©2017 Published in 5th International Symposium on Innovative Technologies in Engineering and Science 29-30 September 2017 (ISITES2017 Baku - Azerbaijan)

2D CARTESIAN ROBOT CONTROLLING BY EYE TRACKING

[1]Burak Karabulut, [1]Ihsan Alperen Yagiz, [1]Kaplan Kaplan, [*1]Ahmet Erdil
[1]Faculty of Engineering, Department of Mechatronic Engineering, Kocaeli University, Turkey

Abstract

Recently, in order to facilitate daily life, eye tracking systems are used with the Cartesian robotic systems. The Eye tracking system provides an accurate and useful solution for users who do not have the ability to use any limb correctly. In this study, two-dimensional printer is designed to track eye position with stepper motors. In addition, real-time control for drawing and painting operations are momentarily performed according to the direction of the eye movements. In this process, the orb is captured to detect eye position, and the eye movement is determined by using image processing algorithm. Eye tracking software is executed in MATLAB environment. The eye positions are sent to stepper motor drivers. Object and shapes are printed with the designed two-dimensional printer according to eye's positions.

Key words: Eye tracking, 2D printer, Cartesian Robot, eye detection.

1. Introduction

Eye tracking systems provide an opportunity to simplify daily life. On the other hand, the Cartesian system provides a great convenience with moving objects. The eye tracking system allows an accurate and convenient solution for users who do not have the ability to communicate outside of their eyes and can not use any limb correctly. This system can change the whole system of disabled persons. On the other hand, the plotting with Cartesian system instantly provides a great convenience in the preparation of project design. Individiual signature, handwritings can be reflected thanks to these systems. In addition, these systems can be seen in military applications. Especially, they can be used in targeting and bomber weapons in order to increase the hit rate accuracy.

The eye-controllable Cartesian robot system has highly variable parameters. It has an algorithm that can detect the eye and transmit its movements directly to the virtual environment. It also distinguishes the movements of the eye as well as the movement direction and the displacement. The Cartesian robot successfully transfers the virtual movement from the eye to the work surface in two axes coordinate space. The Cartesian robot with three degrees of freedom is more functionality and more capability to work in real time writing.

In this study, the eye tracking and 2D Cartesian systems are combined simultaneously. This system has also a camera that will transfer the human action taken from the eye to the computer environment. The information of eyes position are sent to stepper motors in order to write on the surface through the pen. Therefore, human can design or write thoughts with their eyes and so daily life can be continue easier for user who can not use their hands. This study are performed by using these working parameters:

*Corresponding author: Address: Faculty of Engineering, Department of Mechatronic Engineering Kocaeli University, 41380, Kocaeli TURKEY. E-mail address: erdil@kocaeli.edu.tr, Phone: +902623033223

- Workspace with the size of an A4 (210mm x 297mm) paper format for the Cartesian system,
- Edge precision is maximum 1 mm,
- Holding torque of step motors is greater than 40N.cm or 5Kg.cm and one pulse degree is 1.8°,
- The data is processed in real time.

In Fig. 1, the designed system is showed according to these working parameters.

