

# Towards Automating Retinoscopy for Refractive Error Diagnosis

ADITYA AGGARWAL, SIDDHARTHA GAIROLA, and UDDESHYA UPADHYAY, Microsoft Research, India

AKSHAY P VASISHTA, DIWAKAR RAO, ADITYA GOYAL, and KAUSHIK MURALI, Sankara Eye Hospital, India

NIPUN KWATRA and MOHIT JAIN, Microsoft Research, India

Refractive error is the most common eye disorder and is the key cause behind correctable visual impairment, responsible for nearly 80% of the visual impairment in the US. Refractive error can be diagnosed using multiple methods, including subjective refraction, retinoscopy, and autorefractors. Although subjective refraction is the gold standard, it requires cooperation from the patient and hence is not suitable for infants, young children, and developmentally delayed adults. Retinoscopy is an objective refraction method that does not require any input from the patient. However, retinoscopy requires a lens kit and a trained examiner, which limits its use for mass screening. In this work, we automate retinoscopy by attaching a smartphone to a retinoscope and recording retinoscopic videos with the patient wearing a custom pair of paper frames. We develop a video processing pipeline that takes retinoscopic videos as input and estimates the net refractive error based on our proposed extension of the retinoscopy mathematical model. Our system alleviates the need for a lens kit and can be performed by an untrained examiner. In a clinical trial with 185 eyes, we achieved a sensitivity of 91.0% and specificity of 74.0% on refractive error diagnosis. Moreover, the mean absolute error of our approach was  $0.75 \pm 0.67$ D on net refractive error estimation compared to subjective refraction measurements. Our results indicate that our approach has the potential to be used as a retinoscopy-based refractive error screening tool in real-world medical settings.

CCS Concepts: • Human-centered computing → *Ubiquitous and mobile computing*; • Applied computing → *Consumer health*.

Additional Key Words and Phrases: Retinoscopy, retinoscope, myopia, hyperopia, health sensing, refractive error, screening, diagnosis, smartphone, image processing, mobile health, low cost system, optics

## ACM Reference Format:

Aditya Aggarwal, Siddhartha Gairola, Uddeshya Upadhyay, Akshay P Vasishta, Diwakar Rao, Aditya Goyal, Kaushik Murali, Nipun Kwatra, and Mohit Jain. 2022. Towards Automating Retinoscopy for Refractive Error Diagnosis. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 97 (September 2022), 26 pages. <https://doi.org/10.1145/3550283>

## 1 INTRODUCTION

Refractive error is the most prevalent eye disorder [44], where the eye is unable to bend light correctly to focus on the retina, resulting in blurred vision. There are four types of refractive error [63]—myopia, hyperopia, astigmatism, and presbyopia. There are three reasons which can lead to them. First, an eyeball that is too long or too short can make it hard for the eye to focus, leading to myopia (nearsightedness) and hyperopia (farsightedness), respectively. Second, a distortion in the shape of the cornea, which contributes to ~70% of the eye's refractive power, can

---

Authors' addresses: Aditya Aggarwal, aditya097.aggarwal@gmail.com; Siddhartha Gairola; Uddeshya Upadhyay, Microsoft Research, India; Akshay P Vasishta; Diwakar Rao; Aditya Goyal; Kaushik Murali, Sankara Eye Hospital, India; Nipun Kwatra, nkwater@microsoft.com; Mohit Jain, mohja@microsoft.com, Microsoft Research, India.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2474-9567/2022/9-ART97 \$15.00

<https://doi.org/10.1145/3550283>

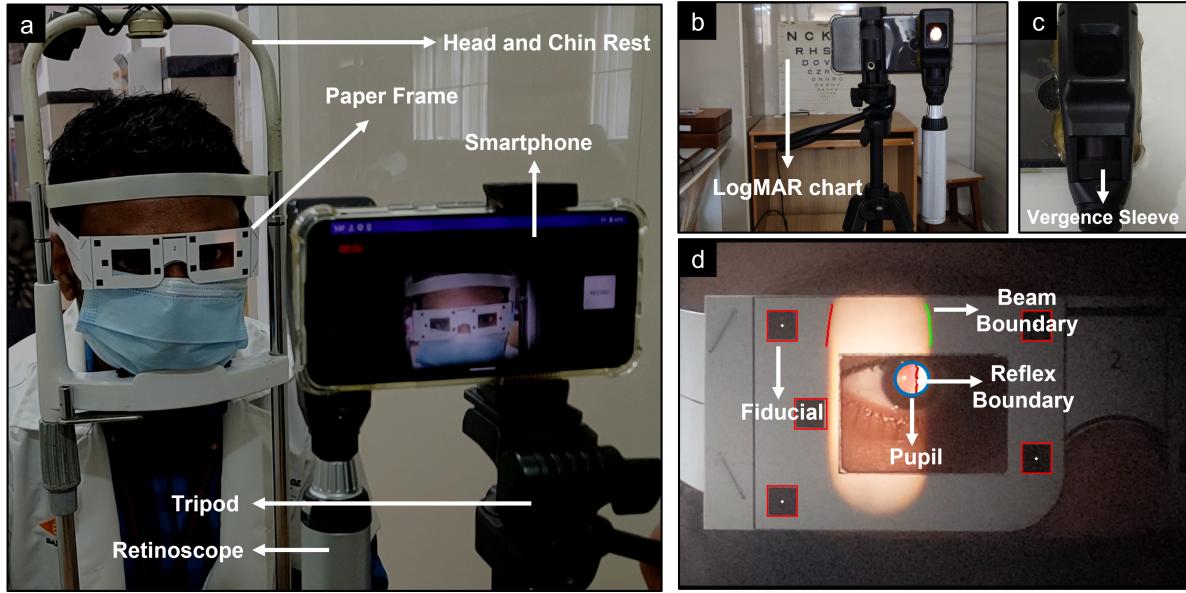


Fig. 1. We propose a smartphone-based automatic retinoscopy solution. (a) Our proposed setup consists of a retinoscope attached to a smartphone, (b) the setup from the patient’s viewpoint facing the logMAR chart, (c) retinoscope and its vergence sleeve, and (d) a single frame of the digital retinoscopy video with automatic detection of fiducials, pupil, reflex edges, and beam edges, using our video processing pipeline.

lead to astigmatism. Third, the eye lens gradually thickens with age and loses its flexibility, making it hard to focus, leading to presbyopia. Refractive error is not a preventable disorder [44]; however, it can be treated using corrective glasses or contact lenses if diagnosed early [63].

