

Semi-autonomous Grasping System Based on Eye Movement and EEG to Assist the Interaction Between the Disabled and Environment

Jiaqi Wang¹, Zihao Wang¹, Zesen Zou¹, Junhao Xue¹, Aojie Zhang¹, Ruochen Tian¹, Shuo Gao^{1,2,*}

¹School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing, China

²Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Interdisciplinary Innovation Institute of Medicine and Engineering, Beihang University, Beijing, China

*shuo_gao@buaa.edu.cn

Abstract—Patients physically handicapped can't take care of themselves. Helping them easily control the objects around them will reduce their psychological burden and social pressure. In this article, a semi-autonomous grasping system based on eye movement and EEG is presented to achieve this goal. Patient just needs to gaze the target object and keep focused, the manipulator will automatically move to its position and grasp it. Experimental results verify the reliability of the system. This system promotes the development of human-computer interaction system based on multi-sensor fusion.

Keywords—eye tracking; EEG; manipulator; laser feedback

I. INTRODUCTION

Eye movement is an important way for patients physically handicapped to interact with environment. By tracking fixation point, both eye-tracking glasses and desktop eye tracker enable people to control devices just by gazing. Eye-tracking glasses are widely used in various scenes due to their portability.

According to system structure and measurement, eye tracking methods are divided into invasive and noninvasive[1]. Pupil Center Cornea Reflection (PCCR) is a noninvasive technique. Supposing the position of camera and infrared source is fixed relative to the eye, and eyeball is a sphere, the absolute position of Purkinje image, i.e. corneal reflection, is fixed when eye rotates. But its position relative to pupil is constantly changing. So the vector(P-CR) from pupil center to cornea center determines the direction of sight line. The methods of gaze estimation consists of polynomial[2][3], neural networks[4] and Gaussian regression[5], etc.

In order to interacts with environment, eye tracker needs to be connected to a manipulator. The systems is used in assisting in surgery[6] and writing and drawing[7]. Combining with other signals can improve its performance. A promising way is to combine with electroencephalogram (EEG)[8][9]. EEG acquisition methods are also divided into invasive and noninvasive. Noninvasive single channel EEG acquisition method is used to control the car[10].

However, most of current systems are still low-level control. Users need to pay attention to the manipulator and give the next

instructions constantly. In this paper, a semi-autonomous grasping system based on EEG and eye movement signals is presented. Users only need to gaze the object and keep focused, then manipulator will automatically move to target position to grasp it. Stability of the system was verified by experiments.

II. SYSTEM DESCRIPTION

The system structure is shown in Fig.1. We designed a pair of eye-tracking glasses to track eye movements using PCCR method. Fixation point is marked in scene image taken by web camera after eye detection and gaze estimation. Laser feedback is introduced to feed back the fixation point recognized by eye-tracking glasses. EEG is obtained by TGAM (ThinkGear AM) module. We extract attention value from it. If five consecutive values are greater than 60, the manipulator will be started. Coordinates of target object are transmitted to manipulator through hand-eye calibration. Manipulator automatically move to target object to grasp it. Details of each module will be introduced in Section III.

III. METHODOLOGY

A. Eye-tracking Glasses

Eye-tracking glasses are mainly composed of glasses frame, a web camera capturing scene image, an infrared camera with infrared light capturing infrared image of eye and a pair of anti-slip rubber sleeves keeping the glasses and head relatively still, as shown in Fig.2. Wavelength of infrared light is 850nm. Its irradiance is about 0.005 W/cm² when forward voltage is 1.3V, which is lower than the threshold harmful to our eyes[11].

Eye detection: Center of cornea and pupil are located in eye image. we crop a region of interest (ROI) containing pupil after eliminate noise by Gauss filtering. ROI is transformed into gray image then converted to binary image based on adaptive threshold binarization algorithm. Eye contour is extracted using Canny edge detection. Considering the ellipse-like outline of pupil in side shooting, we fit it twice by Hough circle transformation and least square ellipse fitting method.

