Episcleral surface tracking: challenges and possibilities for using mice sensors for wearable eye tracking

Frank H. Borsato*
Computer Science Department
UTFPR - Brazil

Carlos H. Morimoto†
Computer Science Department
University of São Paulo - Brazil

Abstract

Video-based eye trackers (VETs) have become the dominant eye tracking technology due to its reasonable cost, accuracy, and easy of use. VETs require real-time image processing to detect and track eye features such as the center of the pupil and corneal reflection to estimate the point of regard. Despite the continuous evolution of cameras and computers that made head mounted eye trackers easier to use in natural activities, real-time processing of high resolution images in mobile devices remains a challenge. In this paper we investigate the feasibility of a novel eye-tracking technique intended for wearable applications that use mice chips as imaging sensors. Such devices are widely available at very low cost, and provide high speed and accurate 2D tracking data. Though mice chips have been used for many purposes other than a computer's pointing device, to our knowledge this is the first attempt to use it as an eye tracker. To validate the technique, we built an episcleral database with about 100 high resolution episcleral patches from 7 individuals. The episclera is the outer most layer of the sclera, which is the white part of the eye, and consists of dense vascular connective tissue. We have used the patches to determine if the episclera contains enough texture to be reliably tracked. We also present results from a prototype built using an off-the-shelf mouse sensor. Our results show that a mouse-based eye tracker has the potential to be very accurate, precise, and fast (measuring 2.1' of visual angle at 1 KHz speed), with little overhead for the wearable computer.

Keywords: eye tracking, sclera database, translation sensor, mouse sensor

Concepts: •**Applied computing** → **Imaging;** •**Computing methodologies** → *Tracking;* •**Human-centered computing** → Interaction techniques;

1 Introduction

Video-based eye trackers (VET) have become the dominant method to track eye movements in the last two decades. Compared to other alternatives such as magnetic search coil and electro-oculography, video-based eye trackers combine reasonable accuracy, quick set up, and comfort [Morimoto and Mimica 2005]. To track eye movements and estimate the point of gaze, VETs rely on the computation of image features such as the center of the pupil, the contour of the iris, the center of corneal reflections generated by external infrared light sources, etc [Hansen and Ji 2010]. Continuous progress in video technology have allowed VETs with higher spatial and temporal resolutions. Top of the line eye trackers such as the EyeLink 1000 can provide eye measures up to 2000 Hz, though most low end trackers provide data at 50 or 60 Hz.

Recently, laser speckle have been used for tracking ocular microtremor (OMT). OMT is the smallest of the involuntary eye movements, with an amplitude of 150 to 2500 nm (equivalent to 12 to 216μ rad) and a frequency range between 20 to 150 Hz. The technique projects a beam of coherent light on the sclera, and the patterns produced are correlated to estimate the translation. Influence of larger movements are removed by filtering low frequency components. Despite the high spatial resolution, this technique is not appropriate to track large eye movements such as saccades, and its cost is still high as EMCCD cameras are used to keep laser emission at safe levels [Kenny et al. 2013].

In this paper we investigate the use of high speed, low resolution sensors largely used in computer pointing devices (mice) to estimate the translation of small regions of the eye. Instead of using the speckle created by a laser, mouse sensors can be adapted to use near infrared LEDs to track small patches of the scleral surface. Advantages of this technique would include a significant increase in speed at a very low cost. Because mouse sensors are embedded with 2D tracking software, the computational complexity of the eye tracker would also be considerably reduced. Nonetheless, several issues must be investigated, such as the appropriateness of the visual texture corresponding to the episcleral vessels for tracking, and the minimum frame rate needed to avoid movement aliasing and its relation to the frame size. There are also many aspects related to the engineering of the solution, such as the size and placement of the sensor, the focusing and depth of field, and the amount of light required by the sensor.

We have used an off-the-shelf mouse sensor to build the prototype used in our experiments. These sensors are low-power, inexpensive, and provide high frame rates. The basic idea is to use the 2D tracking data provided by the mouse to avoid heavy image processing by the wearable platform. Mouse sensors have been used in a variety of applications that are quite different from pointing devices, having demonstrated to be a practical and low-cost solution to challenging tasks [Da Silva et al. 2011; Jackson et al. 2007; Lott et al. 2007; Minoni and Signorini 2006].

Although we are interested in developing a complete eye tracker, the focus of this paper will be on the problem of assessing the sclera surface as a candidate for tracking using low cost mouse sensors. The main contributions are:

- Assessment of mouse sensors as translation measurement devices applied to eye tracking, including optical issues and eye safety;
- Assessment of the sclera as a feature rich surface for tracking and the construction of a database of high resolution scleral images;
- Proposal of a prototype to track the translations of the scleral surface using mouse sensors, capable of sustaining a 1 KHz output rate.

*e-mail: frankhelbert@utfpr.edu.br
†e-mail:hitoshi@ime.usp.br

ETRA 16, March 14-17, 2016, Charleston, SC, USA
ISBN: 978-1-4503-4125-7/16/03
DOI: http://dx.doi.org/10.1145/2857491.2857496

The next section introduces the challenges on the use of the scleral surface to build an eye tracker. Section 3 presents actual implementation details related to the use of mouse sensors to measure eye movements, including optics and safety concerns. Section 4 describe experiments to evaluate the method and Section 5 presents and discusses the experimental results. Section 6 concludes the paper.

2 Optically tracking movements of the sclera

The sclera (the white part of the eye) is formed by a tough connective tissue, providing an opaque protective capsule for the intraocular structures and a stable support during eye movements [Watson and Young 2004]. It may be divided into four layers, but the two outermost are of special interest for this work - the Tenon's capsule and the episclera. The Tenon's capsule is firmly attached to the limbus, but becomes freely mobile over the underlying episclera just 3mm far from the limbal attachment. Its veins run superficially and the arteries come close to the surface only near the limbal area. The episclera forms the superficial aspect of the sclera and is well-vascularized, with the arterial network often going not closer than 4 mm from the limbus [Watson and Young 2004]. Such organization lead to a multi-layered vessel distribution which, despite being stable over time, presents complex non-linear deformation as the result of eye movements [Thomas et al. 2010].

