IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 19, 2020, accepted August 1, 2020, date of publication August 4, 2020, date of current version August 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3014169

A Three-Dimensional Ultrasonic Pen-Type Input Device With Millimeter-Level Accuracy for Human–Computer Interaction

JIAN CHEN, FAN YU, JIAXIN YU, (Graduate Student Member, IEEE), AND LIN LIN
College of Communication Engineering, Jilin University, Changchun 130022, China

Corresponding author: Jian Chen (chenjian@jlu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB1001304.

ABSTRACT In this article, a pen-type device that enables three-dimensional (3D) human–computer interaction is presented. This device is based on a high-precision 3D ultrasonic (US) positioning method, which can achieve millimeter-level accurate 3D positioning and six degrees of freedom (6DoF) within a working range of 4 m × 1.5 m × 1.5 m. The 3D positioning of the electronic pen is realized by a high-precision time-of-flight (TOF) detection method based on dynamic thresholding and zero-crossing detection as well as a multigroup positioning coordinate averaging method. During the 3D positioning process, correlation filtering is adopted to process the received US signals and suppress any interference signals below 20 kHz and above 60 kHz. The positioning method can achieve a high positioning accuracy without any complex calculations. The attitude of the electronic pen is measured using an inertial measurement units (IMUs). To demonstrate the positioning accuracy of the proposed device, a total of 108 test positions are uniformly set in the workspace, and static and dynamic positioning accuracy experiments are carried out at all test positions. The experimental results show static and dynamic positioning errors of less than 0.35 mm and less than 1.0 mm, respectively, which indicate a high positioning accuracy. The proposed device can provide 6DoF and enable writing in 3D. Therefore, the device can meet the common requirements of a 3D human–computer interaction interface.

INDEX TERMS 3D human–computer interaction interface, six degrees of freedom (6DoF), 3D US positioning, dynamic threshold, zero-crossing detection.

I. INTRODUCTION

In recent years, pen-type devices have emerged as input modules for tablet computers, handheld devices, and large interactive display systems [1], [2]. These devices provide free, easy and familiar expression and can be used by people who do not have any special drawing skills, making them a promising class of tools for supporting natural interaction [3], [4]. Traditional windows, icons, menus and pointing devices (WIMPs) are the representative two-dimensional (2D) interactive interfaces designed to be operated by pointing devices, such as a mouse, with two degrees of freedom (2DoF) [5]. Traditional pen-type input devices can provide 2D positional information of the pen tip, as well as other information such as the writing pressure, three-dimensional (3D) orientation and rotation [1]. However, traditional pen-type devices cannot provide six degrees of freedom (6DoF) and can only function on a flat surface, which limits their utility in 3D user interfaces.

Modern computing hardware and software can enable the creation of rich and realistic 3D virtual environments thanks to the rapid development and application of virtual reality (VR), augmented reality (AR), ubiquitous and mobile computing, large interactive displays and other ''off-the-desktop'' technologies. Such 3D virtual environments are popular in gaming, education, training, prototyping, digital 2D and 3D drawing, general 3D manipulation, and any other application where a realistic virtual representation of the real world is useful [6].

Real-time 6DoF and position tracking form the basis for the 3D interaction of spatial input devices, such as controllers, 3D pointing devices, pen-type input devices and whole-hand devices [7]–[9], in the aforementioned 3D applications.

The associate editor coordinating the review of this manuscript and approving it for publication was Maurizio Tucci.

To operate 3D virtual objects, the 3D interaction can be used to track the position, orientation, and/or motion of the user to maintain a proper correspondence between the physical and virtual contents in the 3D space [7]. Consequently, pen-type input devices with accurate tracking and 6DoF are crucial for enabling interactions within 3D applications [7].

Currently, based on the technology used for position and rotation tracking, the major 3D spatial input devices with high accuracy and low latency can be divided into three categories: magnetic, optical, and inertial and acoustic tracking systems [10], [11]. The devices in these categories are discussed in the following paragraphs.

A magnetic spatial input device uses a transmitting device that emits a low-frequency magnetic field. The receiver's position and orientation relative to the magnetic transmitter source is determined using a small sensor. An example of such a device is the Polhemus PATRIOT G4 [12]. This device has 6DoF, a static accuracy of 0.08 inches, i.e., 2.0 mm, and does not suffer from obscuration problems. However, a large transmitter is needed in this system to obtain a large range. In addition, its major disadvantage is the magnetic field distortion caused by any ferromagnetic or conductive (metal) objects present near the transmitter, which can occasionally cause severe accuracy loss.

An optical spatial input device is based on the stereophotogrammetry technique. This technique uses computer vision and optical sensors such as cameras to calculate the 3D coordinates and orientations of different points of physical objects based on the light reflected or emitted by them and the users [13]–[15]. One typical optical tracking system is Optitrack [16]. Although the accuracy of a modern optical tracking system is very high, a drawback of this technology is its high costs due to the presence of many cameras [15].

An inertial tracking system uses a variety of inertial measurement components. These components can include either angular-rate gyroscopes, linear accelerometers and magnetometers or inertial measurement units (IMUs), which can integrate 3-axis gyroscopes, 3-axis accelerometers and even 3-axis magnetometers. Inertial tracking is fast and robust. The inertial measurement components provide derivative measurements of position and orientation. Therefore, these measurements are integrated to obtain position and orientation information. They can be used for motion tracking when fast changes occur; however, in the presence of slow motion, noise and drift can cause errors. Due to this major limitation, inertial tracking is generally used in conjunction with other tracking techniques such as magnetic, optical, and acoustic tracking.

Acoustic tracking systems typically use high-frequency sound waves, such as ultrasonic (US) waves, to measure the distance between a fixed-point station and a mobile target. Multiple synchronized US receivers are needed to implement such a tracking system. The advantage of this tracking technique is the relatively low cost [7]. The disadvantage of the system is the multipath reception that can affect the distance measurements between the emitter and the receivers.

Each tracking technology has its advantages and disadvantages. To increase the accuracy and operational efficiency while reducing latency, hybrid tracking has been adopted to overcome the weaknesses of the individual technologies. For example, Intersense Inc. developed a 6DoF tracker known as IS-900, which uses inertial and US hybrid tracking technologies [17], [18]. Another 6DoF tracker, known as zSpace [19], has four infrared (IR) sensors on the front that emit and receive IR light. The stylus used on the zSpace display emits IR light from its tip and contains an accelerometer, gyroscope and magnetometer. zSpace achieves 6DoF by capturing a visually indicated point and any additional information [20]. However, it has a limited working range because of the sensor's mounting position and angle.

Inertial and US hybrid tracking systems are cheaper than optical systems. In these hybrid systems, US technology is used to obtain the 3D position of the tracker, and the IMU is used to obtain the tracker's orientation and motion. The IS-900 system's cost is lower than that of an optical system, and the IS-900 system is sufficiently compact and lightweight. However, at typical tracking speeds, its spatial accuracy has a root mean square (RMS) error of up to 17 mm [18].

