Advances in Computer, Signals and Systems (2023)
Clausius Scientific Press, Canada
DOI: 10.23977/acss.2023.071006
ISSN 2371-8838 Vol. 7 Num. 10

Research on a Simple Visual Tracking Calibration and Spatial Location Algorithm

Kehua Li[1,*], Yongjie Yao[1], Xiaoxia Yuan[1], Qingmin Wang[1]

[1]*Department of Aviation Medicine, Institute of Naval Special Medical, Second Military Medical University, Shanghai, 200433, China*

[*]*Corresponding author: wqm_8888@hotmail.com*

Keywords: Visual Tracking; Simulation; Calibration; Location; Algorithm

Abstract: A calibration algorithm of line-of-sight (LOS) direction is studied based on the data of eye image taken by camera and pupil recognition. Then, the specific position of LOS on screen is calculated according to the calibrated LOS direction. Detailed methods and formulas are given in the whole process of calibration algorithm and positioning solution, so that in-depth line-of-sight tracking research can be carried out. This research mainly focuses on the location algorithm, but has not carried out in-depth research on pupil recognition and image processing algorithms. This study is currently mainly applied to measuring the focus area of pilots on flight dashboards through pupil monitoring of aircraft, and can also be applied to other fields involving visual tracking.

1. Introduction

The research on how to obtain information from vision originated in the 1990s[1-3]. However, the research of visual tracking technology started late. At first, eye gaze was determined by subjective perception. Later, E.B. Delabarre and E.B. Huery invented the method of recording eye annotation points following the method of recording muscle movements by tattoo drums[4-6]. R. Dodge uses the principle of photogrammetric tambourine and corneal reflex to record eye annotation points.

Eye tracking, also known as line-of-sight tracking technology[7-9], is a technology that uses mechanical, electronic, optical and other detection means to obtain the direction of the subject's current "visual attention". It is widely used in human-computer interaction, helping the elderly and disabled, vehicle driving, human factor analysis, virtual reality and military and other fields.

This article provides a visual image detection method based on the position of the eyeball pupil[10-11], and further calculates the spatial fixation point, thereby obtaining a combined method of visual positioning through computer images.

2. Calibration algorithm

The geometric and optical parameters of the camera are determined by system calibration. The relative position relationship between the camera, the head feature points and the cockpit display screen enables the camera group to carry out three-dimensional space measurement.

Through the sub-module of eye feature acquisition, the detection of pupil and cornea is completed, and the pupil-cornea vector is obtained, which is used as input in the sub-module of eye feature matching with scene.

The matching module of eye feature and scene includes calibration state calculation and working state calculation, and the main solution is to solve the corresponding relationship between corneal-pupil vector POR line-of-sight target in scene as following figure 1.

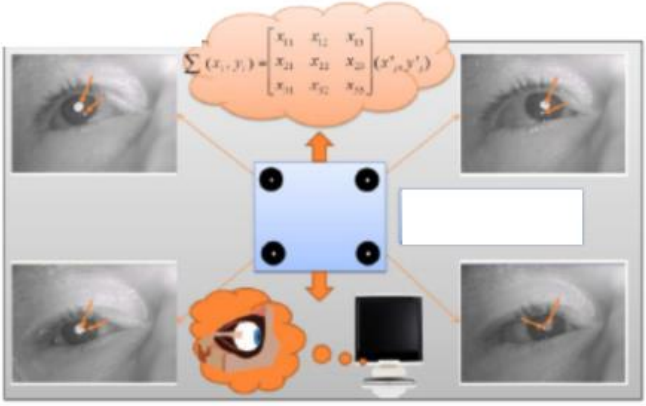

Figure 1: Calibration sketch

According to the fractional definition, a fractional system can be described as

$$x_s = a_0 + a_1 x_e + a_2 y_e + a_3 x_e y_e + a_4 x_e^2 + a_5 y_e^2 \tag{1}$$

$$y_s = b_0 + b_1 x_e + b_2 y_e + b_3 x_e y_e + b_4 x_e^2 + b_5 y_e^2 \tag{2}$$

Where (Xs, Ys) represent the coordinates of the registration points on the target plane, (Xe, Ye) represent the pupil-cornea vector, a0-a5, b0-b5 are the parameters to be determined. Generally, nine points are used for registration, so this is a problem of solving 12 unknowns and 18 equations, which can be solved by least square method.

