A Novel Camera-Free Eye Tracking Sensor for Augmented Reality based on Laser Scanning

Johannes Meyer, Thomas Schlebusch, Wolfgang Fuhl and Enkelejda Kasneci

***Abstract*—Next generation AR glasses require a highly integrated, high-resolution near-eye display technique such as focus-free retinal projection to enhance usability. Combined with low-power eye-tracking, such glasses enable better user experience and performance. We propose a novel and robust low-power eye-tracking sensor for integration into retinal projection systems. In our approach, a MEMS micro mirror scans an IR laser beam over the eye region and the scattered light is received by a photodiode. The advantages of our approach over typical VOG systems are its high integration capability and low-power consumption, which qualify our approach for next generation AR glasses. Here, we present a mathematical framework to estimate the achievable gaze angle resolution of our approach. We further show the viability of the proposed eye-tracking sensor based on a laboratory setup and discuss power consumption and gaze angle resolution compared to typical eye-tracking techniques.**

***Index Terms*—AR glasses, Eye tracking Sensor, MEMS Scanned Laser Eye tracking, retinal projection**

I. Introduction

THE number of commercially available augmented reality (AR) glasses has largely increased in the last years. A new shift in the commercial domain is the integration of eye-tracking sensors into AR glasses to increase user experience by adding new functionalities and introducing new ways of interaction with systems around us [1]. Examples are driver assistance systems [2], human computer interaction (HCI) [3] or smart home control [4].

Additionally, the resolution of AR applications can be perceived at reduced computational effort by utilizing foveated imaging, where high resolution content is only projected sharply on the projection area corresponding to the fovea and lower-resolution content is used in peripheral regions. It is obvious that this technique is dependent on eye-tracking and can significantly reduce the required processing power for image rendering compared to a full-resolution rendering in the full field of view (FOV) [5]. A recent implementation of foveated imaging for VR applications was presented by Tobii [6].

Current eye-tracking sensors for low cost, commercially available systems, e.g. Pupil Labs [7], use video-based oculography (VOG). These eye-tracking sensors rely on a set of infrared (IR) emitters to illuminate the eye and an IR camera to capture eye images. The gaze direction is then determined by image processing algorithms. While camera-based VOG systems are well-established and perform at high accuracy, there is only little potential for significant reduction of power consumption. Further, feasible orientation of the camera modules limits integration into a smart glasses frame and often interferes with the user's view. Other interference-free VOG sensor integration approaches result in large angles between camera sensor and eye [8], which leads to the use of computationally intensive pupil tracking algorithms.

These drawbacks limit the use of VOG based eye-tracking sensors for battery powered AR glasses in consumer applications. The requirements of AR glasses in this segment are low-power consumption and a high degree of integration into the frame temple [9]. Additionally, high stability and reliability are required for everyday use in the wild. A main problem that arises in the outside world is however reduced sensor performance due to artificial or natural light sources [10].

To meet these requirements, we propose a novel camera-free eye-tracking sensor based on retinal projection AR glasses. Our main contribution is a new miniaturized low-power eye-tracking sensor approach using state of the art VOG algorithms for integration into retinal projection AR glasses.

Retinal projection is a very promising near-eye display technique for AR glasses. It is based on direct illumination of the eye's retina using an ultra low-power eye-safe laser projector. Unique aspects of this display technology are a high degree of integratability into a frame temple, providing outstanding design opportunities, as well as near focus-free image projection, providing a sharp projection independent on the user's accommodation state [11]. For most retinal projection systems, eye-tracking is crucial as the scanned beam has to follow the pupil's position to enter the eye [12].

We propose an extension of these retinal projection systems to fully integrate eye-tracking. More specifically, we place an additional IR laser diode in the laser module and a photodiode close to the joint in the frame temple.

Exploiting the scanning device and a holographic optical element (HOE) already available for the projection system,

Johannes Meyer and Thomas Schlebusch are with the Corporate Sector Research and Advance Engineering, Robert Bosch GmbH, Renningen, Germany (e-mail: {Thomas.Schlebusch, Johannes.Meyer2 }@de.bosch.com).

Enkelejda Kasneci and Wolfgang Fuhl are with the Department of Computer Science, University of Tübingen, Tübingen, Germany (e-mail: {enkelejda.kasneci, wolfgang.fuhl}@uni-tuebingen.de).

this enables us to significantly simplify the PSOG setup by using only one IR diode laser and one photodiode for the whole tracking region. At the same time, the spatial resolution is significantly increased. The system output are greyscale images suitable for processing using state of the art VOG algorithms.

This resulting sensor setup is comparable to to photo sensor oculography (PSOG) approaches. PSOG exploits the varying IR reflectivity of different regions of the eye such as sclera, iris, pupil etc. [13].

The remaining of the paper is organized as follows. The next section discusses the state-of-the-art with regard to eye-tracking based on laser scanning. Section 3 presents our eye-tracking technique and its integration into our retinal projection AR glasses concept based on the BML500P [14] as well as a mathematical framework to analyse the theoretical spatial and gaze angle resolution achievable by our technique. Section 4 discusses the data obtained from a laboratory setup and shows the expected viability and resolution of the proposed eye-tracking sensor. For the laboratory setup we used the BML100PI [15] which is a predecessor of the BML500P. Based on this setup the resolution for a head worn sensor is estimated utilizing the mathematical framework derived in Section 3. Afterwards a comparison between our sensor approach and VOG eye-tracking sensors is performed. Section 5 concludes this work and gives an outlook to our future research activities.

II. Related Work

First scanned lasers for retina imaging appeared in the clinical section. Webb et al. introduced a scanning laser opthalmoscope (SLO), which works according to a process of scanning a laser beam over the retina surface to capture an image of the retina [16]. The captured image was used for medical diagnosis of eye diseases. Eye motions during the scan process lead to distortions of the retina images. To resolve this issue, the SLO method was later extended by an eye-tracking system, as for example in [17] who presented a binocular tracking scanning laser ophthalmoscope (TSLO). The beam of a super luminescent diode was collimated and deflected by a $15\,\mathrm{kHz}$ resonant scanner horizontally and by a $30\,\mathrm{Hz}$ mirror galvanometer scanner vertically onto the eye. A photomultiplier tube and a beam splitter were used to detect the reflections of the eye. The photomultiplier tube was sampled based on the position of the mirrors to extract an image. Afterwards, the eye position was calculated offline on a host device. For this techniques, [17] reported a possible tracking speed of $366\,\mathrm{Hz}$ at a resolution of $0.003°$. Such medical scanned laser eye-tracking approaches are characterized by high image resolution, however only during fixation phases of the eye. Furthermore, the optical setups are expensive, large and only suitable for laboratory use. To reduce the size of the optical setups [18] used micro-electro-mechanical systems (MEMS) micro mirrors as scanning devices.