In addition to the above positioning technologies, indoor positioning technologies include radio frequency positioning and visible light positioning technologies. Although radio frequency identification (RFID), ZigBee, and wireless local area network (WLAN) positioning technologies have the advantage of not requiring a line of sight, the accuracy of the above positioning methods is at the meter level [21]–[22] and cannot be adapted to the precision of 3D input devices. Visible light positioning (VLP) technology [23] is a cost-effective indoor positioning technology. Although VLP can achieve centimeter-level accuracy, it still cannot meet the positioning accuracy requirements for air writing, which requires millimeter-level positioning accuracy.

In this article, a new 3D large spatial pen-type input device based on 3D high-precision positioning technology is presented. Compared to existing works, our contributions are as follows.

1. A high-precision time-of-flight (TOF) detection method based on dynamic thresholding and zero-crossing detection is presented. The proposed method overcomes the limitations of the traditional fixed-threshold method, achieves a detection accuracy higher than the sampling period, and has a low computational cost.
2. The pen-type device achieves millimeter-level accurate 3D positioning and 6DoF within a working range of 4 m $\times$ 1.5 m $\times$ 1.5 m and supports air writing. The static and dynamic positioning accuracies of the device are higher than 0.35 mm and 1 mm, respectively.

The rest of this article is organized as follows: Section II describes the architecture of the 3D US pen-type input device. Section III focuses on the 3D high-precision US positioning method. Static and dynamic positioning experiments and

their results are presented in Section IV. Finally, the conclusions and future work are described in Section V.

II. ARCHITECTURE OF THE PROPOSED 3D ULTRASONIC PEN-TYPE INPUT DEVICE

The proposed 3D US pen-type input device uses received IR and US pulses to calculate the absolute time of flight. However, because the beam angle of the US transmitter is not large enough, multiple receivers are needed to ensure that at least three receivers can receive the pulses. Figure 1 shows an overview of the proposed pen-type input device, which consists of six receivers and one electronic pen. In the figure, receivers R1, R2, …, R6 are fixed on the upper and lower edges of the whiteboard. The electronic pen includes one cylindrical piezoelectric polyvinylidene fluoride (PVDF) US transmitter and four infrared emitters. Although the omni-directional horizontal directivity pattern of the cylindrical PVDF US transmitter [24] and the IR emitters allow line-of-sight transmission to each of the receivers, the narrow vertical beam emission pattern of the US and IR transmitters limits their application in a 3D US location system. To measure the position and rotation of the electronic pen in the 3D space, multiple receivers are needed to ensure that at least three receivers can receive the positioning signal in each positioning period. Each receiver includes one microphone (MIC) and one IR photodiode to receive the US and IR pulses transmitted from the electronic pen, respectively.

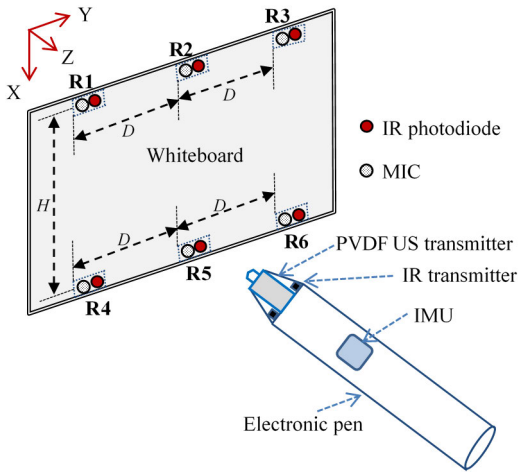


FIGURE 1. Architecture of the proposed 3D US pen-type input device.

The IR and US pulses are simultaneously emitted from the electronic pen for a fixed period. The former is used to determine the beginning of the reference time. The distance between the pen and each receiver is measured based on the difference between the times of arrival of the IR and US pulses at the receiver. After the successful measurement of more than three distances by the corresponding receivers, triangulation can be used to calculate the pen position.

The block diagram of the proposed device is shown in Fig. 2. The device consists of six receivers, one electronic pen, and one controller. The origin of the 3D Euclidean

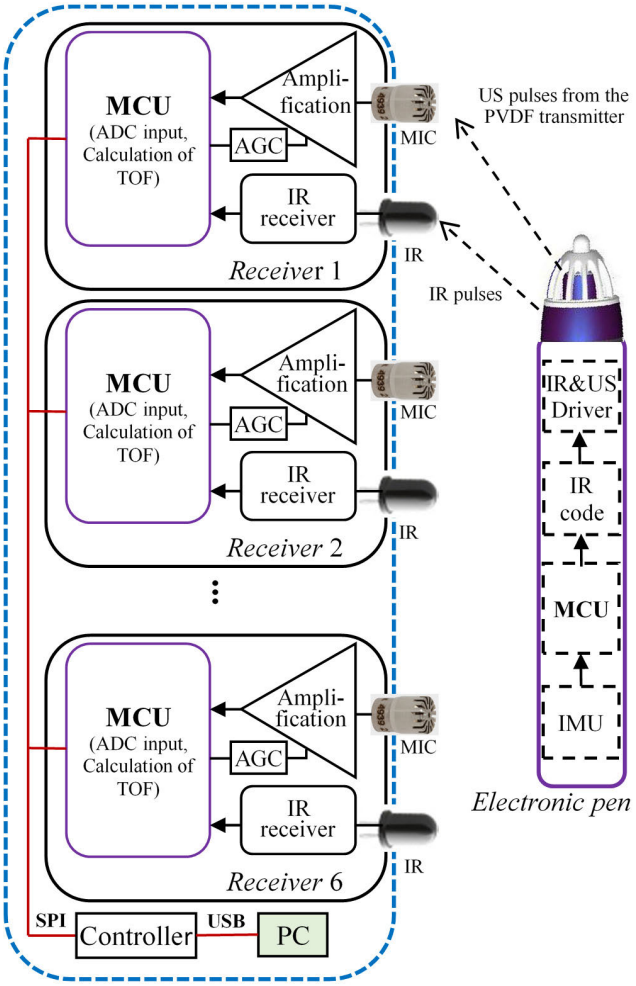


FIGURE 2. Block diagram of the proposed 3D US pen-type input device.

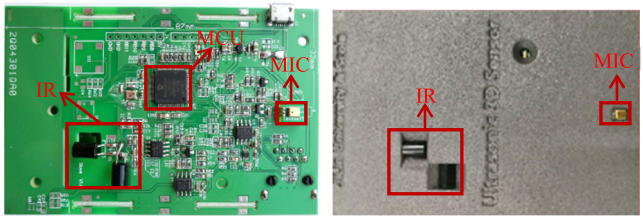
coordinate system is assumed to be at the center of R1. The x-y plane corresponds to the board on which receivers R1, R2, …, R6 are mounted, and the z-axis is perpendicular to the x-y plane. The electronic pen can be used on whiteboards as well as in the air.

A. RECEIVER

Each receiver receives and processes the IR and US pulses emitted by the electronic pen. The IR pulse is used by the receiver as the synchronization signal to measure the TOF, which is then used to calculate the distance from the electronic pen. As shown in Fig. 2, each receiver consists of a microcontroller unit (MCU), a US amplifier with an automatic gain control (AGC) circuit, a silicon microphone (MIC), an IR receiver and two IR photodiodes. The IR photodiodes are connected in parallel to increase the receiving angle of the IR signal. The printed circuit board (PCB) layout and a photo of the receiver module are shown in Fig. 3.

