


Research Article

Application of Virtual Reality Human-Computer Interaction Technology Based on the Sensor in English Teaching

Shuai Jiang,¹ Lei Wang,¹ and Yuanyuan Dong² 

¹Department of Foreign Languages, Cangzhou Normal University, Cangzhou 061001, China

²School of Foreign Languages, Xi'an Shiyong University, Xi'an 710016, China

Correspondence should be addressed to Yuanyuan Dong; dongyuanyuan0125@xsyu.edu.cn

Received 13 August 2021; Revised 16 September 2021; Accepted 5 October 2021; Published 19 November 2021

Academic Editor: Mu Zhou

Copyright © 2021 Shuai Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the online English teaching effect, the paper applies the sensor and human-computer interaction into the English teaching. The paper improves the sensor information by applying Kalman Filter, combines sensor positioning algorithm to trace the students in the English teaching online, and turns the kernels by the skeleton algorithm into corresponding coordinates of space rectangular coordinate system taking the waist as a coordinate origin to get a human-computer interaction skeleton model in the virtual reality. According to the actual needs of English teaching human-computer interaction, the paper builds a new English teaching system based on the sensor and the human-computer interaction and tests its performance. The experiments suggest that the smart system in the paper can effectively improve English teaching effects.

1. Introduction

The traditional English teaching mode has already been unable to fulfill the need of English teaching in the information age, especially in the period of the COVID-19. Online tutoring has already become a new trending. Therefore, it is necessary to conduct a revolution in the English teaching through information technology and smart technology [1]. With the extensive application of the computer, the users have been gradually increased, which are not just the early computer professionals. Generally speaking, the users have different needs and opinions toward the data, kinds of ways to solve the problems as well; and for the same user, he could present changes in different periods [2]. The traditional fixed user interface designed is not to the point, which cannot satisfy the specific need of the user and even the needs in different periods; the traditional Human-computer interaction system is just a system without any “flexibility,” and the users have to follow the interaction way and operate [3]. Sometimes, the user just needs a simple operation, which is obviously with a purpose especially understanding the user on the basis of previous operations, but the system is unable to know the real intention and just treats it as a normal opera-

tion. So this cannot cause a smooth information exchange between the system and the user. The user cannot freely interact with the computer, and the computer cannot deal with the requests from the user just like humans understand their friends. The user-centered human-computer interaction technology is the best effective way to solve the foresaid problems. The human-computer interaction system, to a certain degree, can allow the users to freely talk to the computer in their own familiar way; the computer has a certain ability of understanding and can understand the users to a certain degree or can smartly give the users a specific feedback although the intension is not clear enough [4].

A current research focus of human-computer interaction is to improve the performance of user interaction when users use computers to complete tasks. The improvement of user interaction performance includes a variety of human-computer interaction interface components, such as display media, display content, interface structure and style, and various input mechanisms. These human-computer interaction elements are closely related to user's perception, perception, processing, and reaction to the computer. To get the best user interaction performance, the interactive system must be built on the basis of a full understanding of users.

The first thing to study is how people conduct daily communication in real life, including the way of expressing intentions, the way of receiving outside information, and the way people express their intentions to people in the outside world and people's response models when asking questions. These expression models and response models are related to the types of users, users' experience, abilities, skills, and preferences in the field. At the same time, different users have different computer knowledge, comprehensive capabilities, and various factors that affect their interactions. It is necessary to understand the characteristics of all aspects of user behavior (from behavior to primary perception), establish the corresponding user model, then establish the user model of the corresponding application field, and improve the efficiency of interaction by improving system's understanding of people.

Based on above studies, the paper applies the sensor and virtual reality into the English teaching and builds a smart system so as to improve the English teaching effect.

2. Related Work

The ITS (intelligent tutoring system) is a multidisciplinary field of traditional CAI (computer aided instruction), AI (artificial intelligent), and cognitive science; ITS has internationally become a larger research mode and achieved convincing results [5]. Especially in recent 10 years, ITS grows fast abroad and turns into the commercial stage from research stage [6]. ITS can be established and maintain the student model; on one hand, the system can simulate the teaching decision-making and provide real-time tailored tutor to the students; on the other, the system can automatically analyze the historical data of a set of students in the student model and help teachers objectively understand the overall condition of students so that the teachers can timely adjust their own teaching plans and contents. The student model can tell the knowledge level, learning ability, and cognitive feature of the students; it is essentially a programme based on the algorithm that solves the actual problems in the way that students do [7].

Human-computer interactive learning systems using multimedia ingenuity are widely used at home and abroad. Human-computer interaction learning can provide students with a safe, predictable, repeatable, relaxed, and lively interactive learning environment. The system follows the basic principles of human-computer interaction and scientifically designs learning goals based on related theories and establishes corresponding human-computer interaction learning activities [8]. At the same time, the human-computer interactive learning environment pays attention to the design of interactive experience, is committed to letting students participate in interactive operations in an entertaining way, and focuses on the friendliness and fun of the game interface. The color matching and screen display design can attract students for different student groups. Attention is a system that mobilizes students' interest in learning and guides students to learn actively [9]. Literature [10] designed and implemented a three-dimensional drawing model tool, which can outline and model the Lanwei model. Users can

interact with gestures and pens. Although its accuracy is not as good as professional modeling software, it can satisfy students and nonbasic professional needs. Literature [11] uses an infrared camera to obtain the facial feature data of children, uses a pressure seat to obtain the child's body posture, and combines the learning behavior on child's computer to identify child's rational state. Literature [12] uses computer vision technology to study joint attention (that is, to follow and guide others' attention). The experiment uses a web camera to detect learner's head posture in real time to estimate learner's attention. Literature [13] establishes a multichannel learning environment by using video and audio fusion technology. In this learning environment, the video signal of the lips is obtained through the camera and merged with the voice signal obtained by the microphone. The speech recognition result and the handwriting recognition result are combined for human-computer interaction.