A different scanned laser approach was described by [19]. They integrated an IR laser into an RGB laser module to track the eye position. The IR laser beam was deflected onto a polygon mirror for horizontal deflection and a galvanometer scanner for vertical deflection. The scanned laser beam was then guided by a prism onto the surface of the eye. To capture the reflections of the eye, a semi-transparent mirror was placed in front of the user's eye region. The mirror directs the reflected light from the surface of the eye onto a collective lens on top of a photodiode. Based on the position of both scanning mirrors, the photodiode was sampled to capture an image of the eye. This setup was however rather large and was based on multiple mirrors which interfere with the user's view.

To reduce the power consumption and size of scanned laser eye-tracking sensors, Sarkar et al. [20] introduced MEMS micro mirrors to scan a laser beam vertically and horizontally across the eye-tracking region. Based on this improvement, they integrated the technology into AR glasses [21]. In their setup, a two dimensional resonant MEMS micro mirror and an IR laser were placed on the inside of the frame temple. The MEMS micro mirror deflected the light towards the surface of the eye. To receive reflections from the eye's surface, a photodiode was integrated into the nose bridge of the glasses. To track the horizontal eye position, a linear trajectory was scanned by the MEMS micro mirror over the surface of the eye. The photodiode received the reflected IR light from the surface of the cornea. The maximum light intensity detected by the photodiode was then used to estimate the horizontal eye position. Afterwards, the vertical position of the eye was estimated by a hill climbing algorithm based on the amplitude shifts of the photodiode output in horizontal scanning direction. This approach achieves an angular resolution of approximately $1°$ at a temporal resolution of $3300\,\mathrm{Hz}$ and a power consumption of less than $15\,\mathrm{mW}$ [21].

The main draw back of this approach is the vulnerability to shifts of the glasses. Even small movements of the glasses on the user's face lead to significant drifts and therefore inaccuracy of the eye position estimation [21].

Our approach combines the advantages of robust camera-based eye-tracking systems and the low-power consumption and small size of MEMS scanned laser eye-tracking sensors. Additionally, we integrate an optical element into the spectacle lens to obtain an eye-tracking sensor that is invisible to the user and therefore does not interferes with the users view while providing an advantageous camera perspective directly centered in front of the eye. In addition, the use of retinal projection systems and their integration into lightweight AR glasses was shown in the Intel Vaunt technology [22] later acquired by North Focal and most recently in the Bosch smart glasses system [14]. For this reason, our approach is also very promising from an industrial point of view.

III. Scanned laser eye-tracking

The proposed eye-tracking sensor is based on a retina projector for AR glasses as described in [23]. To integrate the eye tracking sensor into the retina projector an IR laser, an optical receiver and an additional laser beam deflection function of the optical element specific to the IR wavelength of the eye-tracking laser are added as shown in Figure 1.

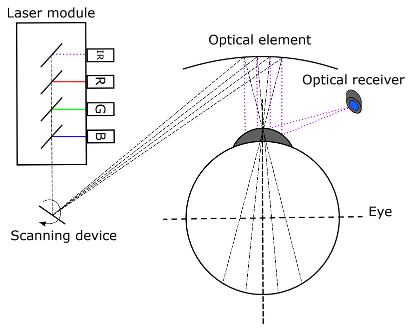


Fig. 1. Principle structure and optical path of a retinal projector with integrated eye-tracking capabilities. An IR laser is integrated into the RGB laser module and an optical receiver is placed near the eye.

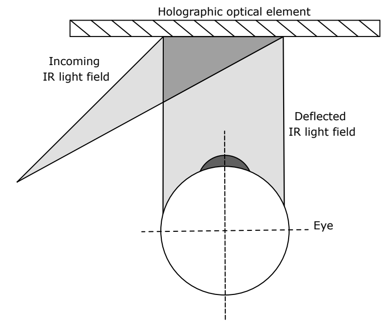


Fig. 2. Description of the function of an holographic optical element (HOE). The Bragg structure inside the HOE diffracts the incoming IR light along parallel lines towards the eye to form a rectangular region in the surface of the eye.

The laser projection unit consists of a laser module and a scanning device. The laser projection module projects IR laser light (purple doted line) onto the optical element. The optical element redirects the laser beam with a defined optical function (e.g. parabolic mirror) onto the surface of the eye. Based on the varying IR reflectivity of the different eye regions, the laser beam is scattered with a different intensity from the surface of the eye. The reflected light is measured by an optical receiver.

The components of the sensor are described in more detail in the following.

A. Laser projection module

The main component of the projection module is the control application specific integrated circuit (ASIC). The ASIC is fed with a digital image via a low voltage differential signaling (LVDS) interface.

Based on the incoming image stream, the lasers are modulated by the RGB values of each image pixel. The IR laser remains in an active state as long as eye-tracking remains active. The modulated and collimated RGB laser beam and the IR laser beam are directed onto the MEMS micro mirrors.

The micro mirror module consists of a fast axis MEMS micro mirror for the horizontal scan direction and a slow axis MEMS micro mirror for the vertical scan direction. The horizontal scan mirror is driven at resonance frequency and oscillates sinusoidal. The vertical mirror is actuated non resonantly by an electrodynamic drive (Lorentz force) with electric coils on the MEMS and an external permanent magnet. The start of a new line scan is reported by sync signals to the control ASIC. For our experiments we used the BML100PI, an of-the-shelf projection module containing the MEMS micro mirrors and the described laser module [15].

B. Holographic optical element

An HOE is integrated into the AR spectacle lens to deflect the incoming scanned IR laser light towards the surface of the eye as shown in Figure 2.

