Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Using Liquid Lenses to Extend the Operating Range of a Remote Gaze Tracking System

T. Poitschke, E. Bay, F. Laquai,
and G. Rigoll
Technische Univerität München
Institute for Human-Machine-Communication
Theresienstr. 90, 80333 Munich, Germany
{poi,bay,laq,ri}@mmk.ei.tum.de

S. Bardins, K. Bartl, J. Vockeroth,
and E. Schneider
Chair for Clinical Neurosciences
University of Munich Hospital
Marchioninistr. 23, 81377 Munich, Germany
{sbardins,kbartl,jvockeroth,eschneider}
@nefo.med.uni-muenchen.de

***Abstract*—Remote eye tracking systems are widely used evaluation tools in many disciplines. Such systems have to deal with free head motions that cause a defocussing of the camera image. Hence, remote eye tracking systems usually have a limited operating range. This contribution presents a novel auto-focus setup for a small and mobile remote gaze tracking system that enables a clear eye image acquisition in a large operating range. Therefore, our setup uses a miniature liquid lens to achieve smallest possible focusing latencies and installation space. Using this technology, we extended the eye tracking system's operating range to distances from 0.6 m up to 1.3 m, with a mean accuracy of 0.45° in 4 exemplary working distances.**

***Index Terms*—liquid lens, auto-focus, focusing, eye tracking, gaze tracking**

I. Introduction and Motivation

Today, eye and gaze tracking systems are widely used in many scientific and application-oriented disciplines. These include areas of medicine such as ophthalmology, where eye tracking is used to analyze ocular motor disorders, or neurology, where neurological control mechanisms of eye movements are investigated. Another application area is the field of human-machine interaction. Here, such systems are used as a novel interaction modality that enables gaze controlled applications in the aid of physically challenged patients. Thus, several handicapped authors managed to write a book using gaze-operated on-screen keyboards [10]. In the field of usability and media research such systems are used, for example, to optimize the design of websites and user interfaces based on the analysis of scan paths of the users' eyes. Also, the cognitive sciences and psychology use such systems to investigate the relationship between eye movements and cognitive factors such as fatigue, attention, and mental workload.

Many applications typically use non-invasive remote eye tracking systems which use remote video cameras in front of the user. Therefore, these systems have to deal with natural occurring head movements. These cause changes in the distance between the camera and the observed object. A best possible image quality is therefore required in order to calculate gaze data with reduced noise. If the user is required to freely move in front of the tracker, a defocussing of the camera system is unavoidable in a traditional fixed-focus setup. Therefore, a fast auto-focus functionality is required in order to enhance the image sharpness even with an unconstrained head. Furthermore, this functionality is a mandatory feature for enabling multi-user scenarios with varying camera-user distances. To maintain the installation space at a minimum, large motor-driven focus adjustments are not to be considered for such functionalities. In this contribution, we propose a tunable focus setup that uses a very fast and small liquid lens to address these problems.

Fig. 1. Typical scenario for gaze-based interaction. The user is tracked by a remote camera, and his eye movements and gaze point are analyzed.

II. Background

A. Eye and Gaze Tracking

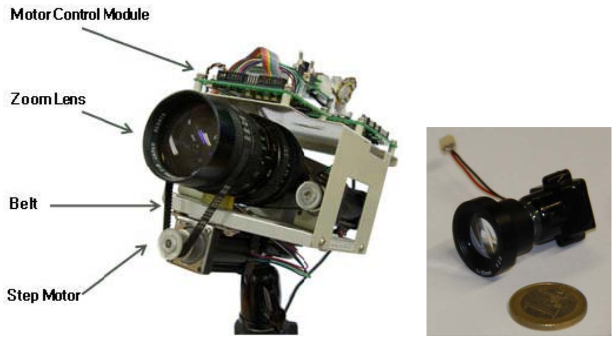
First eye tracking techniques have been developed more than a century ago [19] and modern computer-based video eye trackers are in use since 1974 [13]. The cameras of remote gaze tracking systems are typically placed at a distance of approx. 60-100 cm in front the user. The remote systems use different techniques and algorithms for the computation of the gaze in 3D space. One of the most often used techniques is the detection of the orientation of the eye, using cornea reflections from the eye, e.g., [7]. Another technique is the video-based pupil or limbus detection with additional 3D ellipse reconstruction using stereo camera systems [9].

