

Complete Scanpaths Analysis Toolbox

Piotr Augustyniak, *Member, IEEE*
Zbigniew Mikrut

Abstract—This paper presents a complete open software environment for control, data processing and assessment of visual experiments. Visual experiments are widely used in research on human perception physiology and the results are applicable to various visual information-based man-machine interfacing, human-emulated automatic visual systems or scanpath-based learning of perceptual habits. The toolbox is designed for Matlab platform and supports infra-red reflection-based eyetracker in calibration and scanpath analysis modes. Toolbox procedures are organized in three layers: the lower one, communicating with the eyetracker output file, the middle detecting scanpath events on a physiological background and the one upper consisting of experiment schedule scripts, statistics and summaries. Several examples of visual experiments carried out with use of the presented toolbox complete the paper.

I. INTRODUCTION

SCANPATH analysis is currently one of most vital scientific interest, and is considered as a powerful tool revealing the human pursuit of visual information and thus some aspects of data collection strategy [1], [2]. The scanpath analysis is used to reveal disorders of human sight sense, but new technical applications of scanpath using are recently emerging [3]. Commonly given examples are: perception-based visual information design (e.g. advertisement, websites) and intelligent man-machine interfaces for disabled people. As the scanpath analysis is still under experimental research it needs experimental mathematical and computational tools for data processing and knowledge extraction.

A complete scanpath analysis toolbox was developed in our laboratory dedicated for scientific research on discriminative features in human pattern recognition. The attempt to build our custom collection of procedures was anticipated by detailed studies of existing packages for which the Psychophysics Toolbox extensions [4], [5] for Matlab may be a good reference. Unfortunately, this toolbox was developed mostly for Mac users, and since many subroutines use precompiled libraries (dll's), they cannot be easily converted to the Windows platform.

Manuscript received April 2, 2006. This work was supported by the AGH University of Science and Technology as a research project No. 11.11.120.612.

Authors are with the Institute of Automatics, AGH University of Science and Technology; Al. Mickiewicza 30, 30-059 Krakow Poland, august@agh.edu.pl, zibi@agh.edu.pl

II. EYETRACKER FEATURES

A. Metrological principle of the IR-eyetracker

The principal difference with respect to other software packages is that our toolbox supports the infrared reflection-based eyetracker OBER-2 [6] in its calibration and off-line analysis mode. The on-line analysis and interactive display is currently under development and will be added soon. The head-mounted goggles illuminate each eyeball with four adjacent spots of total power of 5 mW/cm^2 in $80\mu\text{s}$ infrared pulses (wavelength 940nm) repeated at the sampling frequency. Four IR sensors per eye work in a pair wise space-differential configuration and capture two-dimensional trace of each eye at the speed up to 2000 samples per second during the presentation limited by the 32k samples data buffer. Since the sensor captures the visible light as well, a double sampling method is used for the sidelight discrimination. This specific time-differential measurement relates the actual infrared reflection readout to the sidelight background captured ca. 80ms before the LEDs become active. This measurement method eliminates the influence of all common light sources and allows the device to achieve the angular resolution of 0.02 deg .

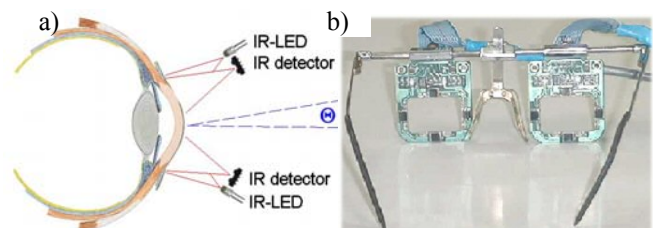


Fig. 1 Details on the Ober-2 eyetracker (a) physical principle and (b) general overview of goggles.