Uncorrected refractive error is one of the leading causes of vision impairment [49]. In 2010, uncorrected refractive errors were found to be responsible for visual impairment in 101.2 million people and blindness<sup>1</sup> in 6.8 million people [43]. Blindness due to undiagnosed refractive error is more prevalent in the global south, e.g., in India, it is 1.06% among people aged >40 years [18], compared to 0.24% in the USA [56] and 0.26% in the UK [51]. Undiagnosed refractive error can also cause other severe eye disorders, such as amblyopia in children, in which high refractive error in one eye causes the brain to ignore the blurred image from the weaker eye. This can lead to permanent vision loss of the weaker eye if not corrected in time. Vision screening programs at a large community scale are thus important to address this preventable source of blindness. However, screenings are difficult in developing countries due to a combination of factors, such as poor access to eye care services, unavailability of affordable solutions, and lack of awareness. The conditions are further exacerbated in low-income regions like Central Africa, where less than six trained eye specialists<sup>2</sup> are available for a million people [8]. Therefore, tools to enable refractive error screening in a low-cost and portable manner, without the need of skilled practitioners, can help in democratizing such mass screening programs. Our work is a step in that direction.

There are multiple ways to measure refractive error. Subjective refraction is considered the gold standard, which comprises complex procedural steps performed by an eye specialist. It is subject to the patient’s ability to

<sup>1</sup>WHO defines blindness as visual acuity worse than 3/60, i.e., a person with blindness needs to be as close as 3 feet to see what a person with normal vision can see at 60 feet [44].

<sup>2</sup>Eye specialists comprise of ophthalmologists, optometrists, and allied ophthalmic personnel.

discern changes in clarity while reading the Snellen/logMAR [22] chart with a combination of corrective lenses to obtain the best visual acuity. The feedback and cooperation required from the patient make subjective refraction challenging for young children and infants, developmentally delayed adults, or individuals whose behavior limits the ability to be cooperative. Moreover, language barrier between the patient and examiner, and the time required for the iterative process, act as deterrents toward frequent mass screening using subjective refraction in remote areas.

To address these issues, objective refraction methods (such as retinoscopy, eccentric photorefraction, and autorefractors) have been increasingly used, where no input is required from the patient. Though autorefractors are popular, they are expensive and bulky, and have been found to significantly overestimate myopia [14]. Moreover, they require some cooperation from the patient, making their use challenging for children under 6 years of age [54]. In contrast, retinoscopy is highly accurate and requires minimal patient cooperation. Thus, retinoscopy is preferred over autorefractors [34], and is the recommended starting point for subjective refraction [3]. Retinoscopy is performed using a handheld device called retinoscope, which is low-cost and lightweight. For children, periodic vision screening using retinoscopy helps in detecting amblyopia risk factors, as amblyopia generally develops from birth up to age 7 years [60]. Moreover, retinoscopy is used to detect scissors reflex for diagnosing keratoconus [2], a serious corneal disease.

Although versatile and accurate, retinoscopy has certain limitations. The first and most important is that it can only be performed by skilled eye specialists. There is an acute shortage of eye specialists in low-income countries [8], making retinoscopy inaccessible. Second, retinoscopy depends on the examiner's perception and is hence prone to inter-observer variability [41, 53]. Third, retinoscopy requires a lens kit of corrective lenses, increasing the cost and bulkiness. These factors limit the use of manual retinoscopy for mass screening efforts.

In this work, we propose an automated way of performing retinoscopy using a smartphone attached to the retinoscope, with the patient wearing a custom pair of paper frames. Our proposed digital approach has a few additional benefits compared to manual retinoscopy. It alleviates the need for a lens kit or skilled examiner, and removes subjectivity from the process, thus helping to democratize retinoscopy for large-scale screening. Moreover, digital retinoscopy allows pausing, (re)playing, and digitally zooming into the recorded retinoscopy video, enabling it to be used as a training device to teach retinoscopy. Finally, the digital record can be used to track the progress of abnormal retinal reflex movements.

To estimate eyes' refractive error, we extended retinoscopy's mathematical model (derived from the optical principles of streak retinoscopy [7]) to work with videos captured from a smartphone camera. Our proposed video processing pipeline automatically extracts the variables needed by the mathematical model and uses it to predict the net refractive error of the patient's eye. We evaluated the proposed video processing pipeline in real-world settings on 185 eyes, and achieved a mean absolute error of  $0.75 \pm 0.67\text{D}$  in net refractive power along the horizontal meridian ( $\text{net refractive error} = \text{spherical} + (\text{cylindrical} \times \sin(\text{axis})^2)$  [9]) compared to subjective refraction measurements. We also output a 4-class classification—normal (-1D to 1D), moderate myopia (-4D to -1D), high myopia ( $<=-4\text{D}$ ), and moderate hyperopia (1D to 4D)—based on the estimated net refractive error, and achieved a sensitivity of 91.0% and specificity of 74.0% on our dataset. We make the video analysis pipeline and paper frame details publicly available at our [project page](#)<sup>3</sup>.

