ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2006. No. 8(72)
ELEKTRONIKA IR ELEKTROTECHNIKA

MEDICINE TECHNOLOGY
T115
MEDICINOS TECHNOLOGIJA

Calibration of Video-Oculographical Eye-Tracking System

N. Ramanauskas
Department of Electronics, Šiauliai University,
Vilnius str. 141, LT – 76353 Šiauliai, Lithuania, e-mail: n.ramanauskas@tf.su.lt

Introduction

In a vision-based user interface, people can interact with computers by pointing with their eyes [1]. The direction of eye gaze expresses the intent of user. The precise mapping from image co-ordinates to computer monitor screen co-ordinates in real time is necessary for efficient computer control.

Advanced remote eye gaze trackers that are being researched today basically try to eliminate two problems, the need of calibration per user session, and the large restriction on head motion. A system suggested by Morimoto et al. [2] estimates eye gaze without calibration and allowing free head motion. But it requires the calibration of the camera with respect to the monitor and light positions, and a model of the user's eye. Experimental results show an accuracy of about 3º using synthetic images. Another system described by Yoo et al. [3] uses four LEDs around the monitor screen to project these corners on the corneal surface and the fifth LED is placed near the CCD camera lens to create a bright pupil image. Using the invariance property of cross ratios under perspective, they compute the point of regard with an accuracy of about 2º. The advantage of this method is that it does not require camera calibration. Newman et al. [4] gives example of system that first computes the face pose in 3D, and then compute the eye gaze. The system runs in real time, but the accuracy is very low, about 5º.

The objective of this study is to compare calibration techniques for pupil/eye corner gaze tracking system.

Calibration techniques

A calibration procedure is required to compute the mapping between the measurements and the eye orientation. A typical calibration procedure presents the user a set of visual targets that the user has to look at while the corresponding measurement is taken. From these correspondences, a mapping or calibration function can be computed.

In this paper we will present four calibration techniques: a standard calibration using linear and second order models; 2D mapping with interpolation; a mapping with developed model, describing eye image formation process.

A standard calibration set consists of 5, 9 or 25 points. The simplest linear calibration model takes into account only 5 calibration points, usually four points in the corners and one in the centre. The mapping between screen coordinates and the measurements is done using following equation:

$$\begin{cases} s_x = a_0 + a_1 \cdot x, \\ s_y = b_0 + b_1 \cdot y, \end{cases} \tag{1}$$

where (s_x, s_y) are screen coordinates and (x, y) is the pupil–corneal reflection vector. The coefficients a_0, a_1 and b_0, b_1 are the unknowns and can be founded using least squares.

A second order polynomial calibration function can be used with a set of 9 or 25 calibration points. The polynomial is defined as:

$$\begin{cases} s_x = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2, \\ s_y = b_0 + b_1 x + b_2 y + b_3 xy + b_4 x^2 + b_5 y^2, \end{cases} \tag{2}$$

where the coefficients a_0 - a_1 and b_0 - b_1 are the unknowns and can be founded using least squares.

Zhu and Yang [5] construct a 2D linear mapping from the vector between the eye corner and the iris centre to the gaze angle. After calibration, the gaze direction is computed by interpolation.

For example, suppose the screen coordinates and the eye corner to iris vector used for calibration in points P_1 and P_2 are respectively {(s_{x1}, s_{y1}), (x_1, y_1)} and {(s_{x2}, s_{y2}), (x_2, y_2)}. Then after the measurement of a corner-iris vector (x, y) is taken, the screen coordinates is computed as follows:

$$\begin{cases} s_x = s_{x1} + \dfrac{x - x_1}{x_2 - x_1}(s_{x2} - s_{x1}), \\ s_y = s_{y1} + \dfrac{y - y_1}{y_2 - y_1}(s_{y2} - s_{y1}), \end{cases} \tag{3}$$

The experimental setup of pupil/eye corner gaze tracking system is shown in Fig. 1. The model was developed to describe eye image formation process [6].

