



Article Visual Perception and Multimodal Control: A Novel Approach to Designing an Intelligent Badminton Serving Device

Fulai Jiang^{1,†}, Yuxuan Lin^{1,†}, Rui Ming^{1,*}, Chuan Qin¹, Yangjie Wu¹, Yuhui Liu² and Haibo Luo¹

- ¹ Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, School of Computer and Big Data, Minjiang University, Fuzhou 350108, China; jiangfulai@stu.mju.edu.cn (F.J.); linyuxuan@stu.mju.edu.cn (Y.L.); qinchuan@stu.mju.edu.cn (C.Q.); wuyangjie@stu.mju.edu.cn (Y.W.); robhappy@mju.edu.cn (H.L.)
- ² School of Fuzhou Melbourne Polytechnic (FMP), Fuzhou 350108, China; lakas8080@gmail.com
- * Correspondence: rming@mju.edu.cn
- [†] These authors contributed equally to this work.

Abstract: Addressing the current issue of limited control methods for badminton serving devices, this paper proposes a vision-based multimodal control system and method for badminton serving. The system integrates computer vision recognition technology with traditional control methods for badminton serving devices. By installing vision capture devices on the serving device, the system identifies various human body postures. Based on the content of posture information, corresponding control signals are sent to adjust parameters such as launch angle and speed, enabling multiple modes of serving. Firstly, the hardware design for the badminton serving device is presented, including the design of the actuator module through 3D modeling. Simultaneously, an embedded development board circuit is designed to meet the requirements of multimodal control. Secondly, in the aspect of visual perception for human body recognition, an improved BlazePose candidate region posture recognition algorithm is proposed based on existing posture recognition algorithms. Furthermore, mappings between posture information and hand information are established to facilitate parameter conversion for the serving device under different postures. Finally, extensive experiments validate the feasibility and stability of the developed system and method.

Keywords: visual perception; badminton serving device; embedded control; posture recognition; multimodal control

1. Introduction

With the enhancement of individuals' material living standards, an increasing number of people are directing their attention toward the holistic development of physical and mental well-being [1]. Badminton, being an exceptionally dynamic and competitive sport, demands athletes to possess advanced serving skills, and the capability to serve with precision and variety is pivotal for securing an advantage on the court [2,3]. In conventional training methods, coaches are required to repeatedly demonstrate various serve techniques over an extended period. As time progresses, the physical fatigue experienced by training coaches may result in the deformation of technical movements. This, in turn, contributes to a decline in serving accuracy and a gradual increase in ineffective feeds, thereby diminishing the overall training effectiveness [4].

The introduction of badminton serving devices offers an effective solution to address the challenges associated with unstable play and insufficient technical proficiency among training coaches. This, in turn, mitigates the limited effectiveness of training and brings about a notable enhancement in training methodologies [5]. The use of badminton serving devices can provide a controlled and adaptable practice environment for athletes, enabling them to fine-tune their serving technique and improve their overall level.



Citation: Jiang, F.; Lin, Y.; Ming, R.; Qin, C.; Wu, Y.; Liu, Y.; Luo, H. Visual Perception and Multimodal Control: A Novel Approach to Designing an Intelligent Badminton Serving Device. *Machines* 2024, *12*, 331. https:// doi.org/10.3390/machines12050331

Academic Editors: Wen-Jer Chang and Cheung Chieh Ku

Received: 9 April 2024 Revised: 25 April 2024 Accepted: 9 May 2024 Published: 13 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). The serving method, serving as the focal point in the design of badminton serving devices, has been a predominant research focus for scholars globally. Currently, two primary mainstream serving modes exist. The first is founded on the fixed-position badminton serving method, wherein trainers modify the hardware structure and placement position of the serving device to alter the serving angle and achieve varying serve distances [6]. This method boasts a relatively uncomplicated mechanical mechanism, facilitating straightforward daily maintenance. However, its drawbacks are conspicuous: each transition to

singular serve mode with lower precision. The second approach involves the serving method based on an embedded control system, wherein an embedded control system is integrated into the traditional serving device. Trainers can adjust serving parameters in response to training requirements, and the serving device, in turn, modifies its hardware structure based on these parameters, enabling fully automated serving [7,8]. Prior literature introduces a Badminton Shuttlecock Feeding Machine that employs trajectory simulation to derive initial parameters, saving them within the Feeding Machine for the automatic launch of four distinct types of initial balls [9]. Prior literature [10] proposes a design of a high-speed, lightweight humanoid badminton robot. Its structure integrates a pneumatic actuator and non-interference multidegree-of-freedom joint to achieve high-precision motion control. In contrast to the first method, this approach utilizes an embedded system to govern the mechanical structure of the ball-launching device, significantly enhancing launch accuracy. The method can also store different parameters of the ball launching patterns, providing trainers with a variety of launching methods. However, a drawback lies in the need for manual adjustment of the embedded system when altering the serve mode, the lack of automatic remote adjustment, and the need for further enhancement in terms of intelligence.

a different serve mode necessitates manual adjustments to the structure, resulting in a