The contributions of our work can be summarized as follows:

- A method to automate retinoscopy using a smartphone attached to a retinoscope, with the patient wearing a paper frame, thus alleviating dependency of retinoscopy on eye specialist and expensive lens kit.
- An extension of the retinoscopy mathematical model to estimate refractive error from retinoscopic beam position and retinal reflex position from smartphone-captured retinoscopy videos; a video processing

---

<sup>3</sup><https://www.microsoft.com/en-us/research/project/auto-retinoscopy-automating-retinoscopy-for-refractive-error-diagnosis/>

pipeline to automate the retinoscopy mathematical model, which takes retinoscopy video as input and provides net refractive error as output.

- An evaluation of the proposed approach on 185 distinct eyes from 128 patients in a real-world setting, compared with refractive error measurements using subjective refraction, autorefractor, and manual retinoscopy.

To the best of our knowledge, ours is the first work that automates the manual process of retinoscopy to enable low-cost, portable, and accurate refractive error estimation without a skilled examiner. We envision that our proposed method and findings will help the community for future research in this direction.

## 2 BACKGROUND

Refractive error comprises three components: spherical error, cylindrical error, and cylindrical axis. Under normal circumstances human eye is spherical in shape, thus refractive errors can be corrected solely using spherical lenses, which refract light in all directions. However, in an aspheric eye, called astigmatism, light is refracted differently at different meridians due to the misshapen cornea. In such cases, cylindrical lenses are prescribed, which are curved along the cylindrical axis and flat along the perpendicular axis. While the focus of a spherical lens is a single point, the focus of a cylinder lens is a line. The cylinder axis is expressed in degrees between 0 and 180. Most prescriptions have some combination of spherical and cylinder lenses. For the purpose of screening, we focus on computing net refractive error, which is a combination of spherical and cylindrical error, along a particular meridian. Here, we describe the principle of retinoscopy for computing spherical error.

In retinoscopy, the examiner shines a linear streak of light (referred to as the retinoscopic *beam*, Figure 1) and *scopes* (move along a particular axis) the beam on the patient's eye using a retinoscope device from a fixed distance (referred to as the *working distance*). The light beam passes through the different refractive surfaces of the eye, including cornea, anterior chamber, eye lens, and posterior chamber, to form an image (referred to as the *retinal reflex*) on the retina. After reflection from the retina, the light travels back through the same refractive surfaces in reverse order to form another image, which is viewed by the examiner through a peephole and a semi-silvered mirror in the head of the retinoscope. The examiner observes four characteristics of the reflex movement—size, brightness, direction, and speed—and tries to achieve *neutralization* by adding corrective lenses. The speed of reflex aids in identifying if the eye is close/far from neutralization [16]. If the examiner observes a slow-moving reflex, it means the eye is far from neutralization, and as corrective lenses are added, the reflex moves more rapidly. Similarly, with respect to size, a narrow streak of reflex means the eye is far from neutralization, and as corrective lenses are added, the reflex broadens, finally filling the pupil at the neutralization. With respect to brightness, large refractive errors form a dull reflex, while the reflex becomes brighter as the refractive error reduces. The direction of reflex movement is categorized as follows:

- **Against Movement:** In this case, the reflex moves in the opposite direction to the retinoscopic beam. The reflected rays from the retina meet at the far point<sup>4</sup> located between the retinoscope and the patient's eye, and the examiner uses negative lenses to correct this case of myopia.
- **With Movement:** Here, the reflex moves in the same direction as the retinoscopic beam. The far point is either located behind the retinoscope or a virtual far point is formed behind the patient's eye. It is corrected by using positive lenses.
- **Neutralization:** This is the end goal of retinoscopy, wherein the reflex completely fills the patient's pupil and is independent of the retinoscopic beam movement. At neutralization, the far point of the patient's eye is located at the peephole of the retinoscope. The corrective lens power and the working distance can then be used to determine the eye's refractive error.

---

<sup>4</sup>The far point of the eye is the maximum distance at which the eye can see the objects clearly. The far point of the normal human eye (emmetropic eye) lies at optical infinity (6m or more) in the relaxed state.

In retinoscopy, instead of determining the actual far point of the eye, neutralization shifts the far point to the known position of working distance. The power of the corrective lenses to achieve neutrality is termed *gross power* ( $P_{gross}$ ). The *net refractive power* ( $P_{net}$ ) of the eye can then be computed by adding a correction term for the known working distance ( $d$ ) as:  $P_{net} = P_{gross} - 1/d$ .

The examiner might perform *cycloplegic refraction*, in which eye drops are added to the patient's eye to relax the ciliary eye muscles in charge of focusing to reduce accommodation<sup>5</sup>. This causes the eye lens to be relaxed, making the eye more hyperopic. Hence a cycloplegic correction is applied, depending upon the used eye drops.

Most retinoscope devices provide two controls—a *streak rotator* to control the orientation of the beam (from 0-179°), and a *vergence sleeve* to adjust the streak from convergent (concave mirror) to divergent (plane mirror). To perform objective refraction, examiners mainly use diverging light (plane mirror). Orientation of the beam is used to identify the axis of the cylindrical power (in the case of astigmatism).

In this work, we capture retinoscopy videos by moving the vertical streak of the beam along the horizontal meridian of the patient's eye. The captured video is passed through a video processing pipeline, which estimates the working distance between the patient's eye and the smartphone, tracks the retinoscopic beam and the position of reflex to output the net refractive error of the eye along the horizontal meridian.