Gray level of the gray image gradually decreases around Purkinje image. Coordinates of Purkinje image center are

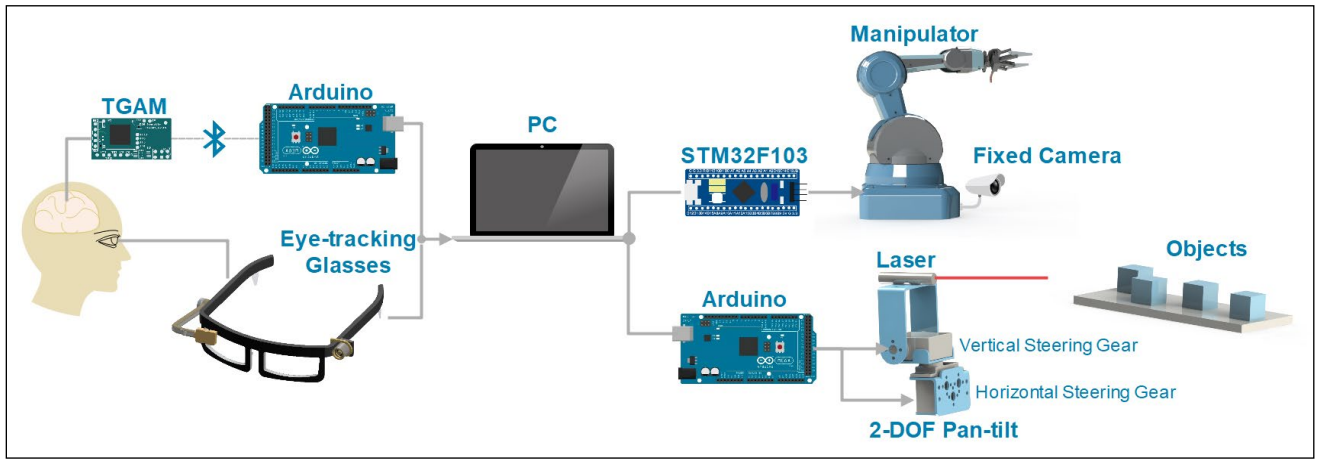


Fig. 1. Architecture of the grasping system.

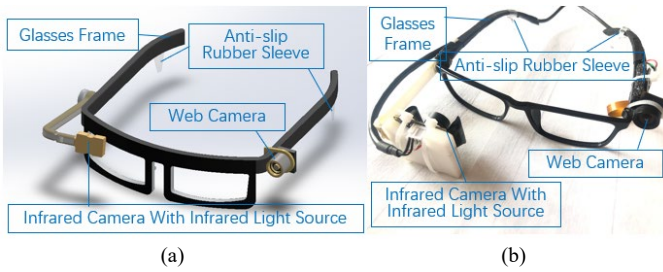


Fig. 2. (a) Design of eye-tracking glasses. (b) Physical Model.

obtained by minimum enclosing rectangle after transforming the gray image into binary image again with higher threshold. The result of eye detection is shown in Fig.3(b).

Gaze estimation: P-CR vector represents gaze direction. We use quadratic polynomial (1) to calculate the mapping between eye coordinate system and field coordinate system.

$$\begin{cases} x = a_0 + a_1u + a_2v + a_3u^2 + a_4v^2 + a_5uv \\ y = b_0 + b_1u + b_2v + b_3u^2 + b_4v^2 + b_5uv \end{cases} \quad (1)$$

(u, v) are coordinates of P-CR vector in eye coordinate system and (x, y) are coordinates of point in field coordinate system. $a_i (i = 0, 1, 2, 3, 4, 5)$ and $b_i (i = 0, 1, 2, 3, 4, 5)$ are unknown constants, which need 12 equations to calculate their value. Each point whose coordinates are known can provide two equations. So at least 6 reference points are needed. We use 9 reference points for calibration to make the mapping relationship more accurate. Subjects, keeping their heads still, gaze calibration points in turn, as shown in Fig. 3(c). Mapping relationship is obtained after fitting (1) by least squares method. Fixation point is marked in scene image, as shown in Fig.3(d).

B. Laser feedback

Laser feedback aims to feed back the fixation point recognized by eye-tracking glasses to users. The device contains a laser and a two degree of freedom (2-DOF) pan-tilt.

The scene image obtained by eye-tracking glasses is processed by graying, morphological processing and threshold segmentation to get the coordinates of laser points. After calculating the difference between laser point coordinates and fixation point coordinates, PID algorithm is used to control