In recent years, with the advancing capabilities of computer vision technology, scholars have progressively integrated computer vision into badminton serving devices [11]. Prior literature [12] explores a badminton serving robot that employs visual recognition technology to identify badminton balls released by the ball feeding mechanism. The robot is equipped with a badminton racket attached to its arm, allowing it to strike the balls with the racket to perform the serving action. Another work [13] introduces a badminton-hitting robot featuring a distance image sensor. This robot detects the flight trajectory of the badminton ball through the sensor, predicts the landing point based on the distance image, and adjusts its position accordingly to strike the ball back with the racket. Additionally, prior literature [14] presents a badminton robot that captures and analyzes athletes' batting videos using a camera on the serving device, thereby enhancing the serving device's intelligence. However, despite these advancements in integrating computer vision technology with badminton serving devices, the method of serving the badminton serving device has not been modified, and the type of serving is still changed by manually adjusting the embedded system.

To address the challenges encountered by existing badminton serving devices, such as the necessity for manual adjustment of the embedded system's serving mode and the lack of automatic adaptation to the player's state, this paper proposes a design method for a badminton serving device based on visual perception and multimodal control. This method involves acquiring the player's posture image through the posture recognition module installed on the badminton ball serving device. The collected image undergoes posture recognition, and the signal control module is then manipulated to adjust the serving device's angle, speed, and serve count based on the recognition results. Alternatively, the angle, speed, and serve count can be modified using the self-developed upper computer module that governs the signal control module. Consequently, this method empowers athletes to practice various strokes within the hitting zone.

In this paper, computer vision's posture recognition technology is seamlessly integrated with the badminton serving device, effectively enhancing the automation and intelligence of the existing system. In comparison to prior research, this paper distinguishes itself in two crucial aspects:

- (1) This paper pioneers the utilization of human posture information as the primary control signal for a badminton serve device. Throughout its usage, the serve device dynamically adjusts the equipment's height, speed, and angle based on the user's distinct posture signals, facilitating a non-contact and automated service mode. This innovative approach not only enhances user experience but also streamlines the process of delivering services, promising significant advancements in the realm of badminton training and gameplay.
- (2) This paper introduces an innovative posture detection process. In contrast to the benchmark detection process, the key point information identified in the image serves as feedback for the subsequent frame's key point detection process. This approach reduces redundant posture mapping, thereby enhancing posture recognition speed.

The structure of the remaining sections in the paper is as follows: Section 2 describes the system design of the badminton serving device. Section 3 describes the overall hardware design of badminton serving device. Section 4 describes the vision based human posture recognition method. Section 5 describes the real serve test conducted to verify the accuracy and reliability of the serve device. Section 6 provides concluding remarks.

2. System Design

The system design and operation flow of the badminton serving device, based on visual perception and multimodal control as proposed in the paper, is depicted in Figure 1. The system comprises several key components, including the upper computer module, the posture recognition module, the signal control module, and the execution module. The upper computer module consists of both software and hardware components. The software component is a self-developed system responsible for selecting the posture recognition type and transmitting service signals to the signal control module. The posture recognition module is composed of a vision module and a microprocessor module. The vision module captures images of the human body posture, while the microprocessor module executes the posture recognition method to identify the posture, subsequently outputting the corresponding signal to the signal control module. The signal control module, in turn, receives signals from both the upper computer module and the posture recognition module. Its primary function is to direct the execution module in adjusting the mechanical structure of the ball-serving device. The execution module comprises a launch structure, a ball-plucking structure, and an angle adjustment structure. Upon receiving a control signal from the signal control module, the execution module dynamically adjusts each of these structures, culminating in the launch of the badminton ball.

When the serving device is in use, it can be controlled either by the posture recognition module based on the user's posture or directly by the upper computer module. In the posture recognition mode, upon posture detection selection, the user's posture image is initially transmitted from the vision module to the microprocessor module. Subsequently, the microprocessor module conveys the recognized posture signals to the signal control module. Ultimately, control signals are dispatched through the signal control module to regulate the mechanical structures of the execution module. Alternatively, in the event the user opts for the upper computer control mode, the initial step involves initiating the upper computer system. Following this, a communication connection must be established with the signal control module through the system interface. Subsequently, the user is required to configure the relevant parameters of the ball-launching device. Upon completing this setup, the badminton launching information is transmitted to the signal control module, allowing the badminton serving device to launch badminton at varying angles and speeds.



Figure 1. Badminton serves device system design and operation flowcharts.

3. Overall Hardware Design

3.1. Mechanical Design of Actuator Modules

The design of the mechanical structure for the executive module of the badminton serving device is illustrated in Figure 2. The module comprises key components: a ball storage structure, a ball plucking structure, a launch structure, an angle adjustment structure, and a support structure. The ball storage structure is tasked with housing the badminton balls and is composed of a cylindrical storage container. The ball-plucking structure extracts badminton balls from the storage cylinder to the ball rest using a configuration primarily comprising a pair of rubber paddles, DC motors, and gears. The launch structure is designed to propel the badminton balls, primarily employing friction wheels, DC motors, and protective shells.