There are numerous commercial remote eye trackers available (e.g., ASL, EyeLink, Tobii, SMI). The accuracy that can be obtained from commercial remote systems is approximately 0.5°. The size of the operating range in which the user is able to make free head movements, is from (width x height x depth) 15x18x20 cm (EyeLink) up to 40x40x30 cm (SMI).

The size of this area is not large enough for systems that are used for example in multi-user human-machine interaction scenarios. Gaze tracking systems containing a mechanical zoom optic to enlarge the operation range are well known [4]. Furthermore, the extension of a remote eye tracker with pivotable cameras increases the user's workspace in the vertical and horizontal directions without loosing camera resolution. Such pivotable tracking systems were also used for face, gaze, and object tracking [4], [21]. Another typical limitation is caused by fixed focal lengths. The combination of focal length, aperture, and short exposure time provides a very poor image quality. With such systems, only a small depth of field is achieved. Hence, an auto-focus functionality has been proposed to extend the workspace of such systems [12].

B. Auto-Focus Setups

There are numerous approaches to realize the auto-focus functionality for camera setups for various application areas. For example, [16] presents a very small setup for mobile phones that uses a voice coil motor to position the lens (movements only orthogonal to the image plane).



(a) (b)

Fig. 2. (a) shows the camera and hardware setup for the auto-focus used in [12]. (b) shows our used 35 mm lens and micro lens mount with included liquid lens.

[11] proposed an auto-focus system based on a video camera with motor-driven optics. The lens used motor-driven focus, zoom, and iris. Furthermore, [12] presents an auto-focus setup that uses an electric motor to drive the focus ring on the optics. [6] presents an auto-focus setup for iris imaging. The auto-focus optic is composed of several individual lens elements, and multiple moving parts.

However, such setups (e.g., see fig. 2(a)) are not suited for the integration into a small setup for a mobile and pivotable remote eye tracker, because of size, weight, and torque constraints.

C. Active vs. Passive Auto-Focus Control

Auto-focus systems can be realized using active or passive control mechanisms. In order to detect the distance to the object, camera systems with an active auto-focus use an additional range sensor (e.g., ultrasonic sound or time of flight differences of a pulsed infrared illumination). Passive auto-focus cameras determine the distance to the object using the contrast or the sharpness of the image sections. Here the scene must have enough light and contrast in order to do this computation. Our gaze tracking system uses a passive auto-focus technique.

D. Detection of Sharpness in Images

There are various approaches to detect sharpness and clarity in images. An approach proposed in [8] uses two edge filters (one for vertical edges E_v and one for horizontal edges E_h), and the average from the two values to determine the sharpness in an image. A new method basing on Weighted Blocking Discrete Cosine Transform (WBDCT) is proposed in [5] to evaluate the image sharpness. This method divides images into a given number of blocks of the same size. Each block is given the corresponding weighted value and is performed by Discrete Cosine Transform. Finally, the sharpness evaluation function is calculated.

III. BASELINE SYSTEM

The remote eye tracker that is being developed in our lab previously provided a performance that was similar to the commercial systems listed above. It computed the human gaze direction by using a camera with a fixed focal length at distances up to 1.0 m (operating range: 20x15x10 cm).

The following section presents this baseline system and it gives an overview on the implemented gaze tracking algorithms.

A. Used Gaze Tracking Algorithm

For the eye and gaze estimation we used an adapted and advanced version of the glint-based single camera eye tracking system proposed in [7]. This approach uses a single camera setup combined with two special illumination sources (see fig. 4). The first illumination (composed of 9 infrared LEDs) is mounted near the optical axis of the camera system. While this source is switched on, it causes the red-eye effect in the corresponding image (bright pupil image, see fig. 3(a)). The second illumination is split up into two infrared LED clusters (dark pupil images). One cluster is mounted above the camera and optics, and the second one is mounted below the optics (each cluster is composed of 6 LEDs). Since this illumination is mounted far from the optical axis, those clusters do not cause the red-eye effect (see fig. 3(b)).

