Indian Journal of Science and Technology, *Vol 9(45), DOI: 10.17485/ijst/2016/v9i45/106770, December 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Communication Board and Visual PerceptionTraining Contents with Gaze Tracking for Augmentative and Alternative Communication(AAC)

Ki-Young Kwon[1], Jae-Yong An[1], Sung-Wook Shin[1], Se-Jin Yoon[2] and Sung-Taek Chung[1*]

[1]Department of Computer Engineering, Korea Polytechnic University, Siheung, Korea;
unitaek@kpu.ac.kr, kkywalk2@kpu.ac.kr,
[2]Remed Co., Korea; ysj@remed.kr

Abstract

This study proposed an gaze tracking system which detects the pupil and the center point of the glints from near-infrared camera images and, in order to use this system, implemented visual perception training contents and augmentative and alternative communication software. The center of the pupil and the center of two glint points of the eye images used for gaze tracking were extracted by using a model with four kinds of simple features. After using the calculated distance between the two centers to locate the relative position of the eye-gaze on the monitor, the x and y coordinates of the screen were mapped to match the gaze and the mouse pointer. The visual perception training contents for using the developed gaze tracking system consists of such components as visual acuity, eye movement speed, visual reaction rate, and visual concentration. Furthermore, software was implemented to enable an alternative means of communication to be used by symbols which convey the meaning of vocabulary being used in spoken language.

Keywords: Augmentative and Alternative Communication, Disabled, Gaze Tracking, Visual Perception.

1. Introduction

With Smart technology, which is representative of the development of interactive computer technologies(ICT), it has become easy to obtain useful information anytime and anywhere. Through a variety of contents, many functional services have been provided. However, in this ageing society, patients with impaired mental and physical capacity are increasing, and it is increasingly difficult for these people to benefit from the various services provided by smart technology[1]. In particular, in the case where patients have upper limb physical impairment or paralysis caused by such conditions as upper limb disorders, subcortical vascular cognitive impairment, or hemiplegic disability due to stroke, it is more difficult for them to use such devices as mobiles or personal computers with existing standardized input devices. Therefore, for people who have difficulty communicating because of impaired motor function or brain disease, alternatives to existing standard input devices are required that enable the human computer interface(HCI) to work[2].

HCI, by bringing together a number of different research areas such software, hardware, User Interface(UI), input and output devices and design, is developing input devices to replace the mouse or keyboard by utilizing such technology as Electroencephalography(EEG), Electromyography(EMG), Inertial Measurement Units(IMU), camera and touch screens.

The HCI field, that studies people with disabilities in its research, utilizes a Brain Computer Interface (BCI) which uses EEG brain waves as the computers input. The HCI also utilizes such methods as detecting the user's particular movements by using EMG and IMU sensors and uses those detected results as the computer's input value. Furthermore, due to the popularity of smart phones in recent years, research is being carried out on the use of Graphic User Interface(GUI) for Intuitive Touch-Based Operations for people with upper limb disorders. Research is being carried out on such techniques as mapping methods that use images from cameras to recognize the specific movements of parts of the body so that this can be used as the computer's input value[3-7].

** Author for correspondence*

Since an gaze tracking interface, which is one kind of technology that recognizes the body's movement, is a very effective tool to use for expressing such things as human attention, initial reactions, interests and satisfaction, a great deal of research is in progress in this area[8]. Generally speaking gaze tracking is a technology which uses historical data for human and computer interaction. After recognizing the center of the pupil which was obtained from a camera image, it uses the monitor's coordinates values which are obtained from what the eye is focused on.

The advantages of this method are that it is more robust to background noise when compared with existing methods which use EMG or IMU sensor values, and it can be implemented at a lower cost. Shih used two stereo cameras to track the gaze with three-dimensional coordinates by using a wide-angle camera to recognize the face and the direction of the face, and a narrow-angle camera to extract a high resolution image of the eye area[9]. Santis detected the pupils by utilizing the Optimal Segmentation method for each frame using a single camera[10]. Hutchinson tracked the relative position of the gaze by using the relationship between the center point of the pupil and the specific glint points reflected on the cornea of the eye by two near-infrared lights[11].

In this study an gaze tracking system was developed based on a near-infrared camera by using the reflection point of the cornea and the pupil as described above. By mapping the user's gaze to the computer's mouse pointer, an input device was developed for patients who cannot freely use their body. This can be a very useful interface for people who cannot move their bodies freely because it is possible to use it with only the eyes.