B. Eyetracker calibration

The eyetracker needs to be calibrated before each use in order to compensate for optical and environmental light errors. The calibration is intentionally as close as possible to the typical visual test. The observer is instructed to gaze at the corners of the displayed rectangle in a specified order. During the calibration the scanpath is acquired and next it is used for correction of the scene-dependent eye trajectories. Considering the crucial role played by the calibration process in an infrared reflection-based eyetracker, we provide two different calibration modes:

- **Orthogonal.** Basic points are rectangle corners and the scale and shift data calculated from the difference between display and gaze coordinates is independently added to the analyzed scanpath. The original calibration rectangle is represented in the corrected domain by a shifted and scaled rectangle,
- **Surface.** Four vectors originating at the central point are drawn to the gaze concentration points corresponding to the rectangle corners. The mismatch between the actual and the perfect gaze points is expressed in scale-angular domain and minimized during an iterative process. This method compensates nonlinear distortions of the projection and recording surfaces as well as rotation.

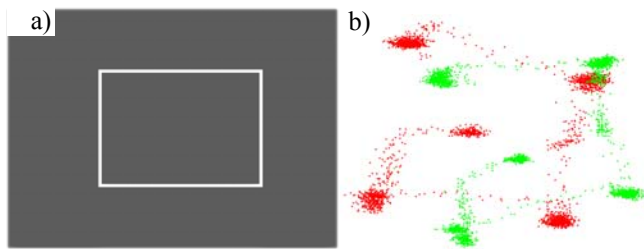


Fig. 2 Calibration process: calibration rectangle (a) and an example raw scanpath (b) red trace represents left eye and green trace – right eye

III. SOFTWARE ARCHITECTURE

The toolbox is organized in three layers. The input layer is the only part dependent on the used eyetracker and its main purpose is translation of the trajectory data to the angular coordinates corrected for optical skewness and technical conditions measured during the calibration. The middle layer contains subroutines for automatic or human assisted detection of most common events in the scanpath, such as saccades and fixation points. The upper layer is dependent on particular visual experiment and aims at display organization, scanpath statistics and file management. The layer philosophy reflects the toolbox adaptation to various experiments: changing equipment affects only the hardware layer, changing the aim of the experiment influences only the upper layer, while the middle layer represents the perceptual physiology and in general remains unchanged.

A. Input layer

The input layer is a set of functions or independent executables enabling frame-accurate display of the scene and data gathering. Certainly many technical features of the display considering time, illumination and color space should be adjusted in the experiment. Various graphic data formats (as bmp, jpeg, gif...) are supported as well as live-originated and animated sequences (avi).

For specific visual experiments there is a set of subroutines preparing the data for displaying:

- sequences of objects with various attributes moving along predefined or random trajectories,

- screens with randomly distorted text complying with a given mixing statistics,
- sequences of screens with added objects of specified attributes.

Currently, the off-line measurement is supported by an independent executable writing the scanpath directly to the hard disk. The input layer reads the file provided by the eyetracker during the measurement and converts the data to the Matlab structure. The calibration data stored in a separate file is also employed at this step to correct the scanpath using one of the methods mentioned above.

B. Detection layer

Regardless the wide range of applications, the scanpath analysis originates from the anatomy and physiology of the retina and the psychophysics of vision. The middle layer of the toolbox is oriented towards automatic recognition of specific eye movements: idle time (fig. 3), fixations, saccades and the visual search end. These movements are characterized by appropriate logical conditions applied to the statistical scanpath parameters. Certainly, accurate detection of basic gaze-related events in the scanpaths requires application of several heuristics. These parameters were precisely selected and carefully adjusted during the experiments completed with our toolbox. In addition to the fully automatic detection, main subroutines can optionally work in the operator-supervised mode for any custom application. This mode enables redefinition of crucial classification parameters of the scanpath determining its sections to fall into saccades, fixation points or other events. The visual research end was defined as the moment when the observer accomplished the visual task (e.g. recognized the object). In order to avoid the aimless gaze wandering, the observers were instructed to put the gaze below the display. This can be easily detected by the software and pursued in order to determine the exact moment of visual search end.