Alternating illumination control and difference imaging is used in order to compute the pupil position from the remaining

(a) Bright pupil image.

(b) Dark pupil image.

Fig. 3. Eye region from the camera images used for the gaze estimation. The images show the bright and the dark pupil, and the tracking results form the pupil. The pupil position is obtained from the difference image.

blobs in the binary difference image. Binarization is achieved with a threshold operation in this difference image. Furthermore, the position of the reflexes on the cornea is computed from the dark pupil images that are acquired with the off-axis illumination. The position of the user's eyes, i.e., the center of the cornea as well as the corresponding line of gaze in 3D space are reconstructed by using the position of the pupil center, the reflexes from the off-axis illumination, and the algorithms from [7].

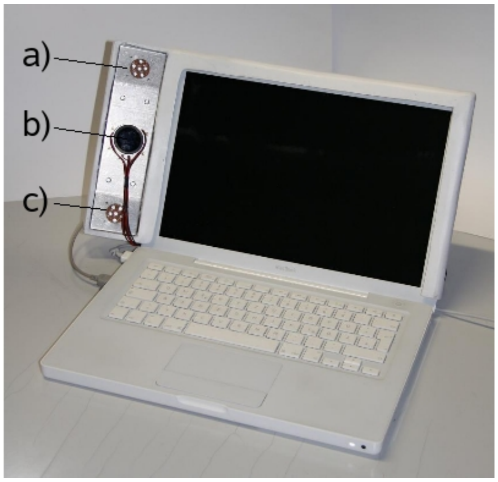


Fig. 4. Hardware setup for the eye tracking system mounted on a laptop computer. On the left of the display, the camera with on-axis illumination (b), and the off-axis illumination (a,c) can be seen. The on-axis illumination, liquid lens, and IR filter are integrated on the camera optics.

A user-dependent calibration of the system is required that mainly consists of the computation of inter-individual differences between the visual and the optical axes of the eyes. We therefore implemented a calibration method that uses the approach proposed in [22]. This is based on the use of a 3D rotation matrix to rotate the optical axis. Thus, the visual axis is obtained from the user calibration. For the presented evaluation, the system was operated with an image rate of 60 Hz which, due to illumination toggling, corresponds to a data rate of 30 Hz. This is twice as much as proposed in [7]. This results in a reduction of motion blur artifacts and allows a more robust extraction of the image features.

B. Hardware

We used an Apple MacBook laptop computer and a custom mount that is attached to the laptop's screen (see fig. 4). The mount contains the camera, the on- and off-axis illuminations, as well as their electronical control circuit. Furthermore, the camera is equipped with a 35 mm optic, an infrared filter, and a liquid lens (Varioptics, see section IV).

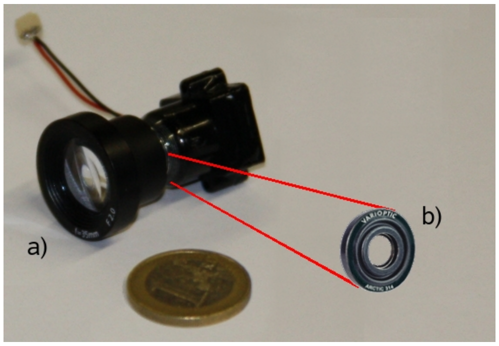


Fig. 5. Varioptic liquid lens integrated into the camera optics. We mounted the liquid (b) lens right behind the 35 mm lens (a).

IV. INTEGRATION OF LIQUID LENSES

Liquid lenses are a very promising technology for realizing optical auto-focus systems. Such lenses don't have the disadvantages of the conventional mechanical auto-focus systems, such as large installation space, latencies induced by moving mechanical parts, and weight. First of all the exploitation of the small latency of the liquid lens is a decisive advantage in multi-user gaze tracking scenarios (see fig. 6). Since the lenses have a very short response time (min to max focus approx. 30 ms, [3]), the optical system can be designed to switch between various users who are located at different distances from the tracking system. Since no actuator components are needed to tune the focus, shock resistance is a further advantage of liquid lenses.