However most patients who want to use an gaze tracking interface have dementia, cognitive impairment, and Parkinson's disease and most of these have problems using language, so for those people a new means for augmentative and alternative communication is needed[12]. This should be able to represent not only traditional oral and written language, but also such methods as sign language, braille, painting or line drawing, by applying a communication method using symbols. Furthermore, because of differences in users' vocabulary levels and usage due to differences in nationality, age, education, and medical history, it should be able to be classified according to people's respective environments[13].

The proposed augmentative and alternative means of communication developed in this study was based on an gaze tracking system which utilized a communication board for people who have problems with verbal communication. The developed augmentative and alternative communication software used symbols that were familiar to users and effort was put into implementing a friendly user interface[14]. Gaze tracking software like this is difficult for patients to use. Therefore in order to solve this problem a variety of visual perception contents were provided to enable training in the use of this software.

2. Main Contents

A. Gaze Tracking using a Near Infrared Camera

Near Infrared USB Camera

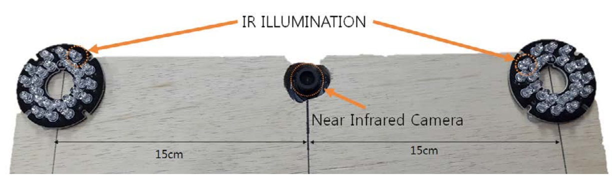


Figure 1. Camera for Gaze Tracking and Infrared Light source.

The gaze tracking system used in this study consists of a single camera with two infrared light sources A and B as shown in Figure 1. Two infrared light sources generate two reflection points called Purkinje images on the surface of the user's cornea[15]. The camera used here is a ELP-USB130W01MT-RL36 model from the ELP company with a resolution of 1280×960 and 1.3 megapixels. The IR LED, which is attached to side of the camera, acts as a light source to the camera lens, and by attaching an infrared filter with an infrared wavelength of 840nm, enabled only the necessary images to be obtained.

Gaze Tracking

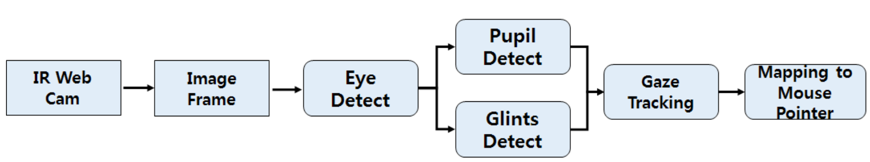


Figure 2. Gaze Tracking Flow Chart.

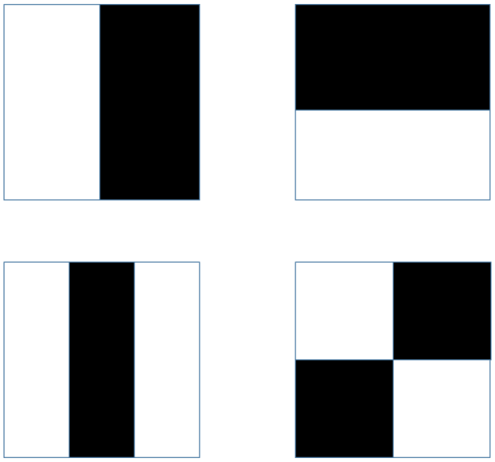


Figure 3. Example Rectangle Features shown Relative to the Enclosing Detection Window.

The overall gaze tracking process is shown in the flow chart in Figure 2. The user's eye image that was obtained from the image frame from the IR Webcam was detected from the Adaboost algorithm. The Adaboost algorithm calculates the difference between the pixel sum of the dark part of the image and the pixel sum of the light part of the image by using the four rectangle features shown in Figure 3, and a number of features are generated. The generated features, as weak classifiers are combined linearly to create strong classifiers with a high detection capability, which in turn are used to detect features of the face or eyes[16].

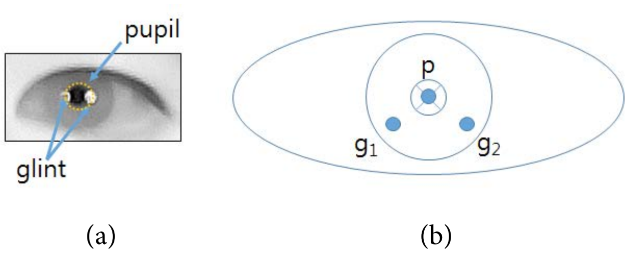


Figure 4. (a) Detected eye image and (b) Pupil and glints.

