

Development of interfaces for impaired users

Abstract. This paper presents EyeSEC, a project being conducted in DEE/ISEC in the area of Assistive Technology, which aims to create alternative technologies to compensate for functional limitations and facilitate the independence of elderly and impaired users. The main goal is to merge the data from eye-tracking, provided by vision sensors embedded in eyeglasses, with head movement's data, provided by inertial sensors. This non-invasive interface that uses low cost sensors can be used to drive electric wheelchairs or as an interface with the computer.

Streszczenie. Zaproponowano technologię interfejsu dostosowaną do możliwości użytkowników niepełnosprawnych. Zastosowano czujniki w obudowie okularów umożliwiające śledzenie ruchu oczu. System może być pomocny przy używaniu wózków inwalidzkich, ale także w obsłudze komputera. (Usprawnienia interfejsów dla niepełnosprawnych użytkowników)

Keywords: assistive technology, eye-tracking, inertial sensors, kalman filter.

Słowa kluczowe: interfejs, użytkownicy niepełnosprawni, śledzenie ruchu oczu, czujnik inercyjny.

Introduction

Traditional input devices such as keyboard, mouse or joystick aren't appropriated for all users. Elderly and impaired people, due to their particular limitations, such as head shaking and limited movement of the hands or arms, present strong difficulties using these input devices. Most existing eye-tracking systems do not allow movements of the head, forcing the user to sit still, moving only their eyes. Tests show [2] that users control the head position and use this feature to compensate for small unwanted movements preventing the use of systems that capture the eye position with fixed camera. In order to eliminate the undesirable effects of head movement, the rotation and tilt of the head can be measured adding MEMS gyroscopes and accelerometers. Thus, with an adequate mathematical treatment will be possible to combine the movements of the head with the eye, resulting in an easier man-machine interaction, giving more freedom to the user, and allowing the use of this equipment in larger scenarios. The global project cost is also an important factor because the main objective is to build robust and low-cost equipment, accessible to a larger number of people.

For detection of eye movement the proposed solution uses a conventional webcam with infrared light placed near the user's eye, using software to detect eye movements in head mounted mode. Since the webcam is placed near the eye it is not required high resolution, but the device dimension should be small in order to minimize the impact it will have on the user's field of vision and weight of the prototype. As a support structure for the webcam and inertial sensors it is proposed to use a frame of eyeglasses (Fig. 1). There are several software programs to detect eye movements, but they are too expensive [3] or complex [4] for our proposed application. For that reason, new processing software was developed under the scope of a BSc project running at DEE/ISEC.

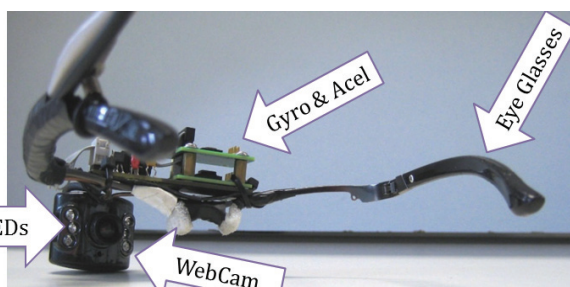


Fig.1. EyeSEC1 prototype with webcam, infrared LEDs, gyroscope and accelerometer MEMS sensors.

Eye-Tracking Software Implementation

This eye-tracking device work by capturing video images of the eye illuminated by an infra-red light source and processing the video frames, outputting the eye's pupil x,y-coordinates relative to the screen being viewed. This eye tracking software for human-machine interaction in real time can be divided into three main stages: image acquisition; image processing and pupil detection. This sequence of image processing is presented in flowchart of Fig. 2.