A. Hardware

For this work, we used the Varioptics Arctic 320 liquid lens [14], [18]. The liquid lens measures 10.5 mm in diameter by 2.5 mm in thickness with an inner and outer aperture of 3.6 mm and 4.0 mm, making it easy to integrate into the micro lens mount.

The control of the lens is realized by using the hardware and software interfaces built into the camera. By writing to a DCAM register, the camera is able to provide the pulse-width modulation (PWM) signal that is needed to tune the lens. Thus, 40 focus steps are possible. The control signals are sent via an IEEE1394 bus to the camera. An additional liquid lens driver circuit (Varioptics DrivIC 60 LL3) is used to transform the

PWM signal to voltage levels that are adequate for adjusting the lens.

B. Software

We decided to use a two step approach for the readjustment of the auto-focus functionality. Therefore, we use the 3D position data provided by the gaze tracking algorithm combined with a sharpness metric. The distance between the user and the camera can be obtained from the 3D reconstruction presented in section III-A (i.e., 3D position of the user's eyes). If the user changes the distance to the camera, the focus is tuned immediately to ensure best possible image quality. Furthermore, the sharpness is calculated in similar way as the approach proposed in [8]. Based on these metrics, the focus can be tuned to achieve a best-possible image quality.


(a) (b)

Fig. 6. Example for multiuser gaze tracking. For this images, we used a 16 mm optic with integrated liquid lens. One user was positioned at a distance of approx. 0.7 m and the second user at a distance of approx. 1.7 m. (a) shows the camera image focusing the front face, while (b) shows the camera image focusing the farther face.

For the presented initial proof of concept, a camera calibration was executed for each measuring distance during the experiment. The obtained calibration parameters are stored in a look-up table and applied depending upon the current settings. Similar approaches were presented in [1], [2]. Approaches that use online computation of the camera parameters were also proposed in [17], [20].

Fig. 6 shows a two-user setup in our laboratory, and the effectiveness of the implemented focusing functionality. Fig. 6(a) shows the camera focusing the first user that was next to the camera. Further, fig. 6(b) shows the camera image focused on the second user further away from the camera.

V. EVALUATION

This section presents evaluation results obtained from a novel gaze tracking system that is based on an auto-focus optics with a liquid lens.

A. Experimental Procedure

In order to determine the accuracy of the extended remote gaze tracker with the built in liquid lens, the user was instructed to fixate nine reference points on one plane from four different measuring distances. The user was instructed to keep his head in a steady position. Nine fixation points with distances of 10 cm (horizontal) and 8 cm (vertical) were presented on the laptop monitor. Users were positioned at different distances from the monitor (1.30 m, 1.00 m, 0.80 m, and 0.60 m).

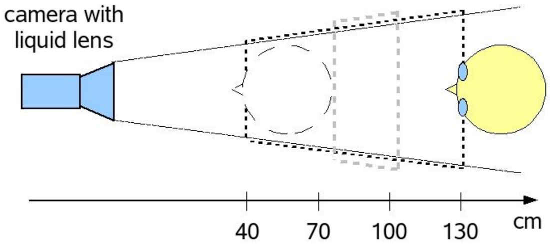
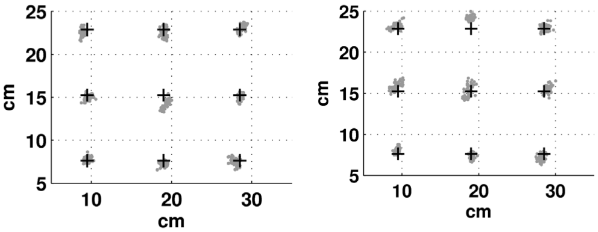


Fig. 7. The user's head in different distances to the camera. The black dotted line denotes the extended action area. The gray dotted line denotes the original action area without auto-focus lens.

