

Toward the design of a low cost vision-based gaze tracker for interaction skill acquisition.

Edwige Pissaloux Andrea Carbone Francis Martinez
Université Pierre et Marie Curie (Paris 6), ISIR, France
E-mail: Edwige.Pissaloux@upmc.fr, {carbone, martinez}@isir.upmc.fr

Abstract

The human gaze is a basic mean for non verbal interaction between humans; however, in several situations, especially in the context of upper limb motor impairments, the gaze constitutes also an alternative mean for human's interactions with the environment (real or virtual). Mastering these interactions through specific tools, requires frequently the acquisition of new skills and understanding of mechanisms which allow to acquire the necessary skills. Therefore the technological tool is a key for new interaction skills' acquisition. This paper presents a tool for interaction skill acquisition via a gaze. The proposed gaze tracker is a low cost head mounted system based on vision technology "only". The system hardware specifications and the status of the gaze tracker design are presented; the dedicated algorithm for eye detection and tracking, and an improvement of G. Zelinsky model for eye movement predication during the search of a predefined object in an image are outlined. Results of the software preliminary evaluation are presented.

1. Introduction

Human gaze interaction is a recent interaction mode. The gaze is 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 surgeon or fighter pilot) are in the obligation 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 [4], fitness, social interaction, team building... are some of the examples. The usage of the gaze tracker as the interaction mean requires the acquisition of special skills such as control of eye movement (dwelling, displacement in precise direction and at giving speed, eyelids blinking at specific time, etc.).

However, the quality of such skill acquisition is strongly influenced by the used technology.

In general, two configurations for eye- and gaze trackers are commercially available : remote eye trackers (RET) and head-mounted eye trackers (HMET). These two configurations are also respectively known as *passive* and *active eye monitoring*.

Remote eye trackers are not well-suited for applications where mobility is required; indeed, subject - remote eye tracker relative speed should be very slow and the maximal displacement very limited. Moreover, a planar surface is used as a mediation tool for interaction. Indeed, one of the two eyetracker cameras is always placed in a manner that it can track the eyes of the user. The camera moves in the same frame of reference 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 [2] or ASL desktop/remote [1] solutions are the most popular desktop/remote eye tracking system.

Head-mounted eye tracking systems (HMET) are experiencing a renaissance due to new challenges [3,4]. 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 the head movements.

However, the all the existing eye/gaze tracker are rather expensive (around 2K€ at minimum), and use the near infra-red technology (IR-A, 780-1400nm). 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 with are still remaining an open question [5]. For all of the above reasons, a design of a low cost only vision-based gaze-tracker is on the progress in the frame of AsTeRICS, FP7 ICT project [6], and its status is rapidly presented in the subsequent sections. Section 2 proposes the system main mechanical and hardware specifications. Section 3 addresses systems algorithmic developments : new

algorithms for eye detection/tracking and an improvement of G. Zelinsky's model for eye movement predictions while exploring an image of 3D scene [7]. Section 4 identifies skills possible to acquire with a gaze tracker for interaction purposes. Section 5 concludes this paper and proposes some future research directions.

2. Low cost vision based gaze tracker system architecture.

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

2.1. Low cost gaze tracker end-user specifications.

The figure 1 schematizes the targeted gaze tracker as a peripheral for a PC supporting ARE (AsTeRICS Runtime Environment). The hardware system is composed from specific frame grabber (a netbook Samsung). It acquires images from two eye cameras and from a scene camera (Sony Ex-View CCD). The software synchronised and conveniently processed images are displayed on the screen of the frame grabber or transferred from there to ARE via the USB.

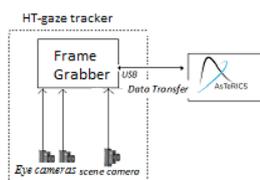


Fig. 1. Schematic architecture of the low cost vision gaze tracker.

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 to wear the correction glasses; c) adjustable size to different head morphologies; d) precision compatible with targeted skills' acquisition; e) lightweight system; f) easy to wear; g) simple calibration procedure.

The system works at the video speed.

2.2. Gaze tracker pre-prototype architecture.

The figure 2 shows the targeted gaze tracker ; its main components are : mechanical support, cameras and (remote, not shown on the picture) frame grabber (processing unit).



Fig. 2. AsTeRICS (UPMC) targeted gaze tracker.

The mechanical support is adjustable with straps on the back of the head. Its two independent boom arms, left and right, support two cameras with adjustable to the end user's eyes distance l and orientation $\square\square\square\square$ DoF - 6 degrees of freedom). The gaze tracker will work in indoor (constant illumination) and outdoor environments (varying illumination). The scene camera is located on the front band attached to forehead. Eye cameras are placed « in front of the eye » in order to reduce the projective distortions of the eye on the eye camera (distortions which should be corrected by the software). Two eye cameras allow depth recovery (thus point of regard estimation).