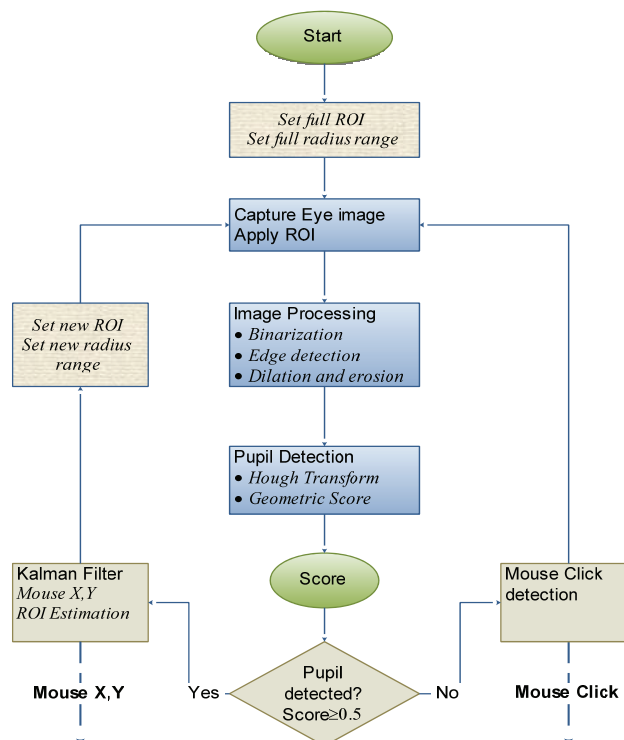


Fig.2. Flowchart for the Eye-Tracking process.

A. Acquisition

The webcam is prepared to capture infrared images minimizing the influence of ambient light variability, and making it possible to obtain images of the eye in the dark. For a daily use of the system the total irradiance level of the IR LEDs should be less than 10mW/cm², so, current LEDs do not produce a potential hazard to the eye according to current safety standards [5].

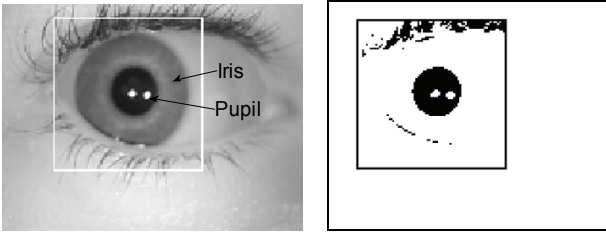


Fig. 3. a) Captured gray-scale image of the eye illuminated by infrared light. The white square represents the ROI (Region of Interest), the area selected to be processed. b) Black and white image resulting from left image processing.

Image capture: The video camera is configured to capture images in gray scale with a resolution of 160x120 pixels (Fig. 3 a). The image resolution can be low because the camera is positioned next to the eye of the user. Frame rate is set up to 10Hz, fast enough for human applications and large enough for software processing tasks.

ROI: We are only interested in pupil detection, so if we can predict its next frame position we can define a Region Of Interest (ROI) around that position (see Fig. 3 a). That way, only the ROI area will be considered for processing. Because it is the most likely area to find the pupil in the next frame, processing algorithms and detection will perform faster improving the final performance. Several examples of ROI application can be seen in the figures delimited by small rectangles.

Similar procedure was applied to radius range used with Hough transform and explained in next section. Knowing the last pupil radius R , which depends mainly from environment light, we can limit the circumference search iterations around a narrow range of values.

When the pupil is not detected during a frame processing due to a noisy image or an eye closing, the ROI and R values are enlarged to capture all possible values in the next frame.

The Kalman algorithm used to filter the extracted position data (x_c, y_c) , and presented in the last section, provides the estimation for the new ROI location and R values (Fig. 11).

B. Processing

Several processing techniques should be applied in order to highlight the contents of an image and reduce the light variability and noise effects.

Binarization: Image binarization converts an image of up to 256 gray levels to a black and white image (Fig. 3 b).

Edge detection: The result of applying an edge detector to an image leads to a set of connected curves that indicate the boundaries of objects that correspond to discontinuities. The gradient is a vector whose components measure how rapidly pixel values are changing with distance in the x and y directions. The gradient of an image $f(x, y)$ at location x, y is the vector:

$$(1) \quad \Delta f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix}$$

and corresponds to convolving the image with the following two masks:

$$(2) \quad G_x = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad G_y = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$$

The gradient magnitude is given by:

$$(3) \quad |\Delta f| = \sqrt{G_x^2 + G_y^2}$$

although typically, an approximate magnitude is computed using:

$$(4) \quad |\Delta f| = |G_x| + |G_y|$$

that is faster to compute. The direction of the gradient can also be defined as follows:

$$(5) \quad \theta = \tan^{-1} \frac{G_x}{G_y}$$

This method, called Roberts method [6], performs fast and is simple to use in 2-D images. This task of edge detection is applied to the black and white image of Fig. 3 b), resulting in a new image where the contours correspond to white pixels, see Fig. 4 a).