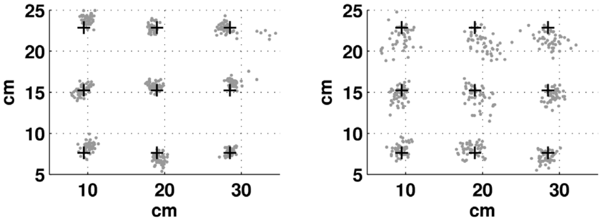
B. Results

The presented approach allowed to extend the system's workspace from 10 cm to 70 cm. The accuracy was defined as the median value of the differences between the gaze point estimates and the 9 target points on the screen. Values of 0.40°, 0.38°, 0.47° and 0.53° (mean 0.45°) were obtained at 0.60 m, 0.80 m, 1.00 m and 1.30 m, respectively (see figs. 9).

The poorer resolution (standard deviation of the gaze point estimates for each fixation) and lower accuracy (see fig. 8) are caused by several factors. First of all, a larger distance to the camera causes a decreasing distance between the two glints in the video image. Additionally, the lighting level decreases if the user is further away from the infrared illumination sources. Also, the lower pixel resolution of the eye in the image leads to a less accurate feature extraction.



(a) Distance: 0.60 m. (b) Distance: 0.80 m.



(c) Distance: 1.00 m. (d) Distance: 1.30 m.

Fig. 8. Recorded raw data (gaze point estimates) of the gaze tracker in the respective distances. The black crosses denote the position of the displayed targets (reference points). The gray dots denote the positions of recorded gaze points on the monitor.

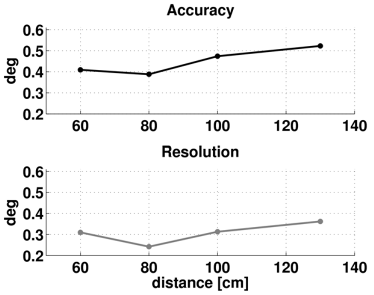


Fig. 9. The accuracy (median value of the differences between gaze point estimates and reference points) and resolution (standard deviation of the gaze point estimates for each fixation) in degrees of the gaze tracker while the user fixates dot pattern at various positions (from 0.6 up to 1.3 m distances from the camera and also the target). The upper black graph shows the accuracy at various tracking distances. The lower gray graph illustrates the resolution at the corresponding distances.

VI. CONCLUSION AND OUTLOOK

The results of our evaluation showed that a gaze tracker with 35 mm optic combined with the auto-focus functionality of a liquid lens provides adequate accuracy and operating range for various human-machine interaction scenarios.

An additional use of zoom optics (as proposed in [4]) might bear the potential to further increase the operating range and also the resolution of the image at large tracking distances. This would reduce noise and increase the accuracy of the gaze tracker. However, high resolution and image sharpness are not solely mandatory for good gaze tracking. Tracking the gaze of the person standing far away is difficult because of a lowered brightness (see fig. 6). Hence, to ensure best quality gaze tracking in both, close and far tracking distances, also the intensity of illumination has to be adapted to the corresponding operating distance. Also, additional off-axis clusters could improve the tracking result at greater distances by increasing the distance between the glints in the video image.

Future work will integrate the presented shock-resistant and fast liquid lens setup into a pivotable camera system (angular velocity above $1000°/sec$ and acceleration of $30.000°/sec^2$), [15]). This will enable single and multi-user gaze tracking for novel human-machine interaction scenarios, e.g. multi joint human-robot interaction use cases. Therefore, robust head tracking approaches are needed to detect faces in defocused images. Furthermore, person verification techniques are required to distinguish between relevant and irrelevant users.

ACKNOWLEDGMENT

This work is supported in part within the DFG excellence initiative research cluster "Cognition for Technical Systems – CoTeSys", see also www.cotesys.org.

REFERENCES

[1] R. Atienza and A. Zelinsky. A Practical Zoom Camera Calibration Technique: An Application on Active Vision for Human-Robot Interaction. In *Proceedings of the Australian Conference on Robotics and Automation, Sydney, Australia.*, 2001.

[2] R. Atienza and A. Zelinsky. Active Gaze Tracking for Human-Robot Interaction. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces.*, 2002.

[3] B. Berge, C. Gabay, G. Dovillaire, and S. Bucourt. Dynamic study of a Varioptic variable focal lens. In *Proceedings of SPIE conference.*, 2002.

