Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 3 (2012) 89 – 96

Procedia Technology

The 2012 Iberoamerican Conference on Electronics Engineering and Computer Science

On the Design of a Low Cost Gaze Tracker for Interaction

A. Carbone[a], F. Martínez[a], E. Pissaloux[a,*], D. Mazeika[b], R. Velázquez[c]

[a]*Université Pierre et Marie Curie (Paris 6), ISIR, France*
[b]*Kaunas University of Technology, Mechanical Engineering Dept., Lithuania*
[c]*Universidad Panamericana, MCS, Mexico*

Abstract

The human gaze is a basic mean for non verbal interaction. However, in several situations, especially in the context of upper limb motor impairment, the gaze represents also an alternative mean for human interaction with the environment (real or virtual). This interaction can be mastered through specific tools and new learned skills. Therefore the technological tool is a key for new interaction models. This paper presents a tool for gaze interaction: a new gaze tracker. The system specifications and the status of the gaze tracker design are presented; the dedicated algorithm for eye detection and tracking as well as an improvement of Zelinsky's model for eye movement prediction during the search of a predefined object in an image are outlined. Results of the first pre-prototype first evaluation with end users are summarized.

© 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

Keywords: Wearable gaze tracker, interaction, upper limb motor impaired, face feature detection, active shape model, computer vision.

1. Introduction: brief status of current eye/gaze tracker technology

Interaction is a key concept of modern society. Human gaze interaction, based on human intention, is a recent interaction mode. The gaze is one of the most promising interaction modes not only in the context of (temporary or permanent) impairments but also in the case where human operators (such as surgeons and fighter pilots) are forced to conduct simultaneously multiple activities some of them, of the highest priority, monopolizing the usual means of interaction such as hands, or where traditional interaction means are not operative or are insufficient (especially true for interactions with virtual environments).

The application domains of gaze interaction are various; human memory enhancement and training, (elderly) people daily activities assistance, games, serious games, simulation, training, robotics, health, accessibility, wellbeing and attention responsive technology [1], fitness, social interaction, team building, etc. are some representative examples. The usage of the gaze tracker as the interaction mean requires the acquisition of special skills such as control of eye movements (dwelling, displacement in precise direction

*Corresponding author. Tel.: +00 33 1 44 27 28 74; fax: +00 33 144 27 51 41.
E-mail address: Edwige.Pissaloux@upmc.fr

2212-0173 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.protcy.2012.03.010

and at giving speed, eyelids blinking at specific time, etc.). However, the quality of such skill acquisition is strongly influenced by the technology used.

In general, two configurations for gaze trackers are commercially available: remote gaze (passive) trackers and head-mounted gaze (active eye monitoring) trackers.

The main advantage of the remote systems is that no devices or sensors have to be mounted on the subject's body, which may increase the comfort for users, especially in long-term use. However, remote eye trackers are not well-suited for applications where mobility is required; indeed, subject/remote gaze tracker relative speed should be very slow and the maximal displacement very limited. From a hardware point of view, one or two gaze tracker cameras are always placed in a manner that they can track the eye(s) of the user. The cameras move in the same reference frame as the user does. Desktop/remotely mounted systems are designed to accurately track pupil diameter and eye position relative to one stationary flat surface (supposed to be parallel to the camera plan). A computer with a built-in or USB connected camera is the most popular implementation. The Tobii is nowadays the most popular desktop/remote gaze tracking system [2].

Head-mounted gaze tracking systems are experiencing a renaissance due to new challenges [3], [1]. They are built around several cameras. The most popular prototypes use one camera for (usually dominant) eye tracking and another for scene scanning; they include additional sensors (such as head tracker) in order to compensate for head movements.

However, all existing gaze tracker are rather expensive (several thousands of euros) and use the near infra-red technology (IR-A, 780-1400 nm). Epidemiological data on IR-A band LED long period exposure do not exist and explicit guidelines have not been yet addressed in any current IR safety standards; potential hazards are still an open question [4]. For all of the above reasons, a design of a low cost only vision-based gaze-tracker is on progress in the frame of AsTeRICS, FP7 ICT project; its status is rapidly presented in the subsequent sections. Section 2 proposes the low cost gaze tracker pre-prototype hardware architecture. Section 3 addresses systems algorithmic developments: new algorithms for eye detection/tracking and an improvement of Zelinsky's model for eye movement predictions while exploring an image of 3D scene [5]. Section 4 proposes the gaze tracker pre-prototype evaluation with end-users, while Section 5 concludes this paper and proposes some future research directions.