The rapid development of multimedia technology has also greatly promoted the progress of the *T* human-machine interface. The human-computer interaction gradually uses multimedia input and output devices such as microphones and cameras [14]. With the rise and development of emerging disciplines such as cognition, science, artificial intelligence, graphics, and image processing, multimedia-based multichannel human-computer interaction has gradually become a research hotspot. Virtual reality technology is a computer system that can create and experience a virtual world. It uses computer technology to generate a realistic virtual environment with multiple perceptions such as sight, hearing, and touch. The user uses various interactive devices to be the same as the virtual environment. The interactive visual simulation and information exchange that produce immersive feelings in the interaction between the entities in the world is an advanced digital human-computer interface technology, and its appearance will change the way of life of human beings [15]. For the research of virtual reality technology, literature [16] searched the WPI database and counted the country/regional distribution of related patents. From the perspective of the development of human-computer interaction, it can be found that the overall trend is towards naturalization and intelligent development. It is believed that the future human-computer interaction design will bring people a more relaxed and comfortable life [17].

3. Sensor Intelligent Fusion Algorithm

Extended Kalman filter (extended Kalman filter, EKF) is suboptimal that performs Kalman filtering after linearizing a nonlinear system.

Assuming that the system equation is represented by a discrete equation that remains constant under stochastic nonlinear conditions, the equation of state described above and the sensor measurement equation of the moving student can be expressed as follows [18]:

$$\begin{aligned}x(k+1) &= f(k, x(k)) + w(k), \\z(k) &= h(x(k)) + v(k),\end{aligned}\tag{1}$$

in which $w(k) \sim (N, Q(k))$ and $v(k) \sim (0, R(k))$ represent the Gaussian white noise sequence.

And $f(x) = [f_1(x), f_2(x), \dots, f_m(x)]^T$, and $h(x) = [h_1(x), h_2(x), \dots, h_p(x)]^T$, assuming that f, h can be expressed as the Taylor formula as followed.

$$f(x(k)) = f(\hat{x}(k)) + J_f(\hat{x}(k))[x(k) - \hat{x}(k)] + \dots \quad (2)$$

$J_f(x)$ is the Jacobi matrix of function f on $\hat{x}(k)$, then [19]:

$$J_f(x) = \left. \frac{\partial f}{\partial x} \right|_{x=\hat{x}(k)} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_m} \end{pmatrix}. \quad (3)$$

Similarly,

$$h(x(k)) = h(\hat{x}(k)) + J_h[x(k) - \hat{x}(k)] + \dots \quad (4)$$

If the Kalman filter incorporates several readings of the above ranging sensors, then $H(x(k))$ can be expressed as $H(x(k)) = d_i^j$

$$d_i^j = \frac{(ax_i + by_i + c)}{\sqrt{a^2 + b^2}}, \quad (5)$$

in which $P(k)$ is the distance from moving student sensor i to a point on the j wall in the environment; thus, the line segment characteristics of the global coordinates in the environment can be expressed as [20]:

$$ax_i + by_i + c = 0. \quad (6)$$

The Kalman algorithm is

(1) Testing phase

$$\hat{x}\left(\frac{k+1}{k}\right) = f\left(\hat{x}\left(\frac{k}{k}\right)\right), \quad (7)$$

$$P\left(\frac{k+1}{k}\right) = F(k) * P\left(\frac{k}{k}\right) * F(k)^T + Q(k). \quad (8)$$

In the above formula, P is the variance matrix of X , and F is the Jacobian matrix of f .

$$f(x(k)) = \begin{bmatrix} x(k) + T * v(k) * \cos \theta(k) \\ y(k) + T * v(k) * \sin \theta(k) \\ \theta(k) + T * (v(k)) * \sin \alpha(k)/l \end{bmatrix}, \quad (9)$$

$$F(k) = \nabla f(k) = \begin{bmatrix} 1 & 0 & -T * v(k) * \sin \theta(k) \\ 0 & 1 & T * v(k) * \cos \theta(k) \\ 0 & 0 & 1 \end{bmatrix}. \quad (10)$$

(2) Observation phase

From formula (7) and formula (10), the predicted measurement equation can be obtained:

$$\hat{z}(k+1) = h\left(\hat{x}\left(\frac{k+1}{k}\right)\right). \quad (11)$$

We use the predicted error of the sensor measurements to correct the predicted equation of state $\hat{x}(k+1)$, and we can obtain the predicted error of the measurements $\gamma(k+1)$, which can be expressed as

$$\gamma(k+1) = z(k+1) - \hat{z}(k+1). \quad (12)$$

The measurement error variance is

$$S(k+1) = H(k)P(k+1)H(k)^T + R(k). \quad (13)$$

H is the Jacobian matrix of $h(x(k))$ in the measurement Equation (7) [21]:

$$H(k) = \nabla h(k) = \left. \frac{\partial h}{\partial x_{x=\hat{x}(k)}} \right| = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \dots & \frac{\partial h_1}{\partial x_m} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \dots & \frac{\partial h_2}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \frac{\partial h_m}{\partial x_2} & \dots & \frac{\partial h_m}{\partial x_m} \end{pmatrix}. \quad (14)$$

(3) Estimation phase

It is seen from the above two steps that $\hat{x}(k+1)$ is the predicted value of the state that is realized by extending the Kalman filter to obtain the prediction error information based on the measured value. Among them, the measured variance $S(k)$ represents the corrected weight factor.