We assume:

$$S_{x,y} = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ & & \cdots & \cdots & & \\ 1 & x_9 & y_9 & x_9 y_9 & x_9^2 & y_9^2 \end{bmatrix}^T \tag{3}$$

And assume $A_x = [a_0 \quad \cdots \quad a_5]^T$, $A_y = [a_0 \quad \cdots \quad a_5]^T$, $E_x = [x_0 \quad \cdots \quad x_9]^T$, $E_y = [y_0 \quad \cdots \quad y_9]^T$, then we have

$$\begin{aligned} E_x &= S_{x,y} \bullet A_x \\ E_y &= S_{x,y} \bullet A_y \end{aligned} \tag{4}$$

And it can be solved with least square method as

$$A_y = (S_{x,y}{}^T \bullet S_{x,y})^{-1} S_{x,y}{}^T \bullet E_y \tag{5}$$

Using Ax and Ay, we can calculate the one-to-one correspondence between the eye line of sight and the scene in the scene, and then use the scene in the scene camera to complete the line of sight tracking through registration parameters.

3. Solution of Spatial Gaze Position

Since the eye's line of sight is a straight line, a straight line is made between the known eye coordinates and the coordinates of the gaze point, and the straight line is extended to the three-dimensional space along the direction from the eye to the gaze point. The three-dimensional model is essentially a set of coordinate points, and the human eye's gaze will only fall on the surface of the three-dimensional object. Therefore, it is necessary to find the coordinate points closest to the eye that have passed the straight line (or are close enough to the straight line). As shown in Figure 2:

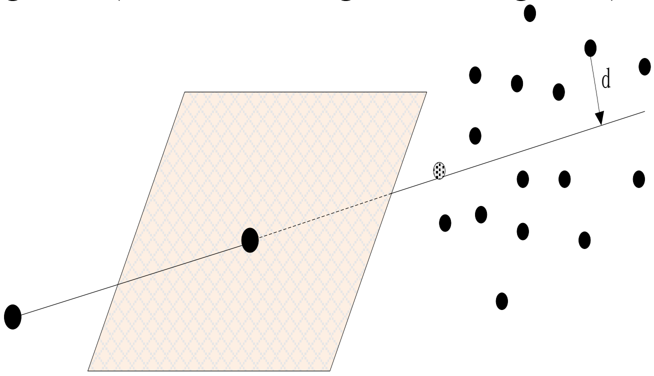

Figure 2: Schematic view of gaze point

In the first step, we need to find the nearest point in the point cloud to the line between the eye and the virtual plane. Formulas for determining a straight line by two points:

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1} \tag{6}$$

In the formula above, eyes and virtual planes focus on coordinate points, which can be brought into the equation to see if the equation is satisfied. Because the coordinates of these point clouds have high numerical accuracy, it is very likely that no point is just above the straight line. Therefore, we need to change this straight line into a "thick" line of sight, which is essentially a straight line, into a "line of sight cylinder", and the point that falls within the "cylinder" is regarded as the point that meets the requirements. So we need to set a threshold, that is, the radius of the cylinder, to calculate the distance between the point in the point cloud and the straight line. If the distance is less than the threshold, then we can judge that the point is in the line of sight.