The angle adjustment structure facilitates the adjustment of the badminton serving device's angle in both horizontal and pitching directions. It comprises the horizontal rotation structure and the pitch adjustment structure.

The horizontal rotation structure comprises a bearing, a baseplate, a rotary plate, and a stepper motor; among them, the four columns on the turntable form an integral structure with the turntable, sharing the weight of the launch platform. The pitch adjustment structure includes a gear strip, a stepper motor, and a putter with a pulley. The support structure is tasked with providing support to the aforementioned four structures and is composed of a tripod.



Figure 2. Structure of the actuator module of the teeing device.

The mechanical structure of the badminton serving devices operates on the principle that the launch angle requires adjustment before ball release. Horizontal angle adjustments are accomplished by a stepper motor that drives the rotary table to rotate horizontally. Furthermore, pitch angle adjustment is facilitated by an additional stepper motor driving the gear strip. This enables the push putter to move back and forth, thereby adjusting the pitch angle of the launch platform. Following the adjustment of the launch angle, it is essential to refine the initial speed of badminton release. This is achieved by modifying the rotation speed of the motor within the launch structure. Subsequently, the rubber paddle is propelled by the rotation of the paddle motor, extracting the badminton ball from the ball storage structure and placing it in the ball holder. Finally, the friction wheel within the launch structure propels the badminton balls.

3.2. Embedded Development Board Circuit Design

To address the functional requirements of the signal control module, this paper designs an embedded development board, the circuitry of which is illustrated in Figure 3. The development board employs the STM32F103C8T6 chip as the main control chip and incorporates A4988 and A4950 chips as motor driver chips.



Figure 3. Embedded development board circuit design diagram.

To accommodate the varied power supply voltages of the motor driver chip and the main control chip, a step-down circuit located on the left side of the development board. The circuit employs the MP4462DN chip along with a low dropout linear regulator. This combination, coupled with a multilayer ceramic capacitor and a low electromagnetic interference capacitor, effectively divides the input voltage to supply 5 V and 3.3 V outputs.

The embedded development board shown in Figure 3 serves as the signal processing module in Figure 1, which is connected to the attitude recognition module through the serial port one, and then connected to the host computer module through the WIFI module or the Bluetooth module, realizing the interaction between the attitude information and the control information. Meanwhile, the pivoting structure, launching structure, and rotating structure of the actuator module in Figure 3 can be adjusted through the interface motor interface (stepper motor interface and DC motor interface) on it.

Upon establishing a wireless connection (WiFi or Bluetooth) between the upper computer module and the development board or when the posture recognition module interfaces through the serial port 1 interface (U12), the master control chip receives signals. The signals can originate either from the upper computer module through serial ports 3 or 2, alternatively, from the posture recognition module through serial port 1. After receiving the signals, the main control chip controls the operation of the two stepper motors (STEPING MOTOR1 and STEPING MOTOR2) in the angle adjustment module, utilizing timer TIM2 channels 1 and 2. Simultaneously, it regulates the two DC motors (U8 and U7) that are associated with the toggle and launching mechanisms using timer TIM3 channels 1 and 2. Ultimately, this leads to the execution of the serve.

4. Posture Recognition Methods

4.1. Posture Detection Principle

Detecting body postures poses a formidable challenge owing to the intricate nature of the human form [15]. In contrast to rigid objects, the human body consists of numerous joints and displays a wide range of degrees of freedom in its limbs [16]. Moreover, human postures exhibit high variability, and human limbs are particularly prone to occlusion and self-occlusion [17].

BlazePose [18] utilizes a detector-tracker setup to extract key point information about human body poses. The detector-tracker is composed of a body posture detector and a posture tracker; when there is an image input, the tracker predicts keypoint coordinates, and when the tracker indicates that there is no human present, re-run the detector network on the next frame. This method effectively improves the accuracy of the recognition of human body poses, and it is currently one of the most widely used methods [19,20]. Although this method can accurately identify information about the human body's pose, the recognition process is run repeatedly for the same posture, resulting in a large computational burden and making it difficult to deploy on embedded computers. Therefore, in this paper, based on the original algorithm, We introduce a posture information comparison process and propose an improved Blazepose algorithm.

As illustrated in Figure 4, the flowchart depicts the improved BlazePose algorithm. The process initiates with the user choosing between body detection or gesture detection. Upon inputting the first frame, it undergoes processing through the target detection model (palm detector or face detector). If target features, such as a face or palm, are present, a candidate region for the target location is generated. Then, keypoint detection is performed, where posture key points are detected by running a keypoint detection model (hand landmark model or pose landmark model) on the candidate region. After successfully detecting posture key points, the corresponding key point information is obtained. Ultimately, action signals are outputted following the matching of posture mapping information.



Figure 4. The improved BlazePose algorithm flowchart.