If the extended Kalman filter gain matrix is K , state is estimated as x , and its covariance matrix is P , and the

estimation value can be expressed in the following formula:

$$\begin{aligned} K(k+1) &= P\left(\frac{k+1}{k}\right)H(k)^T S^{-1}(k+1), \\ \hat{x}\left(\frac{k+1}{k+1}\right) &= \hat{x}\left(\frac{k+1}{k}\right) + w(k+1)\gamma(k+1), \\ P\left(\frac{k+1}{k+1}\right) &= (I - K(k+1) * H(k)) * P\left(\frac{k+1}{k}\right). \end{aligned} \quad (15)$$

In conclusion, the recursive formula of EKF includes the test, observation, and estimation phase. To obtain the optimal estimation value for the system, we can assume that some parameters are not related to the system, including the initial input noise $w(k)$ of the system and the measurement noise $v(k)$ generated during the measurement in the process, then we can handle the system by the irrelevant Gaussian white noise.

The sources of error in information fusion are summarized as follows:

(1) Milodometer system

An odometer is a sensor that uses data obtained from the mobile sensor to estimate the change of the object position over time, usually installed in the internal body position of the moving student. Its working principle is first use the pattern information or coding information on the code plate to obtain the rotation radian information of students' left and right wheel, and the forward direction and speed changes of the students can be calculated through these radian information.

According to the statements above, the odometer can represent the line speed and angular speed of the moving student; thus, odometer movement model can be represented as shown in Figure 1. We take the student as a mobile robot. Based on which, assuming that the student radius of movement is r , photocode disk is p line/turn, in a cycle time ΔT , the movement distance of the left wheel is Δs_l , the moving distance of the right wheel is Δs_r , and the moving distance of the car body Δs , the output pulses of student's photoelectric code disk is M , and the movement distance of the car body is

$$\Delta s = \left[\frac{2M}{p} \right] \cdot \pi r. \quad (16)$$

By the nonlinear regression analysis, it can be assumed that in a sampling cycle time ΔT , the movement track of the moving student is almost a straight line. If the distance between the left and right wheels is a , and then, the moving student is moved from the position $X(k-1) = (x_{k-1}, y_{k-1}, \theta_{k-1})$ to the position $X(k) = (x_k, y_k, \theta_k)$, the distance that the moving student moved is $\Delta s = (\Delta s_l + \Delta s_r)/2$, the angle information that the moving student rotated is $\Delta\theta = (\Delta s_l + \Delta s_r)/a$, and the odometer model input is $\mu(k)$ in the formula $\mu(k) = (\Delta D_k, \Delta\theta_k)^T$ obtained by the control command (v, w) . During a sampling cycle ΔT , the motion path of the moving e student is $\Delta D_k = v \times \Delta T$, and the pose relative angle is $\Delta D_k = v \times \Delta T$.

The odometer has two models: circular arc and straight lines. Since the linear model is simplified from the arc

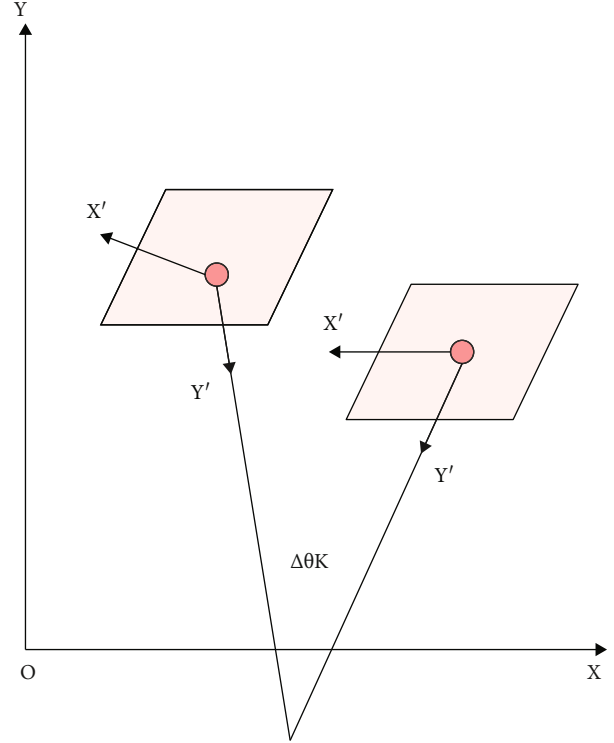


FIGURE 1: Mileometer motion system.

model, the arc mode is generally used in practice, which is more accurate.