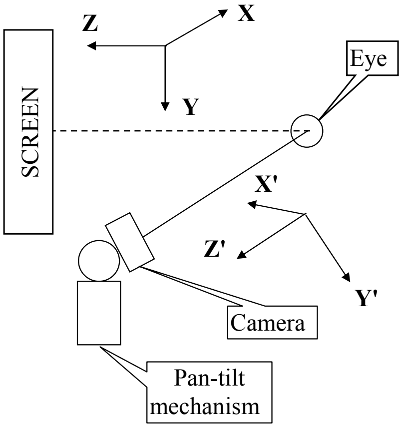

Fig. 1. Experimental setup of pupil/eye corner gaze tracking system and co-ordinates systems, for image formation.

X and Y axis of laboratory co-ordinate system are oriented as the same axis of computer monitor screen. Axis Z is perpendicular to monitor screen. Other co-ordinates system X'Y'Z' is related to image sensor, which is used in video camera. The eye rotations are described in laboratory system. We could obtain the eye image, by linear projection of eye structure to X'Y' plane of camera system.

We aligned a unit vector *e* along eye optical axis. The projections of the *e* vector to plane X'Y' is proportional to measured pupil centre coordinates. We assume that orientation system X'Y'Z' is defined by two angles. One parameter defines proportionality between end of vector *e* and real pupil movements on image sensor.

The gaze point coordinates on screen depend on distance from eye to screen. Also we must take into account that eye sight axis doesn't correspond with eye optical axis, because the last doesn't go through the fovea. For every eye we have one parameter – the angle between optical and sight axis. So minimum number of model parameters is 6.

An example of simulated results for a fixation of 25 calibration points square grid is shown in Fig. 2.

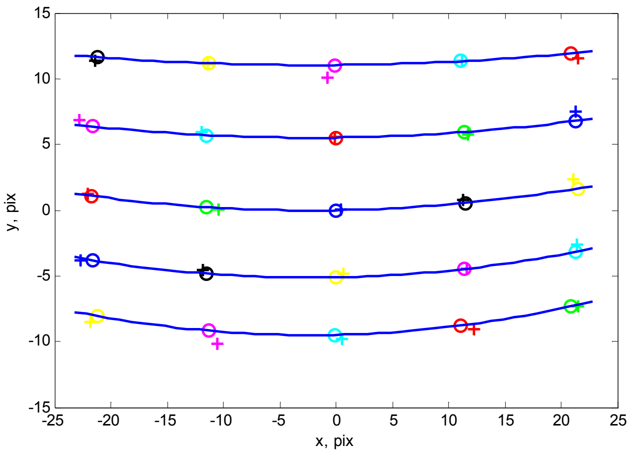

Fig. 2. Experimental results for a fixation of 25 calibration points. "o" - simulated, "*" - measured

Method

The calibration procedure was done using a set of 25 dots distributed evenly on the screen (Fig. 3). Only one dot was displayed at a time at a random order. Saccadic eye movements have a typical duration of 30 – 100 ms and a latency of 100 – 300 ms. The duration of dot appearance was chosen for two seconds, to obtain sufficient calibration data. To pay the maximum user attention, the radius of displayed dot was decreasing.

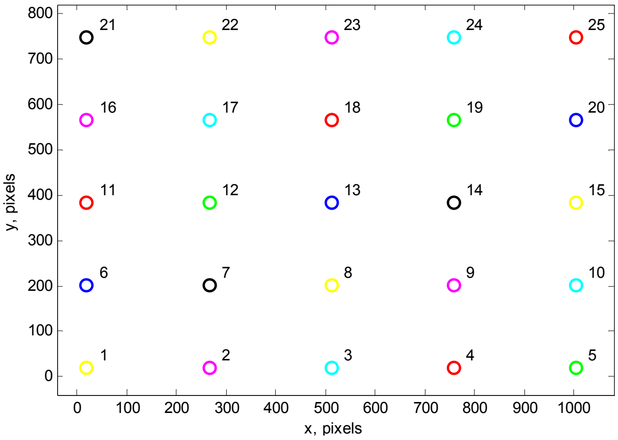

Fig. 3. Calibration points in screen

After calibration procedure the subject had to look at a set of 25 experimental points appearing at a random position of the screen. So each experiment had unique experimental points at random positions. These experimental points were used to test the calibration techniques.