### 3 RELATED WORK

Several portable devices have been proposed over the last few decades for estimating refractive error using a variety of principles, such as self-subjective refraction, eccentric photorefraction, and wavefront aberrometry (Table 1). Our method, on the other hand, is based on the principle of retinoscopy. In this section, we focus on prior work related to refractive error estimation using other principles and demonstrate the relevance of our proposed automatic retinoscopy-based solution.

#### 3.1 Self-subjective Refraction

Subjective refraction, the gold standard for vision screening, relies on an eye specialist. To reduce this dependency, a few works tried estimating refractive error using the patient's judgment of the sharpness/blurriness of a test object, perceived via novel interactive methods. One of the earliest works, NETRA [47], utilized an array of microlens placed over a smartphone display and asked the user to interactively align patterns as seen from a pinhole plane on the display. Later, using the same principle, they introduced NETRA autorefractor, a \$1290 commercially-available device that gamifies the process of measuring refractive error via a series of interactions in a virtual-reality environment. A clinical study with 152 eyes found the NETRA device to estimate refractive error accurately (mean absolute error  $0.69 \pm 0.53$ D) [32]. However, a recent work [1] reported that using the NETRA autorefractor had a steep learning curve for healthcare practitioners and patients alike, and took roughly 10 minutes per patient.

Similarly, Folding Foropter (FoFo) [24] is a ~\$1 paper-based self-refraction tool that aids users in determining their refractive error. It consists of a pair of retractable paper tubes with a lens at each end. The user looks through the paper tubes at a distant target (3m) and adjusts the tube such that the target becomes sharp and in focus. A scale embedded on the sides of the tube depicts the spherical refractive error. USee [4] is another self-refraction tool. It comprises of a glass frame with an adjustable progressive lens system for each eye placed in front of a 3mm rectangular slit. An additional refraction bar is moved up and down to change optical power via a dial on each side of the frame until the characters in the logMAR chart are visible clearly. It can estimate spherical refractive power from -6D to +6D. Though FoFo and USee are affordable solutions, they have not been evaluated in a real-world setting. Moreover, the accuracy of self-subjective refraction devices relies on the patients' cooperation and feedback, making it challenging for young children, infants, and developmentally delayed adults.

---

<sup>5</sup>Accommodation refers to the ability of eye to change its focus from distant to near objects by changing the lens shape via ciliary muscles

Table 1. Summary of approaches for estimating refractive error based on different principles, in contrast to our retinoscopy-based approach. Note: For smartphone-based approaches, ‘Extra Hardware’ states hardware required for diagnosis apart from the smartphone; ‘Cost’ of the solution is categorized into: Low (<= \$500), Medium (\$500 – \$5000), and High (>= \$5000).

Principle	Prior work / Device	Phone	Extra hardware	Cost	Input	Output	Dataset
Self-subjective Refraction	Jeganathan et al. [32]	Yes	Eye-Netra AR	Medium	User feedback	Sph, Cyl, Axis	152 eyes
	USee [4]	No	Lens kit	Low	User feedback	Sph	120 eyes
	FoFo [24]	No	-	Low	User feedback	Sph	-
Eccentric Photorefraction	SPOT [33]	No	-	High	Red reflex image	Sph, Cyl, Axis	134 eyes
	Plusoptix A09 [50]	No	Computer	Medium	Red reflex video	Sph, Cyl, Axis	64 eyes
	Yang et al. [61]	Yes	Trial frame	Low	Red reflex image	Sph	165 eyes
	Fu et al. [25]	Yes	-	Low	Red reflex image	Sph	172 eyes
Wavefront Aberrometry	QuickSee [21]	No	-	High	Wavefront images	Sph, Cyl, Axis	82 eyes
	Ciuffreda et al. [15]	Yes	SVOne AR [62]	High	Wavefront images	Sph, Cyl, Axis	50 people
Retinoscopy	Ours	Yes	Retinoscope Paper frame	Low	Retinoscopic video	Net refr power	185 eyes

### 3.2 Eccentric Photorefraction

Several objective refraction devices based on the principle of eccentric photorefraction [59] have also been proposed. The solutions vary from expensive commercial devices with sophisticated hardware (like Plusoptix A09, SPOT vision screener) to low-cost smartphone-based approaches [25, 61]. These devices estimate refractive error as a function of the size and tilt of the crescent formed on the patient’s pupil when illuminated with (infrared/visible) light from a known working distance. The Plusoptix<sup>6</sup> and SPOT vision screener<sup>7</sup> have been clinically evaluated and are used to detect myopia, hyperopia, astigmatism, and amblyopia in the pediatric age group [33, 57]. Recently, researchers have utilized image processing and machine learning techniques in conjunction with eccentric photorefraction to estimate refractive error from eye images captured by a smartphone. Yang et al. [61] proposed a data-driven approach to estimate refractive error using SVR trained on handcrafted features (like iris radius, pupil radius, and retinal reflex width). In a follow-up work, Fu et al. [25] proposed a CNN-based solution to predict refractive error using eye images captured by a smartphone.

The unique advantage of these solutions is that they require minimal cooperation from the patient, making them ideal for vision screening in schools and pediatric hospitals. However, the eccentric photorefraction principle has a few inherent limitations, limiting its widespread adoption and questioning the reliability of devices based on this principle. First, eccentric photorefraction based devices have been found to be inaccurate in predicting refractive power of specific ethnic groups (especially Caucasian, East African, African, and Indian) [55] due to variations in defocus calibration factor. Second, such devices accuracy is dependent on illumination distribution around the tear film, which in turn is dependent on the time when the eye image was clicked after a blink. Chen et al. [13] reported significant differences in refractive error estimation across multiple eccentric photorefraction images captured at different time lag after a blink. Third, the eccentric photorefraction principle’s accuracy and operating range rely on the distance between the camera and the light source [52], suggesting that the accuracy of smartphone-based approaches [25, 61] is a function of smartphone-specific parameters.

