$\alpha_n = \frac{[N(I_n)]^2 - 1}{2} \tag{7}$$

Among them, $\alpha_n$ is the maximum deviation degree of each experiment, and $N(I_n)$ is the total number of experimental steps in each experiment. For the finally obtained $P(I_k|\delta)$, it is necessary to make a confidence judgment. If $P(I_k|\delta)$ is greater than the intention determination threshold $\Theta$, the user's current most likely experimental target is determined to be $I_k$, otherwise the intention screening state will continue to be maintained.

### 4.3 Navigational Interactive Algorithm Based on Scene Perception

After obtaining the user's experimental intent, the smart glove will determine the user's behavior based on the rules in the knowledge base, and output corresponding voice feedback on the user's experimental behavior to guide the user to complete the experiment in a standardized manner. As shown in the navigational interactive algorithm (Navigational Interactive Algorithm, hereinafter referred to as NIA algorithm) based on multi-modal fusion intention understanding:

Algorithm 2: Navigational Interactive Algorithm Based on Multi-modal Fusion Intention Understanding (NIA Algorithm)

Input: experimental scene image $pig$, sensor information set $Q$, voice command $T$;

Output: user experiment intention, scene response combined with $a$, voice prompt;

1. Obtain information from each channel and perform preprocessing;
2. Use the SPABG Algorithm to process $pig$, get the smart glove's own position information set Oh, and the target object information set $E$;
3. Pass $Oh$ to the Unity platform;
4. Integrate $Q$ and $E$ to obtain the integration result $G_i\{G_{Si}, G_{ni}\}$ of the visual channel and the sensing channel;
5. Determine whether $T$ is empty. If it is not empty, call Baidu Voice API to analyze $T$ and obtain the current voice channel intention $T_i\{T_{Si}, T_{ni}\}$;
6. Use the multi-modal intent matching method to match $G_i\{G_{Si}, G_{ni}\}$ and $V_i\{V_{Si}, V_{ni}\}$:

(1) If the multi-modal intention matching fails, the voice prompt $Y_1$ will be output to ask the user whether to continue the hand operation. If the user confirms that the hand behavior is correct, save $G_i\{G_{Si}, G_{ni}\}$ as the user's operation intention $O_i$ into the experimental behavior set $\delta$, and output $O_i$; otherwise, keep $T_{ni}$ and return to step 1;

(2) If the multi-modal intention matching is successful, the user's operation intention $O_i$ will be output, and $O_i$ will be stored in the experimental behavior set $\delta$.

7. Determine whether the experiment target $I_m$ is empty:

(1)If $I_m$ is empty:

① Combined with the experimental target set $I = \{I_1, I_2, I_3, ....I_n\}$, use formula (5), formula (6) and formula (7) to filter $\delta$ by intention to obtain the calculation result $P(I_n \mid \delta)$;

② Let $P(I_k \mid \delta) = max(P(I_n \mid \delta))$, determine whether $P(I_k \mid \delta)$ is greater than the intention determination threshold $\Theta$;

③ If $P(I_k \mid \delta) > \Theta$, the user's current most likely experimental target is determined to be $I_k$, and a voice prompt $Y_2$ is output to the user to confirm to the user whether the experimental intention is correct:

a) If the user confirms that $I_k$ is wrong, output voice prompt $Y_3$, return to step 1, and update the experimental target set $I = \{I_1, I_2, I_3, ....I_n\}$;

b) If the user confirms that $I_k$ is correct, set $I_m = I_k$ to end the intent screening process.

8. If $I_m$ is not empty, query the rule set $R\{r_1, r_2, r_3..r_n\}$ corresponding to $I_m$ from the experiment library, and judge user behavior according to the corresponding rules in $R\{r_1, r_2, r_3..r_n\}$. Then output the scene response set $a\{a_{0,} a_1, a_{2,}...., a_{n,}\}$ and the corresponding navigation prompt according to the user behavior judgment result until the end of the experiment;

---

In the NIA algorithm, $Y_1$ indicates that the current hand behavior is inconsistent with the voice command, whether to continue the experimental operation; $Y_2$ indicates whether the current experimental target is guessed to be $I_k$ is correct; $Y_3$ indicates that you continue to perform the experimental operation.

# 5 Experiment and Analysis

Experimental hardware environment: CPU: i7-8750H.

Experimental software environment: Win10 64bit + Unity2018.3.8.