(2) Arc model

$$x(k) = f(x(k-1), u(k)) = \begin{bmatrix} x_{k-1} - \frac{\Delta D_k}{\Delta\theta_k} (\cos(\theta_{k-1} + \Delta\theta_k) - \cos\theta_{k-1}) \\ y_{k-1} + \frac{\Delta D_k}{\Delta\theta_k} (\sin(\theta_{k-1} + \Delta\theta_k) - \sin\theta_{k-1}) \\ \theta_k = \theta_{k-1} + \theta_k \end{bmatrix}. \quad (17)$$

$\Delta\theta_k$ in formula (17) indicates the rotation angle difference between the end point and the beginning point of the moving student, in which ΔD_k is the movement displacement of the moving student during the sampling period ΔT time.

(3) Line model

Assuming that the moving students makes a small displacement in a very small period of time, the extremal method of nonlinear regression analysis method can be used for processing, and the model dealing with the problem becomes a linear model, which can be represented by simplified lines, namely, $|\Delta\theta| = 0$, then the model equation can be expressed as

$$x(k) = f(x(k-1), u(k)) = \begin{bmatrix} x_k + \Delta D_k \sin\theta_k \\ y_k + \Delta D_k \cos\theta_k \\ \theta_k + \Delta\theta_k \end{bmatrix}. \quad (18)$$

To calculate the position of moving students, in this paper, linear model and arc model are used. The specific implementation method is the displacement calculation adapts a straight line model, and the arc model is used to calculate the change difference of direction angle. Then, the two models are represented as

$$x(k+1|k) = f(x(k|k), u(k)) = \begin{bmatrix} x_k + \Delta D_k \sin \theta_k \\ y_k + \Delta D_k \cos \theta_k \\ \theta_k + \Delta \theta_k \end{bmatrix}. \quad (19)$$

At the k moment, if line speed and angular speed values of the moving student are known, formula (19) can be transformed as

$$x(k+1|k) = f(x(k|k), u(k)) = \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + v \times \Delta T \sin \theta_k \\ y_k + v \times \Delta T \cos \theta_k \\ \theta_k + W \times \Delta T \end{bmatrix}. \quad (20)$$

The errors of the moving student odometer include two types, namely, nonsystematic error and systematic errors (as shown in Figure 2). Its systematic error accumulation is always existed, while the nonsystematic error accumulation is random and indeterminable.

(4) Sonar calculation method

Generally, most of the sonar applications adapts very cheap Polaroid6500 sonar modules. Among them, the environmental road sign characteristic is used for the sonar sensor to scan the environmental information, extract useful environmental feature information, and correctly calculate the specific location of the environmental characteristics. Assuming that the coordinate position of the line segment in the local coordinate system is (x_l, y_l) , the predicted coordinate position (x_w, y_w) for observation can be estimated and obtained. Among them, the width of the sonar beam and the reflection of the sonar signal are the important factors affecting the sonar sensor. Figure 3 shows the coordinate system transformation diagram.

In Figure 3, the point O' is the origin in this local coordinate system that represents the moving student, the point P represents the wall in the moving student walking area, a is the sonar sensor, and O represents the offset angle of the on-board coordinate relative to the coordinate OXY . The degree of the sonar is represented by the point S , and the distance to the point P is d , and the relative position of each of the sonar is fixed.

Figure 4 shows the interrelationship between the global and local coordinate systems for moving students. The width of the sonar beam and the reflection of the sonar signal are two factors affecting the principle of the sonar sensor.

Assuming that in the moving student environment, the reflected wall is represented by P , the distance from the origin O to the plane and the distance between the plane and

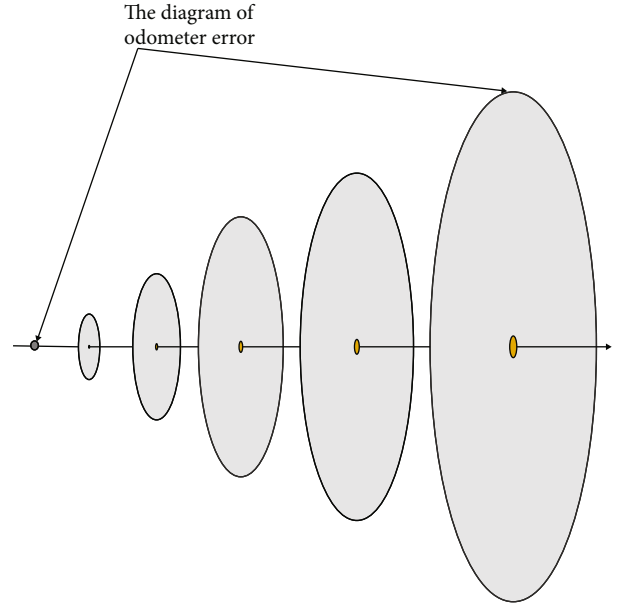


FIGURE 2: The diagram of odometer error.