To track the gaze direction, we propose to exploit the visible texture composed of the episcleral vessels, by estimating the translation of a small patch on the sclera surface using a mouse sensor. The sensor estimates the translation of the underlying surface by correlating the current frame with shifted versions of an internal reference frame [Gordon et al. 2001]. However, the visible portion of the sclera in each side of a typical adult eye is only about 4 mm [Kenny et al. 2013] when the eye is looking straight ahead. Therefore, the sensor must track a small area within the sclera. This might be advantageous since the deformations due to perspective projection and the limited depth of field are aggravated when large areas are imaged (see section 3.1.2). On the other hand, if the tracking area is too small, there are chances that no visible vessels will be inside the field of view, and the mouse sensor will not be able to estimate the translation. In Section 4.2.3 we investigate how the scleral sample size influences the performance of a correlation based algorithm, and we use this result to select an appropriate magnification for the prototype.

Eye movement amplitude is typically measured in units of angular rotation. Studies differ on the top speeds of saccades, in normal individuals, the peak velocity of saccades varies from 500°/s [Rayner 1978] to 700°/s [Sharpe and Wong 2005]. For the purpose of measuring surface translations it is more convenient to express displacement using linear units. The fastest speed of 700°/s, considering a typical eye of 23 mm in diameter, would lead to a top speed of 140 mm/s as perceived at the eye equator. The sampling rate must be fast enough to avoid aliasing and to allow for the fastest eye movements to be tracked. We can relate the magnification of the optical system, the size of the sensor, the speed of the underlying movement, the frame to frame minimum overlay and the minimum frame rate as follows:

$$f_r = \frac{l_s \times m}{sz \times (1 - f_o)} \tag{1}$$

where f_r denotes the frame rate, l_s is the linear speed in mm/s, m is the magnification, sz is the sensor size in mm and f_o is the frame to frame overlay.

For a typical sensor matrix of 30 × 30 pixels distributed along 1.8 mm × 1.8 mm with a unity magnification factor, if the maximum translation between frames is limited to one pixel (which gives $f_o = 0.966$), then the frame rate must be 2333 Hz to track saccades at peak velocity. If the maximum translation between frames is set to 3 pixels, then the minimum frame rate drops to 777 Hz. Observe from Table 1 that such frame rates can be easily achieved using mouse sensors.

Sensor		ADNS-5030	◇	ADNS-3080	◇
Feature	Unit				
Max speed	ips	14	1700°/s	40 (@6400 Hz)	5000°/s
Frame rate	Hz	variable		2000-6469	
Resolution	cpi	500/1000	7.6' @1000cpi	400/1600	4.7' @1600cpi
Acceleration	g	2	> 30000 $°/s^2$	15	$\gg 30000$ $°/s^2$
Sensor size	pixels	15 × 15		30 × 30	
Depth of field	mm	±0.5		±0.5	
Lens magnification		1.25		1.00	
Power	mA	15 (typ)		52 (max)	

Table 1: *Selected features of two LED-based optical mice sensors and associated lens [Avago Technologies 2007; PixArt Imaging Inc. 2008] and; ◇ resulting eyeball surface tracking characteristics under ideal conditions.*

3 Implementation issues

The proposed method is based on the tracking of a small area on the sclera. As the area to be tracked is small, higher frame rate eye trackers can be achieved using low-end sensors which support the definition of ROIs (regions of interest) and binning. Instead, we propose the use of mouse sensors that can directly compute translations of a very small area at high frame rates.

3.1 Mouse sensors to measure eye movements

The mouse was designed to be a pointing device that tracks the underlying surface as its body is translated. Its performance was considerably enhanced with the introduction of optical devices. Latest models can track motion velocities up to 40 inches per second, with frame rates greater than 6 KHz, and a resolution of 3600 counts per inch (cpi). Mass production makes it particularly affordable both in terms of cost and availability. Moreover, most have standardized communication protocols, which makes integration with micro-controllers straightforward. Table 1 shows the specification of two mouse sensors (ADNS-5030 and ADNS-3080) and their theoretical eye tracking performance if the episcleral surface presents favorable characteristics. Observe that the most powerful chip, the ADNS-3080, could potentially track 5000°/s with a spatial resolution of 4.7' when configured at 1600 cpi, and even the lower end ADNS-5030 could perform significantly better than most current high end eye trackers.

Several works using mouse sensors have been published exploiting their high resolution and speed in various tasks, markedly as a low-cost optical displacement sensor. For example, in [Jackson et al. 2007], a mouse sensor is used for vehicle tracking, while in [Cimino and Pagilla 2010] the mouse is used as a robot odometer. In [Lott et al. 2007], the walking of flies were computed using an upside-down mouse with a floating ball with sub-millisecond resolution. Several other applications, including a wood bar strain measurement, a radial artery diameter change, and the off centered rotation of a sphere simulating an artificial eye were also reported [Da Silva et al. 2011]. Regarding works using a mouse as a biological displacement sensor, it is relevant to point out that:

- Previous works do not report the presence of angle snapping in their sensors. It is unclear if and how this feature was disabled.

- Increasing the magnification of the lens leads to reduced depth of field, which makes the adjustment of the device to the subject's eye more difficult;
- Current LEDs used in mice were designed to be safe during occasional eye exposure. New safety guidelines and standards might be required for constant exposure for using mice as eye trackers.

3.1.1 Mouse drift control

The mouse drift control, also known as angle snapping or prediction, is a type of path correction algorithm that are commonly integrated in the mouse. It is an unlisted feature for many sensors such as the ADNS-3060 and ADNS-3080 [PixArt Imaging Inc. 2008]. It works by changing the movement that lies in a small range of degrees from the axes, snapping it to the closest one. Its original purpose was to assist users in drawing straight lines, even with the mouse body slightly rotated. For eye tracking applications, snapping would have serious implications, as all small movements close to an axis would be unpredictably distorted, affecting the accuracy. To verify the presence of angle snapping, we have developed the experiment described in Section 4.1.

3.1.2 Focal length and depth of field

The focal length of the embedded mouse lens is short, ranging from 4 to 6 mm [PixArt Imaging Inc. 2008]. The mouse body is designed to keep the surface in focus at normal operating conditions, which imposes a limitation when the mouse is used for other purposes. A change in the lens may be also required to increase the distance to the surface or allow higher tracking speeds. In [Tunwattana et al. 2009] a telecentric lens with 65 mm working distance is used in the investigation of axial light biasing. [Ross et al. 2012] use a 32 mm focal length lens to track a rough surface while [Jackson et al. 2007] use a 10 mm focal lens to increase the maximum allowed speed. We have employed a reversed mounted 3.6 mm M12 lens in order to reach a magnification of about 2.27×.