Video sequences were recorded with VOG eye tracking system, developed in our laboratory. The recording was done for both eyes simultaneously. Each video sequence was analyzed with the same pupil centre detection method (coordinates averaging) [7]. Elimination of head movements was done using eye inner corner tracking with normalized correlation coefficient [8].

The experimental pupil-eye corner vector measurements were mapped to the screen coordinates using five different techniques:

1) linear calibration model using a set of 5 calibration points;
2) second order polynomial calibration model using a set of 9 calibration points;
3) second order polynomial calibration model using a set of 25 calibration points;
4) interpolation using a set of 25 calibration points;
5) a model based mapping using a set of 25 calibration points.

The mapping error was calculated for each technique at each experimental point:

$$e = \sqrt{(x_e - x_m)^2 - (y_e - y_m)^2}, \qquad (4)$$

where $(x_e\ ,\ y_e)$ – the actual coordinates of experimental point, $(x_m\ ,\ y_m)$ – coordinates of mapped pupil/eye corner vector.

Results

A mapping examples with linear calibration model using a set of 5 calibration points and with second order polynomial model using a set of 25 calibration points is shown on Fig. 4 and 5 respectively.

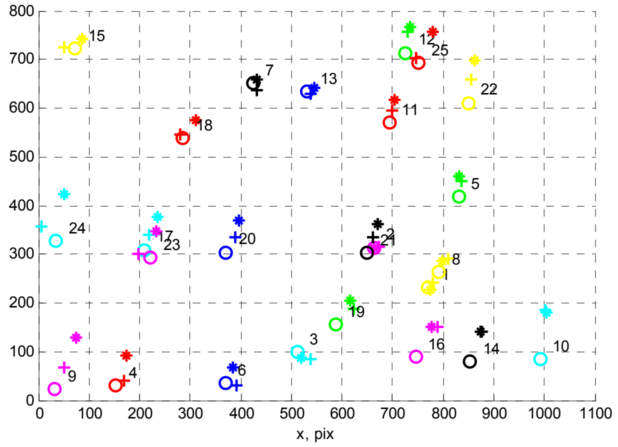

Fig. 4. Experimental points mapped to screen coordinates with linear calibration model. "o" – actual experimental point; "+" – left eye data; "*" – right eye data

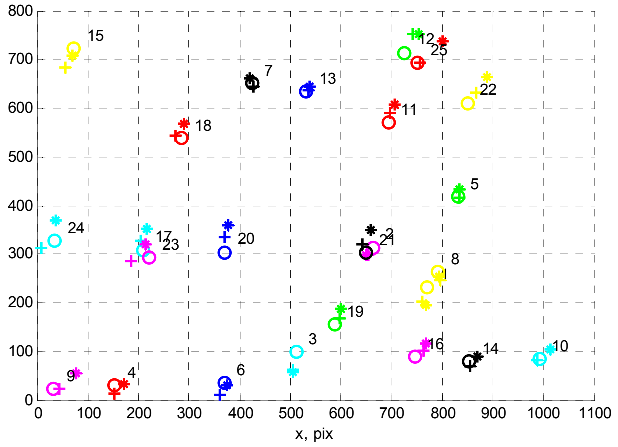

Fig. 5. Experimental points mapped to screen coordinates with polynomial calibration model. "o" – actual experimental point; "+" – left eye data; "*" – right eye data

Fig. 6 represents the mapping errors of different calibration techniques for 5 subjects. The results of mapping errors are presented in the next order for each subject: 1. linear calibration model using a set of 5 calibration points; 2. second order polynomial calibration model using a set of 9 calibration points; 3. second order polynomial calibration model using a set of 25 calibration points; 4. interpolation using a set of 25 calibration points; 5. a model based mapping using a set of 25 calibration points. This numeration of calibration techniques is common in Fig. 7 and in Table 1.

The averaged mapping errors over all experimental points are shown on Fig. 7.

To compare the differences between calibration techniques, a paired Student test was performed. The alternative hypothesis H=1 $\mu_x > \mu_y$ was tested at the significance level 0.05. The results are presented Table 1.