The HOE is realized by recording a mirror function into a photo polymer material. For a more detailed description of the recording process and integration of optical functions into HOEs using Bragg structures, we refer to [24].

The main advantage of HOEs over other optical elements like semi-transparent mirrors is the high selectivity and optical transparency. Ideally, the optical function is only active for the recording wavelength [12].

The HOE function controls the size and contour of the region which is scanned by the IR laser on the eye. For the scanned laser eye-tracking sensor, this scanned surface is referred to as eye-tracking region.

C. Receiving photodiode circuit

To capture a reflection map of the eye-tracking region, a similar receiving photodiode circuit as known from previous approaches like PSOG or the scanned laser approach by [21] is used. Figure 3 shows the system block diagram of the receiving photodiode circuit.

Compared, e.g. to [21], we are not restricted in positioning the photodiode into the nose pad, which further simplifies integration, since no wires have to be routed through the joint of the eyeglasses.

The main component is a photodiode which is sensitive to IR light. To increase the resolution and convert the photodiode current $I_p(t)$ into a measurable voltage $U_p(t)$, a transimpedance amplifier (TIA) circuit is used. The voltage $U_p(t)$ is AC coupled via a high pass filter to remove low frequency disturbances like ambient light. The amplified and filtered signal is then converted into the digital domain by an analogue to digital converter (ADC). A Field Programmable Gate Array (FPGA) converts the digitized samples into an image utilizing the sync signals of the laser projection module. To capture an image of the eye, the photodiode current is sampled at fixed time intervals by the ADC. The interpretation of the photodiode signal as an image allows us to use established (VOG) algorithms to extract the gaze information from the signals.

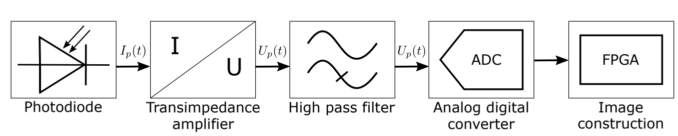


Fig. 3. System block diagram of the receiving photodiode circuit.

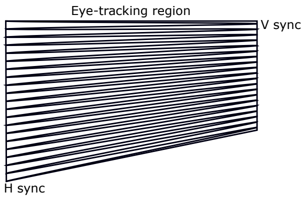


Fig. 4. Simplified sinusoidal scan path of the laser beam over the eye-tracking region. The sync signals are used to reconstruct an image based on the samples captured by the ADC.

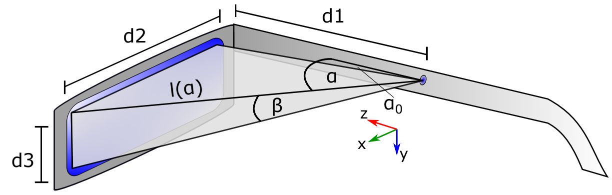


Fig. 5. Integration of the proposed scanned laser eye-tracking sensor into AR glasses. The HOE is integrated into the spectacle lens and the laser projection module is integrated into the frame temple.

The start of a new frame is indicated by the vertical synchronization signal (V sync). With the rise of this signal, the ADC starts to sample the photodiode signal. The start of a new row of the frame is indicated by the horizontal synchronisation signal (H sync). With these two signals, a reflectivity image of the surface of the eye is reconstructed by the FPGA.

Figure 4 shows a simplified scan path (much less lines) of the laser beam in the eye-tracking region. The characteristic bounding box results from the geometric setup as depicted in Figure 5.

We perform a full two-dimensional linescan over the eye-tracking region instead of a single line or Lissajous patterns, as in [21].

The resolution of the image, and therefore the pixel size, depends on the bandwidth of the AFE, the sample rate of the ADC, the time-dependent sinusoidal angular scan speed of the laser projection system and the geometric setup.

D. Spatial resolution

State of the art camera-based eye-tracking approaches utilize pupil edge detection methods to extract the pupil location and calculate the gaze direction [10]. For this purpose, a sufficient spatial resolution in the captured eye region is required. The spatial resolution in VOG eye-tracking approaches is limited by the camera resolution and image distortions due to lens misalignments.

In the scanned laser eye-tracking sensor approach, pixel size and image distortion are related to the scan frequency of the MEMS micro mirrors, the optical path length and the performance of the photodiode circuit [25].

Horizontal distortions occur due to a sinusoidal-like velocity pattern of the laser spot along the scan path shown in Figure 4. Combined with equitemporal sampling of the photodiode circuit, this leads to barrel distortions [25].

The spatial resolution is characterized by analyzing the image data of a high contrast linear or square edged test pattern with known properties in the eye-tracking region captured by the photodiode circuit. The known properties of the test pattern and the resulting number of pixels in the image are used to calculate pixel sizes at various positions in the image. With this, image distortions due to varying optical path length and the sinusoidal angular velocity pattern are quantified.

E. Gaze angle resolution

A gaze angle is described by the pupil center position and eye rotation angles. With rotation of the eye, the pupil center and therefore the pupil edges are rotated as well. To assess the gaze angle resolution of the proposed scanned laser eye-tracking sensor, a mapping from spatial resolution to angular resolution is required. It is dependent on the geometrical system design, the laser projection unit, the sample rate of the photodiode circuit, the HOE function, and the rotation angle of the eye. For the calculations, a parallel deflection function of the HOE as described in Figure 2 is assumed.

Figure 5 shows the geometrical dimensions of the AR glasses setup. The laser projection module is integrated into the right frame temple with a distance d_1 to the hinge between frame temple and spectacle lens. The angles α and β describe the horizontal and vertical scan angles of the laser projection module. The distance d_2 denotes the width of the HOE inside the spectacle lens. The distance d_3 describes the height of the HOE.

As mentioned earlier, pixel sizes vary throughout the eye tracking region. They are dependent on the oscillation frequencies f_h and f_v as well as the maximum scan angles α_{max} and β_{max} of the horizontal and vertical micro mirrors, respectively. In addition, the sample speed f_s of the ADC and the distance d_1 are key parameters.