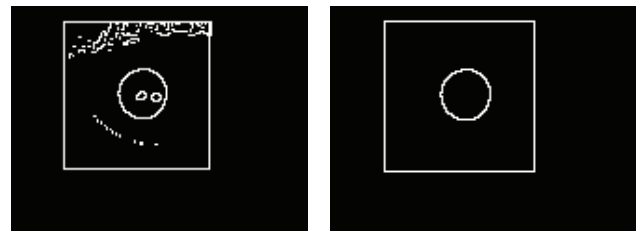


Fig. 4. a) Results of Roberts operator applied to Fig. 3b). b) Results after dilatation and erosion of left figure.

Dilation and erosion: With binary images, dilation connects areas that are separated by small gaps adding pixels to the perimeter of each image object. Erosion completely removes smaller objects and removes perimeter pixels from larger image objects. These two basic morphological operations close well defined curves and provide noise reduction for elimination of small objects. The result can be observed in Fig. 4b).

C. Pupil detection

The pupil is the dark center in the middle of the iris (Fig. 3 a). The pupil determines how much light is let into the eye. It changes size to accommodate the amount of light that is available. The pupil is small in bright light and large in dim light. It appears dark because of the absorbing pigments in the retina. The two main advantages in detecting pupil are that the eyelids do not cover it enabling larger vertical tracking. The second advantage is that the border of the pupil is very sharp. The disadvantages are that the difference in contrast is much lower between the pupil and the iris than between iris and sclera, and it radius changes with light, challenging circle detection.

Due to noise and lighting it is not always possible to get an image with well-defined contours of the pupil in the shape of a circle, this requires that the methods used for detecting circles need to be effective in noisy images.

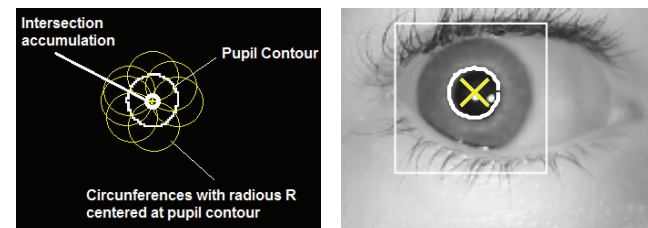


Fig. 5. a) Circular Hough transform for a constant radius R . b) White circumference detected by Hough Transform. The yellow marker 'x' is the centroid of the pupil circle. This image is scored 100%.

Hough Transform: The Hough Transform (HT) was used to detect circumferences because it is a technique that allows recognizing shapes in noisy images or overlapping objects, making it possible to detect the pupil when the eye is not completely open. In order to detect circumferences, a circle-to-point transformation is needed [6]. The parametric description of a circumference is:

$$(6) \quad (x - x_c)^2 + (y - y_c)^2 = R^2$$

where x, y represent the co-ordinates of a circumference pixel, x_c, y_c the center and R the radius.

The process of HT detection of the pupil is very simple and exemplified at Figs. 5 and 6. For each white point of the pupil contour, a circumference with radius R is centered. An array accumulates all the points for all circumferences. If the pupil presents similarities with an R radius circumference, then the point most likely to be the center of the pupil accumulates the largest number of intersections, see Fig. 6a) and 6b).

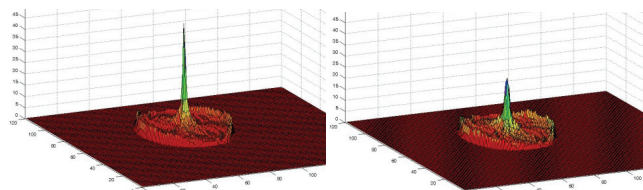


Fig. 6. The local maxima obtained from the accumulator array are the center of the pupil circumference from the image of Fig. 4b). a) On left, the results for HT with $R=13$ and b) on right, for a radius $R=12$. The difference is evident and the choice is clearly $R=13$ from left figure.

Geometric Scoring: Every match is scored. The candidate pupil with maximum accumulated score is chosen as the final solution. The coordinates (x_c, y_c) are the center of the pupil.

In order to reduce false detections a second geometric procedure is applied to validate and score the detection. From the result of the previous procedure we get the radius R , the area A and the center position of the detected pupil. The whole scoring algorithm is described in Fig. 7.