### 3.3 Wavefront Aberrometry

Devices based on the wavefront aberrometry principle project a wavefront of light on the patient’s pupil and record the reflected light from the retina as it passes through the different refractive components of the eye,

<sup>6</sup>Plusoptix: <https://www.plusoptix.com/en-us/products>

<sup>7</sup>Welch Allyn Spot Vision Screener: <https://www.hillrom.com/en/products/spot-vision-screener/>

including the cornea and the eye lens. By analyzing the distortions in that reflected wavefront, a topography map of the eye is created using Zernike decomposition. QuickSee [21] by PlenOptika is a binocular handheld autorefractor that uses this principle. Zhou [62] leveraged the same principle to create SVOne autorefractor. It uses an expensive hardware module to direct light onto the patient's eye, and a smartphone's camera to capture the reflected light. Apart from estimating refractive error accurately, devices based on wavefront aberrometry can also measure higher-order aberrations in the reflected wavefronts, potentially useful for estimating spherical aberrations and secondary astigmatism. However, these devices are expensive, restricting their access in low- and middle-income countries.

To summarize, commercial vision screeners (like Plusoptix, SPOT, QuickSee, Eye Netra, and SVOne AR) though clinically evaluated, are expensive and require a minimum skill level to obtain reliable measurements. On the other hand, low-cost solutions [4, 24, 25, 61] have limited evaluation with respect to refractive error range and/or age group. Most importantly, none of the proposed solutions rely on the retinoscopy principle, which requires minimal patient cooperation, and is considered the best starting point for subjective refraction [34]. Perhaps, the only digital retinoscopy prior work is by Chan et al. [12]. They created a digital retinoscope by attaching a standard retinoscope to a smartphone camera, primarily to teach beginners retinoscopy. However, they neither propose a method to estimate refractive error from the recorded videos nor conduct any evaluation. Therefore, we believe that our current work is the first step towards automating retinoscopy.

## 4 DATA COLLECTION

To evaluate our proposed mathematical model and video processing pipeline, retinoscopy videos were captured using a Pixel 4A android smartphone with a Heine Beta 200 streak<sup>8</sup> retinoscope attached to the smartphone camera at the peephole, with the patient wearing a paper frame (Figure 1a). This allows us to record the beam and reflex movement as observed by the examiner during retinoscopy. Patient volunteers with different refractive powers were recruited from a local eye hospital. Below we describe our data collection procedure, including the paper frame design and data collection app. Next, we elaborate on the diversity of the participant pool and the dataset used for evaluation. All details of our study were approved by the hospital's Institutional Review Board (IRB).

### 4.1 Data Collection Procedure

The data was collected from Sept-Dec 2021 in collaboration with a local eye hospital. The eye hospital is a leading eye care and teaching institution in the country, with more than 15 eye doctors and 10 optometrists. For each patient who agreed to participate in the study, we collected subjective refraction data for both the eyes. Additionally, we collected two sets of medical-grade refractive error measurements, using autorefractor (Topcon KR 8900<sup>9</sup>) and manual retinoscopy, whenever possible. The data was collected by a group of staff optometrists (25-40 years age, 1-10 years of experience). Retinoscopy video data from our setup was collected by two student interns (male, 21, and 23 years of age), who were trained by one of the researchers to use the data collection app.

To collect video data, patients were asked to wear a comfortable size of our paper frame. They were seated on a chair with their chin placed on an adjustable chin-rest and were instructed to focus on the LogMAR chart placed at a 3m distance along the eye's optical axis for relaxing the eye to minimize accommodation. After entering the patient's demographic information (gender, age) and patient ID on the data collection app, the app prompts the data collector to record the video for the left eye, followed by the right eye. For each eye, the data collector needs to rotate the retinoscope attached to the smartphone such that the beam moves from one end of the paper frame

<sup>8</sup>[https://www.heine.com/fileadmin/accessfiles/en\\_GB/download/HEINEBETA200LEDRetinoscopeInstructionforUse.pdf](https://www.heine.com/fileadmin/accessfiles/en_GB/download/HEINEBETA200LEDRetinoscopeInstructionforUse.pdf)

<sup>9</sup><https://topconhealthcare.com/product-category/auto-refractometer>

to the center and back. This is referred to as a *pass*. This process is repeated to obtain four passes (2 left-to-right and 2 right-to-left) for a single eye in a video. The retinoscope-attached smartphone was mounted on a tripod stand to minimize camera jitter while panning the retinoscopic beam across the eye. Several instructions were given to the data collector—vergence sleeve should be at the lowest point to ensure plane mirror effect; to adjust chin-rest height such that the patient’s eye is at the same level as the LogMAR chart; to pan the device slowly from left-to-right and back in order to capture more video frames; ensure that the beam is vertical ( $90^\circ$ ) using the streak rotator, etc.

The hospital staff helped us in recruiting patients as participants. Patients with cataracts or active signs of ocular infection/conjunctivitis/acute eye trauma were excluded from our evaluation. Since retinoscopy-based measurements are possible only when the reflex is visible in the eye, patients with dull reflexes due to corneal haze or very high myopia and small pupil size (less than 3mm) were also excluded. During data collection the participants were given an IRB-approved consent form to read and sign before their data was collected. Every evening, the data collector connected the phone to the hospital’s WiFi network to auto-sync the collected videos along with the metadata to a cloud-based storage. Refractive error readings from autorefractor, retinoscopy and subjective refraction, were recorded from the hospital’s database by matching the patient’s ID saved during data collection. Note: We use the gold standard subjective refraction as the ground truth. Participants were not paid to participate in the study. It took 5-7 minutes per participant for video data collection from our app.