The x-coordinate of a point on the HOE can be expressed by d_1 and α:

$$x(\alpha) = d_1 \cdot \tan(\alpha). \tag{1}$$

The angle α changes sinusoidal in time. For the geometry shown in Figure 5, this is

$$\alpha(t) = \frac{\alpha_{max}}{2} \cdot \sin(\omega_h \cdot t - \frac{\pi}{2})) + \frac{\alpha_{max}}{2} + \alpha_0. \tag{2}$$

α_0 denotes the horizontal offset angle between frame temple and the right edge of the HOE integrated in the spectacle glasses. The pixel size s_x along the horizontal direction can be expressed as the difference between a first horizontal position at angle α and a second position at an angle incremented by $\Delta\alpha$

$$s_x = x(\alpha + \Delta\alpha) - x(\alpha). \tag{3}$$

$\Delta\alpha$ denotes the angle increment between two consecutive samples of the ADC and is dependent on the sample speed f_s of the ADC. Therefore,

$$\Delta\alpha = \alpha(t + \frac{1}{f_s}) - \alpha(t). \tag{4}$$

This can be used to eliminate t in Equation (2):

$$\Delta\alpha = \frac{\alpha_{max}}{2} + \alpha - \frac{\alpha_{max}\cos\left(\frac{2\pi f_h}{f_s} + \operatorname{acos}\left(\frac{2\alpha_0 - 2\alpha + \alpha_{max}}{\alpha_{max}} + 1\right)\right)}{2} \quad (5)$$

Now, the horizontal pixel size s_x from Equation (3) can be expressed by the position on the HOE given by α and known system constants only:

$$s_x(\alpha) = -d_1 \cdot \left(\tan(\alpha) + tan\left(\alpha_0 + \frac{\alpha_{max}}{2} - \frac{\alpha_{max}\left(\cos\left(\frac{\omega_h}{f_s} + \operatorname{acos}\left(\frac{2\alpha_0 - 2\alpha}{\alpha_{max}} - 1\right)\right) - 1\right)}{2}\right)\right) \quad (6)$$

The vertical axis is operated in a sawtooth-like fashion with a linearly increasing $\beta(t)$ until β_{max} and a linear flyback to $\beta = 0$. As apparent from Figure 5, the optical path length increases with increasing deflection of the micro mirrors. With this, the vertical position is dependent on the length of the projection of the optical path between micro mirror and HOE onto the xz-plane. This can be expressed as

$$l(\alpha) = \frac{d_1}{\cos(\alpha)}. \quad (7)$$

The resulting y-coordinate is then calculated using β:

$$y(\alpha, \beta) = l(\alpha) \cdot \tan(\beta). \quad (8)$$

Analog to the horizontal case, the pixel size s_y along the vertical direction is expressed as the difference between a first vertical position defined by α and β and a second position at an angle incremented by $\Delta\beta$. Here, $\Delta\beta$ is the linear angle increment between two consecutive horizontal scan lines and can be expressed as

$$\Delta\beta = \frac{\beta_{max}}{N_h} \quad (9)$$

with the number of horizontal lines N_h scanned over the eye-tracking region. N_h is calculated by

$$N_h = \frac{f_h}{f_v \cdot k_v}. \quad (10)$$

The constant $k_v < 1$ describes the fraction of forward scan to flyback time which is required to drive the horizontal mirror back to the starting position.

With this, the vertical pixel size s_y is

$$s_y(\alpha, \beta) = \frac{y(\alpha, \beta + \Delta\beta) - y(\alpha, \beta)}{\cos(\beta)}. \quad (11)$$

The division by $\cos(\beta)$ results in a projection of the pixel size vector perpendicular to the optical path onto the plane of the HOE.

Now, pixel dimensions can be expressed dependent on mirror positions only (given by α and β). To calculate gaze angle resolution, this spatial resolution has to be mapped to eye rotational resolution.

For calculation, we assume a simplified eye ball model as shown in Figure 6 with radius r_{eye}. It is rotated by θ in the horizontal plane and ϕ in vertical direction around its center. This rotation causes the pupil edge to propagate through different pixels along the reconstructed image. Neglecting methodologies like subpixel interpolation, eye rotation is only noticeable by the eye-tracking system if the pupil edge propagates from one pixel to an adjacent one. Thus, we define an infinitesimal angular increment δ of eye rotation that keeps the pupil edge in the current pixel. In this view, δ is the position-dependent angular resolution of our eye-tracking system.

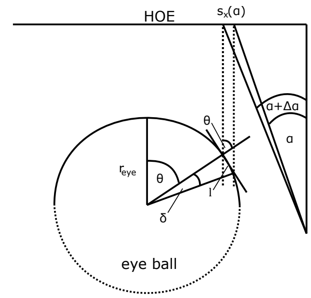


Fig. 6. Mapping of horizontal eye rotation to spatial resolution on the HOE. A parallel deflection function of the HOE is assumed.

Based on Figure 6, δ_h for horizontal gaze angle resolution is calculated by

$$\delta_h(\theta) = \operatorname{atan}\left(\frac{s_x(\alpha(\theta))}{r_{eye} \cdot \cos(\theta)}\right) \quad (12)$$

with $\alpha(\theta)$ as mapping function to map an eye rotation angle θ to a horizontal angle α. With the assumption that the eye-tracking region, and therefore the HOE, is centered over the eye, the mapping function $\alpha(\theta)$ is

$$\alpha(\theta) = \operatorname{atan}\left(\frac{\frac{d_2}{2} - r_{eye} \cdot \sin(\theta)}{d1}\right). \quad (13)$$

The vertical gaze angle resolution δ_v in analogy to Equation (12) is

$$\delta_v(\phi) = \operatorname{atan}\left(\frac{s_y(\alpha(\theta), \beta(\phi))}{r_{eye} \cdot \cos(\phi)}\right) \quad (14)$$

with

$$\beta(\phi) = \operatorname{atan}\left(\frac{\frac{d_3}{2} - r_{eye} \cdot \sin(\phi)}{d1}\right). \quad (15)$$

The overall gaze angle resolution is calculated by the euclidean distance of the horizontal and vertical gaze angle resolution

$$\delta(\theta, \phi) = \sqrt{\delta_v(\theta)^2 + \delta_h(\phi)^2}. \quad (16)$$

