ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2010. No. 5(101)
ELEKTRONIKA IR ELEKTROTECHNIKA

AUTOMATION, ROBOTICS
T 125
AUTOMATIZAVIMAS, ROBOTECHNIKA

Autoassociative Gaze Tracking System based on Artificial Intelligence

T. Proscevičius, V. Raudonis, A. Kairys, A. Lipnickas
Department of Control Technology, Kaunas University of Technology,
Studentų str.48, LT-51367 Kaunas, Lithuania, phone: +370 602 10302, e-mail: tomas.proscevicius@gmail.com
R. Simutis
Department of Process Control Technology, Kaunas University of Technology,
Studentų str.48, LT-51367 Kaunas, Lithuania, phone: +370 37 300296, e-mail: rimvydas.simutis@ktu.lt

Introduction

The establishment and creation of the new technology in the world is developing very quickly. Researchers of all over the world are working on solutions for people with disabilities, aiming to improve living quality and make daily tasks easier. The artificial intelligence [1, 2] is increasingly used with newest devices and technologies. The newly developed device allows solving specific tasks and helping people with specific needs.

Recently the attention was drawn to development of robust and low price automated gaze tracking systems [3–7] for disabled people. The design of such systems includes a wide area of research. Disabled peoples, with high level of disability, usually can use their eyes and tongue. For them it becomes the only natural and effective way to communicate with the external world. The voluntary controlled eye movements can be used to create human - machine interface [8–11] as assistive technique that can be synchronized to control any connected technology. Such technology might be connected to a standard computer, and it will enable to control mouse cursor by eye movements. It is expected that in the near future such systems will be used in cash machine transactions and area protection (airports, laboratories, stores). It also can be used for computer authorization in computer systems, etc.

Many systems that analyze eye movement are often intrusive or use specific hardware, i.e. acquisition cameras and active infrared illumination. In this paper the system of single gaze tracking based on usual USB camera and mapping to computer mouse position on a screen is presented. The use of such system enables impaired user to accomplish such common tasks like text typing, web browsing, message sending and etc. In this paper a relatively inexpensive, real time system for gaze tracking is proposed as a human- computer interface. The system can therefore be run on any available computer without the need for expensive equipment. Ordinary USB webcam is used in the system and it is attached to spectacle frame in order to provide more comfort for the user. User is not additionally loaded with equipment as there are no ancillary devices in this system.

In presented system the mouse cursors on the computer screen is moved synchronously with the eye movement. Cursors control is done by continuous mapping of the eye gaze to the computer screen plane. System is based on effective image processing and application of numerical methods for eye gaze tracking and mapping for various applications. Operation of algorithm consists of three stages: calibration, training and application. First of all the system is calibrated by the user thought observing 9 calibration points on the screen. Each calibration point consists of number of gray scale images obtained by ordinary web camera. After that the pictures are down scaled and used to find characteristic points by principal component analysis. Achieved components are used to train the radial basis function network to map recorded eye gaze positions to calibration positions on the screen. Cared on research of eye tracker usability for computer application control is presented and analyzed in this paper.

Background of Principal component analysis

Not all pixels in frame are informative and important for gaze analysis. Principal components' analysis [12, 13, 14] (PCA) is usually used to reduce amount of image data. PCA is a statistical procedure that consists of covariance structure of variable data analysis. It allows determining main variation directions of the data. It is considered that a matrix **A** consists of elements $n \times n$ [10]. Eigen values are determined as a root of equation

$$(\mathbf{A} - \lambda \mathbf{I}) = \left| (\mathbf{A} - \lambda \mathbf{I}) \right|, \qquad (1)$$

where I – the identity matrix of size $n \times n$.