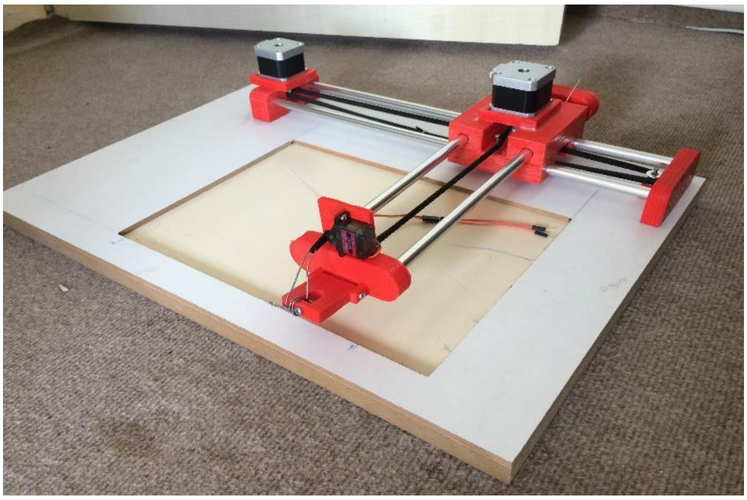

Figure 1. 2D Cartesian plotter

2. Materials and Methods

2.1. 2D-Cartesian Robot System

An industrial Cartesian robot (linear robot) has three linear principal axis. These robots can be controlled in a straight line rather than rotational direction. The three sliding joints represent the moving the wrist throughout up-down, in-out, back-forth way. The most important advantages of these robots are to facilitate the Robot control arm solution. Cartesian coordinate robots that moves in the horizontal directions are named Gantry robots. Gantry robots are generally very big and large mechanical structures which they look like gantry cranes.

Cartesian robots are completely mechatronic devices since they have motors and linear actuators. They can drive a tool to a desired position in work surface. They can move linearly in three axes, X, Y, and Z directions. In some applications that needs high tolerance, full assistance based on X axis is required. On the contrary, some pick and place applications works with less precious. These applications should be supported with base axis in accordance with the actuator manufacturer's recommendations. Cartesian robot movements are limited within the frame boundaries, but the frame can be mounted horizontally or vertically, or in some general gantry configurations [1].

The design and implementation of a 2D Cartesian robot control by using eye tracking are the ideal framework in order to illustrate mechatronic capability over mechanical design, computer science and electronics.

2.1.1. Design of 2D Cartesian Robot

In design part, important factors in Cartesian robot design are illustrated in introduction section. Two moving columns is used since the system was subjected to 2 axis movements. In addition, the end function is limited to a flexible motion scale in order to perform the ON-OFF commands. The hand drawing of design can be seen in Fig. 2 - Fig. 5 [2].