IV. Evaluation

To determine the gaze angle resolutions of the proposed head worn eye-tracking sensor, a laboratory setup is used. It is shown in Figure 7. The subject ① is placed in a distance of $l_2 = 90\,mm$ in front of a semitransparent mirror ④. The semitransparent mirror mimics the HOE with a slightly diverging beam deflection function as compared to Figure 2. The mirror is used to deflect the scanned IR laser beam of

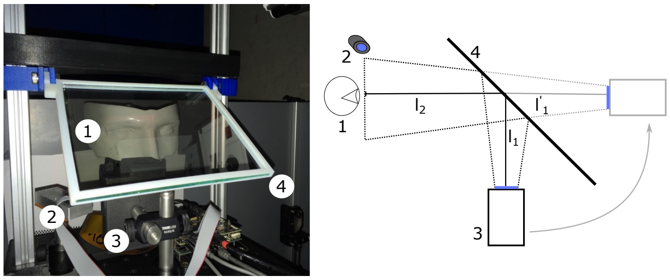


Fig. 7. Laboratory setup to evaluate the accuracy and precision of the proposed eye tracking sensor.

the laser projection module ③. The laser module is placed at a distance of $l_1 = 30mm$ towards the semi-transparent mirror. The photodiode circuit ② is directly oriented towards the subject's eye to receive as much backscattered IR light as possible and thus improve sensitivity. The mirror allows the laser projection module to be virtually rotated so that it points directly at the eye tracking region without disturbing the user's view. In addition, artefacts of the captured images caused by partially retracted eyelids and eyelashes are reduced, which improves the robustness of the sensors.

Table I shows the geometrical and electrical properties of the the laboratory setup. The overall distance d_1 is the sum of l_1 and l_2.

TABLE I

ELECTRICAL AND GEOMETRICAL PROPERTIES OF THE LAB SETUP.

α_0	$\pm\alpha_{max}$	$\pm\beta_{max}$	f_s	f_h	f_v	k_v	d_1	d_2	d_3
0°	15°	9°	22 MHz	21kHz	60Hz	0.83	120mm	69mm	39mm

The laboratory setup is a class 1 laser system according to IEC 60825-1 [26] and therefore does not pose any medical hazard to the eye. To ensure eye safety and limit the IR exposure to the eye, we add an IR filter with 60 % transmission, resulting in an IR laser power of less than 150 μW towards the eye. Laser class 1 would allow 670 μW for a steady beam, even significantly more in scanned operation. Laser safety is therefore ensured for all single error cases as required by IEC 60825-1 as well.

The power consumption of the proposed eye-tracking sensor is estimated roughly at 11 mW using off-the-shelf components. The main components that affect power consumption are the TIA and the ADC. With a higher degree of integration, e.g. by a custom ASIC, further power reductions are expected.

Sarkar et al. [21] report that state of the art VOG eye-tracking sensors consume more than 150 mW of power, which is significantly higher than our sensor approach. Compared to the scanned laser approach by [21], a similar power consumption is achieved.

A. Spatial resolution

To prove the mathematical framework, in Section IV-C the spatial resolution as described in Section III-D is calculated by placing a chess pattern of defined size in the eye-tracking region and compare the measured resolution with the theoretical spatial resolution. The experimental spatial resolution is

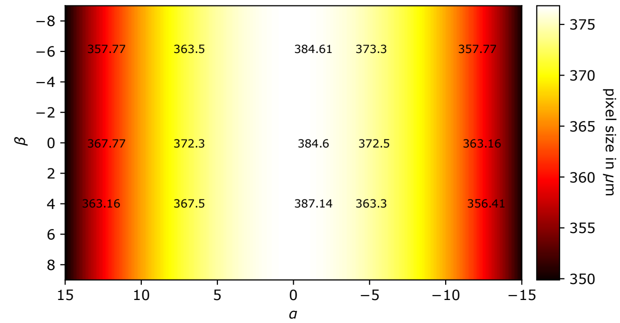


Fig. 8. Calculated pixel size in μm over the whole scan region described by α and β for the horizontal scan direction based on Section III-E with the parameters of Table I. In addition, the measured horizontal pixel sizes from the chess pattern are annotated.

determined based on the known properties of the chess pattern image. The number of pixels per chess field is counted and divided by the known length of a chess field to determine the pixel size.

Figure 8 and Figure 9 show the resulting theoretical and experimental spatial resolutions as color surface and annotated numbers, respectively.

The spatial resolution in horizontal direction shown in Figure 8 is dominated by the non-linear horizontal scan speed of the MEMS micro mirror described by Equation (2). The sinusoidal change of velocity results in smaller pixels in the left and right edge regions of the eye tracking region. In the centre of the eye tracking region, the peak velocity is reached, resulting in increasing pixel size. In consequence, the sensor resolution is higher towards the left and right sides of the eye tracking region.

For vertical direction the pixel sizes increases for increasing angles α and β, as shown in Figure 9. This effect is mainly caused by increasing optical path length of the laser beam for increasing angles α and β. This effect is superimposed by the effect of the sinusoidal change of velocity, which leads to additional distortions in the horizontal direction.

B. Experimental gaze angle resolution

To estimate the achievable gaze angle resolution, an experiment with the laboratory setup is performed. A subject sits approximately 0.5m away from a chart with visual markers M

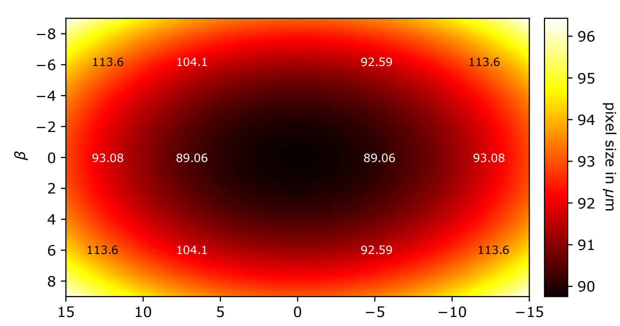


Fig. 9. Calculated pixel size for the vertical direction based on Equation (11) with the parameters of Table I. In addition, the measured vertical pixel sizes from the chess pattern are annotated.

and fixates the markers. For each marker a set of N images is taken. The markers are placed on the chart to cause eye rotation angles θ and ϕ in the range of $\pm 20°$ and of $\pm 10°$ respectively.