A regular depth of field (DOF) for a mouse is about ± 0.5 mm [PixArt Imaging Inc. 2008] and is important as an assembly tolerance. However, when using the mouse in applications which demand the change of the lens, the DOF must be recalculated to assess the suitability of the proposed setup to the intended application. At magnifications greater than $0.5\times$, the DOF depends only on three parameters [Savazzi 2011]:

$$DOF = \frac{2.N_e.C_M}{M^2} \tag{2}$$

where C_M is the maximum allowed size for the circle of confusion of the sensor, M is the actual magnification and N_e is the effective aperture, which can be calculated as follows:

$$N_e = N.(M+1) \tag{3}$$

where N is the aperture and M is the magnification.

Figure 1 shows the effect of magnification on the DOF for specific apertures using (2), (3), and an estimated $C_M = 130\mu$m for the mouse sensor. As the magnification increases, the DOF decreases. A smaller aperture increases DOF but allows less light to enter. We have used a lens with an aperture of $f/1.4$, and the magnification is about $2.27\times$, which gives a DOF of about ± 0.115 mm.

3.1.3 Eye safety

Because a mouse LED can be very bright and occasional retinal exposure may occur, we have investigated if the LED can represent a safety risk to the integrity of the eye, prior to the construction of the prototype.

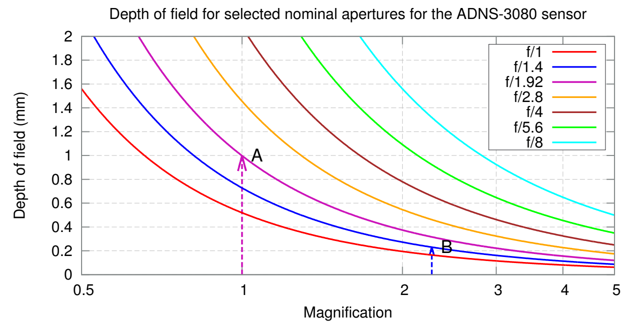


Figure 1: *Depth of field in function of magnification and nominal aperture. (A) indicates the mouse setting while (B) indicates our prototype setting.*

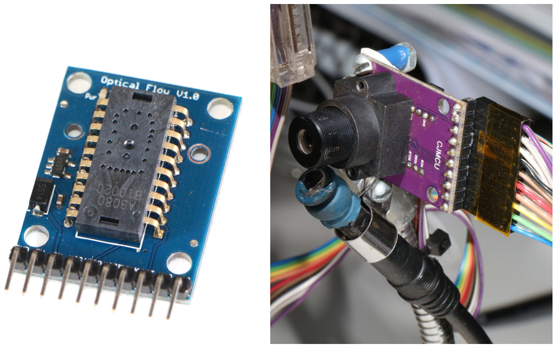
The eye structures that need our attention are the retina and the cornea. The maximum allowed exposure (MPE) of the retina, considering the subtended angle and following the IEC/EN 62471 [IEC 2006], is a radiance of at most 6300 mW/cm^2.sr for long term exposures (longer than 100 s). The corneal limit as defined by IEC/EN 62471 [IEC 2006] is $E_{IR} = 10$ mW/cm^2 for exposures longer than 1.000 s, and the limit defined by the ICNIRP [ICNIRP 2013] guidelines is $E_{IR} = 33.3$ mW/cm^2 for exposures longer than 1 s for sources with a narrow emitting bandwidth and wavelength lower than 1μm. Using the limit defined by the IEC, an LED at $d = 1.157$ cm from the eye can have a radiant intensity $I = 46.71$ mW/sr and a radiance of $L = 237.89$ mW/cm^2.sr. If the ICNIRP MPE is used, those values increase to $I = 155.7$ mW/sr and $L = 793$ mW/cm^2.sr. Taking into account the radiant intensity of the matching LEDs and the radiance specification of the mouse sensor, it is feasible to work bellow those thresholds, i.e., the sensors can be safely used. It is important to note that those are all worst case scenarios, as the device is designed to be pointed to the sclera, where no MPEs are defined. Moreover, the results shown here took long term exposure into account, and therefore, are conservative for both corneal and retinal exposures.

3.2 Prototype

To speed up prototyping, an Arduino Leonardo board was used to interface the mouse sensors with a computer. Arduino is an open-source electronics prototyping platform based on Atmel microcontrollers with a free programming IDE. The board is based on the ATmega32u4 microcontroller, with an integrated 2.0 USB port.The board also has an SPI port on the ICSP header which allows communication with mouse sensors such as the ADNS-3060 and ADNS-3080 [PixArt Imaging Inc. 2008]. It also has a TWI port which allows communication with sensors such as the ADNS-5030 [Avago Technologies 2007].

Interfacing with a mouse involves using the right protocol to send commands and read data. The used the ADNS-3080 [PixArt Imaging Inc. 2008], as it is already soldered on a board with lens mount. Figure 2a shows the board, that is commercially available as an optical flow device for flight stabilization. The coordinate system used in the experiments, unless explicitly stated, is the one centered at the sensor, with the z axis parallel to the optical axis.

The sensor accumulates the movement internally, allowing the host processor to read the sensor asynchronously. The pooling rate must be kept fast enough to avoid two possible conditions: 1) avoid overflow of the registers, in which case the data is lost; 2) avoid translation aliasing, which is the case when rapid movements on opposite directions cancel themselves (i.e., positive and negative values are summed in the register). The raw data given by the sensor contains the amount of translation in both axes (x, y) since the last sample,



(a) (b)

Figure 2: *a) Board with the ADNS-3080 IC. b) Detail of the prototype sensing part, with the lens and light source mounted.*

a measure of the surface quality (called *squal*), the maximum pixel intensity of the last frame, and the frame period, among other parameters.

We have also included a PlayStation Eye Camera [Sony 2014] in our prototype, which we call secondary camera, to record videos of the eye movements. The secondary camera was configured to capture eye images at 30 Hz and VGA resolution, though speeds up to 180 Hz were possible with lower resolutions. The camera was tweaked to output the VSYNC signal[1], which was attached to an interrupt pin on the micro-controller to output synchronized data.