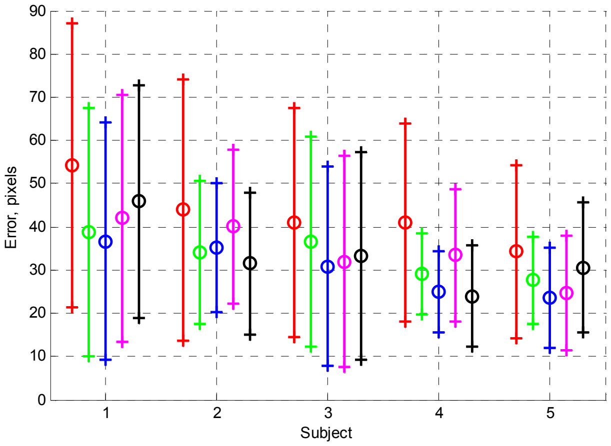

Fig. 6. Mapping errors of different calibration techniques for 5 subjects. "o" – mean value, "+"- standard deviation

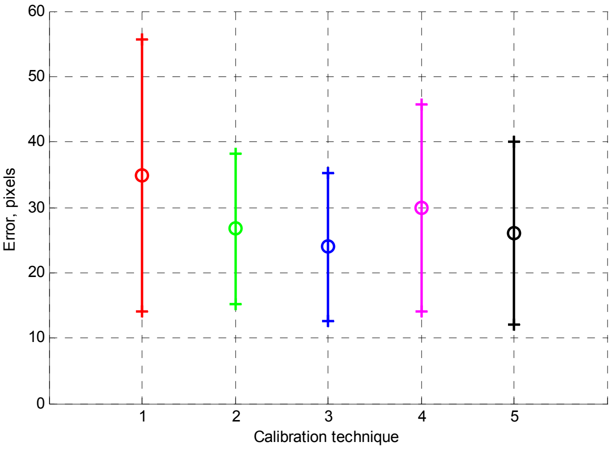

Fig. 7. Averaged mapping errors of different calibration techniques.. "o" – mean value, "+"- standard deviation

Table 1. Paired Student test results (p=0.05)

Hypothesis	H	p-value
$\mu_1 > \mu_2$	H=1	7.3550e-005
$\mu_1 > \mu_3$	H=1	2.7899e-007
$\mu_2 > \mu_3$	H=1	0.0318
$\mu_4 > \mu_3$	H=1	4.5595e-004
$\mu_4 > \mu_2$	H=1	0.0365
$\mu_4 > \mu_1$	H=0	0.9842
$\mu_1 > \mu_5$	H=1	4.8057e-005
$\mu_2 > \mu_5$	H =0	0.3367
$\mu_3 > \mu_5$	H=0	0.8938
$\mu_4 > \mu_5$	H=1	0.0216

Discussion

In accordance with Table 1 we can arrange calibration techniques by the mean error in decreasing order:

- linear calibration model using a set of 5 calibration points;
- interpolation using a set of 25 calibration points;

- second order polynomial calibration model using a set of 9 calibration points;
- a model based mapping using a set of 25 calibration points
- second order polynomial calibration model using a set of 25 calibration points;

There is no significant difference between a mapping with developed model, describing eye image formation process, with both polynomial calibration models.

Calibration techniques based on polynomial models do not allow a wide range of head movements. The system must be recalibrated very often in free head motion. Each time user must follow a difficult calibration procedure, especially using a set of 25 points.

A model based mapping constructs a mathematical model of subject's eye during the first interaction. A model is adaptive to head motions, through changeable parameters. There is no need to follow a whole calibration points for the second and later calibration.

The actual mapping errors expressed in degrees are 1.5^0 and 1^0 for the linear calibration and second order polynomial calibration respectively.

Conclusions

The different calibration techniques were investigated. It was proved that, there is no significant difference between a mapping with developed model and polynomial calibration. An improved model could take an advantage over polynomial mapping.

Acknowledgement

This research is supported by the Lithuanian Science and Studies Foundation (Project title: Environment Control by Gaze System for the Disabled; Registration Number: V-06003) and FP6/IST Network of Excellence project COGAIN (contract 511598).

