MOVIES MADE THROUGH THE EYES OF A MOBILE USER WITH A GAZE-ALIGNED CAMERA

E. Schneider[1*], *K. Bartl*[1], *G. Böning*[1], *T. Dera*[1], *P. Wagner*[2], *S. Bardins*[1], *M. Enzmann*[3], *T. Brandt*[1]

[1]Hospital of the Ludwig-Maximilians University of Munich, Neurology, D-81377 Munich, Germany
[2]Technical University of Munich, Institue of Applied Mechanics, D-85748 Garching, Germany
[3]40° Filmproduktion, D-80469 Munich, Germany

ABSTRACT

Head-fixed camera systems are widely known, but since they are aligned by the head only and not by the eyes they are not able to always look at what the cameraman is looking at, and the image quality is poor if no effort is made to stabilize the camera. Systems like Steadycam, in contrast, focus on image stabilization at the cost of restricting the cameraman in his actions. The prototype of a new mobile head-mounted camera system was developed that is continuously aligned with the orientation of gaze. In doing so, the biological gaze stabilization reflexes are used to keep the video camera stable on target. Applications like movie making through the eyes of an actor or documentary movies for sports and other activities are conceivable. The system was tested by a surgeon who could successfully document his activities.

1. INTRODUCTION

Wearable head-mounted camera systems that document and log not only day-to-day but also complex activities have been a vision for more than half a century [1]. Only with the advent of miniaturized camera systems during the last decade has this vision become more and more a reality in different scientific fields [2, 3]. One major advantage of a wearable head-mounted camera system is that the cameraman can comfortably wear the camera without any restrictions of eye, head, legs, or arm mobility during camera operation tasks and without limitations of his field of view. Thus, different actions can be performed in a natural way.

However, as the human body is usually in motion, the quality of the recorded video stream of a head-mounted camera is fairly poor if no effort is made to stabilize the camera system. Such a stabilization might, for example, consist of an inertial measurement unit that operates in six degrees of freedom (6DoF) and a motion device to counter-rotate the camera around three axes (3DoF). An alternative but heavier approach is implemented for example in the Steadycam systems. However, even a stabilized camera system has the disadvantage that the recorded video stream will not contain the spontaneous view of the user, since the eyes move relative to the head when not restricted by a viewfinder. Therefore two major benefits are gained when human eye movements are used as input signals for the camera motion device: on the one hand, the biological gaze stabilization can be used to keep the gaze direction of the video camera stable, and on the other, the user's gaze direction will control the video camera so that the camera always sees what the cameraman sees.

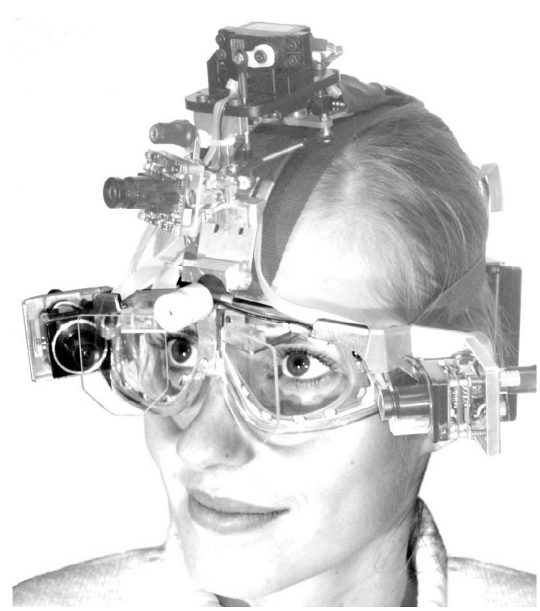

Fig. 1. A gaze aligned head-mounted camera. The system consists of a pair of eye tracking video cameras that are laterally attached to swimming goggles and of a third camera above the forehead that moves with the eyes.

A first 2DoF proof of concept that orients a camera's optical axis parallel to the user's gaze direction was presented in [4]. In that system, eye movements were measured with an eye tracker. Assuming a sufficiently small latency, the human gaze stabilization based on the biological vestibulo-ocular, optokinetic, and smooth pursuit reflexes can be used

*Funding by Bayerische Forschungsstiftung (FORBIAS, Az: 567/03) and Deutsche Forschungsgemeinschaft (GL 342/11)

to stabilize the camera image. In this camera setup the biological equilibrium organ – the inner ear labyrinth – controls a technical system, hence rendering a technical inertial measurement unit obsolete. The camera therefore mimics the natural exploration of a visual scene by using the sensorimotor output of the human eye movement system – that evolved over millions of years – to move an artificial eye. All the capabilities of multi-sensory processing for eye, head, and surround motions are detected by biological sensory systems and used to drive a technical camera system.