4 Experiments

We have conducted three experiments, with the following objectives. The first experiment was to validate the sensors to be used in the initial prototype, by checking if their data is corrupted by angle snapping. The second experiment was to validate the episclera as a feature rich surface for tracking, using pictures from a scleral database, and the third experiment qualitatively evaluates the performance of our first prototype with a human user performing controlled eye movements, such as saccades and smooth pursuits. The prototype used in the experiment is shown in Figure 2.

4.1 Testing for angle snapping

Angle snapping is an embedded feature of many mouse sensors that facilitates human users to draw straight lines but that can significantly distort eye tracking results. Because this feature is rarely reported in the literature (including products' data sheets), our first experiment investigates the presence and influence of angle snapping in a common mouse sensor.

To check for angle snapping, two ADNS-3080 mice chips with identical lens were arranged on the same rigid board, rotated by θ degrees with respect to the other (see Figure 3). Instead of measure θ directly, the angle is calculated by comparing the slope of the lines that best fit the data reported by the sensors. A number of trials are used to find the mean angle and its standard deviation. Those trials consist of a surface moving under the sensors along straight lines. The value of θ should be the same when computed using lines with different orientations. For line orientations close to or parallel to one of the sensors though, the angle might be different due to angle snapping.

4.2 Scleral surface quality

Optical mice typically place a visible LED at an angle with the supporting surface to maximize the shading produced on the sur-

[1] See the camera synchronization report for details.

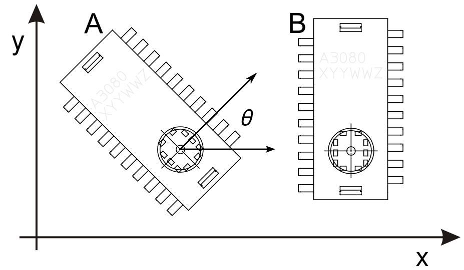


Figure 3: *Top view of the arrangement for testing the angle snapping on the ADNS-3080 IC. The axes shown are for sensor (B).*

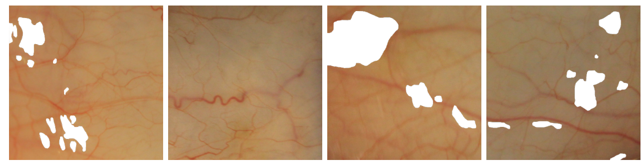


Figure 4: *Sclera database examples. Images* 4 *mm wide from four different volunteers. The white areas are the reflexes masked in the alpha channel.*

face microtexture when captured by the imaging sensor. This technique works very well on surfaces like white paper, manila, and black walnut. However, surfaces with polished glass present difficulties as oblique illumination produces only a few features (generally caused by loose particles). To avoid possible reflections from the tear layer caused by other light sources, we wish to track the episcleral vessels directly using the mouse sensor. As pointed out in Section 2, imaging a small region is preferred. However, the size of such region is limited by the vessel distribution, which is known to vary along the sclera [Watson and Young 2004].

4.2.1 Episcleral database

UBIRIS [Proenca et al. 2010] is a publicly available database containing eye images with resolution 400×300 pixels. UBIRIS has been applied for person identification using the episcleral vessels [Thomas et al. 2010]. Unfortunately, the UBIRIS images are not calibrated and their 15 pixel per millimeter resolution at the sclera is insufficient for our tracking purposes. So we have created our own eye image database for testing the feasibility of the sclera as a tracking surface.

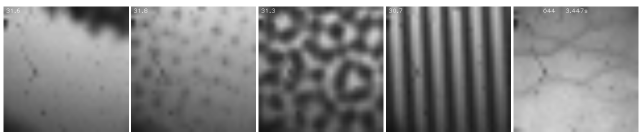
High resolution close-up pictures of the eye were taken using a Canon 600D at 18 megapixels and ISO 800 sensibility, with a 55 mm lens attached to an extension tube of 65 mm. The aperture was set at f/12.9, and exposure to 1/50 s. An LED ring, with 2800 K color temperature, was attached to the lens to lit the eyes evenly. Seven volunteers took part in the image collection. Each person was asked to sit comfortably while photos were taken at three different gaze directions for each eye, frontally, medially and laterally. At each photo session, a 150 line pairs per inch (lp/in) test chart was used for calibration of the pixel/mm resolution, as the focusing needed for a particular person changes the actual focal length. The images were further processed to mask the specular reflections from the LED ring and extract 4 mm patches. No color correction or noise filtering were performed. The normalized resolution for all participants is 210 pixels/mm (i.e. each mm at the sclera corresponds to 210 pixels in the image). Figure 4 shows exemplars of the database, which contains 97 images.

4.2.2 Mouse image sensor noise estimation

The temporal noise in the raw image data of the mouse sensor used in the experiments was estimated using the technique from Foi et al. [Foi et al. 2007] and the images in Figure 5a-c. The method was chosen because it does not require a particular target, and illumination does not need to be uniform or known in advance. Particularly, we considered the mouse raw data output to be as follows:

$$z(x) = y(x) + \sigma(y(x))\xi(x), \quad x \in X, \tag{4}$$

where X is the set of valid pixels, z is the sensor output, y is the ideal output, ξ is zero-mean random noise with standard deviation equal to 1, and σ is a function of y, modulating the standard deviation of the overall noise component. Using 292 captured images, the estimated standard deviation ($\hat{\sigma}$) curve, as a function of the expectation of the pixel value is given in Figure 6.



(a) (b) (c) (d) (e)

Figure 5: *Actual mouse sensor frame captures with approximately* 0.79×0.79 *mm. a-c) used on sensor noise estimation; d)* 200 *lp/in test chart; e) actual sclera snapshot.*

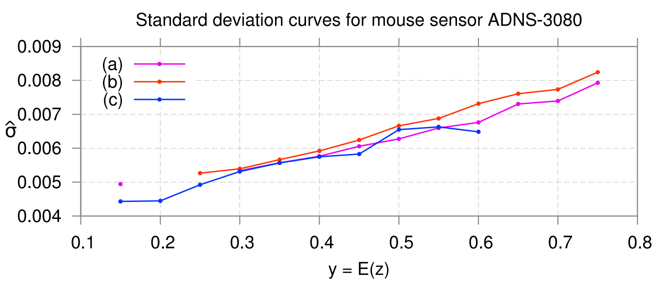


Figure 6: *Signal dependent standard deviation curves for ADNS-3080 raw data and segmentation parameter* $\Delta = 0.05$.