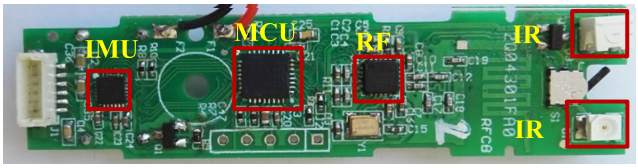
B. ELECTRONIC PEN

The electronic pen generates the US and IR pulses. As shown in Fig. 2, the pen consists of an MCU, an IR driver and an

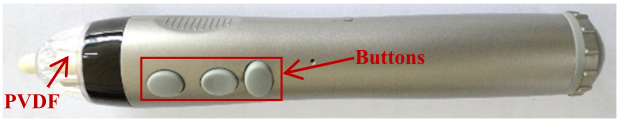


(a) (b)

FIGURE 3. Details of the receiver module. (a) PCB layout. (b) Photo.



(a)



(b)

FIGURE 4. Electronic pen. (a) PCB layout. (b) Photo.

IR transmitter, a US driver, a PVDF US transmitter, and an IMU MPU9250. The PCB layout and a photo of the electronic pen are shown in Fig. 4. In Fig. 4(a), there are a total of four IR transmitters on the top and bottom sides of the PCB. Two transmitters are side-view LEDs, and the other two are top-view LEDs. The purpose of using four IR transmitters is to enlarge the emission angle of the IR pulses.

A cylindrical PVDF US transmitter is used because it is suitable for the structure of the pen and can expand the radiation angle of the US pulse. The IMU MPU9250 IMU provides data from the 3-axis gyroscope, 3-axis accelerometer and 3-axis magnetometer. The radio frequency (RF) circuit transmits these data to the controller, which uses the received data to calculate the attitude of the electronic pen. There are three buttons on the electronic pen: the left mouse button, right mouse button and air writing button. These key functions are encoded by the time interval of the IR pulses emitted by the pen.

C. CONTROLLER

The controller, which includes a 16-bit MCU (dsPIC33EP256MU806), collects the TOF from six receivers and computes the distances between the pen and the six receivers based on the TOF values. Using these distances, the controller calculates the x, y and z coordinates via triangulation. These coordinates define the position of the electronic pen. The controller calculates the attitude of the electronic pen using the IMU data and sends the 3D coordinate data and attitude of the electronic pen to a personal computer (PC).

III. 3D HIGH-PRECISION ULTRASONIC POSITIONING METHOD

The proposed device uses 3D high-precision US positioning technology based on the measured distances to locate the electronic pen in real time. The key factor behind its performance is to achieve a high-precision measurement of the TOF.

The current state-of-the-art acoustic airborne ranging algorithms [25] include simple threshold and cross-correlation techniques. The threshold technique detects the arrival of a US acoustic wave when the receiver voltage exceeds a predetermined threshold value. However, as this method relies on the amplitude of the received signal, it is susceptible to the influences of noise and amplitude variations and has a low measurement accuracy.

The standard TOF estimation technique is the cross-correlation technique [26], which calculates the time delay that maximizes the value of the cross-correlation between the transmitted and received signals. However, cross-correlation performs poorly for the TOF estimation of a single-tone signal because there are several cycles of such a signal that can produce cross-correlation values similar to the maximum cross-correlation value. This technique can improve the accuracy when the signal has a certain bandwidth, such as a linear chirp signal, where this improvement depends on the signal bandwidth. However, in this case, the cross-correlation technique requires a high sampling rate, which necessitates high ADC and DAC circuit configurations, resulting in a high cost.

In this article, a high-precision TOF detection method based on dynamic thresholding and zero-crossing detection is proposed. This method can overcome the shortcomings of the aforementioned TOF detection techniques.

A. HIGH-PRECISION TOF MEASUREMENT METHOD

To reduce the cost and improve the real-time performance, the total amount of data acquired by the device should be low, which means that the sampling frequency should be low. In this section, a high-precision TOF detection method based on dynamic thresholding and zero-crossing detection is proposed, which can obtain a high detection accuracy without a high sampling rate.

The pen emits US and IR pulses simultaneously, which are amplified at each receiver. As the pen and receiver are placed relatively close to each other, the propagation time of the IR pulse can be ignored. Hence, the IR pulses can be used to determine the start time of sampling the US signal, which reduces the complexity of the device. When the receiver detects the IR pulse, the A/D converter starts sampling the US pulses, which are then stored by the MCU in each receiver. A typical US signal, which is sampled and stored in the storage area B_{rcv} by the receiver, is shown in Fig. 5. As the PVDF US transmitter [27] emits a wide bandwidth signal, the received US signal has a short rise time and reaches the maximum value after two cycles. In this

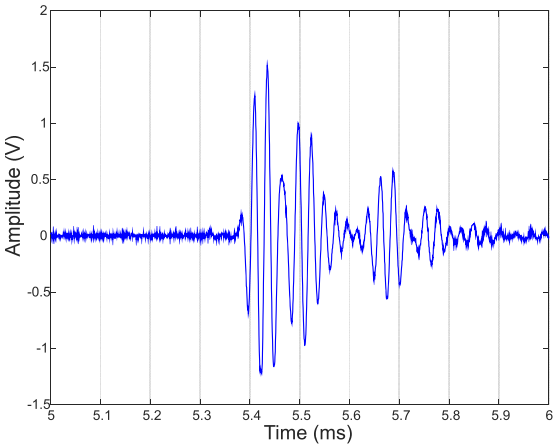


FIGURE 5. Typical US signal received by a receiver.

article, TOF detection is carried out in the time domain based on this feature of the ultrasound signal.

To suppress any interference below 20 kHz and above 60 kHz, the correlation filtering method is adopted to process the received signals. The core concept of this method is the correlation operation between the received signal and a reference signal. If the length of the data over which the correlation is calculated is too large, it will inevitably cause an increase in computational complexity. To reduce this complexity, a monocycle sinusoidal signal is chosen as the reference signal for short-term correlation calculation. This monocycle sinusoidal signal can be expressed as

$$s\,(m) = \sin(2\pi f_0 m T_S), \quad 0 \le m \le M \tag{1}$$

where T_s is the sampling period, f_0 is the center frequency of the received US signal, and $M = T_0/T_s$ (T_0 is the period of received US signal, which is set to an integer multiple of T_s). For example, when the sampling rate is 2 MHz and $M = 50$, the monocycle sinusoidal signal is shown in Fig. 6.

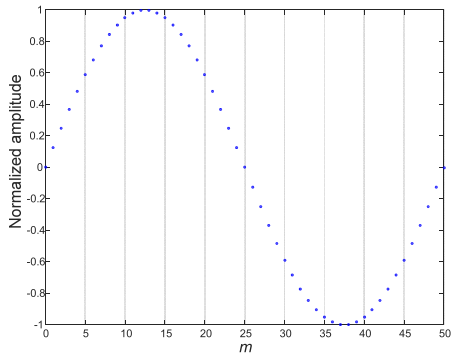


FIGURE 6. Monocycle sinusoidal signal.