2. Low cost gaze tracker architecture

This first approach to low cost gaze tracker design targets the skill acquisition for (body and especially head unconstrained movements) interactions with a surface located at the fixed distance such as computer screen or virtual reality screens. Screen displayed object management (such as access/deletion of existing objects and creation of new objects), access and Internet navigation, access to a PC (as a software development tool or as a basis for e-services, etc.) are targeted skills to be learned/assisted.

2.1. Low cost gaze tracker end-user specifications

In the frame of the EU FP7 AsTeRICS project both wearable and remote (web-cam based) systems are being pursued to support face and gaze based interaction.

Fig. 1 schematizes the targeted head-mounted (HT) gaze tracker as a peripheral for a PC supporting ARE (AsTeRICS Runtime Environment). The hardware system is a specific frame grabber (a Samsung netbook) which acquires images from two eye cameras and from a scene camera (Sony Ex-View CCD). The software synchronises and conveniently processes images that are either displayed on the screen of the frame grabber or transferred to ARE via USB.

The main end-user functional specifications of the gaze tracker first pre-prototype are: a) unobstructed field of view for the end user, b) possibility of wearing the correction glasses, c) adjustable size to different head morphologies, d) precision compatible with targeted skills' acquisition, e) lightweight system, f) easy to wear, and g) simple calibration procedure. The system should be calibrated in order to establish mapping between eye movements and the projection of the gazed point on the camera scene image thus opening the possibility to interact with all the space in the field of view of the camera (subject).

The system works at the video speed.

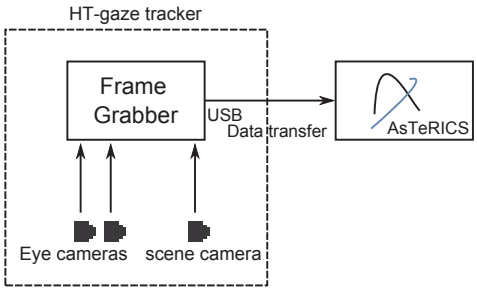

Fig. 1. Schematic architecture of the low lost vision gaze tracker.

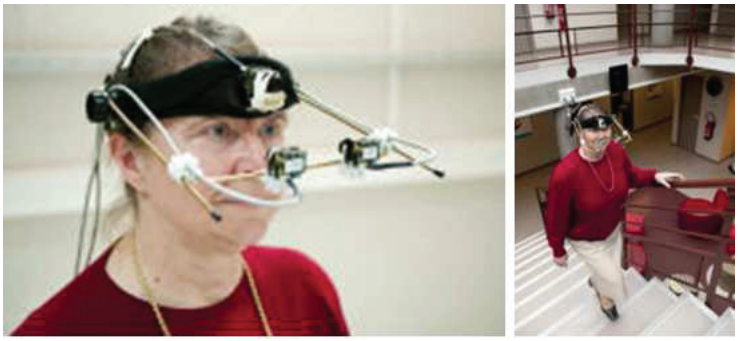

Fig. 2. AsTeRICS (UPMC) vision gaze tracker pre-prototype.

2.2. *Gaze tracker pre-prototype architecture*

Fig. 2 shows the first pre-prototype of the HM-gaze tracker developed for new algorithms design purposes. Its main components are: mechanical support, cameras and (remote, not shown on the photos) frame grabber (processing unit).

The mechanical support is adjustable on the back and on both sides of the head. Its one boom arm supports two cameras with adjustable to the end user's eyes distance and with adjustable orientation (6 DoF - 6 degrees of freedom). The scene camera is located on the front helmet part attached to the forehead. Eye camera should be "in front of the eye" in order to reduce the projective distortions of the eye on the eye camera images (distortions which should be corrected by the software).

The gaze tracker works in indoor (constant illumination) and outdoor environments (varying natural illumination).

3. Some algorithmic developments

3.1. *Algorithms for remote gaze tracker*

The remote gaze tracker is dedicated for example, to simple interactions via nose and chin. Algorithms dedicated to these face characteristics detection (using Haar cascade) and for feature tracking (using the KLT algorithm) have been designed. This software can be used for mouse simulation (mouse displacements and object selection/mouse clicking). The encoding of more complex facial expressions with a support of deformable models for face characteristics detection (namely ASM: Active Shape Models), and the integration of the ASM and 3D point cloud delivered by the Kinect sensor for gesture tracking in a three dimensional space (Fig. 3) are under design.

Fig. 3. Examples of ASM detection and definition.

3.2. Algorithms for HM-gaze tracker

The whole gaze tracking process for HM-system includes several steps. All of them target to recover 3D points from 3 images acquired with the gaze tracker. Here after, vision eye detection and tracking, and eye movement prediction approaches are briefly outlined.