4.2.3 Experiment definition

The database contains images of the sclera which are squares about 4×4 mm wide. The size in pixels of each image is 840×840, with three 8-bit channels (RGB). Each color channel is processed separately. Four different sampling region sizes w at the sclera are tested, squares with side length of 360, 260, 160 and 60 pixels, which correspond, respectively, to a field of view of 1.71 mm, 1.24 mm, 0.76 mm, and 0.28 mm. Additionally, two sensor sizes s are also considered, 30×30 pixels, and 15×15 pixels, corresponding respectively to the sensors used on IC's ADNS-3080 and ADNS-5030.

Starting from the top-leftmost pixel at the sclera image, a region of interest (R_w) is defined with size $w \times w$. Eight overlapping neighbor regions ($R_{v1}..R_{v8}$) of the same size are also defined, at a fixed distance from R_w and directions multiple of $\pi/4$. The distance is defined by the quotient from w and s being used. Using the variance estimated for the sensor in Section 4.2.2, noise is added to R_w to produce R_w^n and R_{v0}^n. Noise is also added to $R_{v1}..R_{v8}$ to produce $R_{v1}^n..R_{v8}^n$. Note that R_w^n and R_{v0}^n are not equal, as uncorrelated noise was applied. Nine correlation coefficients are found using a sum of the squared differences approach between R_w^n and $R_{v0}^n..R_{v8}^n$. A false match is given when the best coefficient is not associated to R_{v0}^n. The occurrence of false matches increase with the lack of features in R_w. The process is repeated by sliding R_w over the whole sclera image, for all images in the database.

This experiment is equivalent to a simulation in which the sclera moves under the mouse sensor by the equivalent of a pixel size. Such situation corresponds to the upper bound of the tracking, if we consider at most one pixel displacement between frames. Smaller moves would be more realistic, as the eye moves independently from the sensor. However, as mice algorithms keep the reference frame for the longest possible time (see [Gordon et al. 2001] and related patents), false matches computed at intermediate positions do not result in tracking failure.

4.3 Tests with a real eye

An optical mouse enhances the surface features using an illumination source at a suitable angle of incidence. Most surfaces, at microscopic level, create a rich set of high contrasting patterns when illuminated by a large angle relative to the surface normal, known as the 'grazing' angle of incidence, However, tracking live tissue using a mouse sensor has been reported only a limited number of times, and with the help of a coherent light source, known to produce high contrasting images [Da Silva et al. 2011].

We describe an experiment using static photos of sclera patches in Section 4.2. Despite being an indicative of the surface quality for tracking, a test with a real human eye is more helpful for showing if the tracking is really feasible. Such test also helps on assessing other aspects, such as usability of the device.

The experiment consists of a common eye tracking task, with targets displayed on a computer screen that must be gazed by the participant. Our prototype measures the translation at the outer sclera of the left eye. Figure 7 shows the target positions. Images of the eye are simultaneously taken by a synced secondary camera. The mouse sensor is sampled at 1KHz and the secondary camera has a frame rate of 30 fps. A computer software stores the data and the instantaneous target screen location for posterior analysis. A 22" monitor was employed with the subject 40cm away, covering a $\pm 20^\circ$ vertical and $\pm 30^\circ$ horizontal field of view.

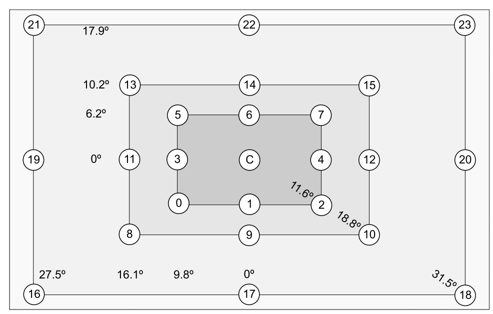


Figure 7: *Distribution of the targets on the screen and approximate gaze angles with the subject* 40cm *away from the monitor.*

5 Results and discussion

5.1 Angle Snapping

To estimate the angle between the two sensors, A and B, data from 21 trials were collected. Each trial consisted of moving a piece of paper under the sensors along a straight line, while data was captured simultaneously from both. The movement was limited to the range of 15° to 50° with respect to both sensors. For each sensor in each trial, a line was fitted to the data using least squares. Figure 8 shows the slopes of the fitted lines and the difference in degrees between the slopes.

A second section was conducted where the sensors were moved within the range of 15° to 30° for sensor B. Figure 9a shows the slopes of each trial. The data from the two sensors are matched by the distance to the polar grid origin and ordered by the distance of the slope to the x axis.

A third section collected data for sensor B moving within -15° and 18°. The result is shown in Figure 9b. The snapping is clearly seen

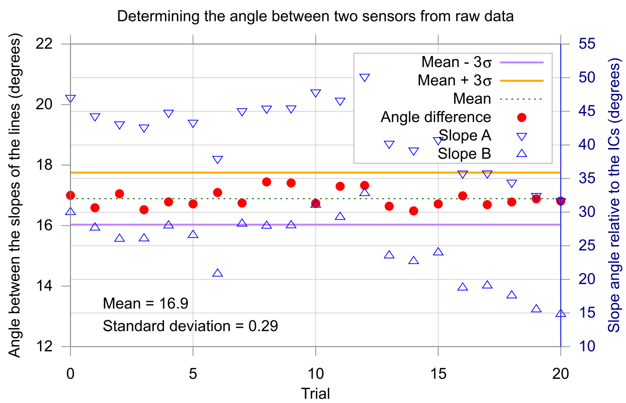


Figure 8: *Estimating the angle between the two sensors. Slope A and B use the scale on the right.*

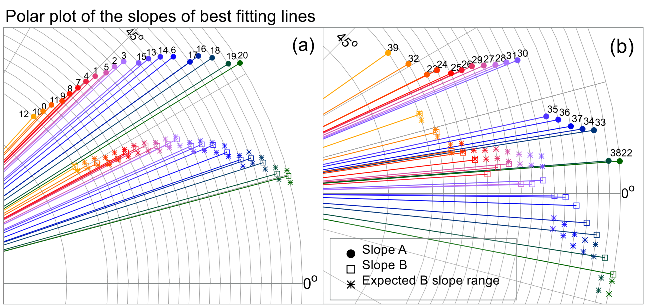


Figure 9: *Testing for angle snapping on the ADNS-3080 IC.*

by the slope of sensor B out of the mean $\pm 3\sigma$ interval with respect to sensor A.