The cross-correlation operation is performed between $s(m)$ and the received US signal over M samples. The operation is expressed as follows:

$$y(n) = \frac{1}{M}\sum_{m=0}^{M-1} x(n+m)s(m) \tag{2}$$

where $y(n)$ is the US data stored by the receiver. If the total length of the received US signal is N, $N - M - 1$ single period correlation calculations are needed. As each correlation calculation is carried out over a small number of samples, the computational complexity will be greatly reduced. The results of the correlation calculations are stored back in B_{rcv}. The waveforms of the US signal before and after the correlation filtering operation are shown in Fig. 7.

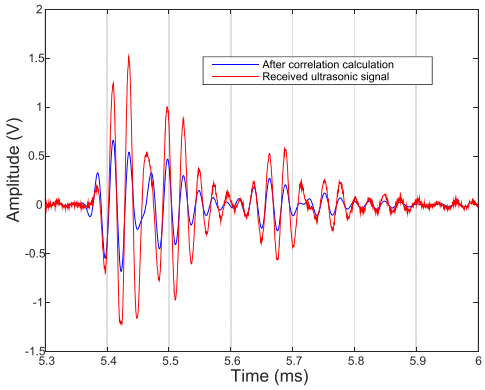


FIGURE 7. US signal before and after the correlation calculation.

Taking the filtered US signal waveform shown in Fig. 7 as an example, the method for detecting the US transit time is described as follows:

Step 1: Search in the filtered data storage area B_{rcv} to find the peak point of each US cycle, represented by '○' in Fig. 8. Store the value and address of each peak point in the new storage areas B_{peak} and B_{rcv}, respectively.

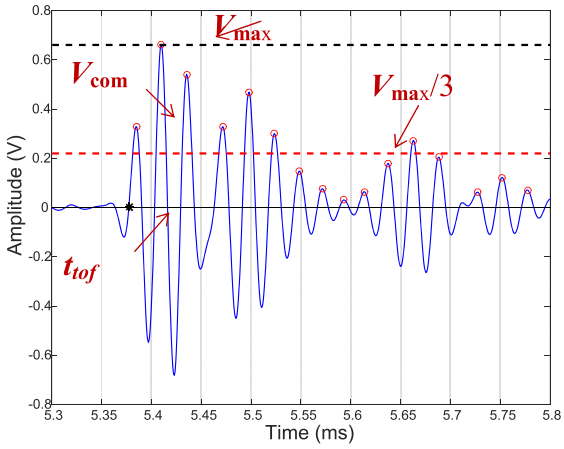


FIGURE 8. TOF detection method based on sampled data.

Step 2: Search the maximum of the peak values stored in B_{peak}, denoted as V_{max} having an address V_{maxad}. This maximum peak value is shown in Fig. 8.

Step 3: Search in B_{peak} to find the number of peak points up to $V_{\max}$, where this number is denoted as N_h. In Fig. 8, $N_h = 2$.

Step 4: Define the threshold value as $V_{th} = V_{max}/(N_h + 1)$, i.e., $V_{th} = V_{max}/3$, and find the first peak point in B_{rcv} that is greater than V_{th}. This peak point is denoted as V_{com} in Fig. 8. In this step, it can be observed that the threshold changes dynamically according to the value of the maximum point in each positioning cycle.

Step 5: In this step, two sample points before and after the first zero crossing are obtained by searching forward from V_{com} in B_{rcv}, which are represented by P_1 and P_2, respectively, in Fig. 9. In the figure, the zero crossing is defined as the intersection of points P_1 and P_2 with the time axis t, represented by P_0. The values V_1 and V_2 and the storage addresses Ad_1 and Ad_2 of the two points are saved. Since the two sample points are adjacent, $Ad_2 = Ad_1 + 1$. In Fig. 9, $t_1 = Ad_1 \times T_s$, $V_1 \geq 0$, $V_2 \leq 0$.

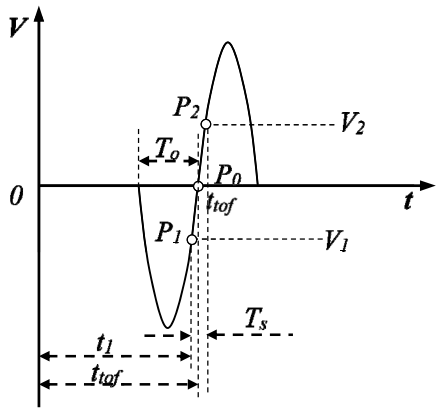


FIGURE 9. TOF calculation based on the zero crossing.

Step 6: The corresponding time location of P_0 is defined as the TOF of the US signal, represented by t_{tof}. By using the values V_1 and V_2 along with the storage addresses Ad_1 and Ad_2, t_{tof} can be obtained by the line fitting method as follows:

$$\frac{t_{tof} - t_1}{T_s} = \frac{|V_1|}{V_2 - V_1} \tag{3}$$

Using $t_1 = Ad_1 \times T_s$ in (3), we obtain the following:

$$t_{tof} = \left(Ad_1 + \frac{|V_1|}{V_2 - V_1}\right) \cdot T_S \tag{4}$$

The TOF of the US signal can be detected based on Steps 1-6. This time is represented by '*' in Fig. 8. It can be seen from (4) that t_{tof} can be determined with a resolution higher than one sampling period T_s.

We assume that the detected TOF can be assumed to be a sinusoidal waveform, whose period is shown in Fig. 9. Its amplitude V is expressed as follows:

$$V = \sin(2\pi f_0 t) \tag{5}$$

The sampling points P_1 and P_2 on this waveform can be represented as $P_1(t_1, 2\pi f_0 t_1)$ and $P_2(t_1 + T_s, 2\pi f_0(t_1 + T_s))$, respectively. The intersection of the line formed by points P_1 and P_2 with the t axis is defined as $P(t_x, 0)$, where t_x can be expressed as

$$t_x = t_1 - \frac{\sin(2\pi f_0 t_1) \cdot T_s}{\cos(2\pi f_0(t_1 + T_s/2)) \cdot \sin(\pi f_0 T_s)} \tag{6}$$

Since the actual zero crossing of the single-period sinusoidal waveform is $P_r(T_0/2, 0)$, the zero-crossing detection error Δ between the calculated and actual zero crossings is as follows:

$$\Delta = \frac{T_0}{2} - t_1 + \frac{\sin(2\pi f_0 t_1) \cdot T_s}{\cos(2\pi f_0(t_1 + T_s/2)) \cdot \sin(\pi f_0 T_s)} \tag{7}$$

It can be seen from (7) that Δ is significantly affected by T_s. A sampling rate of 3-12 times f_0, where $f_0 = 40$ kHz, is selected to analyze this error. The maximum detection error Δ_{max} for sampling frequencies from 120 kHz to 480 kHz is shown in Fig. 10. As the figure illustrates, the error decreases with increasing sampling frequency f_s. When f_s is higher than six times f_0, Δ_{max} is less than 0.1 μs, and the corresponding distance error is less than 0.034 mm. Even in the case where f_s is three times the sampling frequency, Δ_{max} is less than 0.8 μs, and the corresponding distance error is less than 0.27 mm. Therefore, it can be concluded that the TOF detection method proposed in this article has high accuracy.