The mechanical support is under design with SolidWorks¹, and will be prototyped with a Rapid Prototyping tool.

3. Algorithmic developments.

The whole gaze tracking process includes several steps. All of them target to recover a 3D point from 3 images acquired with the gaze tracker. Here after, independent eye detection and eye movements and predictive approaches are briefly outlined.

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

All existing eye tracking approaches can be split into two classes: non probabilistic [8,9] and probabilistic [10,11]; the later seem better simulate the biological mechanism of tracking in close-up images acquired with a low-cost camera and with uncontrolled illumination conditions. The defined approach combines two concepts: the sequential Monte Carlo algorithms (SMC, known also as a particular filter) and the radial symmetry transform.

The SMC algorithm [14] allows formulating multi-hypothesis in order to explore the state space (all probable positions of the eye in the next image) using

¹ <http://www.solidworks.fr/pages/programs/letsdesign/>

the currently acquired image and to find the position of the eye in the next image. As a particular filter converges to the true posterior probability density function (pdf) with the increase of particle number (theoretically, with their infinite number), the SMC is time consuming exploration method. The radial symmetry [13] guides the potential particles selection and improves the temporal performances of particular filter. The radial symmetry has been selected because of eye symmetric shape and a potential eye movement in any direction from the current pixel $p = \{x, y\}$. This transform accumulates contributions of magnitudes and orientations of luminosity function of pixels in the p neighbourhood in different distances (radii) r from p in the gradient orientation.

Figure 3 outlines the proposed approach.

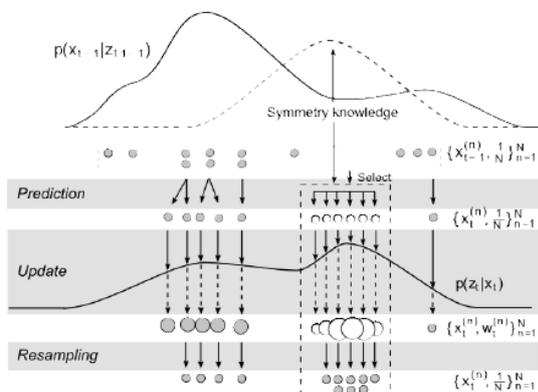


Fig. 3. Radial symmetry guided particle filter (the grey particles are generated according to $p(x_t/x_{t-1})$, while white particles are propagated by $q_{obs}(x_t/x_{t-1}, z_t)$.

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 new set of samples such that sample density better reflects posterior pdf. This eliminates particles with low weights and selects (or generates) particles in more probable regions.

Consequently, the radial symmetry robustifies iris tracking via a particular filter: generates only the correctly predicted next positions of the eye, reduces the volume of calculation, handles abrupt motion and automatically recovers from track loss (due to eyelids occlusion for example). Figure 4 shows the efficiency of our method when comparing the ground true data (red line) with data obtained by the proposed tracking algorithm (blue line).

3.2. Eye movement prediction basic model.

G. Zelinsky's bottom-up model for target acquisition [7] has been investigated as a potential model for 3D point of regard (PoR) or gaze estimation. Indeed, during the interaction with a 3D environment a 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.

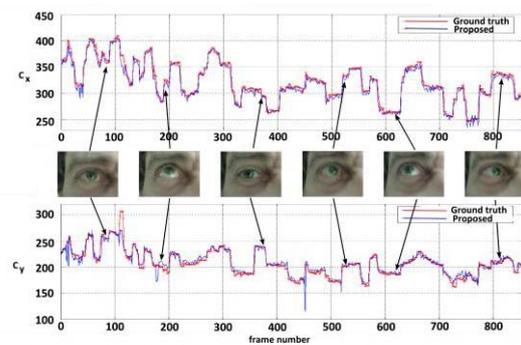


Figure 4. Graphs of center coordinates of the iris (comparison between ground truth and the proposed approach).

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 noisy 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. Figure 5 shows a typical example of target map with current fixation point (green cross-hair), target (cible, blue cross-hair) and hot spot (red cross-hair); the next saccade will move the fovea towards the hot spot.

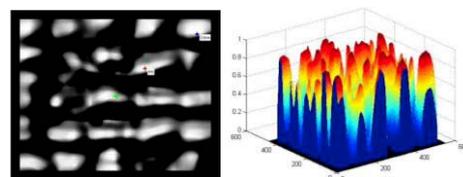


Fig. 5. Simulation of a typical target map in retina (2D foveal (left) and 3D representations).