3.2.1. Eye detection and tracking with radial transform and particle filter

All existing eye tracking approaches can be split into two classes: non probabilistic [6], [7] and probabilistic [8], [9]; the later seems to better simulate the biological mechanism of tracking in close-up images acquired with a low cost vision "only" camera and with uncontrolled illumination conditions.

The defined probabilistic approach [10] combines two concepts: the sequential Monte Carlo algorithms (SMC, known also as a particle filter) and the radial symmetry transform.

The SMC algorithm [11] allows formulating multi-hypothesis in order to explore the state space (all probable positions of the eye in the next image) using the currently acquired image in order to find the position of the eye in the next image. As a particle filter converges to the true posterior probability density function (pdf) with the increase of particle number (theoretically, with their infinite number), the SMC is a time consuming exploration method.

The radial symmetry [12] guides the potential particles' selection and therefore improves the temporal performances of the particle filter tracking. The iris shape is approximated with an ellipse of center (c_x, c_y). Its potential eye movements in any direction from the current pixel p = (x,y) are considered using the radial symmetry which accumulates contributions of magnitudes and orientations of luminosity function of pixels in the p neighborhood in different distances (radii) r from p in the luminosity gradient orientation. Fig. 4 outlines the proposed SMC-radial symmetry approach.

The particles' selection dynamic model is formulated as a Gaussian mixture including observation at time step t given by a radial symmetry detector. Whenever the symmetry knowledge rises above the known pdf, the old set of samples is replaced by a new set of samples such that sample density better reflects posterior pdf. This eliminates particles with low weights and selects (or generates) particles in regions of highest probability for eye detection. Therefore, the radial symmetry a) makes more robust iris tracking via a particle filter as it generates only the correctly predicted next positions of the eye, b) reduces the volume of calculation, c) handles abrupt motions, and d) automatically recovers from track loss (due to eyelids occlusion for example).

Fig. 5 shows the efficiency of the proposed approach when comparing the ground true data (red line) with data obtained by the SMC-radial transform tracking algorithm (blue line).

3.2.2. Eye movement prediction basic model

The eye movement prediction can be used as assistance for an object displayed on a PC screen. Zelinsky's bottom-up model for target acquisition [5] has been investigated as a potential model for 3D point of regard (PoR) or gaze predication (anticipation). Indeed, during the interaction with a 3D environment a

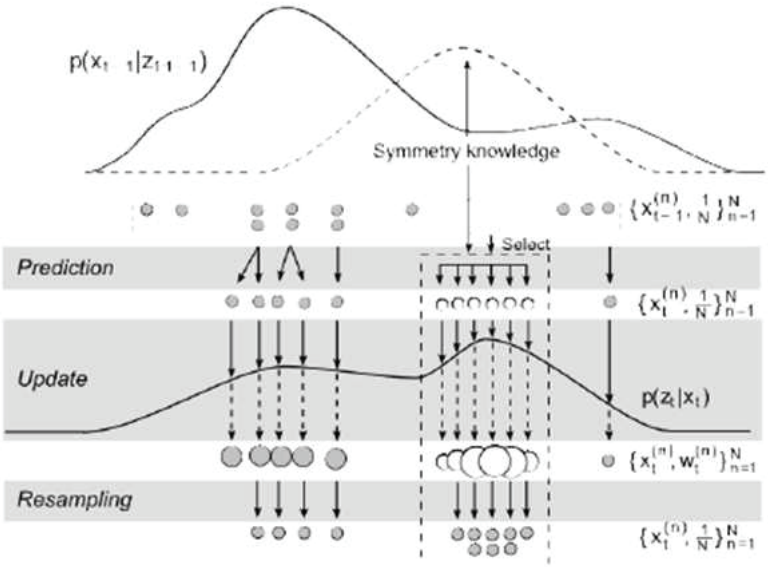

Fig. 4. Radial symmetry guided particle filter (the grey particles x_t are generated at instance t according to the probability $p(x_t/x_{t-1})$, while white particles are propagated thank to the system status z_t. $q_{obs}(x_t/x_{t-1}, z_t)$).