For the second frame input, the target detection process is skipped. Instead, the candidate region for the target position from the previous frame is extended to facilitate key point detection. If the extended candidate region fails to detect the target, the target detection model is reactivated. After obtaining key point detection information for the second frame, a pairwise comparison is initiated with the key point information from the previous frame. If the deviation in the comparison falls below a predefined threshold, it indicates a duplication of the posture action from the previous frame. As a result, the system directly outputs the same posture action as the previous frame without the need for matching the posture mapping information. Alternatively, if the comparison deviation and outputs the corresponding action signal. It is worth noting that in this paper, the target detection model and keypoint detection model are imported through the mediapipe library on Python(The version of Python is 3.7.3), and since this paper mainly addresses the problem of repeated recognition of the same posture, therefore, in this paper, the neural network is not trained and its parameters are not modified.

4.2. Attitude Mapping Creation

The key point information acquired through the aforementioned image key point detection process comprises the two-dimensional coordinates of each key point. Following the processing of the two-dimensional coordinates, they are compared with a pre-set posture action. If the comparison yields a match, the system outputs the action signal corresponding to the key point mapping.

Consider the human body posture of the "cross hand" action in Figure 5 as an example. The posture can be recognized when the keypoint information satisfies the following four conditions or when it satisfies conditions (3) and (4) of higher priority: (1) dx < threshold θ_2 ; (2) dy < threshold θ_3 ; (3) $x_{20} > x_{19}$; (4) $x_{13} > x_{14}$. In this context, the posture is identified as a "cross hand". Here, dx and dy are given by:

$$dx = x_{15} - x_{16} \tag{1}$$

$$dy = y_{15} - y_{16}$$
(2)



Figure 5. Coordinates the key points of the "cross-handed" posture.

In the gesture detection segment, upon acquiring the coordinate information of the hand's key points, the system calculates the distance between the key points and the joint curvature of a single finger. Subsequently, customized semantic judgment is applied to recognize the gesture, achieving system-wide gesture recognition.

The thumb joint nodes 1-2-3-4 are depicted in Figure 6. With the known coordinates of key points 2, 3, and 4, the angle α_1 between and is determined using spatial distance, vector dot product, and the inverse trigonometric function. The formulas for these calculations are presented in Equations (3)–(5).

$$L = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$
(3)

$$dot_product = (x_3 - x_2) \times (x_4 - x_3) + (y_3 - y_2) \times (y_4 - y_3)$$
(4)

$$\alpha_1 = \arccos(\frac{dot_product}{L_{12} \times L_{23}})$$
(5)



Figure 6. Coordinate index diagram of 'four' attitude key points.

If α_1 exceeds the threshold value, the joint is categorized as "straight"; conversely, if α_1 is below the threshold value, the joint is labeled as "bent". Applying the same methodology, the angle α_2 for the key nodes 1-2-3 can be calculated. If α_2 is less than the threshold value, it denotes the bending of joint 2. Additionally, if all other joints are "straight" simultaneously, the recognized gesture action is identified as "gesture 4".

Practical scenarios involving the badminton serving device frequently require the execution of actions like continuous ball serving, serving a near netball, serving a midcourt ball, serving a high long ball, and others. In alignment with the previously outlined principles, the design includes control interaction instructions illustrated in Figures 7 and 8. Figure 7 shows the body posture, which includes four actions, namely raising the right hand (a), the left hand (b), raising both hands (c) and crossing the hands (d), and Figure 8 shows the gesture actions, which we defined a total of nine groups of actions, in addition to the most common 1, 2, 3, 4, 5, 6 actions, adding the thumb action (g), then the gun action (h), and then the heart action (i). These different motions represent different control commands; for example, for the high ball command, there is the raising both hands (c) and the gun action (h), etc.



Figure 7. Definition map of body posture interaction commands (**a**) B1: Speed up the serve speed; (**b**) B2: Slow down the serve speed; (**c**) B3: Launch the high long ball; (**d**) B4: Stop serving.



Figure 8. Gesture interaction command definition diagram (**a**) H1: Turn left 15 degrees; (**b**) H2: Turn right 15 degrees; (**c**) H3: Serve a single ball; (**d**) H4: Tilted down 5 degrees; (**e**) H5: Tilted up 5 degrees; (**f**) H6: Launch the near netball; (**g**) H7: Launch the mid-court ball; (**h**) H8: Launch the high long ball; (**i**) H9: Continuous launch badminton.

4.3. Posture Recognition Accuracy Evaluation

To validate the effectiveness and accuracy of the posture recognition method in this paper, recognition tests were conducted on the defined action commands. Ten experimenters in a badminton court participated in a recognition test for body posture instructions, with each experimenter performing single posture recognition 30 times. To simulate stadium usage, 15 instances involved detecting partial masking of the target, resulting in a total of 300 tests for each posture type and 2900 samples tested overall.

The confusion matrix, based on the test results, is presented in Figure 9. In the matrix, "TIME OUT" denotes results not recognized for more than 2 s. The labels B1–B4 and H1–H9 represent control interaction instructions. The proposed posture recognition method successfully identifies B1, B4, H2, H5, H6, and H8 300 times, showcasing commendable recognition accuracy. However, challenges arise with misrecognition and recognition timeouts during the identification of B2 and B3.