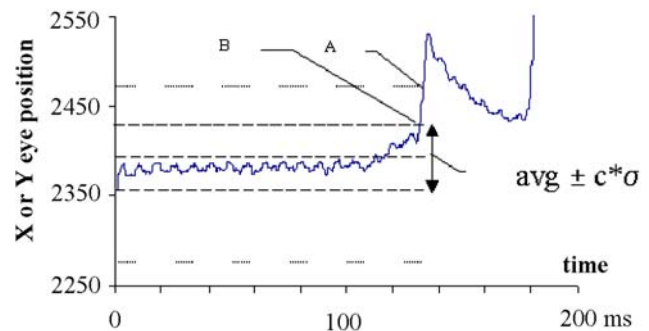


Fig. 3 Iterative detection of the idle time before the observer reaction A - rough estimation of signal stability, B – precise estimation of noise and signal emerging point

C. Experiment layer

A principal advantage of the scanpath analysis is the use

of non-invasive and relatively cheap visual experiments. The crucial part of the experiment is standardization of the observer knowledge and motivation for the visual search for the complementary information in the scene. The willing collaboration is a necessary condition and while it cannot be guaranteed, the attentive human control of the ongoing experiment and careful review of the results are required. In the upper layer the toolbox provides several scanpath analysis-oriented statistics and summaries. The examples of these are:

- fixation points number,
- left (right) oriented saccades,
- x and y axis-distribution of conspicuity,
- dominant eye recognition etc.

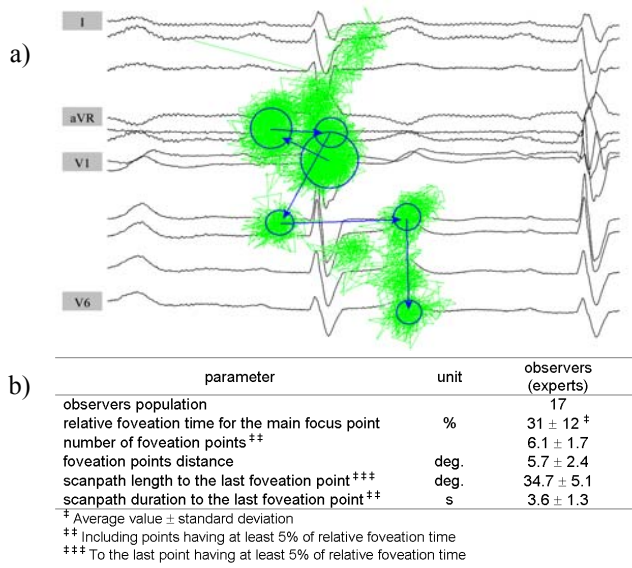


Fig. 4 Scanpath recorded during the human electrocardiogram interpretation (a) and statistics calculated by an experiment layer procedure (b)

The third layer depends on the experiment purpose and assumptions about the expected observer behavior. To some extent it is guided by the initial observer knowledge and displayed scene, therefore the scripts defining the course of experiment and scanpath analysis chain are closely related to the scene generation and analysis procedures. The experiment layer provides several statistics and scanpath oriented procedures for data processing and visualization. However, no graphic user interface is provided, since we assume that for a new experiment design the user should have a basic programming skills.

D. Scene generation and analysis procedures

Each visual experiment uses presentations of prepared scenes. For best accuracy, these images should have exact resolution of the target display. In most cases of visual experiments, the images differing by some controlled parameters are ordered in consecutive sequences displayed

at a specified rate. Two kinds of images are used in visual experiments:

- artificial images fully computer-designed and thus featuring the control over all image parameters (topography, brightness, color space etc.)
- natural images eventually modified by a human using a computer; such images allow the measurement, but not the control of all image parameters.