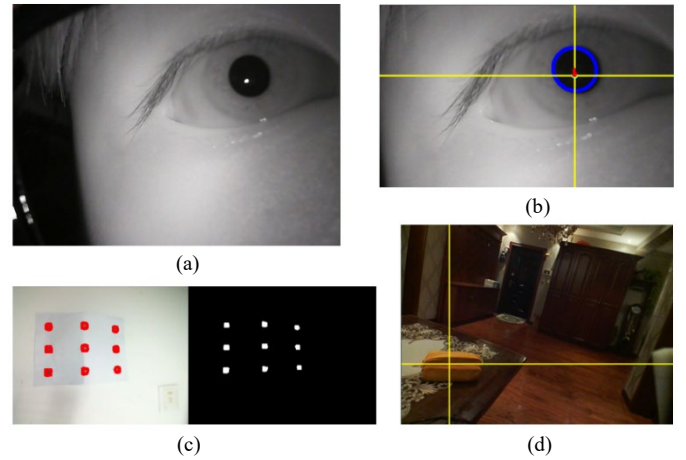


Fig. 3. (a) Original eye image. (b) P-CR vector marked in eye image. (c) The process of gaze estimation. (d) Fixation point marked in scene image.

steering gears and update the coordinates of laser point to realize the real-time feedback of the fixation point.

C. Brain-Computer Interface Model

TGAM, a noninvasive single channel EEG acquisition module, is used to collect attention value. It detects weak EEG signals in frontal lobe and converts it into digital signals.

1) *Acquisition of EEG data:* TGAM module consists of a collection electrode attached to the forehead to collect EEG signals, two reference electrodes clamped to two earlobes to calculate reference potential to reduce noise and a chip filtering and amplifying the EEG and transforming it into digital signals. Processed signals are transmitted to Arduino through Bluetooth.

2) *Processing of EEG data:* TGAM sends 513 packets per second, of which the first 512 packets are small packets and the last packet is large packet containing attention value. The first three bytes of each small packet are "AA AA 04". While the first three bytes of large packet are "AA AA 20". Thus large packets would be distinguished by the 3rd byte. Their 33rd byte is the value of attention.

3) *Transmission of EEG data*: Arduino processes received EEG data in real time, and transmits attention values to PC.

D. Coordinate Transmission and Grasping Model

1) *EEG signal transmission*: Attention value ranged from 1 to 100. The greater it is, the more focused patient is. When people concentrate on something, their attention is above 60. So we use 60 as the threshold to control the manipulator. STM32 captures the attention value from the serial port and judges whether they are greater than 60. If five consecutive values are greater than 60, the manipulator will started.

2) *Eye movement signal transmission*: We installed a camera near the manipulator to recognize the laser points on objects and provide their coordinates in a field coordinate system.

In order to make the manipulator move precisely to target position, hand-eye calibration is required to convert the coordinates from camera field coordinate to robotic arm movement coordinate.

Hand-eye calibration of the camera includes computation of several spatial transformation matrices. Robot pose transformation matrix is calculated by robot kinematic modeling. Base relative position transformation relationship is obtained by measuring the spatially relevant coordinates of the robot base and then by translation and rotation.

The homogeneous transformation from calibration target frame to camera frame is recorded using an affine transformation model from 2D-2D point correspondences[12]. In this hand-eye calibration, since calibration point and target point always remain in the same plane, we set the z-coordinate of space to a constant value to achieve the change from 3D point to 2D point.

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{13} & C_{14} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \quad (2)$$

(x_0, y_0) is the actual coordinate of the target point (the height is set to 1) and (X, Y) are the coordinates of target points in the image. The Random Sample Consensus (RANSAC) method is used to find the rotation and translation matrices with the highest confidence.

Essence of Solving the homogeneous transformation matrix from the camera to the robot arm is solving (3)

$$AX = ZB \quad (3)$$

where A, X, Z, B represent the translation matrices between different coordinate systems. The equation is solved using the Kronecker method [13].

3) *Grasping model*: A commercial robotic arm is used to complete the grasping action. When manipulator is started, the computer converts the position coordinates of the target into a host computer signal and sends it to the STM32 single-chip microcomputer, which controls the movement of the steering gear and goes to the designated place to grab the target.

IV. EXPERIMENTS AND RESULTS

We invited seven subjects, three females and four males, with the average age of 21, to conduct the experiments. In each experiment, subjects were asked to control the system to grasp one object from several. Ten experiments were carried out for each subject, with a three-minute break between adjacent experiments. We recorded the accuracy of the eye-tracking glasses, the success rate of the manipulator starting, the accuracy of the manipulator grasping position and the success rate of grasping (i.e. the ratio of the number of successful grasping to the total number of experiments). We calculated the average values of the them. The results are shown in TABLE I. We chose two of experiments as examples and present the results when they were gazed, as shown in Fig.4.