In conclusion, though there was no mention in the sensor's data sheet about angle snapping, it becomes clear that the mouse sensor tested had it enabled. When used as a displacement measuring device, such feature makes the mouse data unreliable, particularly at movements close to the axes. To correct this feature, we were able to obtain snapping-free firmware directly from the manufacturer[2].

5.2 Scleral results

Table 2 shows the results for the scleral image patch test described in Section 4.2.3. We have also simulated the experiment using a computer generated black image with additive Gaussian noise ($M = 0.5, SD = 1.8 \cdot 10^{-2}$). The results show false matches on 85.3% of the patches ($5.1 \cdot 10^3$), close to the expected $8/9$ given by chance.

Window size (w) / Sensor size (s)	Number of R_w tested	Channel		
		R	G	B
		False matches	False matches	False matches
360 / 30	$3.18 \cdot 10^4$	0.00% (0)	0.00% (0)	0.00% (0)
260 / 30	$1.12 \cdot 10^5$	0.00% (6)	0.00% (0)	0.00% (0)
160 / 30	$3.89 \cdot 10^5$	0.32% (1278)	0.02% (91)	0.01% (53)
360 / 15	$7.17 \cdot 10^3$	0.01% (1)	0.00% (0)	0.00% (0)
260 / 15	$2.28 \cdot 10^4$	0.10% (25)	0.00% (2)	0.00% (1)
160 / 15	$9.46 \cdot 10^4$	0.87% (824)	0.20% (190)	0.06% (66)
60 / 15	$8.11 \cdot 10^5$	13.13% (106593)	6.02% (48876)	4.32% (35079)

Table 2: *Scleral patch false matches*

Using window sizes of 360 and 260 pixels, we obtain very good results with minimal false matches. The error start to grow with a window size of 160 pixels, equivalent to a field of view of 0.76mm. The 0.32% of false matches found (worst case using the red channel and $s = 30$) correspond to a single error every 11° of movement performed by the eye (using 27.8cpd as basis, see section 5.3).

The objective of the experiment was to determine if the episclera contains enough texture to be reliably tracked. Additionally, we were able to estimate the size of a ROI at a typical sclera which still presents enough features to be tracked, despite the acquisition noise. Such size corresponds to a square region of about 0.76×0.76 mm. This finding enabled us to choose the magnification ratio for the optical system to be about $2.27\times$, equivalent to a field of view of 0.79mm. Other factors, however, such as the sensor fill factor, lens distortions and change in contrast due to different wavelengths, can negatively impact the results.

5.3 Eye tracking prototype

We have built a prototype to investigate the feasibility of the technique. Figure 10 shows the sensor data during a smooth pursuit. The figure contains images taken by the secondary camera to check the translation reported by the mouse sensor at particular time instants. To highlight differences between consecutive frames, a dense optical flow was computed using the Gunnar Farneback's algorithm [Farnebäck 2003] and the flow directions were color coded using the hue wheel to improve visibility. Differences between farther frames were highlighted using the square of the frame differences and color coded from dark blue to dark red.

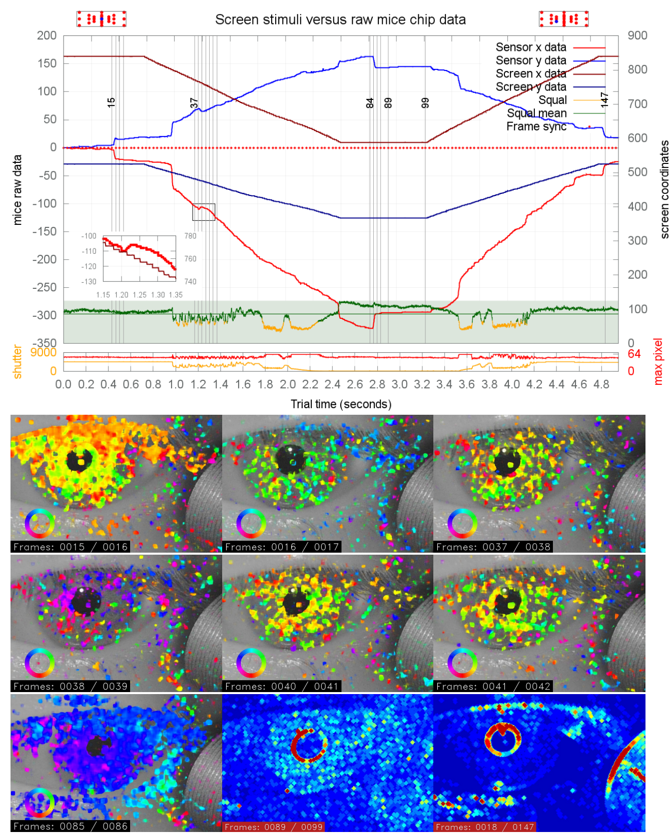


Figure 10: *A lateral downward pursuit from center and back to target 0 and; selected frame flows showing the movement.*

As can be observed on the graph of Figure 10, the mouse successfully estimated the translation of the eye. By comparing the output from the mouse sensor with the target position on screen, we observe the same delays of the smooth pursuit system as reported in

[2]The firmware can be requested free of charge at the company's website and can be uploaded following the datasheet [PixArt Imaging Inc. 2008].