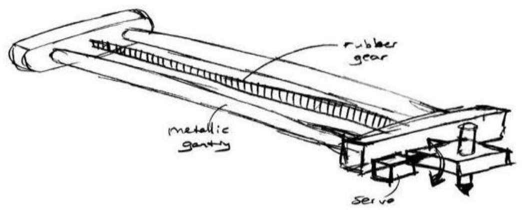

Figure 2. Axis Y

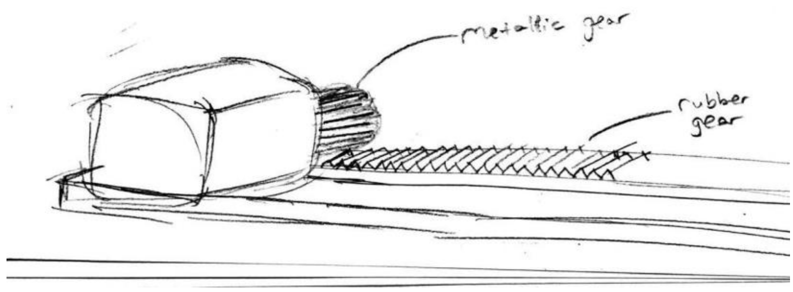

Figure 3. Motor and gear

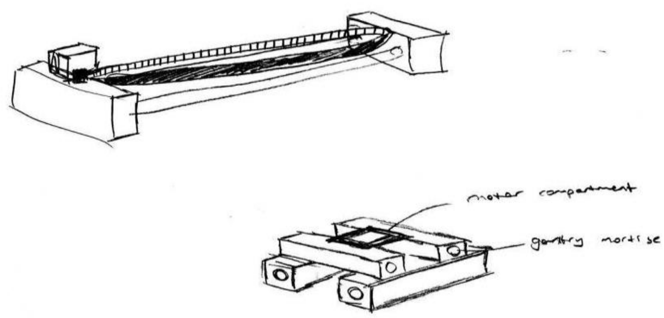

Figure 4. Carrier

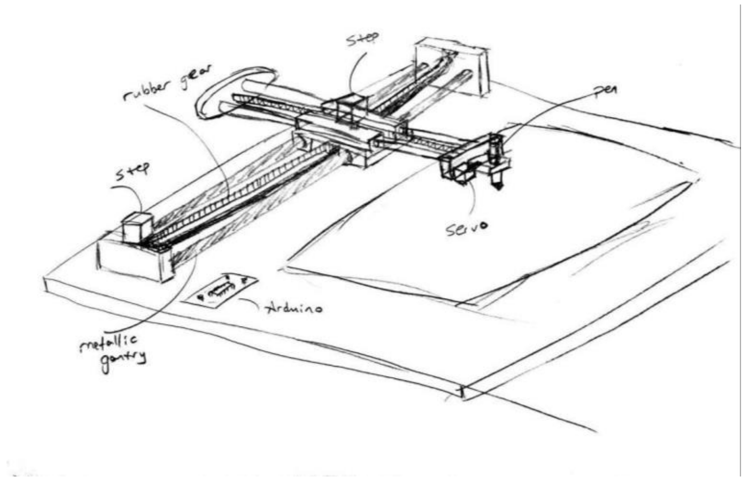

Figure 5. Prototype design of system

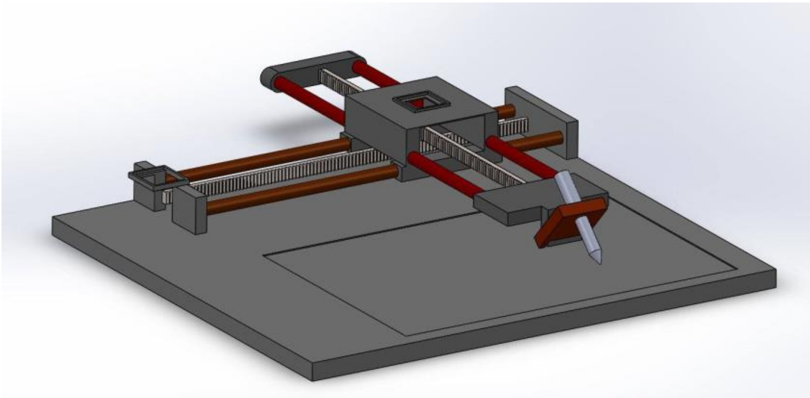

Figure 6. CAD drawing of the System

The CAD drawing above is performed without motor and gear drawings as seen in Fig. 6. The design is realized as a sketch without much detail on it. 2-DOF cartesian robot is choosed because it is very useful both in terms of control and applicability. Non-Cartesian 2-DOF or 3-DOF robots are not so successful in terms of control and applicability. The component used in this system are given in the followings:

- The stepper motor: because the stepper motor is more successful in applications where control is required. A stepper motor is also choosen with a metal gear. The fact that metal gears provide stability in this application and it will prevent from constant motor change.

- Arduino controller card is used as motor driver: Arduino is more advanced than PIC. Therefore, it can be used in real time practically.

- MATLAB as an interface program: MATLAB is more useful than other pocket programs thanks to both control, Simulink and Virtual Reality toolboxes.

- A pencil holder with servo motor control is used as penholder design. It is more useful compared to magnetic and fixed pencil holders. The disadvantage is that the cost is higher than the others are.

- The gear is made from rubber material as the gear. Metal or plastic gears are not flexible and can be a potential drawback in terms of weight. Rubber gears are more flexible than others are, which can save time in extraordinary situations.

- A metallic sled as a sledge. The plastic sled could cause problems in terms of weight (sturdiness).

2.2. Eye Tracking System

Nowadays, most commercial eye trackers systems measure the point of eye by using the corneal-reflection/pupil-center method (Goldberg & Wichansky,2003). These tracker systems usually consist of a standard desktop computer with image processing software to find and identify eye features used for tracking, with an infrared camera mounted to the monitor. During operation, the infrared light from an LED first embedded in the infrared camera makes it easy to monitor them by creating strong impressions on the target eye characteristics (infrared light is used to avoid user's eye from glaring).