This kind of equation is called the characteristic equation and has n roots. Let's take λ as Eigen value of **A**. Then there is a vector **x** that satisfy the equation

$$\mathbf{Ax} = \lambda \cdot \mathbf{x}. \quad (2)$$

Vector **x** is called the Eigen vector of **A** matrix and is associated with Eigen value λ. Usually **x** is normalized and has a value 1, which $\mathbf{xx}^T = 1$ **Error! Bookmark not defined.** [15]. Let's say, that Σ is the covariance matrix of $n \times n$. Therefore there exists the orthogonal matrix **Φ** of size $n \times n$ (columns of the matrix are the Eigen vectors) and diagonal matrix **Λ**. The diagonal elements of **Λ** matrix are Eigen values of Σ**Error! Bookmark not defined.** **Φ** , such as

$$\mathbf{\Phi}^T \cdot \sum \mathbf{\Phi} = \mathbf{\Lambda}. \quad (3)$$

The matrix of Eigen values **Φ** can be taken as a linear transformation, which, for example, the Fig. 2, transforms the axes system *X-Y* to *U-V* axes system

$$p_{uv} = \mathbf{\Phi} \cdot p_{xy}. \quad (4)$$

The principal component analysis was applied to the eye images. User was asked to observe nine calibration points, which were displayed on the computer screen (see fig. 4). The unique projection values are obtained for each class of the eye images, by projecting each eye image to the first three principal axes. The resulting characteristic point's clouds are shown in the Fig. 1. First principal component defines the largest variation in the raw data and any other component describes lesser variation.

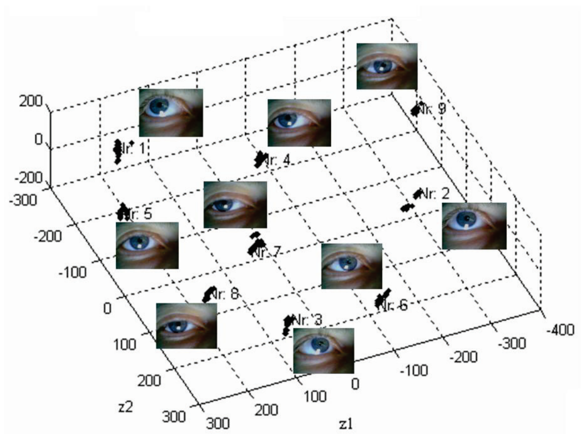

Fig. 1. Nine eye classes plotted on the feature space

Each eye class can be divided by hyper plane, which is obtained using artificial neural network. The distribution of the projection values depends mainly on the image resolution and illumination conditions. More about the ANN is presented in the next section.

The neuron network of the radial base functions.

Artificial neuron networks [16] are one of the numerical intelligence examples, they are made from lots of simple calculation elements that are connected to each other in order to imitate biological processes of the brain. Neuron network (NN) and especialy Radial base functions network (RBF) [17, 18] is used in proposed algorithm. Structure of such RBF is one of the simplest and poses a single hidden layer of radial functions. Since these functions are nonlinear, it is not actually necessary to have more than one hidden layer to model any shape of function: sufficient radial units will always be enough to model any function. For our proposed system the RBF has an output layer containing points with (x, y) of monitor coordinates as identity activation function. The RBF training is done using collected historical data. The main advantage of use of the radial base neuron network is fast training procedure and that the number of selected radial functions is much smaller than the number of available data.

The effectives and the accuracy of the proposed gaze tracking system depend on the results of the principal component analysis and on the structure of the trained RBF network. The optimal structure of the RBF is acquired by using heuristic method. The structures of the network were changed as shown in the table 1 during experimental investigation.

Table 1. Neural network structure

Neural network	Spread of RBF function	Number of radial function	NN training time, s	Recognition quality, %
3	0.1	900	0.9	78
	0.0001	500	1.2	81
8	0.1	900	1.8	85
	0.0001	500	2.2	90

Gaze tracking hardware

Researches are made with the "of the shelf" technical equipments and software. The DPC-4210M micro CCD colour video camera, it capture Pixels NTSC: 510x492, PAL: 500x582. Camera Size 0.5 (L) x 0.5 (W) x 0.75 (H) Inch [13 x 13 x 19 mm].

Data from camera (Fig. 2) is transferred to computer through standard USB connection (Fig. 3). Additional illumination is used to improve the received data quality and to decrease the influence of external illumination. Led supply +5 V.