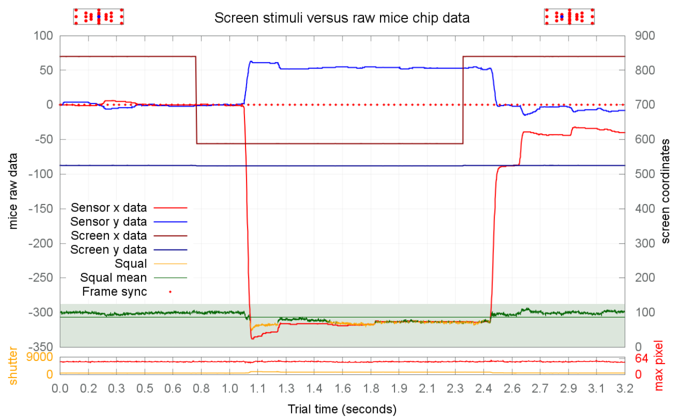


Figure 11: *A lateral saccade from center and back to target 3.*

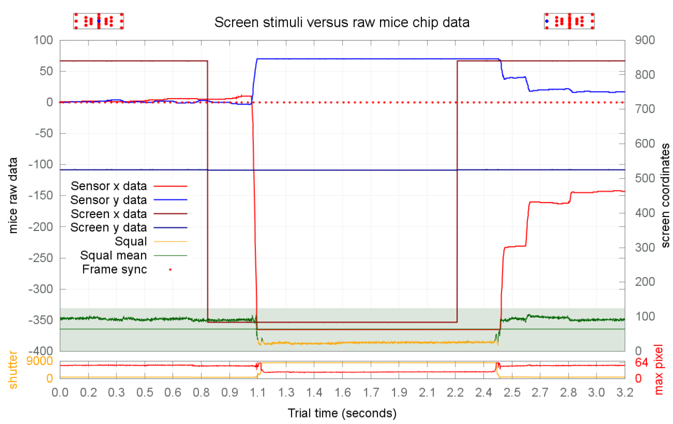


Figure 12: *A lateral saccade from center and back to target 19.*

the literature [Young 1971], i.e. there is a time delay when the target starts to move before the eye starts to move, the same occurring when the target stops.

Figures 11-13 show the results of saccades to targets 3, 19, and 20, respectively. These results are also comparable to those reported in the literature [Fuchs 1971].

The graphs also present the output of the surface quality (Squal) provided by the sensor. The surface quality is proportional to the number of features imaged at a given time [PixArt Imaging Inc. 2008]. The Squal line is encoded by color, with green representing the image quality inside the mean ($M \pm SD$), and orange representing moments where features drop bellow the mean. In the bottom of the graphs, the maximum pixel, in red, and the shutter period of the mouse, in orange, are also shown. The smaller the shutter, the higher the maximum frame rate allowed. The sensor automatically adjusts the shutter period to keep the brightness balanced.

Figure 12 shows that at time instant close to 1.1 s, the tracking is suddenly interrupted. The shutter rises to maximum and the maximum pixel captured by the mouse is reduced. The Squal also drops. The tracking stalls until the eye starts moving back to center. Observe that such behavior is not reproduced in Figure 13, that corresponds to a saccade with similar amplitude. As target 19 is located to the left, the same side used to position the sensor, moving the eye to an extreme angle places the iris under the sensor field of view. Because the depth of field as defined in Section 3.1.2 is critical in this experiment, the small DOF of only $0.2mm$ completely impairs the tracking of the iris.

By taking the raw count to targets 3, 4, 11, 12, and 20, we were able to find a mean horizontal angular resolution of approximately

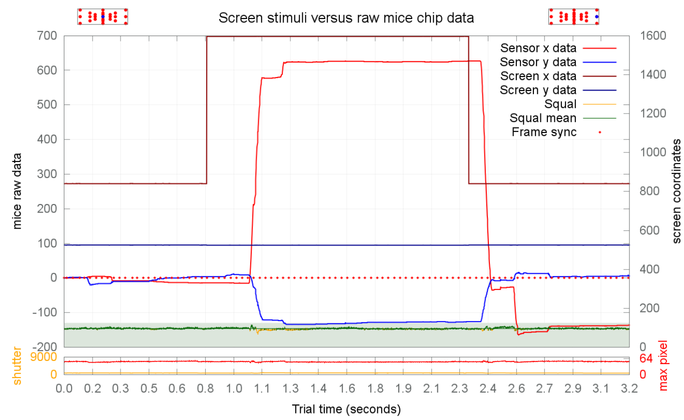


Figure 13: *A lateral saccade from center and back to target 20.*

27.8 counts per degree (cpd) of eye rotation (SD=2.7), equivalent to a resolution of $2.1'$. Correspondingly, using targets 1, 6, 9, 14, 17, and 22, the mean vertical angular resolution was 18.2 cpd (SD=2.3), equivalent to $3.2'$. Supposing the device points at the eye equator, and considering a typical eye of 23 mm, this gives a linear resolution of 3518 cpi (close to the theoretical value of $1600 \times 2.27x = 3632$ cpi).

When comparing the accumulated raw data from start to end, a small difference can be observed. In Figure 10, frames 18/147, the subject is supposed to look at the same point in the screen. However, besides the pupil size, some translation is also visible, represented by the uneven pupil border in the difference image. Indeed, the standard deviation of the eye position across repeated fixations of the same target in healthy subjects is on the order of $1 - 2°$ [Eggert 2007], with an accuracy of about $0.5°$ [Van Opstal and Van Gisbergen 1989].

Our results indicate a horizontal resolution of about 50% higher than the vertical one, when comparing pure horizontal with pure vertical movements. Such difference is expected as a result of the sensor placement but it also depends on the movement performed by the eye. For example, the translation measured at the surface changes with the rotation axis. To complicate the position estimation, physiological evidence shows that the movements performed by the eye include rotations around axes tilted with respect to a common plane [Tweed and Vilis 1990]. Such tilt is responsible for eye torsion, which is likely to occur for the short distances employed. Using the models from [Minken et al. 1995; Van Rijn and Van den Berg 1993], torsion is predicted to be on the order of $\pm 0.77°$ and $\pm 1.54°$, respectively.

5.4 Further challenges

We have used a mouse sensor for estimating the translation of small eye patches. Because the sensor is head mounted, any change on the sensor placement will be considered as an eye movement. Therefore, our method presents problems similar to infrared oculography methods with respect to device adjustment and drift due to the slippage [Eggert 2007].

Blinks are also challenging because they temporarily prevent collecting the eye movement data, which means that the absolute orientation can not be followed. Thus, the current method alone cannot be used for point of gaze estimation, but provides high temporal and spatial resolution data which can be used along with other techniques, such as a video based eye tracker. The resulting hybrid system could benefit from the absolute eye position estimation of the video and the high speed and resolution of the mouse sensor.

The results presented are based on the raw data output by the mouse sensor. The angular resolution was estimated considering a typical eye, with no calibration. Once the eye position can be reliably computed and tracked, current calibration procedures using a number of calibration points to estimate a mapping function could be used to estimate the point of gaze.