References

1. **Jacob R.** The use of eye movements in human-computer interaction techniques: What you look at is what you get // ACM Transactions on Information Systems. – 1991. – V. 9, No. 3. – P. 152–169.
2. **Morimoto C., Amir A., Flickner M**. Detecting eye position and gaze from a single camera and 2 light sources // Proc. of the Int. Conf. on Pattern Recognition. – 2002. –P. 314–317.
3. **Yoo D., Kim J., Lee B., Chung M**. Non contact eye gaze tracking system by mapping of corneal reflections // Proc. of the Int. Conf. on Automatic Face and Gesture Recognition. – 2002. – P. 94–99.
4. **Newman R., Matsumoto Y., Rougeaux S., Zelinsky A**. Real time stereo tracking for head pose and gaze estimation // Proc. of the Int. Conf. on Automatic Face and Gesture Recognition. – 2000. – P. 122–128.
5. **Zhu J., Yang J.** Subpixel eye gaze tracking // Proc. of the 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition. – 2002. – P. 131–136.
6. **Daunys G., Ramanauskas N**. Nonlinear mapping of Pupil Centre Coordinates from Image Sensor to Screen for Gaze Control Systems // Proc. of the ICCHP 2006.– P. 962–965.
7. **Ramanauskas N., Daunys G.** The investigation of eye tracking accuracy using synthetic images // Electronics and Electrical Engineering. – 2003. – No. 4(46). P. 17–20.
8. **Ramanauskas N., Daunys G.** Elimination of Head Shifts in Videooculography // Electronics and Electrical Engineering. – 2005. – No. 8(64). – P. 69–72.

Submitted for publication 2006 05 25

N. Ramanauskas. Calibration of Video-Oculographical Eye Tracking System // Electronics and Electrical Engineering. - Kaunas: Technologija, 2006. – No. 8(72). – P. 65–68.

Video-oculography (VOG) is the most suitable eye tracking method in Human Computer Interaction. The precise mapping from image co-ordinates to computer monitor screen co-ordinates in real time is necessary for efficient computer control. During this study, the calibrations techniques were investigated. It was noticed that, there is no significant difference between a mapping with developed model, describing eye image formation process in the sensor, and with polynomial calibration. Polynomial calibration model do not allow a free head motion and requires frequent recalibration. A model based mapping is tolerant to head motions. An improved model could take an advantage over polynomial mapping. Ill. 7, bibl. 8 (in English, summaries in English, Russian, Lithuanian).

Н. Раманаускас. Калибрирование видеоокулографической системы измерения движений глаз // Электроника и электротехника. – Каунас: Технология, 2006. – № 8(72). – С. 65–68.

Видеоокулография – метод измерения движений глаз, наиболее подходящий для применения в интерфейсах человек–компьютер. Для эффективного управления компьютером угловые координаты глаза нужно точно пересчитать в координаты монитора. В настоящей работе были исследованы методы калибрирования. Замечено, что нет значительной разницы между пересчитанием с созданной моделью, которая описывает формирование образа глаза в сенсоре, и пересчитанием с полиномами второго ряда. Использование полиномов не устойчиво к движениям головы и нужно частое дополнительное калибрирование. Использование модели более толерантно к движениям головы. Усовершенственная модель имеет преимущество по сравнению с полиномным преобразованием. Ил. 7, библ. 8 (на английском языке, рефераты на английском, русском и литовском, яз.).

N. Ramanauskas. Videookulografinės akių judesių matavimo sistemos kalibravimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2006. – Nr. 8(72). – P. 65–68.

Videookulografija – akių judesių matavimo metodas, labiausiai tinkantis žmogaus ir kompiuterio sąveikai. Norint efektyviai valdyti kompiuterį akies kampines koordinates reikia tiksliai perskaičiuoti į monitoriaus koordinates. Šiame darbe buvo tiriami kalibravimo metodai. Pastebėta, kad nėra didelio skirtumo tarp perskaičiavimo naudojant sukurtą modelį, kuris aprašo akies vaizdo susidarymą sensoriuje, ir perskaičiavimo naudojant 2-osios eilės polinomus. Naudojant polinomus galva negali laisvai judėti, dažnai reikia pakartotinai kalibruoti. Naudojant modelį, galvos judesiai labiau toleruojami. Patobulintas modelis būtų dar pranašesnis už kalibravimą naudojant polinomus. Il. 7, bibl. 8 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).