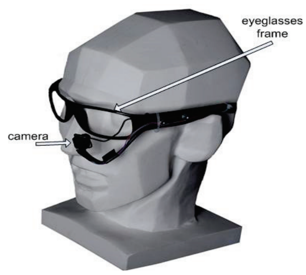

Fig. 2. The gaze tracking system

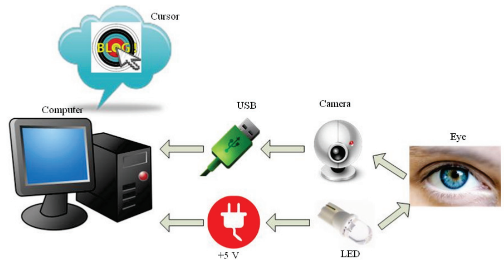

Fig. 3. The functional diagram of the system

Calibration of the system

Data selection is necessary in order to precede calibration between eye direction and mouse cursor. Eye images is collected when user was observing the calibration points, which are shown automatically. Calibration points are shown in Fig. 4. Follow – up points on the monitor system is calibrated. For the calibration of the system 9 monitor points are used (8 points are gradually situated on the borders of monitor and one is in the centre of it).

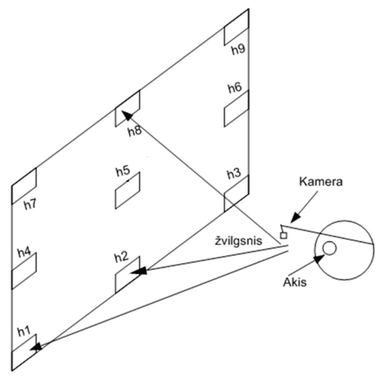

Fig. 4. Calibration points displayed on the monitor screen

Gaze tracking algorithm

In particular, the reduced amount of data is used. This reduction of environmental impact system (lighting change mitigation). Physiology of the eye of the user establishes a working algorithm to compensate for the blinking effect. In order to create precise and stable gaze tracking algorithm, it is necessary to explore eye physiology. When literature was reviewed it was found that the grown-up involuntarily blink up to 15 times a minute, while a child does that up to 10 times a minute [19]. The frequency of blinking depends on fatigue, attention, stress, excitation, sleep duration, etc. Eye motility is made of such types eye movements like: sakado (fast jumps of high and low amplitude from 1° to 90°, speed up to 600 °/s), drift (slow movements, speed is up to 6°/s, amplitude – 3–30 angular seconds), tremor (small vibrations with the amplitude of – 5–15 angular seconds, frequency – 20–150 Hz) [20]. Also selected neuronal network required amount on which the system is satisfied.

Each eye image is transformed to the vector $\Gamma = (x_1, x_2, ..., x_K)$, K- number of data samples.

Algorithm for training of the system

1 step *collect K frames data base of i=1,...,n eye positions by looking to calibration monitor (mouse position)*

$\Gamma = (x_1, x_2, ..., x_K)$;

2 step *close eye and collect k frames of n+1 eye position data*;

3 step *captured data reduction to appropriate size of NxM pixels*;

4 step get parameters for PCA

$C \Leftarrow \hat{G} \Leftarrow \Gamma$;

5 step compute eigenvectors

$\Omega_{tamplate} \Leftarrow \upsilon, \lambda \Leftarrow C$;

6 step *compose data matrix of all n+1 eye position*

7 step *get data sample from test data set*

$\Omega_{new} \Leftarrow \upsilon^T \times G_{new}, G_{new} \Leftarrow \Gamma^i_{test}$;

8 step *train neural network with data from step* 5 *to identify all eye positions*.

Algorithm for online application

1 step *apply trained ANN to manage mouse cursors position according eye position in the frame;*

2 step *mouse click*

If close eye=t_1 here $t_1 = 3s$;

3 step *system turn off;*

if *close eye=t_2 here* $t_2 = 10s$;

end

4 step *acquire new frame and return to the step 1*

Experimental setup and result

System was calibrated using 9 points displayed on the computer screen (Fig. 4), while 10 measurement points were used for testing. User was asked to observe 10 test points, which were displayed in the circle time at the time (Fig.5). Each training point was captured 100 times and each testing point- 60 times. Frame size that was used either for training either for testing is 20x20 pixels. Training results are shown in Fig. 7 to 9. The dispersion of training data set is shown in Fig.s (6, 8) and dispersion of testing data set is shown in Fig.s (7, 9). Results are discussed under the Figs. 5–9.