Figure 9. Chaos Matrix Diagram (**a**) Confusion matrix of gesture detection results; (**b**) Confusion matrix of experimental results for body posture detection.

To further assess the method's strengths and weaknesses, three metrics—precision rate of detection (Pr), recall rate (Re), and accuracy rate (Ac) are applied to the statistical results of the confusion matrix, providing a more standardized measure [21].

Pr denotes the precision rate of recognition, i.e., the ratio of the number of correct recognition to the total number of recognition in each type of posture action recognition [22].

$$Pr = \frac{TP}{TP + FP} \tag{6}$$

TP denotes the number of times the posture action was correctly recognized in this posture action recognition; *FP* denotes the number of times the posture action was recognized as the posture action in other posture action recognition.

Re denotes the recall of recognition, i.e., the ratio of the number of correctly recognized posture actions to the number of times the posture action is recognized.

$$Re = \frac{TP}{TP + FN} \tag{7}$$

FN denotes the number of times other posture actions are recognized in that posture action recognition.

Ac denotes the accuracy of recognition, i.e., the total number of all actions correctly recognized as a percentage of the total number of tests.

$$Ac = \frac{TP + TN}{TP + TN + FP + FN}$$
(8)

TN denotes the number of times a posture action is correctly recognized among other posture action recognition.

The accuracy indices for posture action detection are presented in Table 1. From the table, it can be seen that: the average recognition precision *Pr* of the nine gesture actions is 97.68%, the recall *Re* is 98.66%, and the accuracy *Ac* is 99.59%; the average recognition precision *Pr* of the four body postures is 98.01%, the recall *Re* is 98.34%, and the accuracy *Ac* is 99.08%. In summary, the values of the above three performance evaluation indexes are all above 97%, thus indicating that the method has a good recognition effect. However, when the posture actions are H1, H5, H8 and B4, the accuracy of the above posture actions is lower compared to other actions, and from the aspect of recall, the recall of posture actions H3, H7 and B3 is low, The analysis is due to the detection model in the background environment is complex, the human body or palm segmentation is not accurate and similar action between the discrimination is inaccurate; in addition, the current recognition methods and vision module on the acquisition of action information frame rate is not high and caused by the recognition process of some of the semantic information is lost, resulting in the recognition of inaccurate or recognition of overtime problems, and ultimately will affect the above evaluation indexes.

Posture Action	ТР	FP	TN	FN	Pr%	Re%	Ac%
H1	295	13	2387	5	95.78	98.33	99.33
H2	300	4	2396	0	98.68	100.00	99.85
H3	285	1	2399	8	99.65	97.27	99.67
H4	294	2	2398	2	99.32	99.32	99.85
H5	300	21	2379	0	93.46	100.00	99.22
H6	300	3	2397	0	99.01	100.00	99.89
H7	283	13	2398	17	99.30	94.33	99.30
H8	300	13	2387	0	95.85	100.00	99.52
H9	298	0	2400	0	100.00	100.00	100.00
Haverage	295	7	2393	4	97.68	98.66	99.59
B1	300	0	900	0	100.00	100.00	100.00
B2	293	6	894	7	97.99	97.67	98.92
B3	289	4	896	11	98.63	96.33	98.75
B4	300	15	885	0	95.24	100.00	98.75
Baverage	296	6	894	5	98.01	98.34	99.08

Table 1. Accurate metrics for human posture detection.

5. Experiment

5.1. Experimental Materials and Methods

Prior to conducting experiments with the proposed badminton serving device, it is crucial to complete the module and hardware selection, assemble the mechanical structure, and arrange the experimental site. Table 2 provides a detailed account of the names and parameters of the selected equipment. Noteworthy, the microprocessor module employs the Raspberry Pi 4B; its CPU contains four cores, each with a primary frequency of 1.5 GHz, which allows Raspberry Pi 4B to complete the task of image attitude recognition faster. The vision module employs the Raspberry Pi Camera Module V2; It can capture the image with a maximum resolution of 3280×2464 pixels, meeting the pixel requirements of attitude images in the process of attitude recognition. The signal control module incorporates the embedded development board developed; the development board not only has the motor interface in the execution module but also integrates Bluetooth and WiFi modules to realize wireless communication. In addition, the 1.8 Degree Step Angle of the stepper motor, with high precision characteristics, is suitable for accurate Angle control of the serve.

Table 2. Equipment information sheet.