This paper presents the design and implementation of a first 3DoF prototype of a gaze-aligned head-mounted camera system that serves as an experimental platform as well as a possible solution for a head camera system for consumer products. Other applications like neurophysiological, psychological, and marketing research as well as documentary movies for sports and other activities are also conceivable. As a first application we tested the usability for the documentation and teaching of surgery (see Fig. 2). A user wearing the complete system consisting of the eye tracker and the head camera is shown in Figure 1.


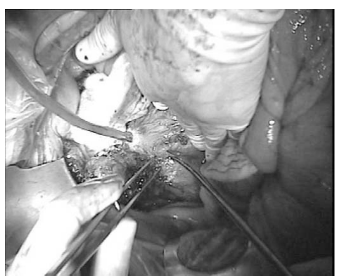

Fig. 2. Surgeon with a gaze-aligned head-mounted camera (left) and the scene from his perspective (right).

2. MATERIALS AND METHODS

Eye movements and their measurement are a thoroghly researched field (for an overview see [5]). Five types of human eye movements can be distinguished: saccades, smooth pursuit, vestibulo-ocular reflex, optokinetic response, and vergence. The eyes can rotate in three dimensions with angular velocities and accelerations of up to 700 $^\circ/s$ and 5000 $^\circ/s^2$, respectively. Saccades and nystagmus quick phases occur at typical rates of 3/s, even during fixations. The latency between a head movement and the corresponding sabilizing eye movement of the vestibulo-ocular reflex is on the order of 10 ms. The implementation of all these eye movements with their respective kinematic properties in a wearable mechanical device can be a challenging engineering task. In contrast, smooth pursuit is less challenging since it operates at velocities of up to 90 $^\circ/s$. The functional significance of other involutary eye movements like tremor, drifts and microsaccades is still under debate. As an alternative to a camera that rotates with the eyes one might think of using a head-fixed wide angle camera in conjunction with a gaze-controlled video generator. However, quick head movements would generate blurred images, rendering such a camera useless in a mobile scenario.

2.1. Eye Tracker

One major aim was to design a lightweight eye tracker at low costs. Lightweight and comfortable swimming goggles (Aqua Sphere Crystal, Technisub, Genova) that were tightly fitted to the head were used as a basis for mechanical support. A flexible, curved flat spring was attached to the goggles to fit the head in the mid-sagittal plane (see Fig. 1). Slippage could therefore be reduced during mobile applications. Each eye was illuminated by five small infrared (IR) LEDs that were distributed along the goggle's lower rim. IR reflecting mirrors allowed the lateral positioning of digital cameras. Similar camera, mirror, and illumination configurations are implemented in many commercially available eye trackers (e.g., SMI, Teltow, and Chronos Vision, Berlin). The exposure of the subject's eyes with IR radiation was at 10% of the maximum permissible value (DIN EN 60825-1: 1997), allowing the use of the camera for many hours.

The eyes can rotate not only around the horizontal and vertical axis (pan, tilt) but also around the line of sight (torsion). Video based eye trackers are a convenient means to measure these 3D eye movements. Since the actuators of the gaze camera require close-to-real-time control, the software was implemented on standard PC-hardware by using well-known eye tracker algorithms that were optimized for short latencies. The video images of the laterally attached cameras are transferred to an off-the-shelf subnotebook computer (Acer TravelMate 382TMi), whose image processing algorithms detect the position of the pupil. In short, pupil detection is initialized by applying luminance thresholds that are determined from the luminance histogram. A connected area labeling algorithm performs a spiral search to detect pupil pixels while removing outliers. Subsequent principal component analysis yields a first estimate of the pupil's ellipse parameters. Finally, an ellipse is linearly fitted to the valid edge pixels.

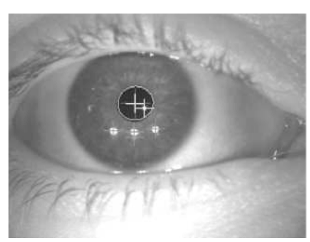
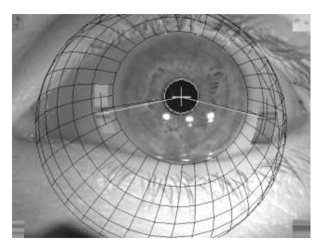

Fig. 3. Analysis of pupil center and corneal reflections denoted by crosses (left) and calculation of eye rotation angles from artificial dark markers on the sclera (right).