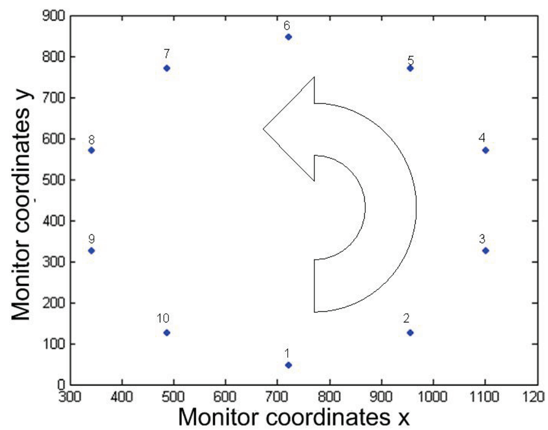

Fig. 5. Scheme tested to collect data

The human and machine parameters of autoassociative system were chosen using the trial and error method, that is, structure and parameters of neuron network and parameters for interpolation of the measuring space. Using PCA informative features from 3 to 8 were selected. Fig. 6 presents results using the RBF network with 100 base units and 3 characteristic features. In the Fig. 6 the dispersion of calibration points is shown and Fig. 7 previews the dispersion of testing points. Results show, that neuron network with such structure and three characteristic points poorly interpolated calibration data, so the results of testing data set are insufficient.