Figure 10 shows eye recordings of a subject fixating different markers. The images are captured with the laboratory setup.

To extract the pupil positions we use the state of the art VOG circular binary features (CBF) pupil detection algorithm by [27]. The estimated pupil center positions are marked with a white cross in Figure 10. The result is very promising and proves the feasibility to apply state of the art VOG algorithms to our low-power eye tracking data. By extracting additional landmarks from the images, a slippage robust pupil extraction for the scanned laser eye tracking sensor with minimal algorithm design effort is possible. In addition, the sensor benefits directly from advances in VOG algorithms.

To calculate the resolution and precision of the sensor, we perform the standard 9-point chart marker based calibration method similar to [7]. For each calibration marker M_c, N images are captured and the corresponding pupil coordinates P_c are calculated using the CBF algorithm. With the known calibration coordinates and the related pupil coordinates a second order polynomial function is fitted using a least mean square optimizer.

In addition to the calibration markers M_c, test markers M_t are placed on the calibration chart. These markers are used to estimate the spatial accuracy and precision of the proposed eye tracking sensor. Figure 11 shows the results of the experiments in gaze angle coordinate space.

To estimate the resolution based on the experiment, the Root Mean Square Error (RMSE) of the average angular distance between the position of the test markers M_t and a set of corresponding locations of fixations P_t is calculated for every test marker. In addition the precision is estimated by the RMSE of successive samples of fixations P_t for a given test marker M_t. The resulting mean resolution of the test markers is $1.31°$ and the precision is $0.01°$ for this setup. A possible source of error, which reduces the achieved gaze accuracy, is that the subject's head is not fixed during the experiment. In addition, a large part of the captured image covers the face around the eye. The pupil information is therefore only contained in a small subset of pixels in the center of the image, thus reducing relevant image resolution.

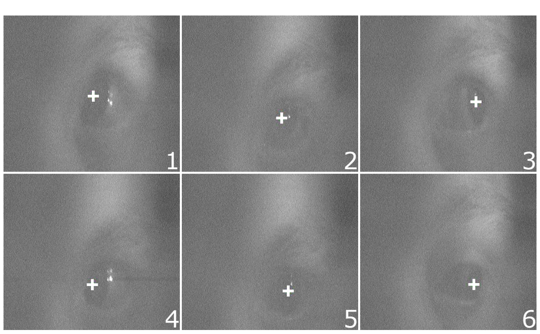


Fig. 10. Images of a subject fixating different markers on a chart.The pupil position is detected by a state of the art VOG pupil detection algorithm [27]. The numbers indicate which marker of the chart is fixated by the subject.

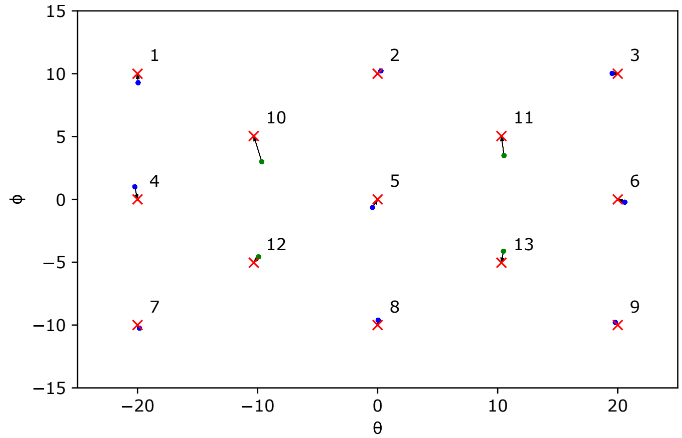


Fig. 11. Result of the accuracy and precision experiment with calibration- (red crosses) $\boldsymbol{M_c}$ (1-9) and test markers $\boldsymbol{M_t}$ (10-13) and the corresponding pupil coordinates $\boldsymbol{P_c}$ (blue dots) and $\boldsymbol{P_t}$ (green dots). The arrow indicates the error between corresponding marker- and pupil coordinates.

C. Theoretical Gaze angle resolution

In addition to the experimental estimation of the gaze resolution of the proposed eye tracking sensor, the the theoretical gaze angle resolution is estimated by Equation (16). The distance r_{eye} in Equation (16) is derived from the Emsley's reduced eye model. Based on this model, the distance r_{eye} between the iris and the center point of the eye ball is 9.77 mm. The result is shown in Figure 12.

Recalling Figure 6, δ is much smaller for given $s_x(\alpha)$ close to $\theta \approx 0, \phi \approx 0$. To some extent, this compensates the lower spatial resolution in the center of the eye tracking region as shown in Figure 8. The theoretical mean single-pixel gaze angle accuracy of the laboratory setups is around $2.3\,°$, which is significant lower as the experimental estimated gaze angle resolution. Thus, our simplified mathematical approach neglecting the effect of pupil tracking algorithms can be understood as an upper boundary estimation for gaze angle resolution.

Based on this assumption, the theoretical gaze angle resolution of the head worn eye tracking sensor is calculated using the electrical and geometric parameters shown in Table II. The main differences between the laboratory setup and the

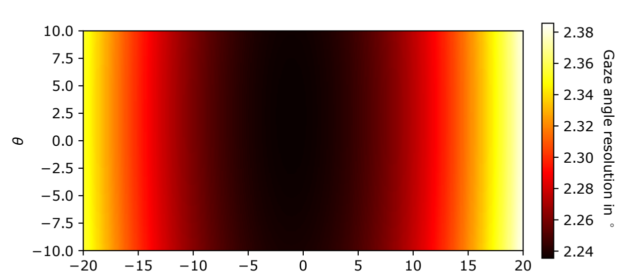


Fig. 12. Calculated gaze angle resolution in degree of the proposed eye tracking sensor based on the laboratory setup utilizing Equation (16).

proposed head worn demonstrator are the reduced distance d_1 towards the HOE, the offset angle α_0, the maximum angles for α_{max} and β_{max} and the parallel diverging beams. Due to the glasses geometry α is in a range between 0 and α_{max} an β is in a range between 0 and β_{max}. The theoretical

TABLE II
ELECTRICAL AND GEOMETRICAL PROPERTIES OF THE AR GLASSES DEMONSTRATOR DERIVED FROM [14].