6 Conclusion

We have presented a preliminary design of an eye tracker based on the measurement of translation at the scleral surface using a low cost mouse sensor. Key aspects of the development were analyzed, including optical parameters and eye safety. A database of scleral images was introduced to study episcleral texture. Small scleral patches of different sizes, with artificially added noise, were tested. The noise was previously estimated using images from the sensor used in our experiments. The results show the potential of using the sclera for eye tracking. The use of a mouse sensor allows high accuracy and high frame rates at a very affordable price. Because the mouse sensor processes images and outputs translations, the computational cost for gaze estimation can also be considerably reduced. The feasibility of the proposed method was confirmed by experiments with a live eye. The amount of translation reported was compatible with the expected value, considering the movement direction, amplitude, and eye radius. Future research towards the construction of a full feature eye tracker will include new tests on different subjects to confirm our results and to investigate other features such as the effect of blinks, the limbus, and the iris.

Acknowledgments

The authors would like to thank Dan Witzner Hansen for his valuable comments during the initial stages of this work. This research was supported by the Fundação Araucária (DINTER project UTFPR/IME-USP) and FAPESP grant number 2011/00267-1.

References

AVAGO TECHNOLOGIES. 2007. Adns-5030 low power optical mouse sensor. No. AV02-0113EN.

CIMINO, M., AND PAGILLA, P. R. 2010. Location of optical mouse sensors on mobile robots for odometry. In *Robotics and Automation (ICRA), 2010 IEEE Int. Conf. on*, IEEE, 5429–5434.

DA SILVA, M. M., DE ALMEIDA NOZELA, J. R., CHAVES, M. J., ALVES BRAGA JUNIOR, R., AND RABAL, H. J. 2011. Optical mouse acting as biospeckle sensor. *Optics Communications 284*, 7, 1798–1802.

EGGERT, T. 2007. Eye movement recordings: methods. *Developments in ophthalmology 40*, 15.

FARNEBÄCK, G. 2003. Two-frame motion estimation based on polynomial expansion. In *Image Analysis*. Springer, 363–370.

FOI, A., ALENIUS, S., KATKOVNIK, V., AND EGIAZARIAN, K. 2007. Noise measurement for raw-data of digital imaging sensors by automatic segmentation of nonuniform targets. *Sensors Journal, IEEE 7*, 10, 1456–1461.

FUCHS, A. F. 1971. The saccadic system. In *The control of eye movements*. Academic Press New York, 343–362.

GORDON, G. B., KNEE, D. L., BADYAL, R., AND HARTLOVE, J. T., 2001. Proximity detector for a seeing eye mouse, Aug. 28. US Patent 6,281,882.

HANSEN, D., AND JI, Q. 2010. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 3 (march), 478 –500.

ICNIRP. 2013. Guidelines on limits of exposure to incoherent visible and infrared radiation. *Health Physics 105*, 1, 74–96.

IEC. 2006. Photobiological safety of lamps and lamp systems. No. IEC 62471:2006.

JACKSON, J. D., CALLAHAN, D. W., AND MARSTRANDER, J. 2007. A rationale for the use of optical mice chips for economic and accurate vehicle tracking. In *Automation Science and Engineering, 2007. CASE 2007. IEEE Int. Conf. on*, IEEE, 939–944.

KENNY, E., COAKLEY, D., AND BOYLE, G. 2013. Ocular microtremor measurement using laser-speckle metrology. *Journal of biomedical optics 18*, 1, 016010–016010.

LOTT, G. K., ROSEN, M. J., AND HOY, R. R. 2007. An inexpensive sub-millisecond system for walking measurements of small animals based on optical computer mouse technology. *Journal of neuroscience methods 161*, 1, 55–61.

MINKEN, A., GIELEN, C., AND VAN GISBERGEN, J. 1995. An alternative three-dimensional interpretation of hering's equal-innervation law for version and vergence eye movements. *Vision research 35*, 1, 93–102.

MINONI, U., AND SIGNORINI, A. 2006. Low-cost optical motion sensors: An experimental characterization. *Sensors and Actuators A: Physical 128*, 2, 402–408.

MORIMOTO, C. H., AND MIMICA, M. 2005. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding 98*, 4–24.

PIXART IMAGING INC., 2008. ADNS-3060, ADNS-3080 High-performance Optical Mouse Sensor / ADNS-2120 Solid-State Optical Mouse Lens, oct.

PROENCA, H., FILIPE, S., SANTOS, R., OLIVEIRA, J., AND ALEXANDRE, L. 2010. The UBIRIS.v2: A database of visible wavelength images captured on-the-move and at-a-distance. *IEEE Trans. PAMI 32*, 8 (August), 1529–1535.

RAYNER, K. 1978. Eye movements in reading and information processing. *Psychological bulletin 85*, 3, 618.

ROSS, R., DEVLIN, J., AND WANG, S. 2012. Toward refocused optical mouse sensors for outdoor optical flow odometry. *Sensors Journal, IEEE 12*, 6, 1925–1932.

SAVAZZI, E. 2011. *Digital photography for science: Close-Up Photography, Macrophotography, and Photomacrography*. Lulu Enterprises, Raleigh.

SHARPE, J., AND WONG, A. 2005. Anatomy and physiology of ocular motor systems. *Walsh and Hoyt's clinical neuro-ophthalmology*, 809–885.

SONY, C., 2014. Playstation eye camera. [Online, 14/jan/2014].

THOMAS, N. L., DU, Y., AND ZHOU, Z. 2010. A new approach for sclera vein recognition. vol. 7708, 770805–770805–10.

TUNWATTANA, N., ROSKILLY, A., AND NORMAN, R. 2009. Investigations into the effects of illumination and acceleration on optical mouse sensors as contact-free 2d measurement devices. *Sensors and Actuators A: Physical 149*, 1, 87–92.

TWEED, D., AND VILIS, T. 1990. Geometric relations of eye position and velocity vectors during saccades. *Vision research 30*, 1, 111–127.

VAN OPSTAL, A., AND VAN GISBERGEN, J. 1989. Scatter in the metrics of saccades and properties of the collicular motor map. *Vision research 29*, 9, 1183–1196.

VAN RIJN, L. J., AND VAN DEN BERG, A. V. 1993. Binocular eye orientation during fixations: Listing's law extended to include eye vergence. *Vision research 33*, 5, 691–708.

WATSON, P. G., AND YOUNG, R. D. 2004. Scleral structure, organisation and disease. a review. *Experimental eye research 78*, 3, 609–623.

YOUNG, L. R. 1971. Pursuit eye tracking movements. In *The control of eye movements*. Academic Press New York, 429–443.