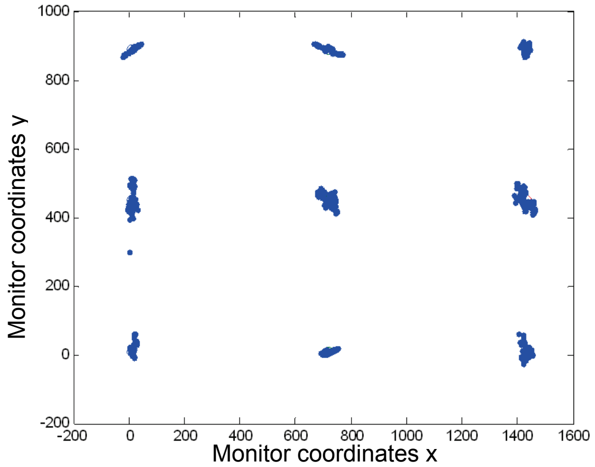

Fig. 6. Training data distribution around, the 3 PCA

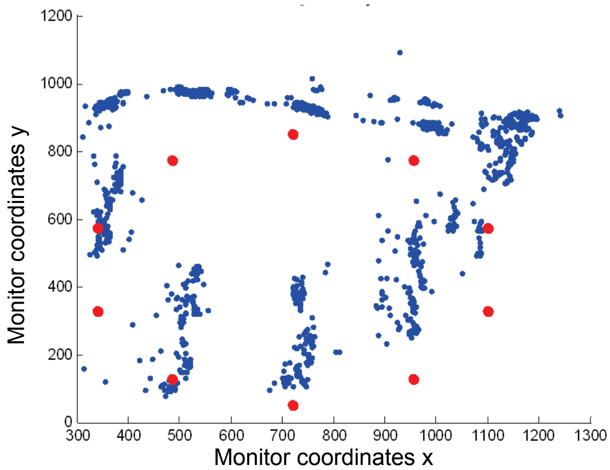

Fig. 7. Distribution of test data variance, with 3 PCA

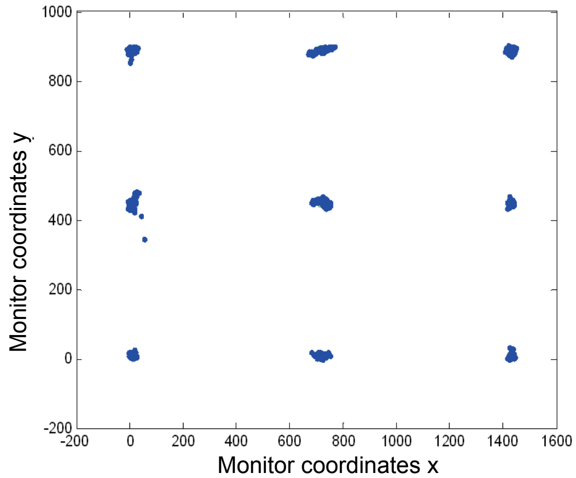

Fig. 8. Training data distribution around, the 8 PCA

Fig. 8 presents results using 500 base nodes and eight characteristic features. Dispersion of calibration points is shown in Fig. 8 while the dispersion of testing points is shown in Fig. 8. Testing points are compactly dispersed and the dispersion is determined by illumination changes as well as physiology of a human eye. Illumination changes occur because user cannot be in stationary state as he moves involuntary.

Results gathered using 500 base nodes and 8 PCA are shown in Fig. 9. As it is seen in test results, magnifying the quantity of characteristic features doesn't show any quality improvement over testing data, though calculation time increased significantly. Conclusion is that 500 base nodes and eight characteristic features are enough for system structure creation.

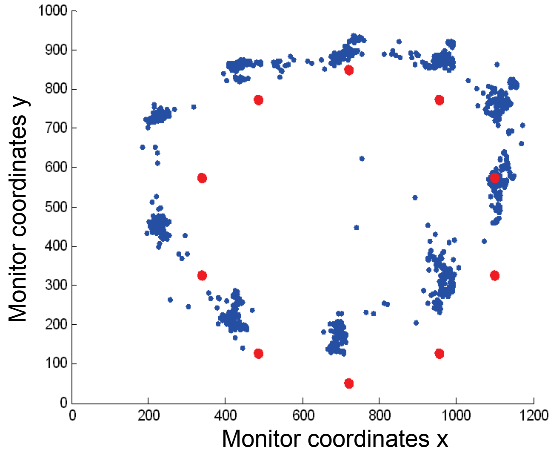

Fig. 9. Distribution of test data variance, with 8 PCA

Future work

In the future the most topical problem to solve is how to compensate the turn of the head. Always moving user calibrates away the system. This problem is very topical because a lot of gaze tracking systems created all around the world requires the user to hold up his head and try to keep it in a straight position. In the paper principles, which can compensate the turn of the head, are proposed. These principles can provide the user with comfortable conditions letting the user use this system.

In this way, Fig. 10 provide the view of one additional camera attached to spectacles.

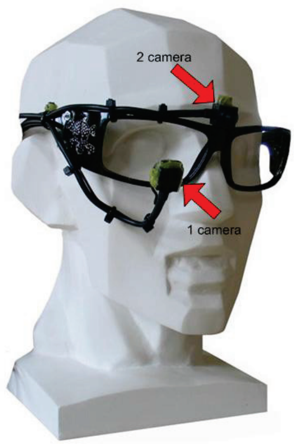

Fig. 10. The head rotation compensation device

This camera will be pointed to the screen used by the user. So the first camera tracks the position of user eye while the second one traces the screen. Each of four corners of the screen has one thermal diode attached and the second camera fixates them. Since thermal diodes are attached stationary, their position will not change. Only the position of the users head or the distance between the screen and the second camera will change. By knowing these parameters, the turning of the users head can be compensated.

Conclusions

In this paper relatively inexpensive and easily adoptable auto associative system, which lets the user to communicate with computer using the gaze, is proposed. The prototype of the system is created and the system is calibrated in order to control the position of the mouse cursor in the computer screen. The structure of radial base neuron network used for the interpolation of computer screen points is determined experimentally. Sufficient amount of required main features is determined and given in the paper as well. Accomplished experiments allow to state that the neuron network of radial base functions can be used for the interpolation of screen points. The system can be calibrated for the cursor control of computer mouse in the screen and that allows the user to control various software applications. The system works in real time so it can provide the disabled people with bigger self-sufficiency and can help while solving the problem of integration into society.