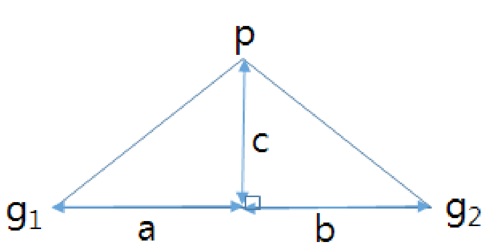


Figure 5. The distance between the center of the pupil and two glints.

In order to detect the center point of the pupil and the reflection points of the cornea in the obtained eye image, first a Gaussian smoothing filter was used to remove noise, and then the Canny edge detection method was used to obtain an edge image. The center point between the pupil and the reflection points of the cornea was extracted by calculating the center point and radius of the most overlapping circles by applying a 2D Hough transform to the obtained edge image[17]. By using the extracted center point of the pupil and the center of the two reflection points as the three feature points, the distance between these points was calculated, and by using this, the mouse pointer and the eye position were mapped to each other on the user's screen.

The mapping method is shown in Figure 4(b) where p is the center of the detected pupil and g_1 and g_2 are the center of the two glints. In Figure 5, c is the length of the line perpendicular to the straight line passing through g_1 and g_2 starting from p, and a and b are the respective distances from each glint point to where line c intersects the line between g_1 and g_2. In Figure 6, when the user looks at the points A and B on the monitor, in order to calibrate the relative location value of the gaze, the distance between the center of the pupil and two glints is used as shown in Figure 5.

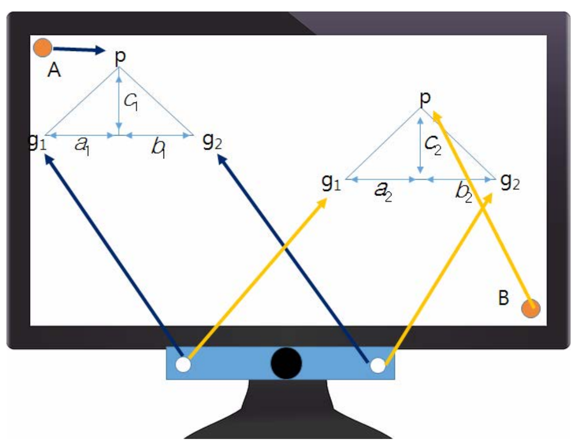


Figure 6. Each feature point projected on to the pupil from the calibration point.

Since the two glints, g_1 and g_2, are projected onto the pupil with a fixed light source, the movement rate of the eye, to the left, right, up and down, can be calculated with the distance between two points, $a+b$. Therefore, when the two points, A and B, are viewed, calibration occurred, as shown in Figure 6, where location information of the mouse pointer was mapped by calculating the x,

y coordinates for the screen using Equation (1) and Equation (2) by using the distance between each of the feature points a_1, b_1, c_1 and a_2, b_2, c_2.

$$x = \left(\frac{a_{frame}}{a_{frame} + b_{frame}} - \frac{a_1}{a_1 + b_1} \middle/ \frac{a_2}{a_2 + b_2} - \frac{a_1}{a_1 + b_1} \right) \times Width \quad (1)$$

$$y = Height - \left(\frac{c_{frame}}{a_{frame} + b_{frame}} - \frac{c_2}{a_2 + b_2} \middle/ \frac{c_1}{a_1 + b_1} - \frac{c_2}{a_2 + b_2} \right) \times Height \quad (2)$$

In equations (1) and (2) a_{frame}, b_{frame}, c_{frame} is the distance between the feature points of the current image frame, and width and height indicate the resolution of the monitor.

B. Visual Perception Training Contents

For users to utilize an gaze tracking system, in order to use it as a means of augmentative and alternative communication, first of all, a training process is required through the development of visual perception contents. Therefore, the training contents developed in this study were to help patients learn in a fun way how to become familiar with using their eyes to control the system. The developed contents allowed users to train such elements as visual acuity, eye movement speed, visual reaction rate, and visual concentration based on recommendations by the American Optometric Association[18-21].

Visual Acuity Training Contents

Figure 7. Each feature point projected on to the pupil from the calibration point.