Part Name	Optional Equipment Name	Specification		
Stepper Motor	42 Stepper Motors	1.8 Degree Step Angle Chip: STM32F103C8T6, A4988 and A4950 Wireless Communication Module: ESP8266 and HC-08		
Signal Control Module	Self-Designed Embedded Development Board			
Microprocessor Module	Raspberry Pi 4B	CPU: Cortex A72 architecture 64-bit 1.5 HGz quad-core		
DC Motor	DC Geared Motor with Encoder	Idling speed 330 rpm Maximum torque 3.1		
DC Battery	2200 Mah Battery Pack	12 V, 3 A		
Vision Module	Raspberry Pi Camera Module v2	OV5647 Sensor SCI Interface Capture up to 3280×2464 Images		

Once the badminton serving device and hardware devices were identified, the overall mechanical structure of the Badminton serving device was assembled using the previously selected modules and hardware. The assembled device is depicted in Figure 10, where the vision module (1) is positioned directly above the device, the signal control module (2) is situated on the horizontal turntable, and the microprocessor module (4) is located at the rear.



Figure 10. Diagram of the location of each piece of equipment in the badminton serving device (**a**) Front diagrams; (**b**) Side diagrams; (**c**) Back diagrams; (**d**) Inside diagrams (Note: (1) is the vision module, (2) the signal control module, (3) the DC battery, and (4) the microprocessor module).

As presented in Table 3, the detailed size parameters of the badminton serve device are provided. The height of the badminton serve device ranges from 179 cm to 219 cm, with a maximum height of 219 cm achieved through the extension of the tripod. Additionally, it's important to highlight that the maximum capacity of 45 balls corresponds to the combined maximum capacity of two assembled ball barrels.

Table 3. Equipment parameters of badminton serving device.

Part Name	Parameters
Height of serving device	179–219 cm
The maximum number of storage balls	45
Width of serving device	24 cm
Size of the storage tank	R: 8 cm; H: 44 cm
Size of friction wheel	R: 4 cm; H: 1.8 cm

The accuracy of the ball served by the badminton serving device proposed in this paper, in accordance with the user's posture, is a crucial evaluation criterion. Consequently, the testing protocol involves the evaluation of three types of balls launched by the ball serving device: the near netball, the mid-court ball, and the high long ball. The experimental environment shown in Figure 11a, where the badminton court is an indoor standard badminton court, was used for testing. Three specified drop zones on the badminton court were employed to document the landing positions of the three aforementioned ball types. These drop zones, ranging from close to the net to away from the net, served as a testing ground for the near netball, the mid-court ball, and the high long ball, respectively. Each drop zone comprises two circles, denoted BC (blue) and RC (red), sharing the same center and position. The radius of BC is 120 cm, while the radius of RC is 80 cm.



Figure 11. Experimental diagram (a) Map of the lab site; (b) Real experimental diagram.

At the outset, we utilized the upper computer module to associate the following three types of balls with specific postures: (1) executing the B3 "hands above the head" action representing instructions for the high long ball, (2) performing the H7 "thumbs up" gesture representing instructions for the mid-court ball, and (3) executing the H6 "six" gesture representing instructions for the near netball.

The experimenter proceeds to conduct the experiment after defining the serving postures. As depicted in Figure 11b, when the experimenter performs the corresponding postures directly in front of the badminton serving device, the device serves the ball based on the associated serving type. Afterward, the recording personnel documented the details of Badminton landing positions. If the badminton ball fails to land within the specified drop zone, it is categorized as a "miss"; conversely, if it lands within the zone, it is labeled as a "hit". Alongside documenting the hit or miss status of each ball, the deviation X_i (where *i* represents the *i*-th ball dispatched by the serving device) is recorded, representing the distance between the drop point and the center of the drop zone.

In accordance with the previously outlined experimental methodology, ten groups were subjected to experimentation for each serve type, launching 16 balls in each group and totaling 160 balls for each serve type, resulting in an overall count of 480 balls. Figure 12 displays a set of landing points for near net balls, mid-court balls, and high long balls launched by the badminton serving device.



Figure 12. Three types of ball drop point diagrams (**a**) The near net ball experimental diagram; (**b**) The mid-court ball data diagram; (**c**) The high long ball experimental diagram; (**d**) The near net ball data diagram; (**e**) The mid-court ball data diagram; (**f**) The high long ball data diagram.

5.2. Experimental Results and Analysis

As shown in Table 4, the badminton serve hit data table shows the average number of hits and the standard deviation of hits for each group. The average number of RC hit is around 15, while the number of RC hit is also around 12–13; from the standard deviation value, the standard deviation of the close ball landing in the RC area is the smallest 0.42, while the standard deviation of the long ball landing in the BC area is the largest 1.20.

Group Number	Number of Serve	Number of RC	Number of BC	RC Hit Average	BC Hit Average	RC Hit Standard Deviation	BC Hit Standard Deviation
The near net ball	160	138	158	13.8	15.8	0.42	1.03
The mid-court ball	160	136	154	13.6	15.4	0.70	0.52
The high long ball	160	129	152	12.9	15.2	0.63	1.20

Table 4. The near-net ball hit the data table.

Illustrated in Figure 13 is the scatter plot depicting the ball drop deviation for the three types of balls mentioned above. The horizontal coordinate of the graph represents the number of badminton drop positions, and the vertical coordinate represents the deviation. The red line represents the RC in the ball drop zone, and the ball falling below the red line represents the ball falling within the RC. Similarly, the ball falling below the blue line represents the ball falling within the BC. If the ball falls above the blue line, it means that the ball does not fall within the drop zone.