Acknowledgment

The authors appreciate the financial support of the Lithuanian state science and studies foundation.

References

1. **Paukštaitis V., Dosinas A**. Pulsed Neural Networks for Image Processing // Electronics and Electrical Engineering. – Kaunas: Technologija, 2009. – No. 7(95). – P. 15–20.
2. **Lipnickas A.** Adaptive committees of neural classifiers // Information technology and control. – Kaunas: Technologija, 2008. – Vol. 37, no. 3. – P. 205–210.
3. **Raudonis V., Simutis R., Narvydas G.** Discrete eye tracking for medical applications // Applied Sciences in Biomedical and Communication Technologies, 2009. – ISABEL'2009, 2nd International Symposium, 2009. – P. 1–6.
4. **Long X., Tonguz O. K., Kiderman A.** A High Speed Eye Tracking System with Robust Pupil Center Estimation Algorithm // Engineering in Medicine and Biology Society, 29th Annual International Conference of the IEEE, 22–26 Aug., 2007. – P. 3331 – 3334.
5. **Lipnickas A., Knyš A.** A Stereovision System for 3-D Perception // Electronics and Electrical Engineering. – Kaunas:Technologija, 2009. – No. 3(91). – P. 99–102.
6. **Lenisa M., Wentzel D., Malinowski A.** Data mining for gaze tracking system // Human System Interactions. - Heidenburg, 2008. – P. 680–683.
7. **Špakov O., Miniotas D.** Processing Gaze Responses to Dynamic 2D Stimuli // Electronics and Electrical Engineering. – Kaunas:Technologija, 2008. – No. 5(85). – P. 89–92.
8. **Miniotas D., Laurutis V.** Relationship between Movement Speed and Accuracy during Eye Gaze Control of Computer Cursor // Electronics and Electrical Engineering. - Kaunas: Technologija, 2000. - No. 2(25). – P. 63–66.
9. **Ramanauskas N.** Calibration of Video-Oculographical Eye Tracking System // Electronics and Electrical Engineering. - Kaunas: Technologija, 2006. – No. 8 (72). – P. 65–68.
10. **Serafinavičius P.** Estimating Characteristic Points of Human **Body** for Automatic Hand Pointing Gesture Recognition // Electronics and Electrical Engineering. Kaunas: Technologija, 2007. – No. 8(80). – P. 83–86.
11. **Porta M., Ravelli A.** WeyeB an Eye – Controlled Web Browser for Hands. – Free Navigation. Human System Interactions (HSI '09), 2nd Conference on 21–23 May, 2009 Catania, Italy. – P. 210 – 215.
12. **Mittal N.; Walia E.** Face Recognition Using Improved Fast PCA Algorithm. Mage and Signal Processing, 2008. CISP '08. Congress on Volume 1, 27 – 30 May 2008. – P. 554 – 558.
13. **Olechnovič J.** Principal components analysis Human face recognize**.** Vilniaus Gedimino technikos universitetas Lithuanian **Young** scientists conference „Study – Lithuanian future", Vilnius 2007 04. 12. – P. 252 – 259.
14. **Smith L. I**. A tutorial on Principal Components Analysis. 2002. 27. – P. 189–193.
15. **Theodoridis D. C., Christodoulou M. A.** A New Neurone – FDS Definition for Indirect Adaptive Control of Unknown Nonlinear Systems Using a Method of Parameter Hopping Boutalis // Neural Networks, IEEE Transactions. – Vol. 20, iss. 4, April 2009. – P. 609–625.
16. **Dragašius E., Lipnickas A.** Intelligent Mechatronic Systems. – Vilnius: Pedagogika, 2008.
17. **Bishop C. M.** Neural Networks for Pattern Recognition. Institute for adaptive and Neural Computation Division of Informatics Edinburgh University.
18. **Wang C., Lang J., Xiang W.** RBF neural network image reconstruction for electrical impedance tomography // Proceedings of Machine Learning and Cybernetics, 2004. – Vol. 4.. – P. 2549–2552.
19. **Michael J.** Consideration of Three Types of Spontaneous Eye blinks Activity in Normal Humans: during Reading and Video Display Terminal Use, in Primary Gaze, and while in Conversation // Optometry and Vision Science. – Glasgow, Scotland 2001. – No. 10(78) – P. 712–725.
20. **Pivrikas V.** Vestibulometrija. – Klaipėda: S. Jokužio publishing printing, 2006. – 12 p.