## 4.2 Paper Frames

Computation of refractive error using the retinoscopy mathematical model (Section 5) needs the position of beam and reflex w.r.t a *fixed* world co-ordinate system. However, since in our setup the video is captured via a moving camera, we need a fixed target that can be used as a reference point. For this, we use a custom pair of paper frames (without any lens) with a unique pattern of known dimensions (Figure 2) as a robust and low-cost solution. The glasses also allow us to compute the working distance between the smartphone camera and eye accurately, which is again needed by the retinoscopy mathematical model. A single size of glasses for patients with different facial sizes posed a challenge, thus we created three paper frame sizes with varying reflex search space dimensions (width  $\times$  height)—(1) small:  $3.0\text{cm} \times 2.0\text{cm}$ , (2) medium:  $3.5\text{cm} \times 2.0\text{cm}$ , (3) large:  $4.0\text{cm} \times 2.0\text{cm}$ . We further number-coded them for easy identification. We tried a variety of unique patterns to estimate the working distance accurately and track the retinoscopic beam (which is falling partially on the top part of the paper frame). Here we describe three patterns and the issues we identified.

**4.2.1 Colored Stripes.** Inspired by [40], we tried colored stripes of fixed width along the edges of the frame (Figure 2a). For processing the video frames, we leveraged color thresholding and known strip width to uniquely identify and segment the colored stripes. However, the fixed color threshold values failed to detect the colored stripes robustly in varying ambient lighting of the room.

**4.2.2 BW Stripes with Square Fiducials.** Instead of multiple colored stripes, we switched to black-and-white stripes (Figure 2b) as it needed only a single color threshold value. We also added black squares of fixed known size at the corners (similar to QR code corners) as fiducial markers. The size of fiducials helped us to estimate the working distance, and the stripes helped us to track the retinoscopic beam. However, the processing failed if even a single fiducial marker among the three was undetected due to glare, cropping, or any error. Moreover, the beam light intensity interfered with the stripes, resulting in erroneous beam edge detection.

**4.2.3 Square Fiducials.** We removed the black-and-white stripes and defined the empty space between fiducials as the search space for beam edges (Figure 2c). We added two more fiducials along the vertical sides of the frame to ensure that the video processing is robust against missing fiducial(s). The fiducials were placed such that the relative position of any fiducial w.r.t any other fiducial is unique. During the video processing, this allowed us to

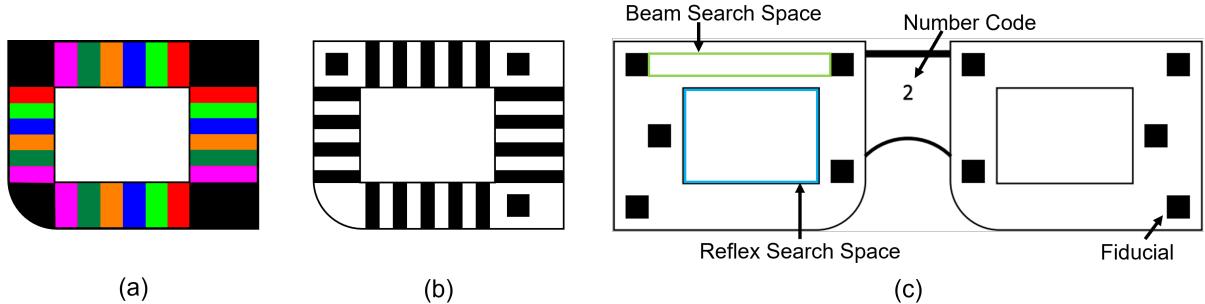


Fig. 2. Different paper frame designs we tested: (a) frame with colored stripes, (b) frame with black-and-white stripes and square fiducials, and (c) the final frame design with five fiducials.

uniquely identify the detected fiducials if any two (out of 5 for each eye) fiducials were detected. This made the frame localization step of our video processing pipeline very robust.

#### 4.3 Data Collection Parameters

During the initial stage of data collection, we observed that external factors (like room brightness, working distance), video recording settings (like FPS, video resolution), and retinoscope settings (like vergence sleeve position) significantly affected the captured video quality. Hence, we followed an iterative approach to identify their impact on our video processing pipeline and recommend the most suitable setting.

**4.3.1 Room brightness.** Generally, retinoscopy is performed in a dark setting, as it increases the pupil size to allow more light to enter the eye, enabling the examiner to observe a clear reflex movement. Hence we followed a similar dark room setting for capturing retinoscopy videos. However, videos in the low-light setting are grainy, and at times, all the five fiducials were not visible. We solve these problems via our paper frame design and by leveraging image processing techniques like median filtering for noise reduction and adaptive histogram equalization for contrast adjustment. Moreover, capturing videos with no light source in the room helped us avoid unwanted reflections in the eye.

**4.3.2 Working distance.** Retinoscopy is performed with the examiner sitting at an arm's length (50 cm to 1 m) away from the patient. This allows the examiner to quickly add/remove corrective lenses on the trial frame. As our proposed approach does not require external lenses, we used a smaller working distance (approximately 30-40 cm) to obtain a larger resolution in our region of interest (i.e., pupil). Its impact on our proposed mathematical model and results has been discussed in Section 7.2.

**4.3.3 FPS and Video resolution.** We had a few options for frames/second (30, 60, or 120 fps). Although higher fps provided more precise tracking of beam and reflex movements, it resulted in noisier video in low-light settings as exposure time decreased with increasing fps. Hence we selected 30 fps for recording the videos. For video resolution, we selected the highest available option ( $3840 \times 2160$ ), as an increase in video resolution resulted in capturing more pixels in our region of interest, which helped us in more accurately detecting the beam and reflex position.