With this setup a mean pupil detection resolution of 0.04° ± 0.01° was achieved (25 fixations, one subject). In addition, corneal reflections are analyzed and the vector difference between their coordinates and the pupil center are used to compensate for the residual camera slippage during fast head movements (see Fig. 3, left). For the current prototype the torsional component of the eye position is determined from two artificial markers, that are applied to the sclera by using a dark cosmetic pigment (see Fig. 3, right). This method is only suitable for testing the current prototype and certainly not for the final consumer product. However, the marker method is well established in the field of oculomotor research and it achieves good quality data with resolutions down to 0.01° [6]. For future implementations, ocular torsion calculations from iris signatures are preferred, however, the data quality is expected to have an inferior resolution in the range of 0.1° [6].

2.2. Camera Motion Device

The major purpose of the newly designed camera motion device is to rotate the camera's viewing direction around three perpendicular axes with kinematic properties that are similar to those of the human oculomotor system. The construction has to be extremely compact and lightweight in order to affect the user as little as possible.

The construction of a small 3DoF camera motion device based on a serial design was already proposed in [7]. Instead, we decided to use a parallel structure at the cost of more complex kinematics [8]. Other parallel motion devices for cameras can be found in [9, 10], however, these are not as lightweight as the system shown in Fig. 4.

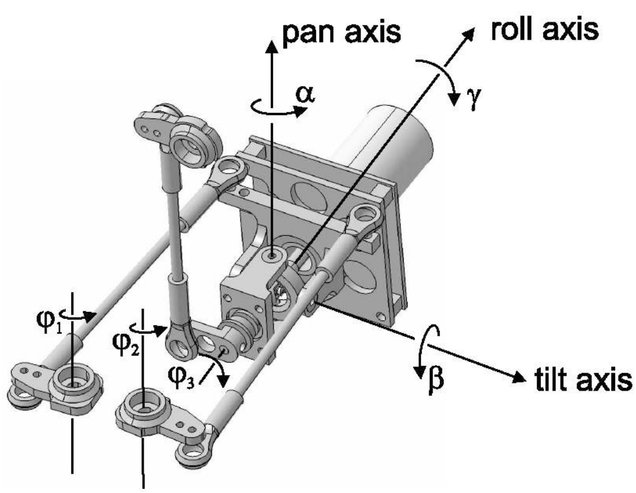

Fig. 4. Kinematics of the 3DoF camera motion device. The CAD model shows the actuator angles ($\varphi_1, \varphi_2, \varphi_3$) and the Cardan angles (α, β, γ).

2.3. Drive Specification

The use of small cameras allows the choice of servo actuators designed for miniature aircraft models that must be able to reach high velocities and accelerations in order to perform like the human oculomotor system. Other important requirements are low weight, small dimensions, and high reliability. The chosen actuators (Graupner DS 3781, Kirchheim, Germany) are among the fastest servo actuators that still have a reasonable size and weight (1000 °/s, 28 g). With these components a camera motion device can be demonstrated with dynamic properties that come close to or even excel those of the human eye.

2.4. Calibration

Once the difference vector between the pupil and the corneal reflection is detected in the coordinate system of Fig. 3, it is then transformed into the coordinate system defined by the angles φ_1 and φ_2 from Fig. 4. This transformation is accomplished with a set of variables that are determined in a calibration procedure. This procedure was especially designed to minimize user interaction and was described previously together with the involved algorithms [4]. In short, a small laser pointer that is rigidly attached to the camera module projects a bright dot that appears centered in the head camera image. The user is asked to fixate it as it jumps to predefined φ_1 and φ_2 positions. The eye position values are then linearly fitted by two 3rd order 2D-polynomials to the corresponding motor commands φ_1 and φ_2.

2.5. Electronical Setup

In order to ensure image stabilization and wearability the electromechanical components were all chosen on the basis of weight considerations and of a low over-all system latency . A pair of IEEE-1394 digital cameras (Flea, Pointgrey, Vancouver) were used for eye tracking. With pixel binning the cameras can operate at 100 Hz with a pixel resolution of 320x240. A PIC18F4431 microcontroller generates the PWM position signals for the three off-the-shelf servo actuators. An above-standard sampling rate of 300 Hz is used for the PWM signal to reduce latency. The servo controller is connected to a RS232 interface working at 115200 Baud.

The controller additionally serves as a power supply system for the actuators, the calibration laser, and the head-mounted camera. A switching voltage regulator and a DC/DC-converter generate the required voltages from the power supply of the IEEE-1394 connector of the laptop. Only with these prerequisits was it possible to use the laptop battery as the power supply for the whole system. A Firewire-hub (Macally FH-220) is used to connect both eye tracker cameras. It is assembled together with the servo controller in a small plastic case and attached to the back of the head mount.