After light come to retina, a large part of it is reflected back. The pupil (called as the brilliant-pupil impact) becomes visible as a bright and well-defined disc. The corneal reflection (or the first Purkinje image) is also produced by infrared light and is seen as a small but sharp shining [3].

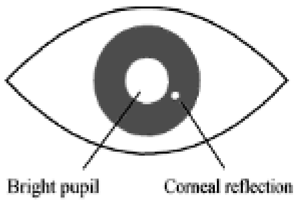

Figure 7. The infrared camera image of corneal reflection and bright pupil.

2.2.1 Design of the eye tracking algorithm

Paul V. & Michael J. presented an effective object detection algorithm by using Haar feature-based cascade classifiers in their paper [4]. Machine learning based this algorithm is a cascade function that is trained with a lot of positive and negative images. Therefore, it can be implemented to detect objects in other images.
The Viola-Jones algorithm utilize from Haar-like features. a scalar product between the image and some Haar-like templates is applied to obtain Haar-like features. The Haar-like features associated with pattern P of image I can be calculated by Eq. 1 where I and P are an image and a pattern, respectively. The images are both in the same size N × N.

$$\sum_{1\le i\le N}\sum_{1\le j\le N} I(i,j)1_{P(i,j)\ is\ white} - \sum_{1\le i\le N}\sum_{1\le j\le N} I(i,j)1_{P(i,j)\ is\ black} \tag{1}$$

In order to perform face detection, firstly the model must be trained with a lot of positive images (images of faces) and negative images (images without faces). Then the features of face must be extracted. Haar-like features are like convolutional kernel. When substracted sum of pixels under white rectangle from sum of pixels under black rectangle, a single value is obtained for each feature [5].

Therefore, many features is calculated by using all possible sizes and locations of each kernel (with 24x24 windows over 160000 features). For each feature account, the sum of the pixels under white and black rectangles are needed to find. In order to solve this problem, they presented integral images. In order to increase algorithm speed, the calculation of the sum of pixels must be simplified since the number of pixels in a process that contains only four pixels.

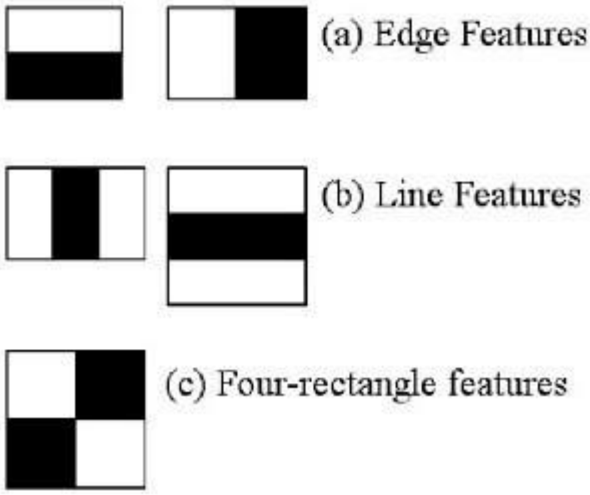

Figure 8. Black-white shapes

There are insignificant features among these features. As seen in the Fig. 9, top rows has two good features. The first features is that eyes regions is generally darker than the nose and cheeks regions. The second feature represent darker eyes than the bridge of the nose. When the same window is applied on cheeks or another place, it become insignificant feature. In order to select useful features, Adaboost machine learning is used [6]. Firstly, the algorithm calculates the best threshold in order to classify the faces to positive or negative. There can be some errors or misclassifications as disadvantage. Therefore, the features must be selected with minimum error rate. So the features characterize the face and non-face images accurately. This process continues until algorithm finds the misclassified images. At the beginning of algorithm, an equal weight appoints to each image. After classification, weight of misclassified images is increased. This process is repeated when the error rate is minimum. Finally, necessary features are obtained.