Artificial images are produced by a specialized procedure from given patterns, motion vectors and time-color functions. Image analysis procedures are applied to original images to reveal the parameter changes implied by added or relocated details. As the altered image must remain ‘natural’, the changes of parameters are not linear.

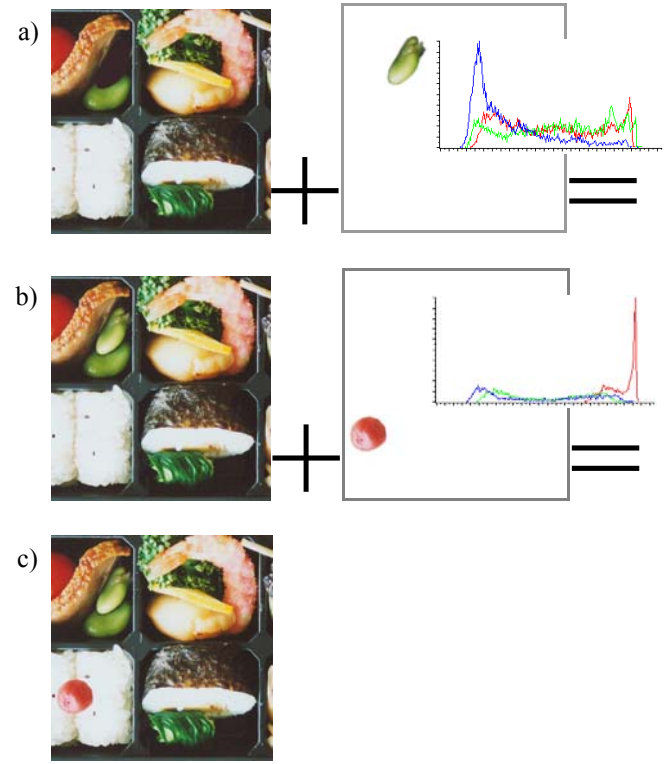


Fig. 5 Example of image sequence analysis (a to c); aiming at the estimation of color influence on details conspicuity, natural details are successively added to presented scene and the detail parameters are measured in RGB color space (only histograms are presented).

IV. VISUAL EXPERIMENTS REVIEW

The backbone of the toolbox was developed on the basis of various visual experiments. Selected examples are listed below in order to highlight possible applications area:

A. Character sequence recognition

Investigation of correctness of character sequence memorization in a time-limited show [7]. This experiment correlates the percentage of correct immediate reproductions with the presentation time (200 to 800 ms), with the

sequence length (3 to 10 characters) and with the display position (± 24 deg. horizontally, ± 18 deg. vertically). The correctness of reproduction increased with the observation time and saturates on the average value of 85% at a display time depending on the sequence length: 500ms for three-characters sequences and 800ms for over eight-characters sequences. We did not observe a significant decrease of the correct reproduction percentage when moving sequence locations far from the display center.

B. Handwritten digits recognition

Human recognition of handwritten digits [8]. The experiment aimed at detection of the most conspicuous part of the image of handwritten digit and at the assessment of the segment contribution to the final decision. The scanpath analysis revealed several approaches the observer had to the recognition of each digit, however some gaze strategies dominated significantly. This difference was particularly interesting in case of similar shapes as 1 and 7, in result we determined the regions of hierarchical importance specific for each digit. The target application for such research is an automated handwritten character recognition system matching the input with patterns only where it is important, in a manner similar to a human observer.

C. Distorted text reading

Pursuit of the eyegaze during the reading of distorted text [9]. The displayed text was progressively distorted with use of programmable character replacement procedure. The beginning of page was quite easy to read while the end was very hard to guess. Although the reading is usually linear from the left to the right, the prolonged gaze and returns to the middle of the line indicate difficulties. The experiment aimed at studying the human capabilities of text reconstruction and classification of the words by their importance in message understanding. This experiment had two applications: adaptive error correction in a message transmission system and objective skills assessment in foreign language reading.