**4.3.4 Vergence sleeve position.** In retinoscope device, the position of the light source is fixed, but the position of the condensing lens inside the retinoscope is controlled by the examiner using the vergence sleeve [16] (Figure 1c). By changing the sleeve position, the retinoscopic beam vergence is altered, thus manipulating its focal point. In

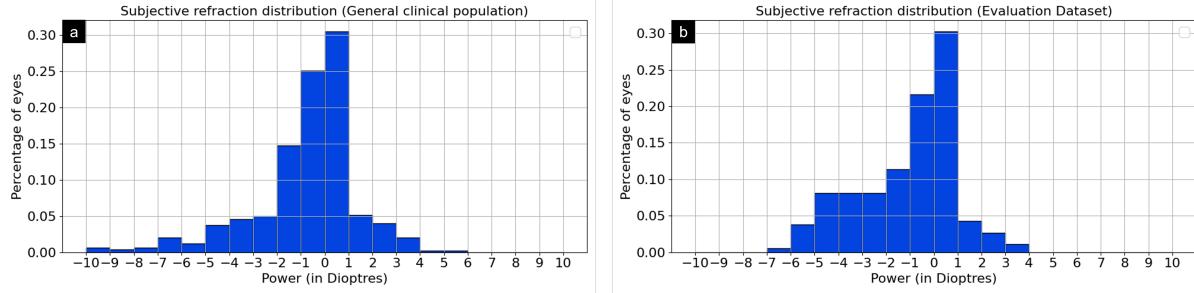


Fig. 3. Normalized distribution of the percentage of refractive power (in Dioptries) for (a) historical data collected at the local eye hospital, and (b) our dataset. Both show a similar distribution with majority of samples (~95%) lying between -6D to 4D.

the sleeve up position, the streak emanates as a converging beam (concave mirror effect), and the beam is focused and very bright, thus making it difficult for the video processing pipeline to track the retinal reflex movement. When the retinoscopy sleeve is lowered, the streak emanates as a diverging beam (plane mirror effect), which the examiners use most often. We also recorded our videos with the plane mirror setting.

#### 4.4 Data Collection App

To collect data, we developed an Android application wherein, for each patient, the data collector enters the age, sex, and hospital-assigned patient ID. The patient ID is the only link between the captured video and the medical-grade refractive error measurements (using autorefractor, retinoscopy, and subjective refraction methods) of the patient. Next, the data collector has an option to choose data collection parameters, such as working distance, vergence sleeve position, fps, and video resolution. Our recommended values for these parameters are selected by default. The app also allows the data collector to select the eye drops used in case of cycloplegic refraction. This helps our refractive error estimation algorithm to compensate for the accommodation due to the cycloplegic effect. Next, the app prompts the data collector to start recording video for the left eye. After recording, the app provides an option to recapture the left eye, start recording video for the right eye, or end the session. After completing the session, the app asks the data collector to enter refractive error measured using an autorefractor, retinoscopy, and/or subjective refraction method for both eyes. For each patient, the app stores a metadata file containing the patient's demographic information, data collection parameters, medical-grade refractive error, etc. The collected data is auto-synced to cloud-based storage in a best-effort manner. The videos were recorded using a Google Pixel 4A smartphone, utilizing its primary camera (12MP, f/1.7 aperture, 4.4mm focal length, 5.6mm × 4.2mm sensor size).

#### 4.5 Participants and Dataset

128 patients (74 male, 54 female) with an average age of  $30.2 \pm 11.1$  years (7-58 years age range) volunteered for the data collection. (Note: Due to COVID-19, we were able to recruit very few patients, resulting in a relatively small dataset). Out of the 256 eye videos collected (128\*2), 71 videos were removed due to these factors: (1) retinoscope device movement was very fast, (2) out-of-focus/blurry video, (3) patient was blinking throughout the session, and (4) very small pupil size (happened mostly with old-age patients). Our final dataset comprises 185 distinct eyes (12 dilated and 173 non-dilated eyes), for which we have both smartphone-captured retinoscopic video data and ground truth subjective refraction measurements. The subjective net refractive error along the horizontal meridian ranged from -6.25D to 3.23D (Figure 3b shows normalized distribution), with 74 myopic (<-1D), 15

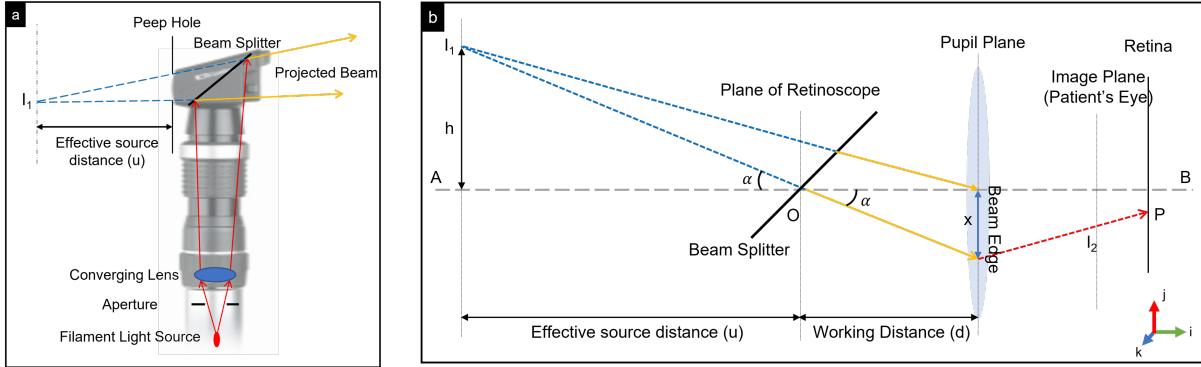


Fig. 4. Retinoscopy working principle. (a) Different parts of a retinoscope device, and (b) ray-diagram for a case when the retinoscope is rotated about  $\hat{k}$  axis by an angle  $\alpha$ .

hyperopic ( $>1D$ ), and 96 normal (-1D to 1D) eyes. Among these 185 eyes, we have autorefractor measurement data for 161 eyes and retinoscopy measurement data for 130 eyes.