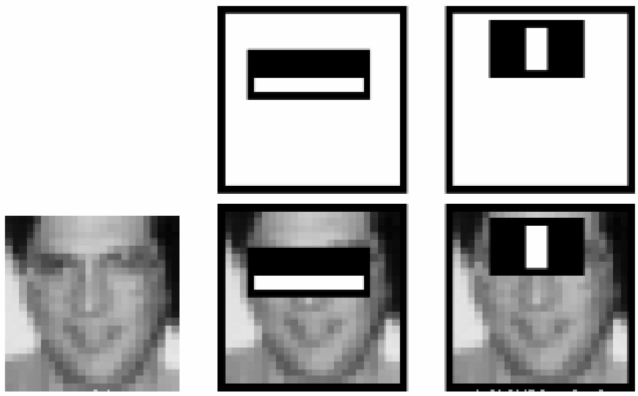

Figure 9. Catching eyes

When a weighted sum of these weak classifiers is combined with the others, they forms a strong classifier. Dan W. H. & Arthur E. C. P. performed detection with obtained 200 features with 95% accuracy. Their final data set had around 6000 features. (they reduced 160000+ features to 6000 features to decrease calculation cost) [7]. In an image, most of the image region belongs to non-face region. A non-face window can detect face region basically. So it is important to focus on the area where is possible face region.

Kim, K. N., & Ramakrishna, R. S. (1999) studied on the concept of Cascade of Classifiers. Instead of applying all 6000 features to a window, group features into different stages of classifiers and apply them individually. (Normally the first few stages consist of far fewer features). If a window fails in the first place, throw it away. This feature is not necessary. If so, apply the second step and continue the process. The window passing through all stages is a face regions [8].

The author's detector had 6000+ features of 38 stages with 1, 10, 25, 25 and 50 features in the first five stages. (The two features in the picture above are actually the two best features from Adaboost). According to the authors, on average, 10 features out of 6000+ per child window were evaluated.

2.3 Image processing and recognition

Image processing and recognition systems have gained popularity today. Thanks to the use of computer cameras, they can process the images and receive from the outside world, and after they translate them into their own understandable language, they can perform the relevant task.

The advantage for designed system is that it recognizes the books as well as the position of the books. By the way, they stand on the shelves to provide a more effective process. However, the image processing system is the slowest and most complex of all recognition processes [3].

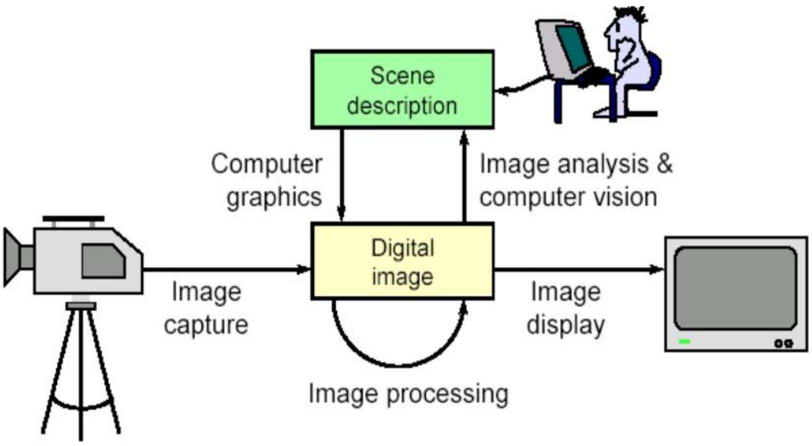

Figure 10. Image Processing Recognition

2.4 Electronic System Design

2.4.1 Drive System of Stepper Motors

In order to keep the stepping motors, the Arduino communication with Matlab was provided. After this communication, the coordinates taken from eye tracking system transferred to the Arduino. On the Arduino, the ready-made 'Nema 17' stepper motor with the library 'Arduino IO'. The data was not sent to stepper motors at the same time since angular motion is not desired.
Therefore, the data is sent to the 'X' and 'Y' motors in short intervals with the calculation of the gradient. diagonal movement is achieved between these coordinates and avoided cornering.