α_0	α_{max}	β_{max}	f_s	f_h	f_v	k_v	d_1	d_2	d_3
15°	30°	18°	22 MHz	21kHz	60Hz	0.83	22.5mm	22.5mm	7.3mm

gaze angle resolution of the proposed head worn eye tracking sensor is shown in Figure 13. The offset angle α_0 adds an increasing optical path length with increasing angles α which leads to a decrease in resolution for extreme eye rotational angles especially towards the lower left edge, as apparent from Figure 13. The theoretical mean single-pixel gaze angle accuracy for this setup is around $0.28\,^\circ$ and is therefore a upper boundary estimation of the gaze angle resolution. A further increase in gaze angle resolution by the effect of VOG algorithms is expected.

Compared with the scanned laser approach of [21], the calculated gaze angle resolution of the proposed laser based eye-tracking sensor for the glasses geometry is higher especially for relevant eye rotation angles around the center. Compared to state of the art VOG eye-tracking sensors like [7], our approach is capable to achieve higher gaze angle resolution for the glasses geometry. Based on these results, a less complex pupil tracking algorithm can be used to reduce the computational effort and therefore power consumption, while maintaining comparable gaze angle resolutions to VOG sensors.

D. Limitations

The temporal resolution of the eye-tracking sensor is limited by the frame rate of the retinal projection system, which currently is 60 Hz. To overcome this limitation, a faster micro mirror could be used. However, as a collimated laser beam requires sufficient micro mirror aperture, miniaturization of the micro mirror is restricted. Therefore, frequencies up to about 120 Hz are technically feasible.

The low-power consumption is achieved because components already contained in the projection system such as the micro-mirror module are not considered in the eye tracking power budget. This is valid as long as the sensor is combined with a retinal projection system.

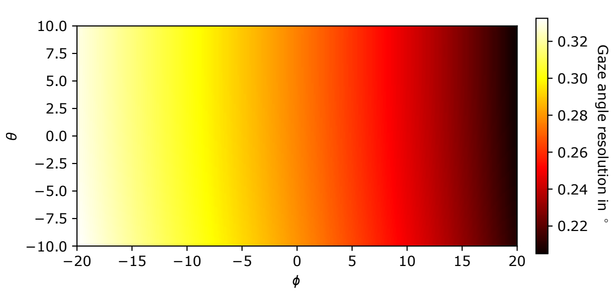


Fig. 13. Calculated gaze angle resolution in degree of the proposed eye tracking sensor for the glasses geometry based on Equation (16).

External light may lead to disturbances in the captured images, which will affect the accuracy of VOG algorithms. A detailed discussion of this limitation is presented in [28].

V. CONCLUSION

We presented a novel and nearly invisible eye-tracking sensor for integration into our existing retinal projection AR glasses prototype. Compared to VOG eye-tracking sensors, our technology achieves a significant reduction in power consumption at comparable gaze angle resolution and full integration into a frame temple.

The achievable gaze angle resolution is evaluated experimentally based on a laboratory setup. It shows comparable gaze angle resolution in comparison with VOG sensors. In addition, the theoretical gaze angle resolution of the proposed head worn eye tracking sensor is calculated, which leads to even better gaze angle resolution as the laboratory setup.

Based on these results, the next step in our future work is the s integration of the proposed eye- tracking sensor into our AR glasses and the evaluation of the whole setup under real-worlddemonstrator to perform experiments and evaluate the sensor under real conditions.

Additionally, the eye- tracking algorithms used must be transferred to an embedded platform and integrated into the demonstrator. This work has paved the way to always-on data glasses at negligible power consumption, which - we believe - will open up many novel applications in human-computer interaction.

REFERENCES

[1] S. Bazrafkan, A. Kar, and C. Costache, "Eye gaze for consumer electronics: Controlling and commanding intelligent systems." *IEEE Consumer Electronics Magazine*, 2015.

[2] C. Braunagel, W. Rosenstiel, and E. Kasneci, "Ready for take-over? a new driver assistance system for an automated classification of driver take-over readiness," *IEEE Intelligent Transportation Systems Magazine*, 2017.

[3] P. M. Corcoran, F. Nanu, S. Petrescu, and P. Bigioi, "Real-time eye gaze tracking for gaming design and consumer electronics systems," *IEEE Transactions on Consumer Electronics*, 2012.

[4] A. Bissoli, D. Lavino-Junior, M. Sime, L. Encarnacao, and T. Bastos-Filho, "A human machine interface based on eye tracking for controlling and monitoring a smart home using the internet of things," *Sensors*, 2019.

[5] J. Kim, Y. Jeong, M. Stengel, K. Akşit, R. Albert, B. Boudaoud, T. Greer, J. Kim, W. Lopes, Z. Majercik, P. Shirley, J. Spjut, M. McGuire, and D. Luebke, "Foveated ar: Dynamically-foveated augmented reality display," *ACM Trans. Graph.*, 2019.

[6] *Implementing Technology as Mother Nature intended to be - its impact on a wide range of experiences and market opportunities*, Tobii AB, Danderyd, Sweden, 2019.

[7] M. Kassner, W. Patera, and A. Bulling, "Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 2014.

[8] M. Tonsen, J. Steil, Y. Sugano, and A. Bulling, "Invisibleeye: Mobile eye tracking using multiple low-resolution cameras and learning-based gaze estimation," *ACM Interact. Mob. Wearable Ubiquitous Tech.*, 2017.

[9] D. Katrychuk, H. K. Griffith, and O. V. Komogortsev, "Power-efficient and shift-robust eye-tracking sensor for portable vr headsets," in *11th ACM Symposium on Eye Tracking Research & Applications*, 2019.

[10] W. Fuhl, M. Tonsen, A. Bulling, and E. Kasneci, "Pupil detection for head-mounted eye tracking in the wild: an evaluation of the state of the art," *Machine Vision and Applications*, 2016.