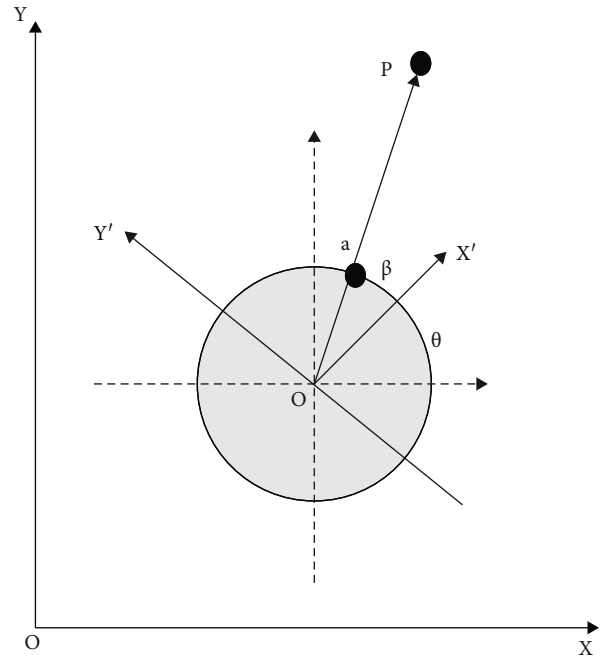


FIGURE 3: The coordinate system transformation of sonar data.

the sonar sensor in the environment are represented by P_r^j , P_n^j , P_l^j . Since the relative position information of the sonar in the local coordinate system is known, which can be represented by (x_i, y_i, θ_i) , formula (22) and formula (23) can be used to calculate and coordinate position of d_i^j and P point on the wall in the environment.

In the figure above, j is the number of planes, $d_i^j(k)$ is the number of sonar sensors, and the distance from the obstacle

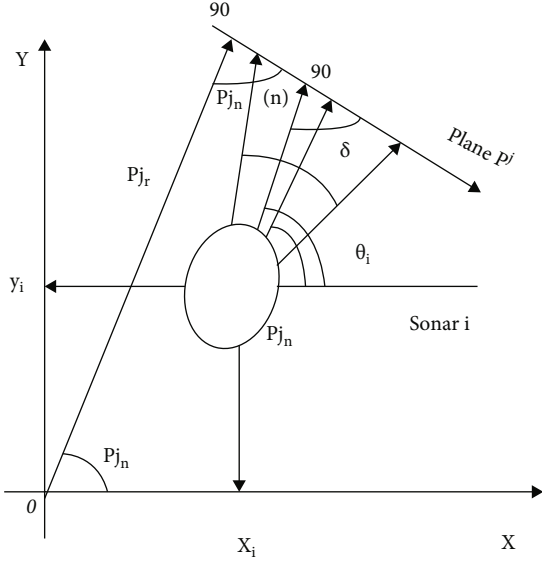


FIGURE 4: The position of sonar in the global coordinate system.

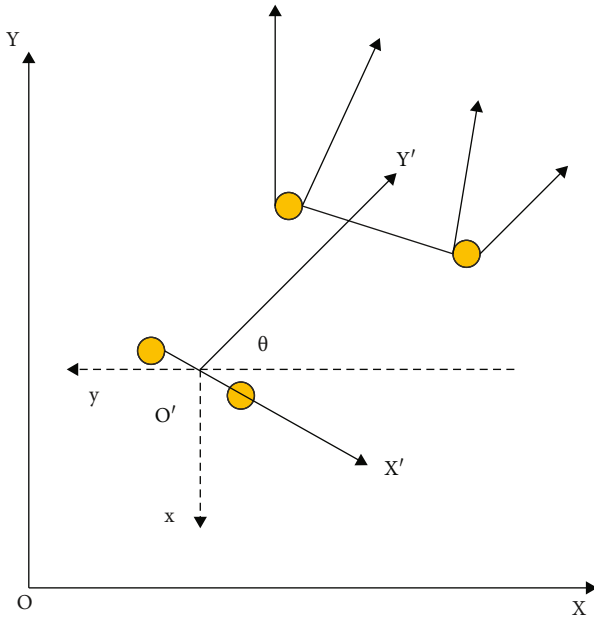


FIGURE 5: Odometer and coordinates.

plane p^i to the sonar i at the k moment is $d_i^j(k)$, which can be simplified as

$$d_i^j(k) = P_r^j - x_i(k) \cos(P_n^j) - y_i(k) \sin(P_n^j). \quad (21)$$

Set $P_n^j = \theta_l$, $P_r^j = D$, in $P_n^j \in [\theta_l(n) - \delta/2, \theta_l(n) + \sigma/2]$, δ is the sonar beam width, and $d_i^j(k)$ may be represented based on formula (21) above as

$$d_i^j(k) = D - x_i(k) \cos(\theta_l) - y_i(k) \sin(\theta_l) = \frac{ax_i + by_j + c}{\sqrt{a^2 + b^2}}. \quad (22)$$

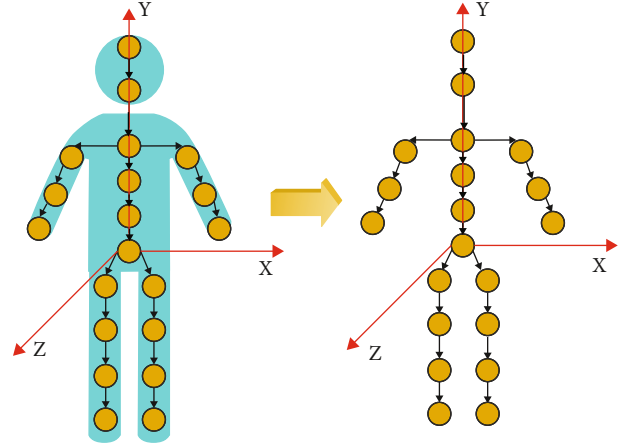


FIGURE 6: Coordinate system of English human-computer interaction.