The visual acuity content in Figure 7 is intended to train visual acuity skills with moving objects. When the user watches the screen, three different kinds of birds move around the screen in any direction. If the user continues to follow a bird, which is the same as the bird in the upper right hand corner of the screen, with their gaze for more than three seconds they score points. Once the gaze tracking is completed a new bird appears in the upper right hand corner and the degree of difficulty of the game is controlled by increasing the speed of the birds little by little.

Eye Movement Speed Training Contents

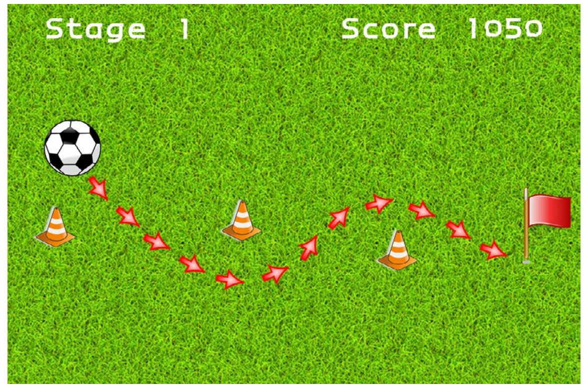


Figure 8. Eye Movement Speed Training Contents.

This content enables users to continuously track an object depending on the speed of the moving object in order to improve eye movement speed. In Figure 8, as the soccer ball moves along the arrows, users continue to track the soccer ball until it reaches its destination. Points are scored in proportion to how long users can maintain their gaze on the soccer ball. It is possible to control the degree of difficulty of the task by continuously increasing the soccer ball's speed.

Visual Reaction Rate Training Contents

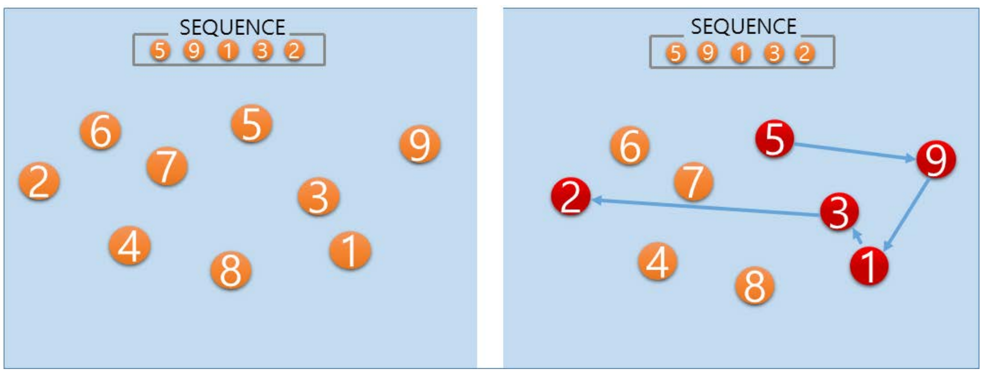
Figure 9 shows the visual reaction rate training contents and is concerned with reactions to visual stimuli. It consists of six boxes and one red circle. After users fix their gaze on the red circle in the middle, one box can change color at any time. The faster that users shift their gaze to the box which has changed color, the higher they score.

Visual Concentration Training Contents

Visual concentration contents are for improving the ability to maintain visual focus in one place. This is a necessary function for choosing any symbols as a means of communication, and according to how well users are trained, the time needed to deliver the message to be communicated can be reduced. The visual concentration content is shown in Figure 10, and by using gaze tracking, users track the numbers according to the SEQUENCE of the numbers given in the box at the top of the screen. The gaze must be maintained on each number for more than three seconds, and in the case where the wrong order or number is selected, users must go back to the beginning

(a) (b)

Figure 9. Visual reaction rate training contents .