2.4.2 Computer Interface Design

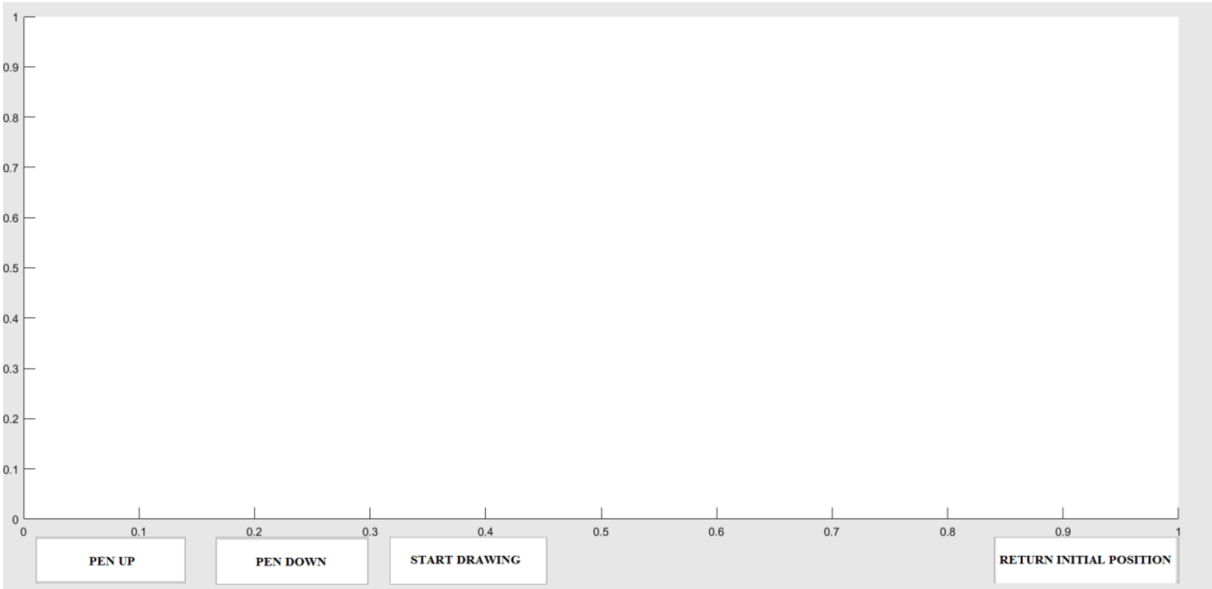

Figure 11. Computer GUI

the interface is designed through Matlab/GUI and used axes to display the graphics in this interface and the coordinates are displayed from here. The pen and button downloads the pen holder at the end of the step motor. If the pen is desired to remove, the pen holder at the end of the step motor is lowered. The start drawing button starts the drawing. The return initial position button allows you to return to the first position.

3. Results

In this study, Cartesian robot control is performed by processing the image taken on this camera using a basic camera. The image taken from the camera was processed to determine the eye region and then the eye was determined by multiplying the determined eye region by the normalized correlation coefficient. The location information of the specified zone has been received. The positional information of the receiver is firstly transmitted to the Nema 17 stepper motors via the IO library via the Arduino and calibrated position data. The system has worked properly with the designed driver system. As a result, when the eyes are focused on a region that are interested on the screen, the Cartesian robot's end function also moves to that region. In this case, it is very suitable to be used only in the areas where dangerous and rapid reaction is required without using any organ other than eye. There is no problem in the system if proper light, proper screen and proper calibration are provided for operation. Eye tracking system is very important for improvement and future. There are many companies that work on eye-tracking and it will be an indispensable point in the future. In the Fig. 12, an sample real time eye tracking of pen and eye tracking coordinates plot is given.