D. Response and recognition time

Research on the initial response and shape recognition time [10], [11]. Various shapes were displayed and the observers' task was the detection and discrimination of the shape having specified features (e.g. rounded corners). This experiment correlated the idle time with the illumination and contrast conditions and had a principal contribution to the rules for determination of observation beginning and end. These rules and appropriate signal processing subroutines allowed exact measurements of observation time from the scanpath.

E. Added details conspicuity

Investigation of the contribution of objects' features to their conspicuity. A series of progressively modified images

were presented and the observer was asked to detect a change (fig. 5). Images were modified by adding a natural object of increased amount of selected feature: size, distance to the display center, brightness and color in both RGB and HSL spaces. The physiologically derived statements on the prevalence of long-wave cones were confirmed by greater conspicuity of red changes in the scenes. However, we obtained also some unprecedented results e.g. greater conspicuity of objects with increased distance to the center.

V. CONCLUSION

In authors' opinion, supported by other colleagues using the developed scanpath analysis toolbox, a valuable software package for visual experiments analysis was created. The toolbox is designed to help the human visual system research in recognition and measurements of basic events, preparation and interpretation of experiments as well as eyetracker support.

ACKNOWLEDGMENT

Authors thank their students: L. Chornikowski, L. Czajczyk, R. Kaminski, B. Kluza, P. Koprowski, A. Korczynski, F. Korta, L. Maselek, M. Miekus, J. Piekarz and J. Winiarska for their involvement in the visual experiments described in this paper, in the software tests and for their fruitful remarks.

REFERENCES

- [1] G. Boccignone, "An Information-theoretic Approach to Active Vision" *Proc. 11th Int. Conf. on Image Analysis and Processing* 2001.
- [2] J. B. Pelz, R. Canosa, "Oculomotor behavior and perceptual strategies in complex tasks" *Vision Research*, vol. 41, 3587-96, 2001.
- [3] J. K. Ober, J. J. Ober, M. Malawski, W. Skibniewski, E. Przedpelska-Ober, J. Hryniewiecki, "Monitoring Pilot's Eye Movements during the Combat Flight-The White Box" *Biocybernetics and Biomedical Engineering*, vol. 22, (2-3), pp. 241-64, 2002.
- [4] D. H. Brainard, "The Psychophysics Toolbox", *Spatial Vision* vol. 10, pp. 443-446, 1997.
- [5] D. G. Pelli, "The VideoToolbox software for visual psychophysics: Transforming numbers into movies", *Spatial Vision* vol. 10 pp. 437-442, 1997.
- [6] J. J. Ober, J. Hajda, J. Loska, M. Jamicki "Application of eye movement measuring system Ober2 to medicine and technology" *Proc. of SPIE* 3061 (1), 327-32, 1997.
- [7] P. Augustyniak, Z. Mikrut "Scanpath-based localization of principal image components" (in Polish) *13-th Conference on Biocybernetics and Biomedical Engineering*, pp. 931-936, 2003.
- [8] P. Augustyniak, Z. Mikrut, "Detection of Object Salient Features Based on the Observer Scanpath Analysis", *IFMBE Proceedings* vol. 11 (1), paper 1416F, 2005.
- [9] P. Augustyniak, Z. Mikrut, "Investigation of Natural Perception Rules and their Application in Intelligent Visual Systems" (in Polish) *Conference on Artificial Intelligence in Biomedical Engineering SIIB*, Krakow, 2004.
- [10] Z. Mikrut, P. Augustyniak, "Estimation of Execution Time for Tasks of Objects Counting and Localization Using the Ober2 Device" *IFMBE Proceedings*, vol.2, pp. 144-145, 2002.
- [11] P. Augustyniak, Z. Mikrut, "Correlating the Degree of Observer's Preoccupation and the Observation Time: Visual Tasks with OBER2 Eyetracker" *JMIT* vol. 3, pp. MT3-MT10, 2002.