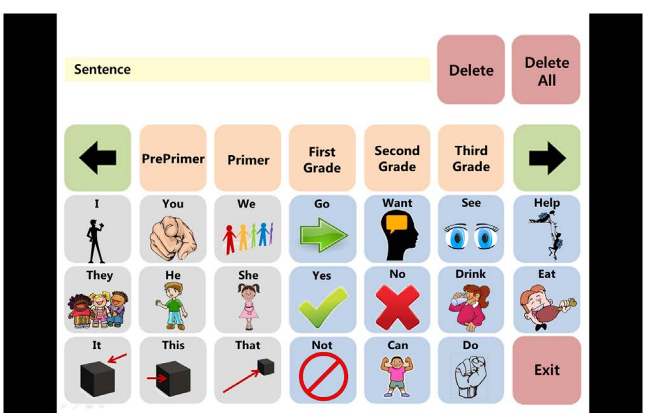


(a) (b)

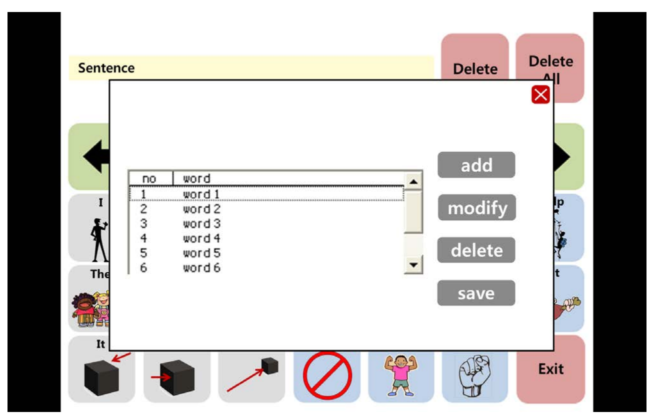
Figure 10. Visual concentration training contents

and start again. The time it takes to select all the numbers in order as shown in Figure 10(b) is measured. The degree of difficulty of the task is controlled by using other symbols besides numbers.

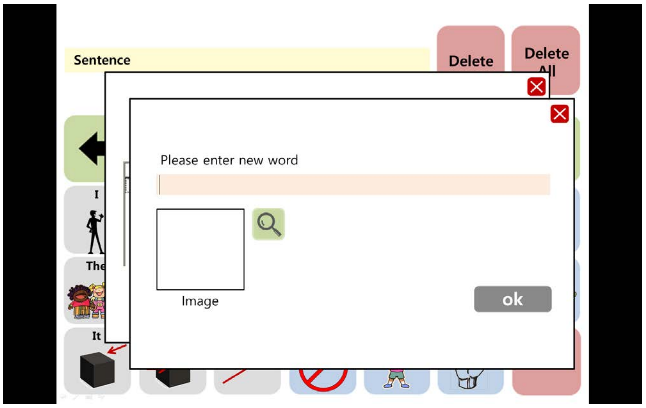
C. Augmentative and Alternative Communications Software



(a)



(b)



(c)

Figure 11. EImplemented augmentative and alternative software.

Augmentative and alternative communication was implemented based on the communication board proposed by Yeongchi in the form shown in Figure 11 in conjunction with an gaze tracking system[14]. Basic vocabulary configuration was applied by a method proposed by Dolch using a list of English words often used in elementary education[22].

The most important factor in augmentative and alternative communication are the symbols which represent the vocabulary's meaning - in other words, the kind of vocabulary that users can use. The vocabulary needed is different depending on the user's nationality, age, education levels and medical history[13]. Accordingly in this study, the contents were designed so that vocabulary can be added any time as needed. Figure 11(b) shows the current list of registered words, and by clicking on 'add' or 'modify', new words can be input in the pop-up window as shown in Figure 11(c). Symbols of new words can also be registered by uploading an image.

3. Conclusion and Discussion

This research proposed an gaze tracking system by extracting the pupil and the center point of the glints from near infrared camera images. The Adaboost algorithm was applied to the eye images for gaze tracking by using a model with four kinds of simple features. By using the extracted center point of the pupil and the center of the two reflection points, which were made in this way, as the three feature points, the distance between these points was calculated and the relative location value of the gaze on the monitor was calibrated. Based on this, the value of the screen's x, y coordinates were used to map the gaze and the location value of the mouse pointer.

To use the developed gaze tracking system visual perception training contents were proposed and augmentative and alternative communication software was implemented. The proposed visual perception contents made it possible to train for such elements as visual acuity, eye movement speed, visual reaction rate, and visual concentration.

Finally, augmentative and alternative communication software was implemented by using symbols which can convey the meaning of words used in spoken language. In addition, the symbols, which were used in the augmentative and alternative communication software, were easily uploaded by users through the server. By classifying frequently used vocabulary according to the user's environment, it can be easily used by other users. It was designed to enable social participation of patients who have difficulty in communication due to disabilities, by using an gaze tracking system which solves these communication problems. Furthermore, if serious games or contents, or a variety of supporting software are developed using gaze tracking systems, then people who have difficulties using the human computer interface can receive information more efficiently.