## 5.1 Feasibility Verification of Smart Gloves

In order to verify the effectiveness of the smart gloves designed in this article, this article selects one typical experiments in middle school experimental teaching for testing. The user wears gloves to verify whether the intelligent navigational interaction paradigm

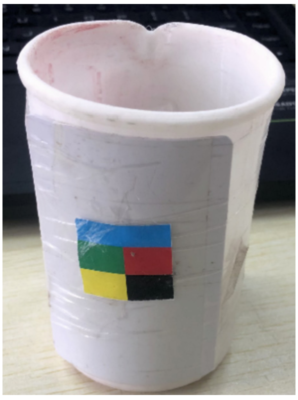

**Fig. 5.** Experimental equipment diagram

based on scene perception proposed in this paper can realize the perception of the experimental scene. In this paper, 3D printing technology is used to make the experimental supplies according to the size of the experimental supplies used in the actual experiments, and a label representing the category of the experimental supplies is set on the outside of the experimental mold. Some experimental equipment is shown in Fig. 5.

### 5.2 Experiment: Explore the Experimental Process of Concentrated Sulfuric Acid Dilution Experiment

The scene shown in Fig. 6 is a scene where users conduct experiments in a chemical experiment platform. Figure a shows that the user wears smart gloves to select experimental items. The items in the red boxes in Fig. 6b and Fig. 6c are the experimental reagents and instruments selected by the user; By judging the user's operating behavior, the smart glove speculates that the user wants to do a concentrated sulfuric acid dilution experiment. If the user confirms that the experiment intention is correct, the name of the experiment currently in progress and the steps of the experiment will appear on the screen (the information marked in the green box in Fig. 6d). At the same time, the smart glove recognizes that the reagent currently held by the user's hand is an aqueous solution by identifying the user's operating behavior. The smart glove reminds the user that if the experiment operation is continued, there will be danger. The information marked in the yellow box in the picture e is that the user injected water into the concentrated sulfuric acid solution, and the two reacted violently. The information marked in the green box in the figure f is the user's experimental results. The system also uses voice to explain the user's wrong operation and prompts the user to retry the experiment.

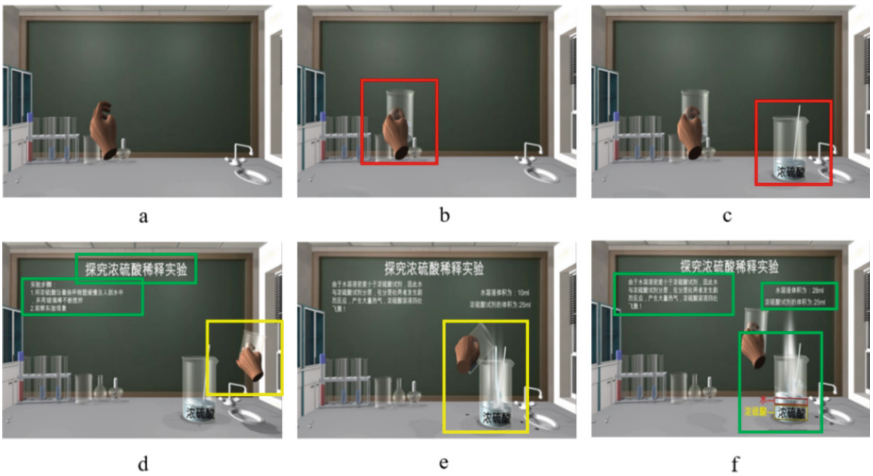


**Fig. 6.** Concentrated sulfuric acid dilution experiment diagram (wrong operation). (Color figure online)

As is shown in Fig. 7, after the user chooses to re-experiment, he will re-select experimental reagents (the information in the red box in Fig. 7b). The information marked in the yellow box in Fig. 7c is that the user pours the concentrated sulfuric acid solution in front of the aqueous beaker. Figure 7d shows that due to the user's dumping position is too high (information marked in the yellow box), the concentrated sulfuric acid solution splashed, causing corrosion of the desktop (information marked in the green box). Under the guidance of the smart glove, the user finally conducts the concentrated sulfuric acid dilution experiment at the correct experimental location, and uses a glass rod to continuously stir.