Figure 13. Distribution plot of mean deviation and standard deviation (**a**) The near net ball deviation scatter plot; (**b**) The mid-court ball deviation scatter plot; (**c**) The high long ball deviation scatter plot.

As depicted in Figure 13, the overall deviations in ball drop for the near netball and the mid-court ball are small, while the deviation for the high long ball, particularly for lofted balls, is larger. Additionally, fewer balls fall within the 0–10 cm deviation range compared to the other two types of balls. For a more in-depth analysis of the deviation data, the deviations for each type of ball were averaged, and standard deviation analysis was conducted, as indicated by the calculation formulas in Equations (9) and (10).

$$Aver = \frac{\sum |X_i|}{n} \tag{9}$$

 A_{ver} represents the average deviation, calculated as the average of the sum of all deviations. X_i denotes an individual deviation, and n represents the total number of deviations counted.

$$S = \operatorname{sqrt}(\sum \frac{(X_i - A\operatorname{ver})^2}{n})$$
(10)

S represents the standard deviation, indicating the dispersion of all deviations from the mean deviation. A_{ver} is the mean deviation as defined in Equation (9).

9

Illustrated in Figure 14 are the histograms representing the average deviation and standard deviation of the falling balls for three types. The figure demonstrates a gradual increase in the average deviation for all three ball types, with the near netball having the smallest deviation at 25.60 cm and the high long ball exhibiting the largest deviation at

33.26 cm. In terms of standard deviation, the mid-court ball displays a larger deviation of 15.84, while the near-net ball has a smaller deviation of 15.28. This suggests that the launching device achieves more accuracy when launching the near netball; however, the launching process is unstable, resulting in a more discrete landing point. The instability is attributed to the insufficient stability of the friction wheel's speed during each ball launch in the ball-serving device.



Figure 14. Histogram of mean and standard deviation.

6. Conclusions

In this paper, we propose a design method for a badminton serving device based on visual perception and multimodal control, comprising five modules: the upper computer module, the posture recognition module, the signal control module, and the execution module. We individually design the angle adjustment structure, launch structure, and ball plucking structure of the execution module to meet the specific requirements of the ball serving device. To address the needs of the signal control module, we independently design an embedded development board integrating the motor drive chip, wireless communication module, and step-down circuit. In the posture recognition method section, we introduce the image posture detection process. The key point information recognized in the previous frame of the image is input into the detection process of the next frame, preventing the repeated recognition of the same posture and enhancing the speed of posture recognition.

The experimental tests for the proposed badminton serving device focus on two main aspects: accuracy in recognizing posture actions and the device's performance in ball delivery. Conducted 300 trials for each posture recognition test, achieving a consistently high posture recognition accuracy exceeding 98%. Conducted 160 launches of each ball type for evaluation and over 150 hits within the drop zone for all three tested ball types. In forthcoming endeavors, our focus will be on refining the badminton serving device through practical usage, which entails upgrades such as replacing the execution module drive motor and enhancing posture recognition methods. These enhancements are poised to elevate the device's performance and user experience to new heights.

Author Contributions: F.J.—conceptualization, design and implementation, data analysis, writing and editing the manuscript. Y.L. (Yuxuan Lin)—conceptualization, design and implementation, data analysis, writing and editing the manuscript. R.M.—Supervision, conceptualization, design and implementation, fund acquisition, writing, and editing of the manuscript. C.Q.—data acquisition. Y.W.—data acquisition, formal analysis. Y.L. (Yuhui Liu)—editing the manuscript. H.L.—data acquisition, formal analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 32201679); in part by the Science Foundation of Fujian Province of China (Grant No. 2022J05230); and in part by the Open Project Program of Guangdong Provincial Key Laboratory of Agricultural Artificial Intelligence (Grant No. GDKL-AAI-2023008); and in part by the Minjiang University Talent Introduction Technology Project (Grant No. MJY22012).