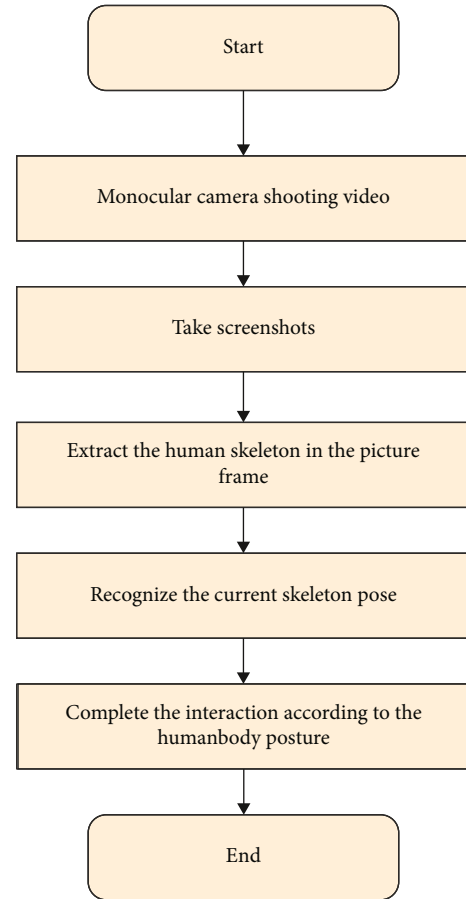


FIGURE 7: Process of human-computer interaction in English teaching.

According to the view of Figure 5 above, assuming we are sampling at the k time, the position of the moving student is $X(k) = (x(k), y(k), \theta(k))$. After a rotating coordinate transformation, we can convert the coordinate numbers (x', y', θ') of i in $OX'Y'$ system to (x_i, y_i, θ_i) in the OXY

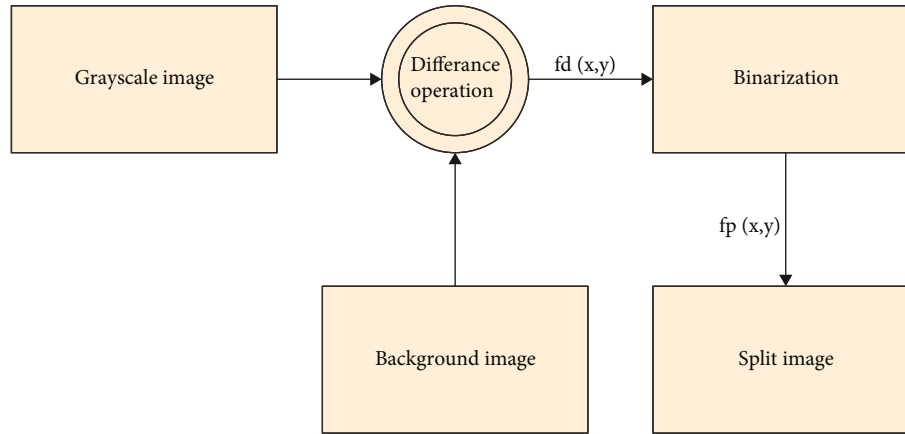


FIGURE 8: Principle process of background subtraction.

TABLE 1: Performance appraisal of English teaching system based on sensor and human-computer interaction technology.

Number	Sensor performance	Human-computer interaction effect	Number	Sensor performance	Human-computer interaction effect
1	88.66	91.65	25	92.27	93.08
2	94.53	93.03	26	96.85	80.75
3	87.85	80.98	27	86.34	93.00
4	87.70	94.24	28	95.74	81.96
5	94.49	86.22	29	89.47	86.43
6	90.02	79.25	30	95.93	90.59
7	91.49	84.28	31	96.30	88.78
8	96.77	91.27	32	91.53	91.93
9	96.69	93.75	33	86.41	89.75
10	96.10	84.68	34	95.50	86.21
11	89.65	94.55	35	86.79	87.97
12	94.61	88.78	36	91.01	79.27
13	94.37	81.09	37	93.39	85.66
14	86.32	82.13	38	87.51	93.74
15	89.84	89.90	39	96.28	87.71
16	92.45	90.37	40	88.11	91.88
17	88.14	83.13	41	87.76	79.93
18	87.21	81.10	42	91.21	85.00
19	94.42	79.82	43	94.64	85.07
20	86.71	87.15	44	94.73	82.44
21	96.06	90.64	45	88.49	81.41
22	94.55	91.73	46	96.01	94.98
23	94.37	93.82	47	89.35	85.58
24	92.06	93.69	48	95.78	82.30

coordinate system, then the model of the sonar sensor is expressed as the following formula:

$$\begin{cases} x_i(k) = x(k) - x'_i(k) \sin(\theta(k)) + y'_i(k) \cos(\theta(k)), \\ y_i(k) = y(k) - x'_i(k) \cos(\theta(k)) + y'_i(k) \sin(\theta(k)), \\ \theta_i(k) = \theta(k) + \theta_i. \end{cases} \quad (23)$$

As the P point on the wall in the sonar coordinate system in Figure 5, the point P , a , O' is on the same line, and the position detected by moving the student odometer and the sonar sensor can be represented as

$$\begin{aligned} X_p &= x_0 + x_p = x_0 + x_i \times \cos(\theta) - y_i \times \sin(\theta) + d_i^j \times \cos(\theta), \\ Y_p &= y_0 + y_p = y_0 + x_i \times \sin(\theta) - y_i \times \cos(\theta) + d_i^j \times \sin(\alpha + \theta). \end{aligned} \quad (24)$$