As shown in Figure 3, if a point coordinate $P(x^*, y^*, z^*)$, eye coordinate $P_1(x_1, y_1, z_1)$, line direction vector $l=(x_2-x_1, y_2-y_1, z_2-z_1)=(\Delta x, \Delta y, \Delta z)$ and unit direction vector of any point cloud $l_0=\frac{l}{\|l\|}$ are set, then the distance formula from point $P(x^*, y^*, z^*)$ to line is as follows:

$$d=\left\|\overrightarrow{P_1P}\times l_0\right\| \tag{7}$$

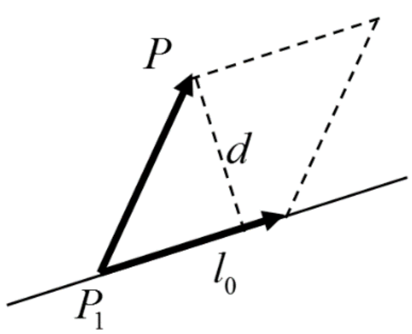

Figure 3: Schematic view of gaze point

Firstly, the plane perpendicular to the line of sight is found. From the line of sight equation, the plane equation can be obtained.

$$\Delta x \cdot x + \Delta y \cdot y + \Delta z \cdot z + D = 0 \tag{8}$$

D is a constant. Because of the crossing point P of the plane, a constant D can be obtained by taking the point into above equation:

$$D = -(\Delta x \cdot x^o + \Delta y \cdot y^o + \Delta z \cdot z^o) \tag{9}$$

The intersection of the line of sight and the plane is the projection point to be found.

4. Conclusion

Eye tracking involves the comprehensive application of multidisciplinary and interdisciplinary, and the research at home and abroad is not yet fully mature. Based on the pupil data extracted from the pupil photographs, this paper studies the calibration algorithm, and finally gives the formula for calculating the position of the spatial gaze point, which provides the technical basis for the accuracy of line-of-sight tracking. This paper also gives the completed calculation formula, which lays a solid foundation for further research.

References

[1] Ma Sugang, Sun Siwei, Hou Zhiqiang, et al. A Visual Tracking Algorithm Based on Multiple Attention and Dual Template Updating [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023. DOI: 10.13700/j.bh.1001-5965.2023.0334

[2] Liu Qianqiang, Zhong Xiaoyong. Design of Visual Tracking System for Quadrotor Unmanned Aerial Vehicle Based on OpenMV [J]. Automation and Instrumentation, 2023, 38 (5): 5-8

[3] Hu Shuo, Yao Meiyu, Sun Linna, et al. Accurate Visual Tracking by Integrating Attention Features [J]. Computer Science and Exploration, 2023, 17 (4): 11. DOI: 10.3778/j.issn.1673-9418.2107010

[4] Chu Xinjian, He Dan, Zhang Fu. Research on binocular visual positioning based on improved GA optimized BP network [J]. Modular Machine Tool and Automation Processing Technology, 2023 (2): 28-30

[5] Zhu Dongying, Zhong Yong, Yang Guanci, et al. Research progress on motion segmentation for visual positioning and mapping in dynamic environments [J]. Computer Applications, 2023, 43 (8): 2537-2545

[6] Shen Ning. Surface foreign object detection of belt conveyors used in coal preparation plants based on binocular vision [J]. Industrial and Mining Automation, 2023, 49 (S1): 82-85

[7] Chen Shoubin. Research on Key Technologies of Laser LiDAR/Visual Fusion SLAM (LV SLAM) [J]. Journal of Surveying and Mapping, 2023, 52 (1): 1. DOI: 10.11947/j.AGCS.2023.20210184

[8] Li Xiaoyue. Research on Key Technologies for Bus Passenger Count Statistics Based on Multiple Visual Sensors [D]. Jilin University [2023-10-17]. DOI: CNKI: CDMD: 21013.195427

[9] Li Huiyi, Li Xi, Xie Jingwei. Research on Visual Positioning Method for Transmission Line Deicing Robot [J]. Mechanical Design and Manufacturing, 2023 (3): 198-202.

[10] Wang L, Tang Y, Xu H Z. A fast pupil localization method based on Harris corner detection [J]. Liquid Crystal and Display, 2008, 23 (1): 4. DOI: 10.3969/j.issn.1007-2780.2008.01.019

[11] Wang Xianmei, Yang Ping, Wang Zhiliang. Pupil localization algorithm in multi pose eyeballs [J]. Journal of Computer Aided Design and Graphics, 2011, 23 (8): 6. DOI: CNKI: SUN: JSJF0.2011-08-020