3. RESULTS

3.1. Latencies

To measure the latency of the complete system, we attached an artificial eye to an additional actuator. The eye tracker detected the artificial eye's pupil as it would with a real eye and moved the head-mounted camera accordingly. Two gyroscopes were attached to the artificial eye and to the head-mounted camera in order to synchronously measure their angular velocities. The time difference between the two (sinusoidal) movement profiles was used as a measure for latency (see Fig. 5). When we used our standard camera with a frame rate of 100 Hz an over-all mean latency of 36.3 ms was measured at 1, 2, 3.3, 5, and 10 Hz. This value is relatively close to the latency of the human vestibulo-ocular reflex, which is on the order of 10 ms [5]. In addition, we also tested a custom-made eye tracker camera (courtesy of Chronos Vision, Berlin, Germany) at 200 Hz and this one yielded a latency of 20 ms.

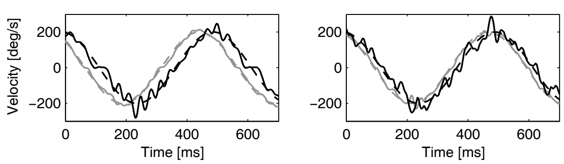

Fig. 5. Over-all system latency for a standard IEEE-1394 camera (left) and for a custom-made low-latency VOG camera (courtesy of Chronos Vision, right). Velocities of the head-mounted camera (black) and of the artificial eye (gray) as well as the corresponding sine fits (dashed) are shown.

3.2. Precision and Resolution

To measure the precision of the camera positioning, we measured the angular distance between 25 predefined fixation dots on a wall (1.18 m away) and the corresponding fixation positions that were indicated by the luminous dot of the camera-attached calibration laser. The fixation targets were regularly aligned on a rectangular grid (8.5 deg distance). This measurement yielded a mean precision of 0.5 deg and a resolution (standard deviation) of 0.25 deg.

4. CONCLUSION

Video recordings from the subjective perspective of a mobile user are possible with this 3DoF prototype of a gaze-aligned head-mounted camera. The accuracy of the whole system – including the user – is in the range of 0.5 deg, which is well within the 2 deg of foveal vision. Such a camera system can already be used for health care, quality assurance, movie making, and even sports. A consumer product based on this prototype is also conceivable. The camera was tested in a surgical environment and yielded movies seen through the eyes of the surgeon. The recorded video streams show that the camera is well able to log and document surgical activities for teaching purposes or liability questions. Future developments will focus on improving the latency even further with faster eye tracker cameras and on optimizing the ergonomic constraints.

5. REFERENCES

[1] V. Bush, "As we may think," *The Atlantic Monthly*, vol. 176, pp. 101–108, 1945.

[2] J. B. Pelz and R. Canosa, "Oculomotor behavior and perceptual strategies in complex tasks.," *Vision Res*, vol. 41, pp. 3587–3596, 2001.

[3] A. E. Patla and J. N. Vickers, "How far ahead do we look when required to step on specific locations in the travel path during locomotion?," *Exp Brain Res*, vol. 148, pp. 133–138, 2003.

[4] E. Schneider, K. Bartl, S. Bardins, T. Dera, G. Boening, and T. Brandt, "Eye movement driven head-mounted camera: It looks where the eyes look," in *Proceedings of the IEEE Conference on Systems, Man and Cybernetics (SMC2005)*, Hawaii, USA, October 2005, pp. 2437–2442.

[5] R. J. Leigh and D. S. Zee, "The neurology of eye movements," Oxford University Press, New York, Oxford, 1999.

[6] E. Schneider, S. Glasauer, and M. Dieterich, "Comparison of human ocular torsion patterns during natural and galvanic vestibular stimulation.," *J Neurophysiol*, vol. 87, pp. 2064–2073, 2002.

[7] W. W. Mayol, B. Tordoff, and D. W. Murray, "Wearable visual robots," in *Proceedings of the International Symposium on Wearable Computing*, 2000.

[8] S. Riebe and H. Ulbrich, "Modelling and online-computation of the dynamics of a parallel kinematic with six-degrees-of-freedom," Archive of Applied Mechanics, Vol. 72, No. 11-12, Springer-Verlag, Berlin, Heidelberg, New York, July 2003.

[9] H. Truong, S. Abdallah, S. Rougeaux, and A. Zelinsky, "A novel mechanism for stereo active vision," in *Proceedings of the Australian Conference on Robotics and Automation (ACRA2000)*, Melbourne, Australia, August 2000.

[10] C. M. Gosselin and J.-F. Hamel, "The agile eye: a high-performance three-degree-of-freedom camera-orienting device," in *Proceeding of the IEEE International Conference on Robotics and Automation*, San Diego, 1994.