TABLE I. AVERAGE RESULTS OF EXPERIMENTS

Index	Result
Accuracy of the eye-tracking glasses	0.58°
Success rate of the manipulator starting	97.14%
Accuracy of the manipulator grasping position	1.22mm
Success rate of grasping	85.71%

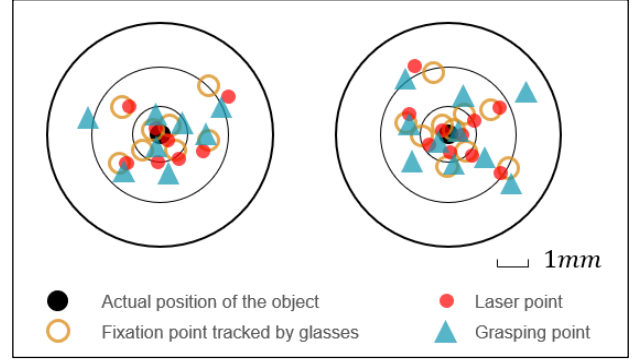


Fig. 4 Experiments results of two objects.

Length, width and height of the objects are all greater than 2 cm. Range of motion of the manipulator is a circle with a radius of 0.5m. Therefore, average error of the eye-tracking glasses and manipulator will not affect the grasp. The grasping system has high success rate from the results.

V. CONCLUSION

We have presented a semi-autonomous grasping system based on eye movement, visual feedback and EEG for users physically handicapped. Users just need to gaze the target object and keep focused, the manipulator will automatically move to target position and grasp the object. Experimental results successfully demonstrate the feasibility of the proposed technique.

REFERENCES

- [1] C. Jian-nan, Z. Peng-yi, Z. Si-yi, Z. Chuang and H. Ying, "Key Techniques of Eye Gaze Tracking Based on Pupil Corneal Reflection," in *2009 WRI Global Congress on Intelligent Systems*, 2009, pp. 133-138: IEEE.
- [2] S. Rattarom, N. Aunsri and S. Uttama, "Interpolation based polynomial regression for eye gazing estimation: A comparative study," in *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2017, pp. 1-4: IEEE.
- [3] Z. R. Cherif, A. Nait-Ali, J. F. Motsch and M. O. Krebs, "An adaptive calibration of an infrared light device used for gaze tracking," in *IMTC/2002. Proceedings of the 19th IEEE instrumentation and measurement technology conference (IEEE Cat. No. 00CH37276)*, 2002, pp. 1029-1033: IEEE.
- [4] Zhu, Z., Ji, Q. "Eye and gaze tracking for interactive graphic display," *Machine Vision and Applications*, vol. 15, pp. 139-148, 2004.
- [5] Sesma-Sanchez. Laura, Sesma-Sanchez. L, Zhang. Y, Bulling. A and Gellersen. H, "Gaussian processes as an alternative to polynomial gaze estimation functions," in *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, 2016, pp. 229-232.
- [6] H. M. Yip, D. Navarro-Alarcon and Y. Liu, "Development of an eye-gaze controlled interface for surgical manipulators using eye-tracking glasses," in *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2016, pp. 1900-1905: IEEE.
- [7] S. Dziemian, W. W. Abbott and A. A. Faisal, "Gaze-based teleprosthetic enables intuitive continuous control of complex robot arm use: Writing & drawing," in *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2016, pp. 1277-1282: IEEE.
- [8] H. Wang, X. Dong, Z. Chen and B. E. Shi, "Hybrid gaze/EEG brain computer interface for robot arm control on a pick and place task," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 1476-1479: IEEE.
- [9] D. McMullen et al., "Demonstration of a Semi-Autonomous Hybrid Brain-Machine Interface Using Human Intracranial EEG, Eye Tracking, and Computer Vision to Control a Robotic Upper Limb Prosthetic," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014, pp. 784-796: IEEE.
- [10] L. Zhang, Q. Lv and Y. Xu, "Single channel brain-computer interface control system based on TGAM module," in *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017, pp. 1-5: IEEE.
- [11] Sliney, David, et al. "Adjustment of guidelines for exposure of the eye to optical radiation from ocular instruments: statement from a task group of the International Commission on Non-Ionizing Radiation Protection (ICNIRP)," *Applied optics*, vol. 44, pp. 2162-2176, 2005.
- [12] H. Lin, P. Du, W. Zhao, L. Zhang and H. Sun, "Image registration based on corner detection and affine transformation," in *2010 3rd International Congress on Image and Signal Processing*, 2010, pp. 2184-2188: IEEE.
- [13] Shah, Mili. "Solving the robot-world/hand-eye calibration problem using the Kronecker product." *Journal of Mechanisms and Robotics*, vol. 5, no. 031007, 2013.