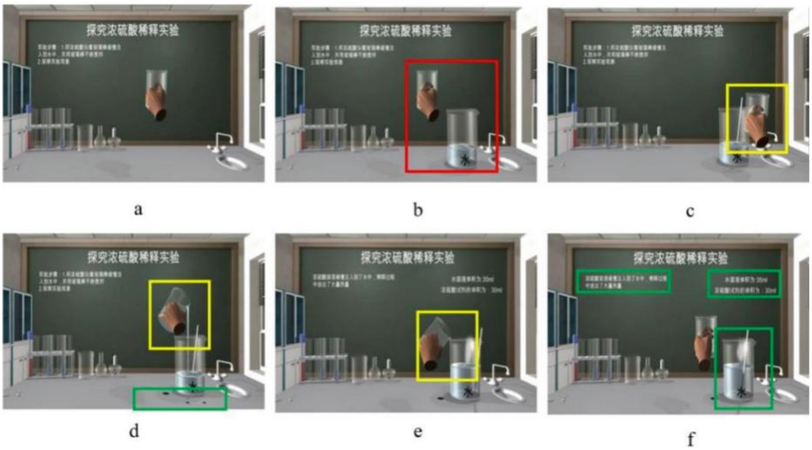


**Fig. 7.** Concentrated sulfuric acid dilution experiment diagram (right operation). (Color figure online)

## 6 Conclusion

Aiming at the existing problems in the experimental teaching process of primary and secondary schools, this paper designs a smart glove with scene perception ability, which can perceive the object information in the experimental scene and the corresponding position relationship in real time. On the basis of scene perception, this paper also proposes an intelligent navigational interaction paradigm based on multi-modal fusion to infer the user's experimental intention, so as to give the user corresponding feedback and guidance during the experiment, so that the user is in the smart glove Complete the experimental operation under the guidance.

The smart gloves designed in this article have the following advantages: 1. It can perceive multi-modal information, such as voice, scene information, etc.; 2. The smart glove kit combines human, machine, and material, allowing students to operate real experimental objects, which can improve students' experimental immersion and practical ability; 3. The smart glove kit designed in this paper can monitor and guide the user's behavior, and to a certain extent solve the problem that the teacher cannot guide every student in the classroom.

**Acknowledgment.** This paper is supported by the National Key R&D Program of China (No. 2018YFB1004901), and the Independent Innovation' Team Project of Jinan City (No. 2019GXRC013).

## References

1. Mur-Artal, R., Tardós, J.D.: ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Trans. Robot. (2017)
2. Redmon, J., Farhadi, A.: YOLOv3: An Incremental Improvement. arXiv e-prints (2018)
3. Han, T., Kennington, C., Schlangen, D.: Placing objects in gesture space: toward real-time understanding of spatial descriptions. In: AAAI conference of Artificial Intelligence (2018)
4. Yongda, D., Fang, L., Huang, X.: Research on multimodal human-robot interaction based on speech and gesture. Comput. Electr. Eng. **72**, 443–454 (2018)
5. Ismail, A.W., Billinghurst, M., Sunar, M.S., Yusof, C.S.: Designing an augmented reality multimodal interface for 6DOF manipulation techniques. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) IntelliSys. AISC, vol. 868, pp. 309–322. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01054-6_22
6. Jung, J., et al.: Voice+Tactile: augmenting in-vehicle voice user interface with tactile touchpad interaction voice user interface; tactile feedback touchpad; in-vehicle user interface. In: 2020 CHI Conference on Human Factors in Computing Systems (CHI 2020) (2020)
7. Sidenmark, L., Gellersen, H.: Eye, head and torso coordination during gaze shifts in virtual reality. ACM Trans. Comput. Hum. Interact. **27**(1), 1–40 (2019)
8. Sullivan, S., et al.: Middle school students' learning of mechanics concepts through engagement in different sequences of physical and virtual experiments. Int. J. Sci. Educ.
9. Gabunilas, L., et al.: Utilizing Portable Virtual Reality in Teaching Chemistry (2018)
10. Fidan, M., Tuncel, M.: Integrating augmented reality into problem based learning: the effects on learning achievement and attitude in physics education. Comput. Educ. **142**, 103635.1-1036351.9 (2019)
11. Yip, J., et al.: Improving quality of teaching and learning in classes by using augmented reality video. Comput. Educ. **128**, 88–101 (2018)
12. Salinas, P., Pulido, R.: Visualization of conics through augmented reality. Procedia Comput. Sci. **75**, 147–150 (2015)
13. Carlos, C.A., et al.: An approach of training virtual environment for teaching electro-pneumatic systems - ScienceDirect. IFAC-PapersOnLine **52**(9), 278–284 (2019)
14. Liu, D., et al.: Integration of virtual labs into science E-learning. Procedia Comput. Sci. **75**(2015), 95–102 (2015)
15. Hodges, G.W., et al.: An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. Comput. Educ. **122**, 179–193 (2018)
16. Aljuhani, K., Sonbul, M., Althabiti, M., Meccawy, M.: Creating a Virtual Science Lab (VSL): the adoption of virtual labs in Saudi schools. Smart Learn. Environ. **5**(1), 1–13 (2018). https://doi.org/10.1186/s40561-018-0067-9
17. Sotomayor-Moriano, J., Pérez, G., Soto, M.: A Virtual laboratory environment for control design of a multivariable process. IFAC-PapersOnLine **52**(9), 15–20 (2019)