Zelinsky's model of target acquisition fails when the target orientation changes in 3D (changement of a scene viewpoint).

3.3. Improvement of the eye movement prediction model.

We improved Zelinsky's model by making invariant to the point of regard (i.e. to the rotation in 3D) the potential targets' map. Indeed, this letter has been replaced by a map of saliency established from

object features' appearance probabilities in 3D (a Gaussian distribution); each element of saliency map is a probability that the corresponding location contain a given target. Figure 6 gives an example of objects used for tests (borrowed from the Lübeck University library) which can be localized with the improved model.



Fig.6. Example of object tested for target acquisition algorithm validation.

4. Skills' acquisition with assistance of a gaze tracker.

The gaze tracker system is designed in order to allow motor (upper limb) impaired people to acquire new skills for a full exploration of a PC possibility as an interaction and communication tool. Operations such as object (basic icons) search, detection, and selection on a PC screen will be evaluated through specific protocols (under design); the access to the internet and navigation will be considered.

New skills will be acquired by repetition of action with a feedback on a computer screen.

The system reaction time is adaptable, but not dynamically, to the end-user needs.

The gaze tracker presentation to the end user will be done by secondary users (nurses, para-medical staff, etc.). The built system will be tested with end users affected by quadriplegia, cerebral palsy, stroke, amyotrophic lateral sclerosis (ALS), multiple sclerosis, muscular dystrophy in three European countries (Spain, Poland, Austria).

5. Conclusion.

This paper introduces gaze tracker as tool for new skill (via repetition) acquisition for interactions with a computer screen.

The hardware and software design of a low cost and head-mounted vision technology only based gaze tracker has been addressed. The mechanical support is adjustable for end-user specific head morphology and current interaction capabilities. The system precision of should be qualified. The next prototype of the system will evaluation the possibility of inclusion a "spy" cameras within the gaze tracker having a mechanical support like usual glasses and head movement compensation mechanism.

The system has been designed in order to allow motor (upper limb) impaired people to acquire new skills for a full exploration of a PC possibility with adjustable system reaction time. Two futures research directions target to dynamically adjust system reaction

time (via learning from end-user) and provide a support for skill learning for access to the internet (and to other screen based applications).

Finally, the low cost vision based only gaze tracker opens new possibilities for gaze-guided (visual attention) computer vision, interaction with multiple screens in ubiquitous 3D environments, svisual lifelog computing and controlled skill for interaction learning.

6. Acknowledgement

The presented here research is supported by the European project AsTeRICS (FP7 ICT).

We thank you the MIT students who, in the frame of France-MIT program, have contributed to implementation of the mechanical support for gaze tracker and some software implementations. We thank you our Master student, M. G. Jaza for his implementation of the eye movement models.

7. References

- [1] <http://asleyetracking.com/site/>
- [2] <http://www.tobiiglasses.com/scientificresearch/>
- [3] Hansen, D.W., Qiang, J. "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze", IEEE PAMI, 32(3), pp. 478-500, 2010.
- [4] Shi, F., Gale A.G., Purdy, K., A New Gaze-Based Interface for Environmental Control, in C. Stephanidis (ed.), LNCS, vol. 4555, Springer, 996-1005
- [5] Mulvey, F., Villanueva, A., Sliney, D., Lange, R., Cotmore, S., Donegan, M. (2008), Exploration of safety issues in Eyetracking. Communication by Gaze Interaction (COGAIN), IST-2003-511598: D 5.4.
- [6] <http://www.asterics.eu/>
- [7] Zelinsky, G.J., A Theory of Eye Movements during Target Acquisition, Psychol. Rev. 2008 October, 115(4), 787-835
- [8] Ryan, W.J., et al., Matching-moving for area-based analysis of eye movements in natural tasks, Proc. ETRA, 2010, 234-242
- [9] Li, D., Parkhurst, D., Open-source software for real-time visible spectrum eye tracking, COGAIN, 2006, 18-20
- [10] Hansen, D.W., Pece, A.E.C., Iris tracking with feature free contours, proc. AMFG, 2003, 208-214
- [11] Wu, H., et al., Tracking iris contour with a 3D eye model for gaze estimation, ACCV, 2007, 688-697
- [12] Martinez, F., Carbone, A., Pissaloux, E., Radial Symmetry guided Particle Filter for Robust Iris Tracking, Proc. CAIP 2011, Spain
- [13] Loy, G., Zelinsky, A., Fast radial symmetry for detecting points of interest, IEEE PAMI, 25 (2003), 959-973
- [14] van der Merwe, R.et al., The Unscented Particle Filter, Advances in Neural Infor. Proc. Systems (NIPS13), T.G. Dietterich, T.K. Leen, V. Tresp (eds), Dec 2000