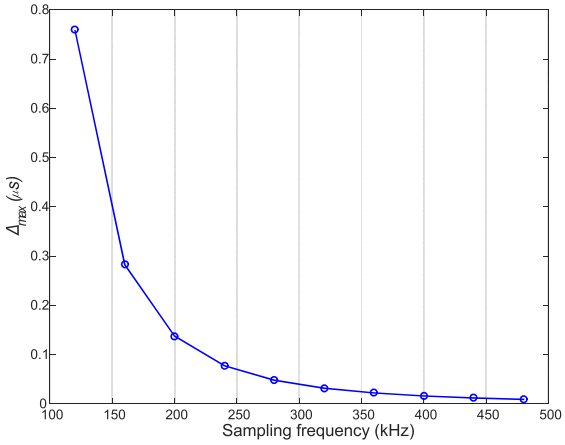


FIGURE 10. Zero-crossing detection error versus the sampling frequency.

Considering the US signal waveform shown in Fig. 5, the simulation accuracy of the zero-crossing detection methods is compared at different signal-to-noise ratio (SNR) values, which are defined as SNR $= P_s/P_n$, where P_s denotes the average power of the US signal in Fig. 5, and P_n denotes the power of the white Gaussian noise added to the US signal shown in Fig. 5 in the simulation. The root mean square error (RMSE) is calculated by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(TOF - TOF_o)}, \tag{8}$$

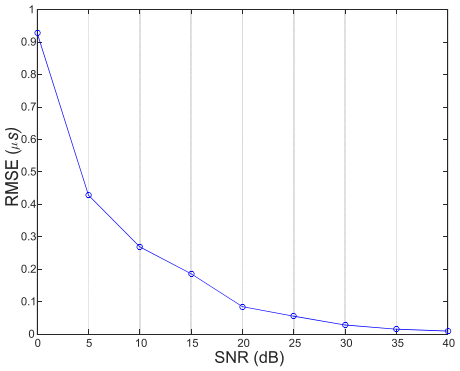


FIGURE 11. RMSE of the proposed TOF detection method versus SNR.

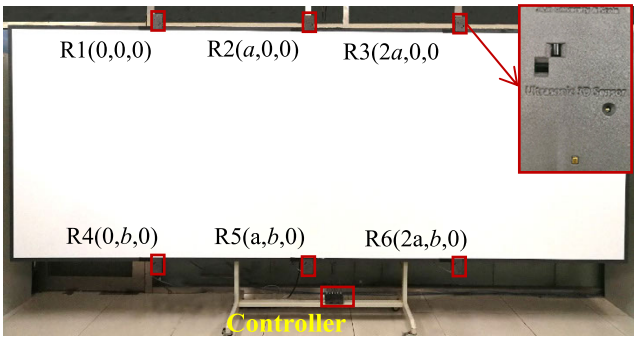


FIGURE 12. Proposed device with six receivers and one controller. The receivers are installed on the top and bottom borders of the whiteboard.

where TOF denotes the TOF value obtained by zero-crossing detection, TOF_0 denotes the true TOF of the US signal displayed in Fig. 5, and N represents the number of repeated detections at each SNR value. In the simulation, N was set to 100, and the sampling frequency was set to 120 kHz. The RMSE results are displayed in Fig. 11. As shown in Fig. 11, the proposed TOF detection method has higher accuracy; even when the SNR is 0 dB, the RMSS is less than 0.93 μs, and the corresponding distance error is less than 0.29 mm.

Based on the aforementioned procedure, receiver i calculates t_{tofi} $(i = 1, \ldots, 6)$ of the received US pulse, which is then used to calculate the distance l_i $(i = 1, \ldots, 6)$ as

$$l_i = c \cdot t_{tofi} \tag{9}$$

where c is the sound velocity.

B. 3D HIGH-PRECISION POSITIONING METHOD

Next, the distances $l_1, l_1, \ldots, l_6$ obtained using (9) are used to calculate the position of the electronic pen. As mentioned earlier, the six receivers R1, R2, …, R6 of the proposed pen-typed input device are installed on the top and bottom borders of the whiteboard, as shown in Fig. 11. As displayed in Fig. 12, the coordinates of the six receivers are set to R1 (0, *0*, 0), R2 (*a*, *0*, 0), R3 (2*a*, *0*, 0), R4 (0, b, 0), R5 (*a*, *b*, 0), and R6 (2*a*, *b*, 0) for the convenience of calculation, and they can be set arbitrarily in practical applications. As a and b increase, the impact of distance measurement error on positioning accuracy decreases [28], that is, the positioning accuracy increases. Therefore, in a given working area, on the premise that positioning signals emitted by the electronic pen can be received by three receivers that are not located on the same horizontal line, the distance between receivers should be longer. The width and height of the whiteboard are 4 m and 1.5 m, respectively. The controller is connected to these receivers via the SPI bus.

Based on the distances $l_1, l_1, \ldots, l_6$, the controller can use triangulation to obtain the x, y and z coordinates of the electronic pen. The distances between the pen and the six receivers can be written in terms of the following six spherical equations:

$$\begin{cases} x^2 + y^2 + z^2 = l_1 \\ (x-a)^2 + y^2 + z^2 = l_2 \\ (x-2a)^2 + y^2 + z^2 = l_3 \\ x^2 + (y-b)^2 + z^2 = l_4 \\ (x-a)^2 + (y-b)^2 + z^2 = l_5 \\ (x-2a)^2 + (y-b)^2 + z^2 = l_6 \end{cases} \tag{10}$$

To calculate the 3D position of the electronic pen based on triangulation, three receivers that are not located on the same horizontal line are selected. This means that the three receivers on the top border, or the three receivers on the bottom border, cannot be selected simultaneously. As only three receivers are used for the calculations, if the distances from all six receivers are available, the total number of receiver combinations available for positioning is $2 \cdot C_3^2 \cdot C_3^1 = 18$, where C_n^k represents the number of k combinations in a set of n elements. Therefore, in this case, a total of 18 position values for the pen can be determined. These positions are given as $(x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_{18}, y_{18}, z_{18})$.

To obtain accurate positioning results and reduce the influence of the direction of the electronic pen, the mean value of these 18 values over each of the three axes is calculated as

$$\begin{cases} x = \dfrac{1}{18}\displaystyle\sum_{n=1}^{18} x_n \\ y = \dfrac{1}{18}\displaystyle\sum_{n=1}^{18} y_n \\ z = \dfrac{1}{18}\displaystyle\sum_{n=1}^{18} z_n \end{cases} \tag{11}$$

The averaging shown in (11) minimizes small errors present in the positioning results. The average calculation in (11) only includes 51 addition instruction operations and 3 division instruction operations. These calculations take only a few microseconds to complete for the MCU (operating at 70 MIPS). In other words, the positioning coordinate averaging method has a high real-time performance.