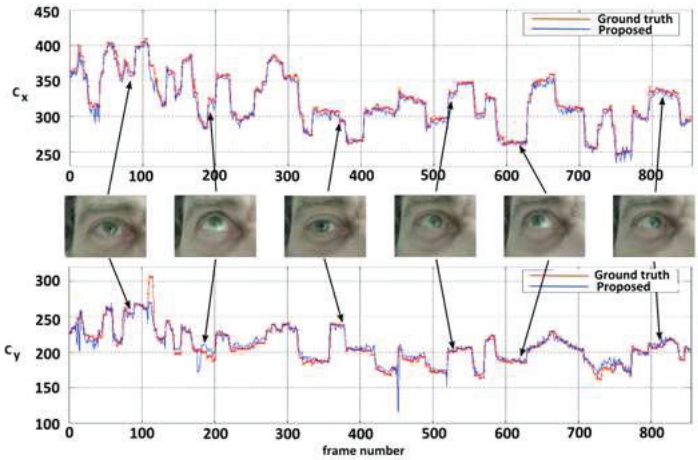

Fig. 5. Graphs of iris centre coordinates (comparison between ground truth and the SMC-radial transform approach).

search for a predefined object is frequently a basic operation; the search of the 3D object is a fundamental approach for skill acquisition in rehabilitation contexts where perception-action are usually associated.

Zelinsky's model proceeds in the following main steps:

1. generation of the potential targets' map (or HS map = hot spots' map) in retinal image associated to an image of a 3D scene (with random noise adjunction for multiple targets differentiation);
2. potential target selection (adaptive thresholding of the cross-correlation), and
3. eye movement generation (as a function of threshold and saccade amplitude) or
4. rejection of a false target.

Fig. 6 shows a typical example of target map with current fixation point (green cross), target (cible, blue cross) and hot spot (red cross); the next saccade will move the fovea toward the hot spot.

Zelinsky's model of target acquisition fails with scene illumination changes.

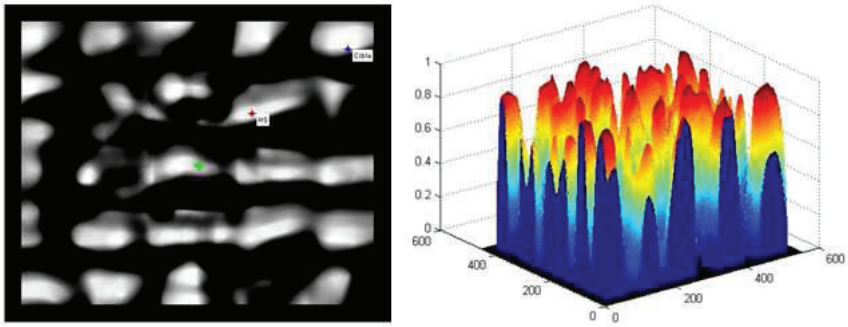

Fig. 6. Simulation of a typical target map in retina (Left: 2D foveal. Right: 3D representation).

Fig. 7. Example of object tested for target acquisition algorithm validation.

3.2.3. Improvement of the eye movement prediction model

We improved Zelinsky's model by making the potential targets' map invariant to illumination. Indeed, this later has been replaced by a map of saliency established from object features' appearance probabilities in 3D (a Gaussian distribution) obtained via learning using specific image base (The Amsterdam Library of Object Image, [13]); each element of saliency map is a probability that the corresponding location contain a given target. Fig. 7 gives an example of objects used for tests which can be localized with the improved model.

4. Evaluation a gaze tracker with end users

Before training end-users for new skills acquisition [14], our gaze-tracker has been evaluated with primary users affected by quadriplegia, cerebral palsy, stroke, amyotrophic lateral sclerosis (ALS), multiple sclerosis, and muscular dystrophy. All the end users have the perceptual abilities for gaze-based training such as vision, hearing, perceptual memory, sense of time, and low-level cross modal integration (ie. vestibulo-ocular and opto-kinetic responses).

All expected potential benefits from gaze tracker usage have been discussed with primary and secondary users through administrated questionnaires and analyzed with two metrics: Positive Affect and Negative Affect Schedule (PANAS) which estimates the emotional balance; and Satisfaction With Life Scale (SWLS); both metrics allow to obtain a subjective measure of well-being.

The statistical data collected from 50 subjects in Austria, Poland and Spain show that:

1. end users can be divided into two groups: spatially immobile (bedridden subjects) and spatially mobile (including wheelchair users);
2. more than 97% of end-users access to the PC via unconventional peripherals such as joystick, buttons, specific mechanical pointers;
3. the most common interaction modes with a PC is access to the Internet (chat, email, search: 42%), electronic games (36%), and sometimes even work (22%).

Our gaze tracker system is designed in order to allow new skills acquisition for a full exploration of a PC's possibilities as an interaction and communication tool.

During the first qualitative user evaluations of the AsTeRICS system (June-August 2011), different sensors and sensor-combinations have been tested. The remote (web camera-based) system was already available for these tests and reached the highest level of acceptance in the evaluations of users with given fine motor control of the head.