The resulting score will be used in the Kalman filter for the measurement covariance matrix R_k that models the certainty in the measurements (Eq. 12).

A. Pupil detection results

The accuracy of pupil detection in the eye region is largely affected by the accuracy of the image edge detection. If there are many round edges that do not belong to the pupil, the Hough Transform algorithm might get confused and cause a detection error. A detection error is defined as any detected eye center with an accumulated count lower than a defined threshold. If higher than a given threshold, the detected object presents enough similarity to be accepted at least with a 50% score.

Fig. 5 b) presents a successfully detected pupil eye image whereby the crosshair is located at the centre of the eye. This is the final result for Fig. 3 a) image processing. HT detects maximum accumulation for $R=13$ (Fig. 6a) and the pupil contour detected in Fig. 4b) presents a similarity metric=0.96 so, the pupil is full scored with 100%.

The next example is for a partially covered pupil. The eyelashes over pupil fragment the edges and they become more difficult to detect resulting on incomplete eye edges (see Fig. 9a and 9b).

The procedure of edge detection creates several small objects removed by erosion (Fig. 9b). The circumference of the pupil will not close completely preventing the

implementation of geometric algorithms previously mentioned. Therefore, it can be classified by the HT method, obtaining a 50% score.

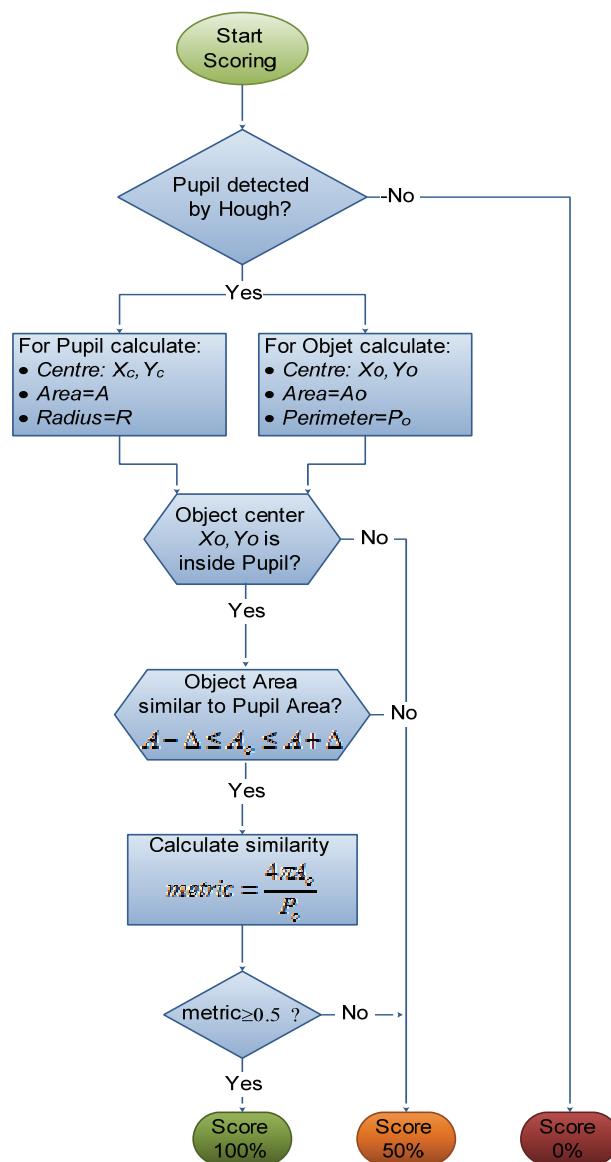


Fig. 7. Pupil detection score algorithm.

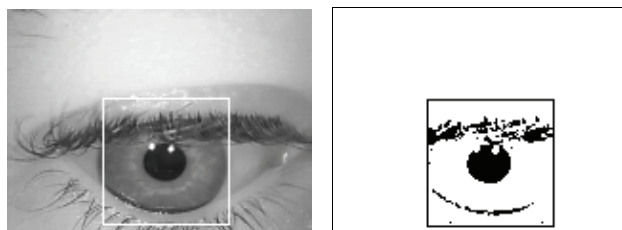


Fig. 8. a) Pupil partially covered by the eyelashes. b) As result, the back and white image is noisy.