To improve the robustness of the positioning method, it is necessary to evaluate the correctness of the measured

distance. Due to the influence of the attitude change of the electronic pen and the transmitting beam angle of the US transmitter, all six receivers cannot always successfully obtain the distance to the electronic pen. The success of the distance measurement is based on the maximum speed of hand movement supported by the device. If this maximum speed is v_o, the maximum distance change between two consecutive sampling circles is given by

$$\Delta d_{\max} = T_s \cdot v_o \tag{12}$$

If the difference between the distance calculated in the current sampling period and that calculated in the previous sampled period is less than $\Delta d_{\max}$, the distance measurement is considered successful; otherwise, the measurement is considered a failure. Therefore, when only successfully measured distances are considered, the number of combinations available for calculation of the 3D position of the electronic pen will be less than 18. Thus, by using more than three receivers, the problem of ranging failure caused by the small beam angle of the electronic pen US transmitter can be mitigated, and the stability of positioning results can be improved through averaging.

C. ATTITUDE MEASUREMENT

The microprocessor in the pen converts the data from the IMU MPU9250 into quaternions and periodically sends the data to a controller. The controller calculates the yaw angle ψ, pitch angle θ and roll angle ϕ of the pen based on the quaternions using the following relationship:

$$\begin{pmatrix} \psi_i \\ \theta_i \\ \phi_i \end{pmatrix} = \begin{pmatrix} \arctan \dfrac{2(q_{i0}q_{i1} + q_{i2}q_{i3})}{1 - 2(q_{i1}^2 + q_{i2}^2)} \\ \arcsin(2(q_{i0}q_{i2} - q_{i3}q_{i1}) \\ \arctan \dfrac{2(q_{i0}q_{i3} + q_{i1}q_{i2})}{1 - 2(q_{i2}^2 + q_{i3}^2)} \end{pmatrix} \tag{13}$$

where q_{i0}, q_{i1}, q_{i2}, and q_{i3} are the quaternions read in the ith sampling period and ψ_i, θ_i, and ϕ_i are the corresponding attitude angles.

Using the above method, the 3D x, y, and z coordinates and the attitude angles ψ, θ, and ϕ of the electronic pen can be calculated, and the 6DoF of the electronic pen are realized.

IV. EXPERIMENTAL RESULTS

In this section, we present the details of static and dynamic precision experiments carried out to evaluate the positioning accuracy of the proposed device.

A. EXPERIMENTAL SETUP

The experimental setup used for measuring the positioning accuracy of the proposed device is shown in Fig. 13. The experimental apparatus is constructed using a metal frame 2.2 m high with a horizontal linear slide guide shaft 4.2 m long. This shaft can be adjusted up and down. During the experiment, the electronic pen is placed on the slider, as shown in Fig. 13, and moved horizontally according to the experimental requirements. The adjustable upper and lower

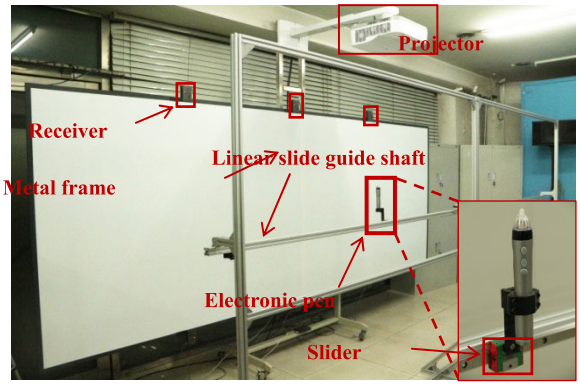


FIGURE 13. Experimental setup.

ranges of the linear slide guide shaft are equal to 2 m, and the slider can move over a range of 4 m. Therefore, by adjusting the height of the linear slide guide shaft and the position of the slider, the position of the electronic pen can be adjusted within the range of 2 m × 4 m. During the experiment, the coordinates a and b in Fig. 12 are set to 1 m and 1.55 mm, respectively.

B. STATIC EXPERIMENTS

In static positioning accuracy experiments, the electronic pen is fixed to the slider and moved to 714 different test positions, as shown in Fig. 14. In the figure, the static test positions are represented by 'o' and are numbered. The lower left corner of the whiteboard is fixed as the origin of the frame, the x-axis points to the right of the whiteboard, the y-axis points to the top of the whiteboard, and the z-axis points perpendicular to the whiteboard. At each test position, the static position of the pen is measured 100 times.

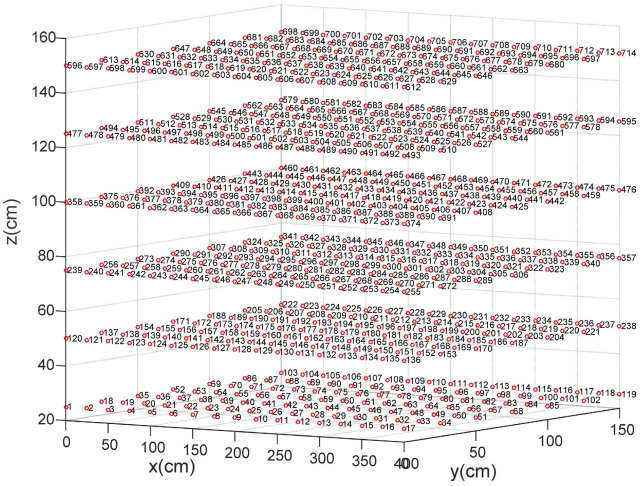


FIGURE 14. Test point settings for the static experiments.

The static positioning error of the proposed device is defined as

$$\sigma_p = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \tag{14}$$

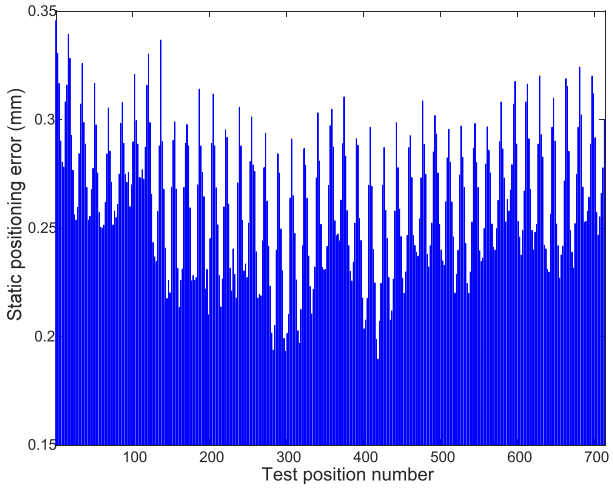


FIGURE 15. Static positioning errors corresponding to all test positions.

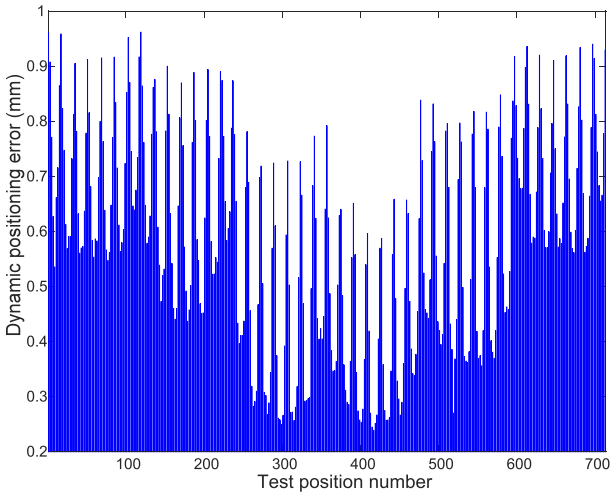


FIGURE 17. Dynamic positioning errors at all test positions.

where σ_x^2, σ_y^2 and σ_z^2 are the deviations of the x, y and z coordinates, respectively. The positioning errors at all test positions are calculated using (14) and shown in Fig. 15. The figure illustrates that the maximum positioning error is less than 0.35 mm, which indicates that the system has a high static positioning accuracy.