Operations such as object (basic icons) search, detection, and selection (dwell time) on a PC screen was evaluated through specific protocols, and with a great success as almost all users have succeeded control in some way (without any training) the PC mouse with a nose and with a chin.

Access to the Internet, Internet navigation and Internet-based services (games, e-library, e-shopping, e-health, e-rehabilitation, e-learning, etc.) were not tested.

However, for gaze tracker interactions, spasms and involuntary head movements represented a big problem, preventing precise pointing or computer mouse control. The application of tremor reduction approaches to mouse simulation should be integrated in new version of software (via image stabilization, for example) and once again evaluated with end-users.

It should be added that present system reaction time is adaptable, but not dynamically, to the end-user needs.

5. Conclusion

This paper has introduced the concept of low cost gaze tracker as a tool for new skills acquisition for gaze interactions with a computer screen for upper limb motor impaired. The principles of hardware and software design of a low cost remote and head-mounted gaze tracker have been addressed. The very first pre-prototype of a mechanical support for gaze tracker is adjustable for end-user specific head morphology, his/her current interaction capabilities and to different processing configurations (one or several, IR or visible spectrum cameras). The designed new algorithms for features' tracking and eye movement predictions are executed at the speed closed to video rate. However, the system precision of 5 degrees should be improved.

Moreover, a new lighter and more flexible mechanical support using "spy cameras" should be designed. The new system should be able to compensate head movement, tremor reduction and should adapt the point of regard Zelinsky's prediction approach to work in natural scenes.

Furthermore, two future applicative research directions are: dynamically adjustment of system reaction time (via learning from end-user) and support for skill learning for access to the Internet and Internet supported services (games, e-library, e-shopping, e-health, e-rehabilitation, e-learning, etc.).

Finally, the low cost gaze tracker opens new possibilities for gaze-guided (visual attention) computer vision, visual lifelong computing, and controlled skill for interaction learning.

Acknowledgments

This research is supported by the European project AsTeRICS (FP7 ICT). We thank the students (from MIT who, in the frame of France-MIT program, have contributed to implementation of the mechanical support for the gaze tracker and from UPMC who have participated in software implementation).

References

[1] F. Shi, A. Gale, and K. Purdy, "A New Gaze-Based Interface for Environmental Control". In C. Stephanidis (ed.), LNCS 4555, Springer, pp 996-1005, 2007.
[2] Tobii Glasses. Updated information available at: http://www.tobiiglasses.com/scientificresearch/
[3] D. Hansen and J. Qiang, "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze", *IEEE PAM*I, 32(3), pp 478-500, 2010.
[4] F. Mulvey, A. Villanueva, D. Sliney, R. Lange, S. Cotmore, and M. Donegan, "Exploration of Safety Issues in Eye Tracking", *Communication by Gaze Interaction (COGAIN)*, EU IST-2003-511598, 2008.
[5] G. Zelinsky, "A Theory of Eye Movements during Target Acquisition", *Psychol. Rev.*, 115(4), pp 787-835, 2008.

[6] W. Ryan et al., “Matching-Moving for Area-Based Analysis of Eye Movements in Natural Tasks”, *ETRA 2010*, pp 234-242, 2010.
[7] D. Li and D. Parkhurst, “Open-Source Software for Real-Time Visible Spectrum Eye Tracking”, *COGAIN 2006*, pp 18-20, 2006.
[8] D. Hansen and A. Pece, “Iris Tracking with Feature Free Contours”, *AMFG 2003*, pp 208-214, 2003.
[9] H. Wu et al., “Tracking Iris Contour with a 3D Eye Model for Gaze Estimation”, *ACCV 2007*, pp 688-697, 2007.
[10] F. Martinez, A. Carbone, and E. Pissaloux, “Radial Symmetry Guided Particle Filter for Robust Iris Tracking”, *IAPR CAIP*, 2011.
[11] R. van der Merwe et al., “The Unscented Particle Filter”, *Advances in Neural Information Processing Systems (NIPS13)*. In T. Dietterich, T. Leen, V. Tresp (eds), 2000.
[12] A. Loy and G. Zelinsky, “Fast Radial Symmetry for Detecting Points of Interest”, *IEEE PAMI*, 25 pp 959-973, 2003.
[13] J. Geusebroek, G. Burghouts, and A. Smeulders. “The Amsterdam Library of Object Images”, *Int. J. of Computer Vision*, 61(1), pp 103-112, 2005.
[14] G. Evreinov, “Skills vs. Abilities”, *Introduction to the Specific Thematic Session - ICCHP’08*, pp 1160-1165, 2008.