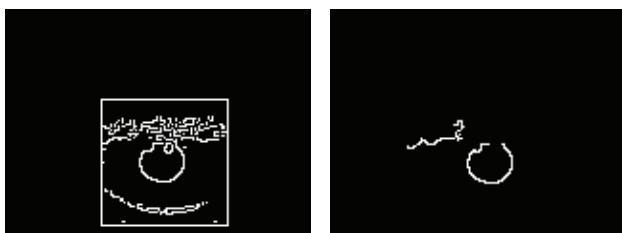


Fig. 9. a) Edge detection. b) Small object removed.

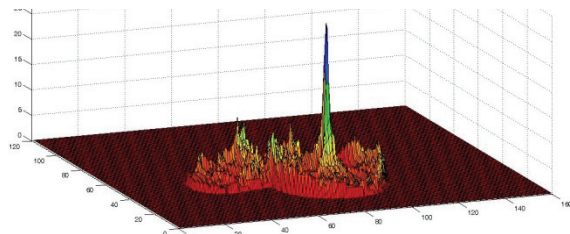


Fig 10. HT applied to the image of Fig. 9b). The pupil is partially covered by eyelashes that do not allow the closure of the circular object, although it is clearly detected by the HT transform.

Even in the presence of noisy and open circumferences, the Hough transform is effective (Fig. 10). As the area of image processing is restricted by the predicted ROI, the probability of finding another circumference, even without geometric proofs, is very small. Moreover, these results are applied to a Kalman filter that, according to the model will avoid large trajectory deviations.

Table 1 shows the eye detection rate of errors for different lighting scenarios. We obtained recognition rates of over 90% in normal ambient light or at night where the eye is illuminated by infrared. In the presence of direct sunlight, the worst condition for image processing, the average recognition rate decreases slightly.

Table 1. The pupil detection error rate

Score	Total Frames	Negative		Positive			% Accuracy
		True	False	True	False	True	
		0%	0%	100%	50%	100%	
Test A sunlight	100	13	19	66	2	0	81%
Test B Sunlight	100	20	19	52	6	3	78%
Test C indoor	100	2	2	94	2	0	98%
Test D Night & light	100	17	7	39	37	0	93%
Test E night	100	0	0	84	16	0	100%

The Kalman filter was introduced in order to prevent false detections that arise from bright light conditions with the advantage of tolerating small occlusions of the eye.

Kalman Filtering and Tracking

Visual tracking can be described as the process of determining the location of a feature in an image sequence over time. Various factors such as lighting and occlusions can affect the appearance of the target, thus making accurate tracking difficult. Despite the infrared immunity to the ambient light, variation in illumination, and partial or full occlusion of the pupil contribute to the position measurement error when tracking a target. The Kalman filter has been used successfully in different prediction applications or state determination of a system [7, 8, 9] and one advantage of the Kalman filter for the vision tracking is that it can be used to tolerate small occlusions.

In our preliminary experiment we are considering an application as interface with the PC. In this example the objective is to follow an object in the screen using only the eye-tracker extracted information.

The Kalman filter model requires a representation of the tracker in state space. In this application gaze tracker is used for tracking an object in a PC screen with a linear uniform movement, which could be described by the following system equations:

$$(7) \quad X_{k+1} = F_k X_k + W_k$$

$$(8) \quad \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ vx_{k+1} \\ vy_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ vx_k \\ vy_k \end{bmatrix} + W_k$$

where two obvious parameters are the x and y coordinates of the tracker. In addition, the state will include the velocity components vx and vy . W_k is the process noise which is assumed to be drawn from a zero mean multivariate normal distribution with covariance Q_k

$$(9) \quad W_k \sim N(0, Q_k)$$

At time k an observation (or measurement) Z_k of the true state X_k is made according to

$$(10) \quad Z_k = H_k X_k + V_k$$

$$(11) \quad \begin{bmatrix} Z_{xk} \\ Z_{yk} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ vx_k \\ vy_k \end{bmatrix} + V_k$$

where V_k is the observation noise which is assumed to be zero mean Gaussian white noise with covariance R_k :

$$(12) \quad V_k \sim N(0, R_k)$$

The Kalman filter algorithm provides recursively an estimate $X_{k|k-1}$ in terms of the previous estimate $X_{k-1|k-1}$ and the most recent observation Z_k . It involves a cycle of prediction, observation, and updating (there is often an additional data validation step prior to updating). The flowchart in Figure 11 summarizes the Kalman filter process. In the last step is calculated the estimate that will be used to determine the ROI position in the next frame.