C. DYNAMIC EXPERIMENTS

Dynamic positioning accuracy experiments are performed to evaluate the dynamic positioning accuracy. These experiments are carried out by fixing the pen on the slider at each test position shown in Fig. 14 and moving the slider back and forth on the linear guide rail over a distance of 20 cm.

The least squares method is used to fit a straight line in three dimensions through the position coordinates of the electronic pen, and the distance of each coordinate from the straight line is calculated. The maximum distance is defined as the dynamic positioning error (DPE). For example, at test position no. 417, the DPE is 0.2379 mm, as shown in Fig. 16.

The dynamic positioning errors of the electronic pen in all test positions are calculated based on the above method. The calculated results are shown in Fig. 17, from which it can be observed that the DPE is less than 1 mm. Although the DPE is slightly higher than the static positioning error, it is still less than 1 mm, which indicates the high dynamic positioning accuracy of the system.

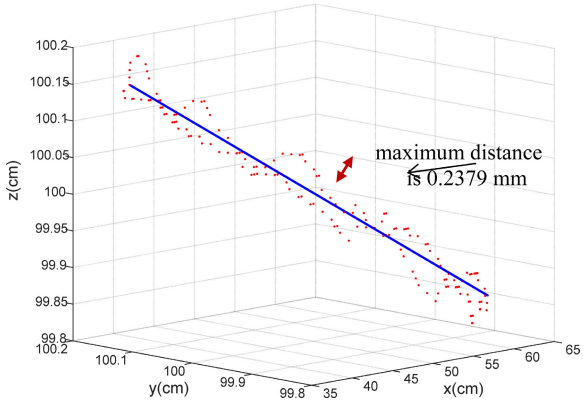


FIGURE 16. Calculation of the dynamic positioning error.

TABLE 1. Characteristics of 6DoF devices.

Device Type	Cost	Working range	Static accuracy	Dynamic accuracy
PATRIOT G4	High	1 m [b]	2 mm RMS	-
Optitrack [a]	Very high	Up to 30 m	Submillimeter	Submillimeter
Vicon [a]	Very high	Up to 22 m	Submillimeter	Submillimeter
zSpace	High	1 m or so	-	-
IS-900	Medium	2.5 m × 2.5 m × 3.0 m	2 mm RMS	17 mm
The proposed device	Low	4 m × 1.5 m × 1.5 m	0.35 mm RMS	1 mm

[a]Needs installation of several reflective balls on the pen, and the pen volume is large.

[b]When the range is more than 1 m, the positioning accuracy decreases significantly.

As Figs. 15 and 17 show, the positioning accuracy at the boundary of the test area is relatively low, especially around the corners, i.e., test positions nos. 1, 17, 103, 119, 596, 612, 696 and 714. The main reason for this low accuracy is that these positions are far from the receiver, and due to the influence of the pen direction, some receivers do not receive the US signals, or the received signals have low SNRs.

To verify the 3D dynamic positioning characteristics of the proposed device, we hold the electronic pen and draw spiral lines freely in the air. These spiral lines, shown in Fig. 18, are continuous and smooth, which indirectly proves that the proposed device has a high dynamic positioning accuracy. This dynamic feature also demonstrates that the device supports writing in the air.

A comparison between the proposed device and existing 6DoF devices in terms of their positioning range, accuracy and cost is shown in Table 1. The table shows that the

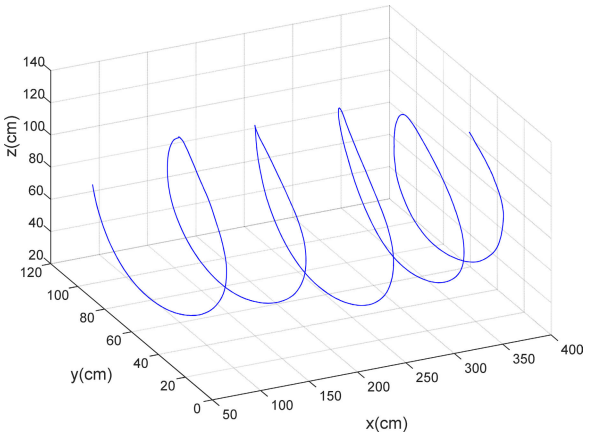


FIGURE 18. Drawing 3D spiral lines using the proposed device.

proposed device has a high positioning accuracy and low cost. The proposed device provides a cost-effective pen-type input device solution for human–computer interaction.

V. CONCLUSION

In this article, a pen-type input device using a high-precision 3D US positioning method was proposed. The device was designed for a large-scale 3D human-computer interaction interface and could achieve millimeter-level tracking accuracy and 6DoF within a working range of 4 m × 1.5 m × 1.5 m. To achieve 6DoF, high-precision 3D positioning of the electronic pen was realized by a high-precision TOF detection method based on dynamic thresholding and zero-crossing detection. In addition, accurate attitude measurements of the electronic pen were performed using an IMU MPU9250 and a multigroup positioning coordinate averaging method. To evaluate the positioning accuracy of the proposed device, static and dynamic positioning accuracy experiments were carried out. The experimental results showed that static and dynamic positioning accuracies higher than 0.35 mm and 1 mm were obtained, respectively. Thus, the proposed pen-type input device could provide a high positioning accuracy and attitude detection, which are two of the common requirements for realizing 6DoF in a 3D human-computer interaction interface.

In this article, we used six receivers to achieve high-precision 3D positioning of the electronic pen. However, in practice, a higher number of receivers can be used to increase the working range. The proposed device has two major limitations. One limitation is that it does not support multiple pens to work at the same time; that is, it does not support multiple access, which will result in multiple pens not working at the same time. The other limitation is that the ultrasonic transmitter is not an omnidirectional transmitter, which requires the device to adopt multiple receivers, which affects the positioning performance of the device due to the arbitrary attitude change of the electronic pen and causes the system positioning accuracy to be uneven. In view of the above limitations, we will focus on multiaccess location technology and small ultrasonic omnidirectional transmitters for 3D pen-type input devices in the future.