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors wish to sincerely thank the editors and anonymous reviewers for their critical comments and suggestions for improving the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- 1. Steels, T.; Van Herbruggen, B.; Fontaine, J.; De Pessemier, T.; Plets, D.; De Poorter, E. Badminton activity recognition using accelerometer data. *Sensors* 2020, *20*, 4685. [CrossRef] [PubMed]
- Zhi, J.; Luo, D.; Li, K.; Liu, Y.; Liu, H. A novel method of shuttlecock trajectory tracking and prediction for a badminton robot. *Robotica* 2022, 40, 1682–1694. [CrossRef]
- 3. Mori, S.; Tanaka, K.; Nishikawa, S.; Niiyama, R.; Kuniyoshi, Y. High-speed and lightweight humanoid robot arm for a skillful badminton robot. *IEEE Robot. Autom. Lett.* **2018**, *3*, 1727–1734. [CrossRef]
- 4. Xie, J.; Chen, G.; Liu, S. Intelligent badminton training robot in athlete injury prevention under machine learning. *Front. Neurorobotics* **2021**, *15*, 621196. [CrossRef] [PubMed]
- Kumar, A.; Vhatkar, P.; Shende, H.; Chavan, A.; Mahapatro, K. Real-Time Trajectory Prediction and Localization of Omnidirectional Badminton Robot. In Proceedings of the 2022 IEEE Pune Section International Conference (PuneCon), Pune, India, 15–17 December 2022; pp. 1–8.
- Aslam, M.U.; Bashir, A.; Draz, W.U.; Raja, H.M. Optimized Shuttlecock Propulsion Machine to Facilitate Badminton Training. In Proceedings of the 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Swat, Pakistan, 24–25 July 2019; pp. 1–6.
- 7. Depraetere, B.; Liu, M.; Pinte, G.; Grondman, I.; Babuška, R. Comparison of model-free and model-based methods for time optimal hit control of a badminton robot. *Mechatronics* **2014**, *24*, 1021–1030. [CrossRef]
- Chen, Z.; Li, R.; Ma, C.; Li, X.; Wang, X.; Zeng, K. 3D vision based fast badminton localization with prediction and error elimination for badminton robot. In Proceedings of the 2016 12th World Congress on Intelligent Control and Automation (WCICA), Guilin, China, 12–15 June 2016; pp. 3050–3055.
- De Alwis AP, G.; Dehikumbura, C.; Konthawardana, M.; Lalitharatne, T.D.; Dassanayake, V.P. Design and development of a badminton shuttlecock feeding machine to reproduce actual badminton shots. In Proceedings of the 2020 5th International Conference on Control and Robotics Engineering (ICCRE), Osaka, Japan, 24–26 April 2020; pp. 73–77.
- 10. Mori, S.; Tanaka, K.; Nishikawa, S.; Niiyama, R.; Kuniyoshi, Y. High-speed humanoid robot arm for badminton using pneumaticelectric hybrid actuators. *IEEE Robot. Autom. Lett.* **2019**, *4*, 3601–3608. [CrossRef]
- 11. Cao, Z.; Liao, T.; Song, W.; Chen, Z.; Li, C. Detecting the shuttlecock for a badminton robot: A YOLO based approach. *Expert Syst. Appl.* **2021**, *164*, 113833. [CrossRef]
- 12. Wang, K. Design and Research of Badminton Robot; China University of Mining and Technology: Beijing, China, 2017.
- 13. Mizuno, N.; Makishima, T.; Tsuge, K.; Kondo, S.; Nonome, T.; Kurebayashi, H.; Otake, S.; Shibata, D.; Yamakawa, S. Development of automatic badminton playing robot with distance image sensor. *IFAC-PapersOnLine* **2019**, *52*, 67–72. [CrossRef]
- Ordoñez-Avila, J.L.; Pineda, A.D.; Rodriguez, J.D.; Carrasco, A.M. Design of badminton training robot with athlete detection. In Proceedings of the 2022 7th International Conference on Control and Robotics Engineering (ICCRE), Beijing, China, 15–17 April 2022; pp. 26–31.
- 15. Huang, X.; Wang, F.; Zhang, J.; Hu, Z.; Jin, J. A posture recognition method based on indoor positioning technology. *Sensors* **2019**, 19, 1464. [CrossRef] [PubMed]
- 16. Zhu, M.; Sun, Z.; Chen, T.; Lee, C. Low cost exoskeleton manipulator using bidirectional triboelectric sensors enhanced multiple degree of freedom sensory system. *Nat. Commun.* **2021**, *12*, 2692. [CrossRef] [PubMed]
- Cheng, Y.; Yang, B.; Wang, B.; Yan, W.; Tan, R.T. Occlusion-aware networks for 3d human pose estimation in video. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 723–732.
- 18. Bazarevsky, V.; Grishchenko, I.; Raveendran, K.; Zhu, T.; Zhang, F.; Grundmann, M. Blazepose: On-device real-time body pose tracking. *arXiv* **2020**, arXiv:2006.10204.
- 19. Liu, W.; Liu, X.; Hu, Y.; Shi, J.; Chen, X.; Zhao, J.; Wang, S.; Hu, Q. Fall detection for shipboard seafarers based on optimized BlazePose and LSTM. *Sensors* **2022**, *22*, 5449. [CrossRef] [PubMed]
- Mroz, S.; Baddour, N.; McGuirk, C.; Juneau, P.; Tu, A.; Cheung, K.; Lemaire, E. Comparing the quality of human pose estimation with blazepose or openpose. In Proceedings of the 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), Paris/Créteil, France, 8–10 December 2021; pp. 1–4.

- 21. Feldman, S.; Stadther, D.; Wang, B. Manilyzer: Automated android malware detection through manifest analysis. In Proceedings of the 2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems, Philadelphia, PA, USA, 28–30 October 2014; pp. 767–772.
- 22. Shi, H.; Chen, J.; Si, J.; Zheng, C. Fault diagnosis of rolling bearings based on a residual dilated pyramid network and full convolutional denoising autoencoder. *Sensors* **2020**, *20*, 5734. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.