4. Acknowledgment

This work was supported by the Technology Innovation Program(10062378, Development of fusion parts for 3.5 Tesla 6 channel magnetic stimulation and information feedback for a cure for intractable brain disease) funded By the Ministry of Trade, Industry & Energy (MI, Korea).

5. References

1. Organization for Economic Cooperation and Development (OECD), Health Data 2013. 2014.
2. KoppenhaverDA, SteelmanJD, PiercePL, YoderDE and StaplesA. Developing Augmentative and Alternative Communication Technology in Order to Develop Literacy. Technology and Disability.1993; 2(3):32-41.
3. MuCS and MingTC. Voice-controlled human-computer interface for the disabled. Computing & Control Engineering Journal.2001October; 12(5):225-30.
4. Gao X, Xu D, Cheng M and GaoS. A BCI-based environmental controller for the motion-disabled. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2003June; 11(2):137-40.

5. Xiong A, Chen Y, Zhao X, Han J and LiuG. A novel HCI based on EMG and IMU.International Conference on Robotics and Biomimetics.2011December;p. 2653–57.
6. MoonHJand KimSI. Touch Button Design for Smartphone Considering People with Upper Limb Disabilities. Korean Society for Rehabilitation of Persons Disabilies.2015; 19(2):181-203.
7. Zelinsky A and HeinzmannJ. Real-time visual recognition of facial gestures for human-computer interaction. Proceedings of the Second International Conference on Automatic Face and Gesture Recognition.1996; p. 351-56.
8. Park JW, Kwon YM and Sohn KH. Gaze Tracking System Using Feature Points of Pupil and Glints Center. The Korean Society of Broad Engineers.2006 March; 11(1):80-90.
9. Shih SW and LiuJ. A novel approach to 3-D gaze tracking using stereo cameras. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics).2004February;34(1):234-45.
10. Santis AD and IacovielloD. Robust real time eye tracking for computer interface for disabled people. Computer Methods and Programs in Biomedicine.2009October; 96(1):1-11.
11. Hutchinson TE, White KP, Martin WN, ReichertKC and Frey LA. Human-computer interaction using eye-gaze input.IEEE Transactions on Systems, Man, and Cybernetics.1989November; 19(6):1527-34.
12. Illes J, Metter EJ, Hanson WR and IritaniS. Language production in Parkinson's disease: Acoustic and linguistic considerations. Brain and Language.1988January; 33(1):146-60.
13. Yorkston K, Honsinger M, Dowden P and MarrinerN. Vocabulary selection: A case report. Augmentative and Alternative Communication.1989; 5(2):101-8.
14. Wu YC and Voda JA. User-friendly communication board for nonverbal, severely physically disabled individuals. Archives of Physical Medicine and Rehabilitation.1985; 66(12):827-28.
15. Cornsweet TN and Crane HD. Accurate two-dimensional eye tracker using first and fourth Purkinje images. Journal of the Optical Society of America.1973; 63(8):921-28.
16. Viola P and Jones MJ. Robust Real-Time Face Detection. International Journal of Computer Vision. vol. 57, issue 2, pp. 137-154.
17. Ioannou D, Huda W and Laine AF. Circle recognition through a 2D Hough transform and radius histogramming. Image and Vision Computing.1999January; 17(1):15-26.
18. Herdman SJ, Tusa RJ, Blatt P, Suzuki A, Venuto PJ and RobertsD. Computerized dynamic visual acuity test in the assessment of vestibular deficits. American Journal of Otology.1998November; 19(6):790-96.
19. Williams HG and HelfrichJ. Saccadic eye movement speed and motor response execution. Research Quarterly, American Alliance for Health, Physical Education and Recreation.1977; 48(3):598-605.
20. Tuch DS, SalatDH, WiscoJJ, ZaletaAK, HeveloneND and RosasHD. Choice reaction time performance correlates with diffusion anisotropy in white matter pathways supporting visuospatial attention. Proceedings of the National Academy of Sciences of the United States of America.2005August;102(34):12212-17.
21. Harris DV and HarrisBL.The athlete's guide to sports psychology: Mental skills for physical people. Human Kinetics.1984; 1.
22. Dolch EW.Garrard Press:Problems in reading. 1948.