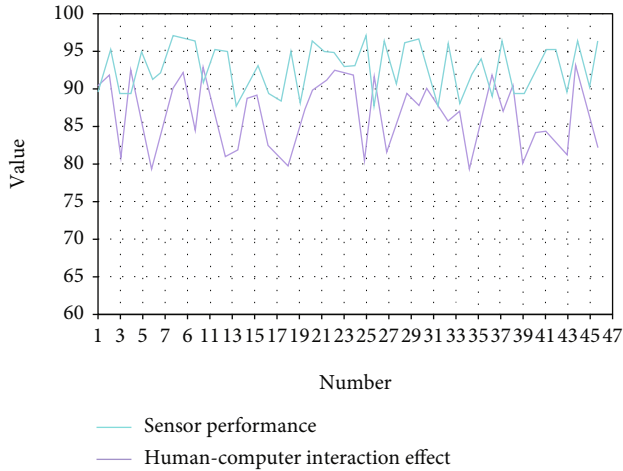


FIGURE 9: Statistics of English teaching system performance based on sensor and human-computer interaction.

Mobile students must rely on the collected external environment information to achieve more accurate navigation and positioning. However, in the real environment, there are various uncertainties and intertwined complexity in the environment itself, which leads to the absence of theoretically reliable environmental information. In addition, the data measured by the mobile student's own sensors will also be disturbed by the environment. As a result, the measured environmental information is not our ideal information, and there are often extremely complex uncertain factors. In order to reduce or even eliminate these interferences in the observation information of mobile student sensors, we can establish a reliable noise simulation system to express the uncertain factors described above. However, this is only our ideal state. In fact, due to the nonlinear and incomplete constraint system of moving students, it is not easy to establish an efficient and accurate mathematical model. In this case, we use a model (also called an error model) that introduces noise that obeys the Gaussian function model to approximate it.

In this paper, the sonar sensor is used to detect the road signs of environmental information. Due to the characteristics of the sonar sensor, if an observation point cannot be selected correctly in the process of selecting an observation point, it will lead to the environment corresponding to it. It is impossible for the feature information to perform correct data association, which will cause the divergence of the filter. The divergence of the filter will cause all the predicted observations to be wrong. In order to make the road sign data scanned by the sonar sensor match correctly, a data association model can be introduced to solve this problem.

The analysis shows that the position estimation inferred by the mobile student through the odometer and sonar sensor is inaccurate, so the correct position estimation of the mobile student cannot be obtained. In this paper, a two-dimensional vector road sign is selected as the observation in the position of the global coordinate system. Since the environment in the space is variable and complex, but there is a certain correlation between the environmental characteristics, it can be

estimated; how to eliminate the interference of the correlation between these variables is the problem we want to solve. A large number of studies have verified that Mahalanobis distance can well eliminate these interferences, so Mahalanobis distance is selected as the measure of data association in this article. In addition, it should be noted that the correlation gate used in this article is an elliptical correlation gate.

Mahalanobis (Mars distance) is defined as $d_k^2(x) = (x - m_k)s_k^{-1}(x - m_k)$, $d_k^2(x)$ represents the mean distance of the measured value within group k , the vector m in the formula represents the mean of the variable in group k , x represents the value of the variable in the environment observed by the sensor model, and s represents the covariance matrix within the group.

4. English Teaching System Based on the Sensor and Human-Computer Interaction

The kernels of the skeleton extracting algorithm are universally turned into corresponding coordinates of space rectangular coordinate system taking the waist as a coordinate origin. The right part of the body is x -axis, the vertical upper part is y -axis, the front of vertical to the body is z -axis, which is shown in Figure 6.

The paper proposes human-computer interaction method based on monocular camera. First extract the physical skeleton from the video taken by monocular camera and complete the interaction with virtual environment through the human body pose. The process is shown in Figure 7.

After cutting out the current frame, first of all, pretreat the picture. That is because the profiles in the video are of sizes, if the body in the frame takes up smaller, which can greatly reduce the precision of the later kernels. At the same time, much irrelative information in the frame could cause too much useless information in the network training process and affect the precision of identification result. Therefore, we need to find out the boundary of human profile and get it out from the frame to identify aiming to improve the precision in the following stages. There are many ways to extract from the RGB images. As the video in the paper is shot by a fixed camera and there is only one person in the frame, our background is fixed, and we use background subtraction to extract the human profile from the video; the principle is shown in Figure 8.

The paper builds English teaching system based on the sensor and human-computer interaction technology and tests the performance. According to actual needs, the paper conducts a test on the basis of the real-time information transmission in English teaching by the sensor and human-computer interaction. The simulation experiment is used to test the performance of the system, transmission effect of sensor, and human-computer interaction (by the way of experts' appraisal); the statistics is shown in Table 1 and Figure 9.

From above experiment, the English teaching system in the paper based on sensor and human-computer can effectively improve the English teaching effect and have some positive influence on the following English teaching revolution.