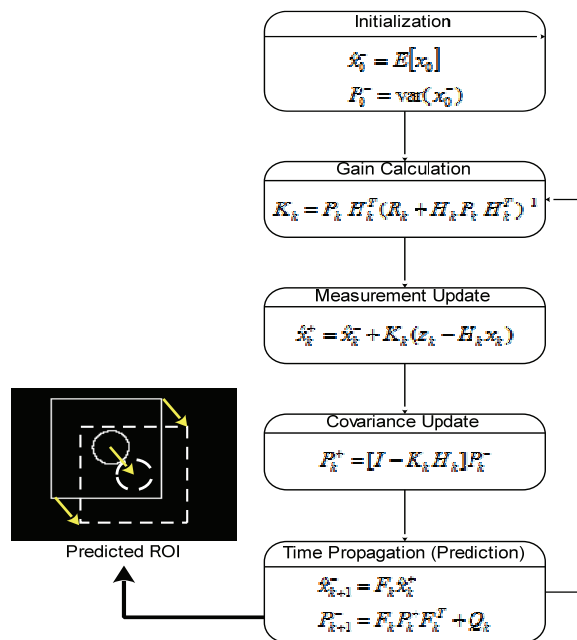


Fig. 11. Discrete Kalman filter equations.

A. Experiments

To validate the eye tracking system, a trajectory was planned on a computer screen. As the image moved along

this path at a constant speed, the user must follow it with his eyes wearing the eye-tracker.

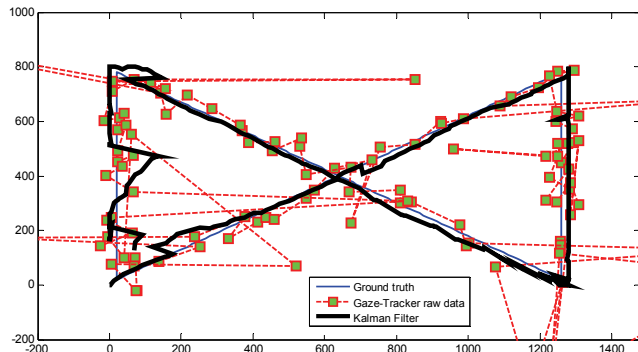


Fig 12. The blue line represents the ground truth. The eye was pointing to a circle moving from left-down screen corner to right-top corner, and looping back at a constant speed. Green squares represent raw data measurements. The black line represents the Kalman Filtered data.

The first tracking experiment was carried out using a third party open source software. Fig. 12 presents the filter results for a loop controlled by eye-tracking movements on a 1280,800 resolution screen. As can be observed, the raw data is very irregular in time, highly noisy and frequently jump off the screen.

The next experiments were carried out on as PC interface using our developed software EyeSEC.

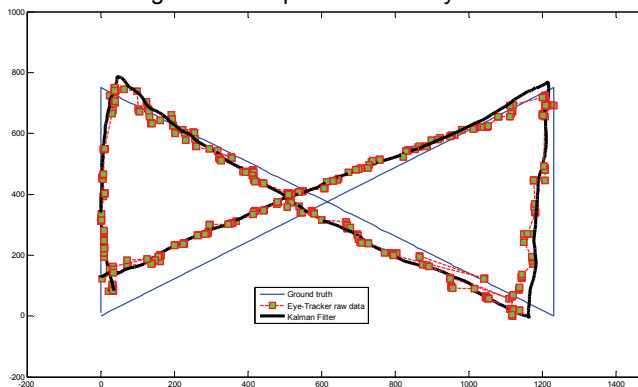


Fig 13. Blue line represents the ground truth. The eye was pointing to a circle moving from left-down screen corner to right-top corner, and looping back at a constant speed. Green squares represent EyeSEC Tracker software raw data measurements. The black line represents the Kalman Filtered data.