[4] M. F. D. Beymer. Eye gaze tracking using an active stereo head. In *IEEE Computer Vision and Pattern Recognition (CVPR), vol. II*, pages 451 – 458, 2003.

[5] K. Han, X. mei Jiang, Q. Feng, and X. chang Zhu. A New Auto-Focusing Method Based on WBDCT for Many Object Situations. In *Proceedings of 9th IEEE International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, ACIS.*, pages 481 – 484, 2008.

[6] Y. He, J. Cui, T. Tan, and Y. Wang. Key Techniques and Methods for Imaging Iris in Focus. In *Proceedings of 18th International Conference on Pattern Recognition. ICPR 2006.*, pages 557 – 561, 2006.

[7] C. Hennessey, B. Noureddin, and P. Lawrence. A single camera eye-gaze tracking system with free head motion. In *Eye Tracking Research & Applications 2006*, pages 87 – 94, 2006.

[8] S. H. Jin, J. U. Cho, and J. W. Jeon. FPGA based Passive Auto Focus System using Adaptive Thresholding. In *Proceedings of SICE-ICASE International Joint Conference.*, pages 2290 – 2295, 2006.

[9] S. Kohlbecher, S. Bardins, K. Bartl, E. Schneider, T. Poitschke, and M. Ablassmeier. Calibration-free eye tracking by reconstruction of the pupil ellipse in 3d space. In *Proc. of ETRA 2008, Savannah, USA, March 25-28, 2008. CD ROM.*, 2008.

[10] LC-Technologies. Books written with the Eyegaze Communication System. http://www.eyegaze.com/content/books-written-eyegaze-communication-system, 2009.

[11] J.-S. Lee, S.-J. Ko, Y. Kim, and A. Morales. A video camera system with adaptive zoom tracking. In *IEEE International Conference on Consumer Electronics, ICCE, Digest of Technical Papers.*, pages 56 – 57, 2002.

[12] S. J. Lee, S. H. Jin, and J. W. Jeon. FPGA based Auto Focus System using Touch Screen. In *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems.*, pages 350 – 353, 2008.

[13] J. Merchant, R. Morrissette, and J. L. Porterfield. Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Trans Biomed Eng*, 21:309–317, 1974.

[14] L. Saurei, G. Mathieu, and B. Berge. Design of an autofocus lens for VGA 1/4" CCD and CMOS sensors. In *Proceedings of SPIE conference.*, 2003.

[15] E. Schneider, T. Villgrattner, J. Vockeroth, K. Bartl, S. Kohlbecher, S. Bardins, H. Ulbrich, and T. Brandt. EyeSeeCam: An eye movement-driven head camera for the examination of natural visual exploration. *Ann N Y Acad Sci*, 1034:in press, 2008.

[16] S.-M. Sohn, S.-H. Yang, S.-W. Kim, K.-H. Baek, and W.-H. Paik. SoC Design of An Auto-Focus Driving Image Signal Processor for Mobile Camera Applications. In *IEEE Transactions on Consumer Electronics, Volume 52, Issue 1*, pages 10 – 16, 2006.

[17] J. Tzeng and T. Q. Nguyen. Image Enhancement for Fluid Lens Camera Based on Color Correlation. In *IEEE Transactions on Image Processing, Vol. 18, No. 4* , 2009.

[18] Varioptics. Varioptics – The Liquid Lens Company. http://www.varioptic.com, 2009.

[19] N. J. Wade and B. W. Tatler. The moving tablet of the eyes. Oxford University Press, New York, Oxford, 2005.

[20] R. Willson. Modeling and Calibration of Automated Zoom Lenses. In *Proceedings of SPIE 2350: Videometrics III, Boston MA,*, 1994.

[21] A. Zelinsky, Y. Matsumoto, J. Heinzmann, and R. Newman. Towards human friendly robots: Vision-based interfaces and safe mechanisms. *Lecture Notes in Control and Information Sciences*, 250:487–498, 2000.

[22] Z. Zhu and Q. Ji. Novel Eye Gaze Tracking Techniques Under Natural Head Movement. *IEEE Transactions on Biomedical Engineering, VOL. 54, NO. 12, DECEMBER 2007*, 54(12):2246 – 2260, December 2007.