Reiceived 2010 02 15

T. Proscevičius, V. Raudonis, A. Kairys, A. Lipnickas, R. Simutis. Autoassociative Gaze Tracking System based on Artificial Intelligence // Electronics and Electrical Engineering. – Kaunas: Technologija, 2010. – No. 5(101). – P. 67–72.

Real-time auto-associative interface between the users gaze and a computer is presented and analyzed in this paper. This interface can provide more self-sufficiency to the disabled person and may help them while dealing with the problem of public integration. That is the reason for creation of real time and free positioning gaze tracking system which is applicable to control the computer application. The gaze tracking precision, computer processing rate and robustness of the system were explored experimentally. The artificial neuron network method and principal components analysis are used in the presented system for the user gaze and computer screen auto-association. The applied methods reduce the amount of the received video data by filtering out unimportant information either reduce the total computation burden of the system. Proper structure of neuron network and the number of the principal components were estimated through heuristic approach. The presented system of gaze tracking was tested with computer applications in real-time. Ill. 10, bibl. 20, tabl. 1 (in English; abstracts in English, Russian and Lithuanian).

Т. Просцявичюс, А. Кайрис, В. Раудонис, А. Липницкас, Р. Симутис. Автоассоциативные взглядом системы слежения на основе искусственного интеллекта // Электроника и электротехника. – Каунас: Технология, 2010. – № 5(101). – С. 67–72.

В этой работе представлен и анализирован автоассоциативный интерфейс между взглядом пользователя и монитором компьютера, работающий в реальном времени. Этот интерфейс может предоставить больше самостоятельности паралитикам и инвалидам, помогая им при решении проблемы общественной интеграции. Именно поэтому создана система отслеживает взгляд пользователя для управления компьютерными прикладными программами. Точность отслеживания взгляда и быстъродействие системы было экспериментально проверена. Для анализа данных, в системе использована неровная сеть и анализ главных компонентов. Метод анализа главных компонентов уменьшил количество полученных видеоданных путем фильтрации ненужной информации. Структура нейронной сети и число основных компонентов были получены путем эвристического метода. Представлена система отслеживания взгляда была проверена с компьютерными программами в режиме реального времени. Ил. 10, библ. 20, табл. 1 (на английском языке; рефераты на английском, русском и литовском яз.).

T. Proscevičius, V. Raudonis, A. Kairys, A. Lipnickas, R. Simutis. Autoasociatyvioji stebėjimo žvilgsniu sistema pagrįsta dirbtiniu intelektu // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2010. – Nr. 5(101). – P. 67–72.

Pateikiama ir analizuojama vartotojo žvilgsnio ir kompiuterio realaus laiko autoasociatyvioji sąsaja. Ji suteikia neįgaliajam daugiau savarankiškumo, padeda integruotis į visuomenę. Sukurta realaus laiko nepozicionuojama žvilgsnio stebėjimo sistema, kuri taikoma kompiuterinėms programoms valdyti. Žvilgsnio sekimo tikslumas, kompiuterinio vaizdų apdorojimo sparta ištirti eksperimentiškai. Sekimo sistemos programinę dalį sudaro dirbtiniai neuroniniai tinklai, pagrindinių komponenčių analizės metodai. Taikant šiuos metodus sumažėja vaizdo duomenų kiekis, nes neinformatyvūs duomenys nufiltruojami. Tai paspartina skaičiavimus. Euristiniu metodu parinkta neuroninio tinklo struktūra ir pagrindinių komponenčių skaičius. Pristatoma žvilgsnio sekimo sistema testuota realiuoju laiku. Il. 10, bibl. 20, lent. 1 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).