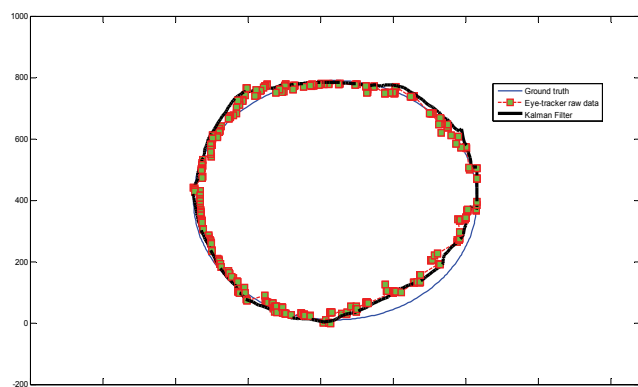


Fig 14. The Blue line represents the ground truth. The eye was pointing to a small circle describing a circumference in the screen at a constant speed. Green squares represent EyeSEC Tracker software raw data measurements. The black line represents the Kalman Filtered data.

The obtained results show a significant reduction of tracking errors, due to improved image processing

algorithms optimized and adapted to the human eye. Despite the improvements in the proposed approach it is still necessary to add multiple sensors.

Conclusion and Future Work

Noninvasive methods for vision based eye tracking that involve the monitoring of the pupil movements were investigated in the conducted work. The results obtained from the proposed EyeSEC prototype demonstrate its ability as an alternative to traditional computer interfaces. More robust experience with different users and different environments are essential to validate these techniques before to be employed as a man machine interface.

The performed experiments so far have used only the information computed by finding correspondences between the pupil and a circle. Improvements were expected, as soon as data provided by the gyroscope and accelerometer are integrated and merged in processing, making the system independent of head movement.

New biological inspired Kalman models suited to merge information from the multiple sensors will be tested, improving the performance of the final tracker. The new system intends to be lighter by using a micro-camera and small inertial sensors embedded in the eyeglass frame. Processing will be simplified and local, with wireless communication between the frame and the master system resulting in a practical portable system.

Besides being able to be used as a mouse, it is expected that this device will be used in a broader range of assistive technologies, such as, an electric wheelchair or autonomous robot assistance and we believe that our EyeSec eye-tracker will give people with disabilities not only the advantages of an interface with a computer, but also a tool to support and expand their participation in society.

REFERENCES

- [1] Moita F., Oliveira R., Santos V., and Silva M., "EyeSec Project, Development of interfaces for impaired users", 12th Portuguese-Spanish Conference on Electrical Engineering, XIIIEEE, Ponta Delgada - Azores, Portugal, July, 2011.
- [2] G. S. Agustin, J., Skovsgaard, H., et al, *Evaluation of a low-cost open-source gaze tracker*, in Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, p. 77-80, 2010, Austin, Texas.
- [3] At: <http://www.tobii.com> (2010)
- [4] ITU Gaze Tracker, *Open source eye tracking and gaze interaction Project*, Retrieved September, 2010, from <http://www.gazegroup.org>
- [5] Mulvey, F., Villanueva, A., Sliney, D., Lange, R., Cotmore, S., Donegan, M. (2008) *D5.4 Exploration of safety issues in Eyetracking. Communication by Gaze Interaction (COGAIN)*.
- [6] Bong D. B. L., Lim K. H., "Application of Fixed-Radius Hough Transform in Eye Detection", Int. Journal of Intelligent Information Technology Application, 2(3), 121-127, 2009.
- [7] Cuevas, E., Zaldivar, D., & Rojas, R. (2005). *Kalman filter for vision tracking.Measurement*, (August), 1-18.
- [8] Bozic, S.M. (1994). *Digital and Kalman filtering*. Butterworth-Heinemann.
- [9] Bento L., Nunes U., Moita F. and Surrício A., "Sensor Fusion for Precise Autonomous Vehicle Navigation in Outdoor Semi-structured Environments" ITSC '05 - 8th International IEEE Conference on Intelligent Transportation Systems, Vienna, Austria, 13 16 September 2005.

Authors: Fernando Moita, E-mail: moita@isec.pt; Rúben Oliveira, E-mail: rsoliveira@live.com.pt; Victor Santos, E-mail: vsantos@isec.pt; Marco Silva, E-mail: msilva@isec.pt Polytechnic Institute of Coimbra, Institute of Engineering of Coimbra, Dept. of Electrical Engineering, Rua Pedro Nunes - Quinta da Nora, 3030-199 Coimbra, Portugal