5. Conclusion

In the teachers-centered link, teachers play roles of traditional instructor, teaching monitor, after-school tutor, homework assigner and corrector, and document provider. With further development of teaching-related theory and interaction design theory, the human-computer interaction between teachers and students has attracted more and more attention among people, from the feasibility of human-computer interaction in English teaching in the light of theory and study, to the theory foundation of creating learning environment. In order to improve the effect of human-computer interaction in English teaching, the paper studies the sensor technology and puts a Kalman Filter Algorithm which stresses the precision to solve the problems from the human-computer interaction information by the positioning in sensor information. The virtual reality is used to build in English teaching online interaction system. The system goes through the performance test, and the study suggests that the system in the paper has certain effects.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

Authors' Contributions

Yuanyuan Dong completed the analysis and thesis writing as the main writer. Shuai Jiang and Lei Wang completed the collection and sorting of experimental data and jointly participated in the revision, polishing, and typesetting of the article, so I decided to add them to the list of authors of the article. Shuai Jiang and Lei Wang assisted in the analysis through constructive discussions based on their professional expertise.

Acknowledgments

This study is sponsored by “The 2020 year project of 13th Five-year plan of Educational Science in Shaaxi Province, China (grant no. SGH20Y1131)”.

References

- [1] L. Delgaty, “Twelve tips for academic role and institutional change in distance learning,” *Medical Teacher*, vol. 37, no. 1, pp. 41–46, 2015.
- [2] M. Stefanovic, D. Tadic, S. Nestic, and A. Djordjevic, “An assessment of distance learning laboratory objectives for control engineering education,” *Computer Applications in Engineering Education*, vol. 23, no. 2, pp. 191–202, 2015.
- [3] B. P. Remesh, “Developing open and distance learning programme in labour and development: results of a needs assessment study,” *Journal of natural history*, vol. 196, no. 29, pp. 265–291, 2015.
- [4] P. Wu, S. P. Low, J. Y. Liu, J. Pienaar, and B. Xia, “Critical success factors in distance learning construction programs at Central Queensland University: students' perspective,” *Journal of Professional Issues in Engineering Education and Practice*, vol. 141, no. 1, 2015.
- [5] E. A. Willis, A. N. Szabo-Reed, L. T. Ptomey et al., “Distance learning strategies for weight management utilizing social media: a comparison of phone conference call versus social media platform. Rationale and design for a randomized study,” *Contemporary Clinical Trials*, vol. 47, no. 6, pp. 282–288, 2016.
- [6] H. J. Ye, D. C. Zhan, and Y. Jiang, “Fast generalization rates for distance metric learning,” *Machine Learning*, vol. 108, no. 2, pp. 267–295, 2019.
- [7] Y. Luo, Y. Wen, T. Liu, and D. Tao, “Transferring knowledge fragments for learning distance metric from a heterogeneous domain,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 4, pp. 1013–1026, 2019.
- [8] T. Chuk, A. B. Chan, and J. H. Hsiao, “Understanding eye movements in face recognition using hidden Markov models,” *Journal of Vision*, vol. 14, no. 11, pp. 8–8, 2014.
- [9] H. Bouchech, “Selection of optimal narrowband multispectral images for face recognition,” *Monthly notices of the royal astronomical society*, vol. 402, no. 4, pp. 2140–2186, 2015.
- [10] R. Ramachandra and C. Busch, “Presentation attack detection methods for face recognition systems,” *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–37, 2017.
- [11] Z. Huang, S. Shan, R. Wang et al., “A benchmark and comparative study of video-based face recognition on COX face database,” *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5967–5981, 2015.
- [12] C. P. Wei and Y. C. F. Wang, “Undersampled face recognition via robust auxiliary dictionary learning,” *Image Processing IEEE Transactions*, vol. 24, no. 6, pp. 1722–1734, 2015.
- [13] N. L. Ajit Krishna, V. K. Deepak, K. Manikantan, and S. Ramachandran, “Face recognition using transform domain feature extraction and PSO-based feature selection,” *Applied Soft Computing*, vol. 22, pp. 141–161, 2014.
- [14] R. Weng, J. Lu, and Y. P. Tan, “Robust point set matching for partial face recognition,” *Image Processing IEEE Transactions*, vol. 25, no. 3, pp. 1163–1176, 2016.
- [15] Z. Cui, H. Chang, S. Shan, B. Ma, and X. Chen, “Joint sparse representation for video-based face recognition,” *Neurocomputing*, vol. 135, pp. 306–312, 2014.
- [16] P. J. Phillips and A. J. Otoole, “Comparison of human and computer performance across face recognition experiments,” *Image and Vision Computing*, vol. 32, no. 1, pp. 74–85, 2014.
- [17] D. Tang, N. Zhu, F. Yu, W. Chen, and T. Tang, “A novel sparse representation method based on virtual samples for face recognition,” *Neural Computing & Applications*, vol. 24, no. 3–4, pp. 513–519, 2014.
- [18] R. Mehta, J. Yuan, and K. Egiazarian, “Face recognition using scale-adaptive directional and textural features,” *Pattern Recognition*, vol. 47, no. 5, pp. 1846–1858, 2014.
- [19] M. Pinxten, C. V. Soom, C. M. Peeters, T. De Laet, and G. Langie, “At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies?,” *European journal of psychology of education*, vol. 34, no. 1, pp. 45–66, 2019.

- [20] M. Brehm, S. A. Imberman, and M. F. Lovenheim, "Achievement effects of individual performance incentives in a teacher merit pay tournament," *Labour Economics*, vol. 44, no. 5, pp. 133–150, 2017.
- [21] R. J. McGill and A. R. Spurgin, "Assessing the incremental value of KABC-II Luria model scores in predicting achievement: what do they tell us beyond the MPI?," *Psychology in the Schools*, vol. 53, no. 7, pp. 677–689, 2016.