[11] J. Lin, D. Cheng, C. Yao, and Y. Wang, "Retinal projection head-mounted display," *Frontiers of Optoelectronics*, 2017.

[12] C. Jang, K. Bang, S. Moon, J. Kim, S. Lee, and B. Lee, "Retinal 3d: Augmented reality near-eye display via pupil-tracked light field projection on retina," *ACM Trans. Graph.*, Nov. 2017.
[13] I. Rigas, H. Raffle, and O. V. Komogortsev, "Photosensor oculography: Survey and parametric analysis of designs using model-based simulation," *IEEE Transactions on Human-Machine Systems*, no. 6, Dec 2018.
[14] *Smartglasses Light Drive*, Bosch Sensortec, 2020, online.
[15] *Interactive Projection Module*, Bosch Sensortec, 2018, online.
[16] R. H. Webb, G. W. Hughes, and O. Pomerantzeff, "Flying spot tv ophthalmoscope," *Appl. Opt.*, 1980.
[17] S. Stevenson, C. Sheehy, and A. Roorda, "Binocular eye tracking with the tracking scanning laser ophthalmoscope," *Vision Research*, 2016.
[18] D. Carpentras and C. Moser, "See-through ophthalmoscope for retinal imaging," *Journal of Biomedical Optics*, 2017.
[19] Y. Shoji, "Device for tracking pupil of eyeball using intensity changes of reflected light from eyeball and image display using the same," U.S. Patent Patent, US20 070 705 471, Dec. 29, 2009.
[20] N. Sarkar, D. Strathearn, G. Lee, M. Olfat, A. Rohani, and R. R. Mansour, "A large angle, low voltage, small footprint micromirror for eye tracking and near-eye display applications," in *18th International Conference on Solid-State Sensors, Actuators and Microsystems*, 2015.
[21] N. Sarkar, B. O'Hanlon, A. Rohani, D. Strathearn, G. Lee, M. Olfat, and R. R. Mansour, "A resonant eye-tracking microsystem for velocity estimation of saccades and foveated rendering," in *IEEE International Conference on Micro Electro Mechanical Systems*, 2017.
[22] G. Mickael, M. Christophe, and T. Eric, US Patent 9 846 307B2, 2016.
[23] M. Sugawara, M. Suzuki, M. Ishimoto, K. Hasegawa, N. Teshima, K. Yasui, and N. Miyauchi, "Every aspect of advanced retinal imaging laser eyewear: principle, free focus, resolution, laser safety, and medical welfare applications," in *MOEMS and Miniaturized Systems XVII*, 2018.
[24] O. Sakhno, Y. Gritsai, H. Sahm, J. Stumpe, and M. Wegener, "Bragg polarization gratings used as switchable elements in AR/VR holographic displays," in *Digital Optics for Immersive Displays*, 2018.
[25] W. R. Folks, J. M. López-Alonso, B. Monacelli, A. R. Weeks, G. Zummo, D. Mullaly, and G. D. Boreman, "Characterization of digital-micromirror device-based infrared scene projector," *Optical Engineering*, 2005.
[26] *IEC 60825-1:2014 Safety of laser products*, International Electrotechnical Commission Std., 2014.
[27] W. Fuhl, D. Geisler, T. Santini, T. Appel, W. Rosenstiel, and E. Kasneci, "Cbf: Circular binary features for robust and real-time pupil center detection," in *ACM Symposium on Eye Tracking Research & Applications*, 2018.
[28] J. Meyer, T. Schlebusch, H. Spruit, J. Hellmig, and E. Kasneci, "A novel -eye-tracking sensor for ar glasses based on laser self-mixing showing exceptional robustness against illumination," in *ACM Symposium on Eye Tracking Research & Applications*, 2020, pp. 1–5.

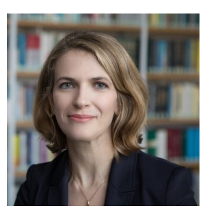
Johannes Meyer was born in Haselünne, Germany in 1994. He received his B.Eng. degree in electrical engineering and M.Eng. degree in industrial informatics from the University of Applied Science Emden-Leer, Germany, in 2017 and 2018 respectively.

From 2017 to 2019, he was a research assistant with the University of Applied Science Emden-Leer. In 2019 he was a research scientist at the Rosen Group, Germany. Currently he is a PhD student with the Human Computer Interaction group at the University of Tübingen, Germany and the Optical Microsystems team, department of Advanced Technologies and Microsystems at the Corporate Research and Advance Engineering division of Robert Bosch GmbH, Germany. His research focuses on low power sensor systems and novel eye tracking sensor systems.

Thomas Schlebusch is team lead for Optical Microsystems, department of Advanced Technologies and Microsystems at the Corporate Research and Advance Engineering division of Robert Bosch GmbH, Germany. He studied Electrical Engineering at RWTH Aachen University, Germany, and at NTNU Trondheim University, Norway. He received his Dr.-Ing. in the field of Medical Information Technology at Helmholtz Institute for Biomedical Engineering, RWTH Aachen. His research focuses on new sensing principles for applications of optical and medical microsystems.

Wolfgang Fuhl was born in Freudenstadt, Germany in 1983. He received the Dipl.-FH degree in computer science, M.Sc. degree in computer science and his Phd from the University of Tübingen Germany, in 2011, 2014, and 2019 respectively.

Currently he is a Postdoc at the University of Tübingen Germany in the Human Computer Interaction group. His research focuses on the methods of machine learning as well as computer vision. He is also interested in the application of machine learning methods to industrial challenges and in the field of human computer interaction.



Enkelejda Kasneci is a Professor of Computer Science at the University of Tübingen, Germany, where she leads the Human-Computer Interaction Lab.

As a BOSCH scholar, she received her M.Sc. degree in Computer Science from the University of Stuttgart in 2007. In 2013, she received her PhD in Computer Science from the University of Tübingen. For her PhD research, she was awarded the research prize of the Federation Südwestmetall in 2014. From 2013 to 2015, she was a postdoctoral researcher and a Margarete-von-Wrangell Fellow at the University of Tübingen. Her research evolves around the application of machine learning for intelligent and perceptual human-computer interaction. She serves as academic editor for PlosOne and as a TPC member and reviewer for several major conferences and journals.