REFERENCES

[1] F. Tian, L. Xu, H. Wang, X. Zhang, Y. Liu, V. Setlur, and G. Dai, "Tilt menu: Using the 3D orientation information of pen devices to extend the selection capability of pen-based user interfaces," in *Proc. 26th Annu. CHI Conf. Hum. Factors Comput. Syst.*, 2008, pp. 1371–1380.
[2] X. Bi, T. Moscovich, G. Ramos, R. Balakrishnan, and K. Hinckley, "An exploration of pen rolling for pen-based interaction," in *Proc. 21st Annu. ACM Symp. User Interface Softw. Technol.*, 2008, pp. 191–200.
[3] Y. Xin, X. Bi, and X. Ren, "Acquiring and pointing: An empirical study of pen-tilt-based interaction," in *Proc. Annu. Conf. Hum. Factors Comput. Syst.*, 2011, pp. 849–858.
[4] T. P. Moran, P. Chiu, and W. van Melle, "Pen-based interaction techniques for organizing material on an electronic whiteboard," in *Proc. 10th Annu. ACM Symp. User Interface Softw. Technol.*, 1997, pp. 1–10.
[5] R. Gonzalo, M. Boulos, and R. Balakrishnan, "Pressure widget," in *Proc. Conf. Hum. Factors Comput. Syst.*, Apr. 2004, pp. 487–494.
[6] D. Kim, "Three-dimensional user interaction," U.S. Patent. 12 939 891, May 10, 2012.
[7] J. J. LaViola, Jr., *3D User Interfaces: Theory and Practice*, vol. 6, 2nd ed. Boston, MA, USA: Addison-Wesley, 2017.
[8] P.-C. Wu, R. Wang, K. Kin, C. Twigg, S. Han, M.-H. Yang, and S.-Y. Chien, "DodecaPen: Accurate 6DoF tracking of a passive stylus," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2017, pp. 365–374.
[9] S. Zhai, "User performance in relation to 3D input device design," *ACM SIGGRAPH Comput. Graph.*, vol. 32, no. 4, pp. 50–54, Nov. 1998.
[10] B. Sobota, "Virtual reality technologies as an interface of cognitive communication and information systems," in *Proc. IEEE 2nd Int. Conf. Cognit. Infocommun.*, Budapest, Hungary, Jul. 2011, pp. 1–5.
[11] P. Koutsabasis and V. Panagiotis, "Empirical research in mid-air interaction: A systematic review," *Int. J. Hum.-Comput. Interact.*, vol. 35, no. 18, pp. 1747–1768, 2019.
[12] *Polhemus PATRIOT G4*. Accessed: Apr. 4, 2020. [Online]. Available: https://polhemus.com/motion-tracking/all-trackers/g4
[13] G. Nagymáté and R. M. Kiss, "Application of OptiTrack motion capture systems in human movement analysis," *Recent Innov. Mechtron.*, vol. 5, no. 1, pp. 1–9, Jan. 1970.
[14] L. Mainetti, L. Patrono, and I. Sergi, "A survey on indoor positioning systems," in *Proc. IEEE 22nd Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, Split, Croatia, Sep. 2014, pp. 1–10.
[15] H. Kaufmann and D. Schmalstieg, "Designing immersive virtual reality for geometry education," in *Proc. IEEE Virtual Reality Conf.*, Aug. 2006, pp. 51–58.
[16] *Optitrack and Vicon*. Accessed: Apr. 4, 2020. [Online]. Available: https://www.optitrack.com/
[17] D. Wormell and E. Foxlin, "Advancements in 3D interactive devices for virtual environments," in *Proc. Workshop Virtual Environ.*, 2003, pp. 47–56.
[18] S. J. Gilson, A. W. Fitzgibbon, and A. Glennerster, "Quantitative analysis of accuracy of an inertial/acoustic 6DOF tracking system in motion," *J. Neurosci. Methods*, vol. 154, nos. 1–2, pp. 175–182, Jun. 2006.
[19] *Tracking, Lighting, And How It All Works*. Accessed: Apr. 6, 2020. [Online]. Available: https://www.optitrack.com/
[20] D. A. Chavez, M. Paranjpe, and J. C. Tu, "Three-dimensional tracking of a user control device in a vol. using, orientation sensing," U.S. Patent 8 970 625. Jan. 28, 2016.
[21] M. Omer and G. Y. Tian, "Indoor distance estimation for passive UHF RFID tag based on RSSI and RCS," *Measurement*, vol. 127, pp. 425–430, Oct. 2018.
[22] Y. Zhang, X. Tan, and C. Zhao, "UWB/INS integrated pedestrian positioning for robust indoor environments," *IEEE Sensors J.*, early access, Jun. 1, 2020, doi: 10.1109/JSEN.2020.2998815.
[23] W. Guan, S. Chen, S. Wen, Z. Tan, H. Song, and W. Hou, "High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning," *IEEE Photon. J.*, vol. 12, no. 2, pp. 1–16, Apr. 2020.
[24] M. Toda, "Cylindrical PVDF film transmitters and receivers for air ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 49, no. 5, pp. 626–634, May 2002.

[25] J. C. Jackson, R. Summan, G. I. Dobie, S. M. Whiteley, S. G. Pierce, and G. Hayward, "Time-of-flight measurement techniques for airborne ultrasonic ranging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 60, no. 2, pp. 343–355, Feb. 2013.
[26] M. O. Khyam, L. Xinde, S. S. Ge, and M. R. Pickering, "Multiple access chirp-based ultrasonic positioning," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3126–3137, Dec. 2017.
[27] J. Chen, J. Zhao, L. Lin, and X. Sun, "Truncated conical PVDF film transducer for air ultrasound," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8618–8625, Oct. 2019.
[28] X. Yan, Q. Wu, X. Wang, and X. Sun, "Semicool temperature compensation algorithm based on the double exponential model in the ultrasonic positioning system," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 995–1010, Apr. 2020.

JIAN CHEN received the B.S. degree in electrical and information engineering, the M.S. degree in signal and information processing, and the Ph.D. degree in communication and information systems from Jilin University, Changchun, China, in 2001, 2004, and 2007, respectively. From 2007 to 2014, he was a Lecturer with the College of Communication Engineering, Jilin University. From 2014 to 2015, he was a Visiting Scholar with the School of Electronic and Electrical Engineering, Institute of Integrated Information Systems (I3S), University of Leeds, Leeds, U.K. He is currently an Associate Professor with the College of Communication Engineering, Jilin University. His current research interests include ultrasonic positioning systems, array signal processing, and sensor applications.

FAN YU received the B.S. degree in communication engineering from Jilin University, Changchun, China, in 2018, where he is currently pursuing the M.S. degree in electronics and communication engineering. His research interests include ultrasonic positioning systems and sensor applications.

JIAXIN YU (Graduate Student Member, IEEE) received the B.S. degree in communication engineering from Jilin University, Changchun, China, in 2016, where he is currently pursuing the Ph.D. degree in communication and information systems. His research interests include 3D human–computer interaction and multimodal interaction with AR/VR environments.

LIN LIN received the B.S. degree in measurement and control technology and instrumentation, the M.S. degree in communication and information systems, and the Ph.D. degree in communication and information systems from Jilin University, Changchun, China, in 2001, 2004, and 2007, respectively. From 2007 to 2019, she worked as a Lecturer at the College of Communication Engineering, Jilin University. From 2014 to 2015, she was a Visiting Scholar with the Department of Computer Science, The University of Sheffield, Sheffield, U.K. She is currently an Associate Professor with the College of Communication Engineering, Jilin University. Her current research interests include ultrasonic positioning systems, array signal processing, and digital signal processing applications.