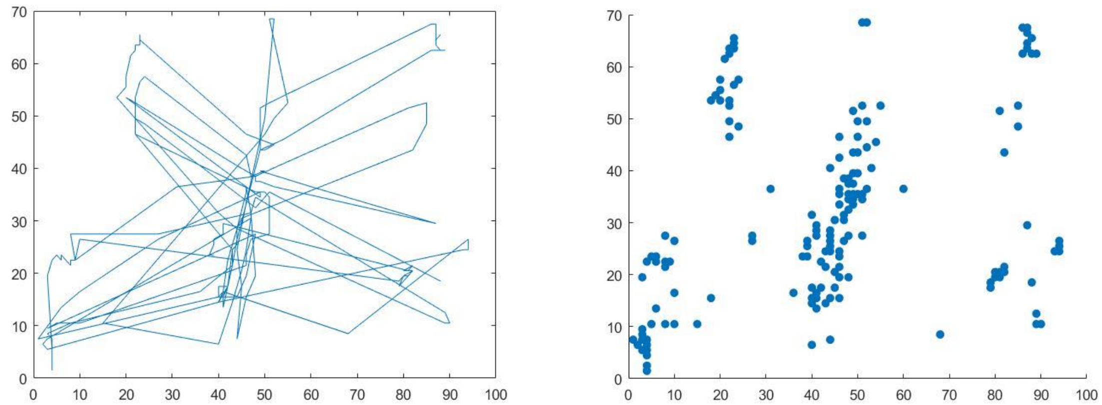

Figure 12. Real time eye tracking of pen and eye tracking coordinates plot

4. Discussions

In this section, some problems and solutions are presented during system testing. First of all, when the data from the camera is processed and the eye region is found, the algorithm used is a slowing algorithm. Also, after finding the eye zone, it is still needed to crop the image that is obtained when working on a single eye. After all this, the resolution of the image is very low. Decreasing resolution also reduces the sensitivity of the data. In order to solve this problem, the eye will be fixed and only the eye target system will increase the speed of the system and more sensitive values will be taken in order to determine the position.

Another problem is that the right eye looks to the right while the left eye moves less while looking to the left. If it is taken reference to the right eye, there is a considerable difference between the motion when looking to the left of the right eye and the motion when you look to the right. This makes differences in the position values that is achieved and so it becomes difficult to reach precious values. In order to prevent this, it must be incorporated both eyes. The position can be obtained more precisely because the eye moves more if taken in reference to the eye in the opposite direction and the same operation is applied to the other eye in the other direction.

The position data can not be sent to the stepping motors while being sent to the motors. Therefore, using the inclination of the position data, motion is provided by sending short step data to the motors. But there is another problem here. It also does not get productive results at sharp points, for example, points where position data changes suddenly. Smaller motors can be used for this purpose.

References

[1] Simon Baker, Ralph Gross, and Iain Matthews. Lucas-kanade 20 years on: A unifying framework: Part 4. Technical Report CMU-RI-TR- 04-14, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, February 2004.

[2] http://2dofpenplotter.blogspot.com.tr/2014/02/2-dof-articulated-pen-plotter-beng.html

[3] Daniel Cheng and Roel Vertegaal. An eye for an eye: a performance evaluation comparison of the lc technologies and tobii eye trackers. In Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications, pages 6161. ACM Press, 2004.

[4] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001.

[5] F. Dornaika and J. Ahlberg. Model-based head and facial motion tracking. In Computer Vision in Human-Computer Interaction: ECCV 2004 Workshop on HCI, Prague. Proceedings, pages 221-232, 2004.

[6] Fadi Dornaika and Jörgen Ahlberg. Active appearance model search with motion compensation for 3d face tracking. In Proceedings of the 4th European Workshop on Image Analysis for Multimedia Interactive Applications (WIAMIS), pages 359-364. IEEE, 2003.

[7] Dan W. Hansen and Arthur E. C. Pece. Iris tracking with feature free contours. In Proc. workshop on Analysis and Modelling of Faces and Gestures: AMFG 2003, October 2003.

[8] Kyung-Nam Kim and R.S. Ramakrishna. Vision-based eye-gaze tracking for human computer interface, 1999.