As we found our dataset refractive error distribution to be skewed towards myopic eyes, we also collected all patients data (autorefractor, retinoscopy, and subjective refraction measurements for right eye) for 2 days from the hospital's database, to compare the refractive error distribution of our dataset with the general clinical population. In 2-days, 506 patients (255 male, 251 female) with mean age  $40.7 \pm 21.7$  years had undergone subjective refraction, and the mean spherical power was  $-0.42 \pm 2.58D$ . Comparing the normalized distribution of subjective power for the general clinical population (Figure 3a) with our dataset (Figure 3b), we found the Kullback-Leibler divergence [35] to be 0.1, thus it is evident that our dataset is a good representation of refractive error among the clinical population.

## 5 OPTICS BEHIND RETINOSCOPY

In retinoscopy, the examiner observes the speed, direction, size, and brightness of the retinal reflex to achieve neutralization. In this section, we describe the optical principle behind manual retinoscopy, its mathematical formulation only using speed and direction of beam and reflex [7] (note: the formulation encapsulates information provided by the reflex size and brightness [7]), and extend it to estimate refractive error from smartphone-captured retinoscopy video.

### 5.1 Mathematical Modelling

In streak retinoscopy, light from a linear filament bulb passes through a small aperture via a strong converging lens (Figure 4a). (Note: The bulb and lens are located in the handle of the retinoscope.) The position of the lens with respect to the bulb is controlled by the examiner using the *vergence sleeve*, which changes the vergence of the projected beam. A beam splitter at a  $45^\circ$  angle present at the top of the retinoscope, changes the direction of light by  $90^\circ$  towards the patient's eye. Extrapolating these rays backwards produces a virtual image ( $I_1$ ) which acts as a light source for the patient (Figure 4a). However, only the rays travelling from  $I_1$  through the peephole form a beam and reach the patient's eye. We refer to the distance between  $I_1$  and retinoscope as the effective source distance ( $u$ ; which is a retinoscope specific parameter). We captured videos using a Heine Beta 200 retinoscope at the lowest sleeve position (i.e.,  $u = 40cm$ ).

During retinoscopy, an examiner rotates the retinoscope about the  $\hat{k}$  axis from a fixed working distance  $d$  (Figure 4b), where working distance is defined as the distance between the apex of the eye and the camera. The retinoscope's rotation moves the beam along the  $\hat{j}$  vertical axis on the pupil plane, sweeping a small angle with the optical axis  $AB$ , and shifting the virtual light source vertically. For example, Figure 4b shows the beam at an angle  $\alpha$ , and the virtual light source ( $I_1$ ) shifted by a distance  $h$  above the optical axis  $AB$ . The rays from the virtual light source  $I_1$  exit the peephole of the retinoscope in the form of a beam, and enter the patient's eye through the pupil at a distance  $x$  from the optical axis  $AB$ . After entering the eye, light undergoes multiple refractions through the cornea, anterior chamber, eye lens and the posterior chamber. For simplicity, we assume all those refractions to take place at the *pupil plane*. These rays then form an image  $I_2$  (Figure 4b) on a plane, whose position is dependent on the refractive power of the patient. Note: For emmetropic eye (i.e., eye without any refraction error), the image plane will coincide with the retina; for myopic and hyperopic eye, the image plane will be before and after retina, respectively<sup>10</sup>. However, since  $I_2$  is not directly visible to the examiner, it cannot be used to estimate the refractive error. The rays finally illuminate a region on the retina at point P.

After the diffused reflection from retina, point P acts as a light source (Figure 5a), and rays emanating from P undergo multiple refractions to emerge out of the pupil. A subset of these rays, at distance  $y$  from optical axis  $AB$ , pass through the peephole of the retinoscope into the examiner's eye, creating a perception of the presence of a bright region in patient's pupil (referred as the reflex). These emergent rays finally meet and form an image  $I_3$  at the *far point*, at a distance  $f$  from the pupil plane. For illustration purposes we only show the central ray emerging from the light source and subsequent images.

Based on the similar triangles in Figure 4b:

$$\frac{h}{x} = \frac{u}{d} \quad (1)$$

Similarly in Figure 5a:

$$\frac{z}{h} = \frac{f}{u+d} \quad (2)$$

$$\frac{y}{z} = \frac{d}{f-d} \quad (3)$$

Multiplying these 3 equations we get:

$$\frac{y}{x} = \frac{u * f}{(f-d) * (u+d)} \quad (4)$$

In equation 4, the values  $x$  and  $y$  represent the position of beam and the reflex, respectively, in the 3D world co-ordinate system. A version of this formulation has been proposed earlier in [7]. To extend this formulation for refractive error estimation using smartphone captured videos, we further model the formation of the reflex and beam's image on the smartphone camera image plane. Figure 5b shows the image formation on a camera of focal length  $F_c$ , where the reflex at a distance  $y$  from the optical axis  $AB$  forms an image at  $y_p$  on the camera's image plane. The beam at position  $x$ , will similarly form a corresponding camera image at  $x_p$ . Again, from similar triangles:

$$\frac{x}{d} = \frac{x_p}{F_c}; \frac{y}{d} = \frac{y_p}{F_c} \quad (5)$$

<sup>10</sup>This is for the *gross power*,  $P_{gross}$  before adding the working distance correction.

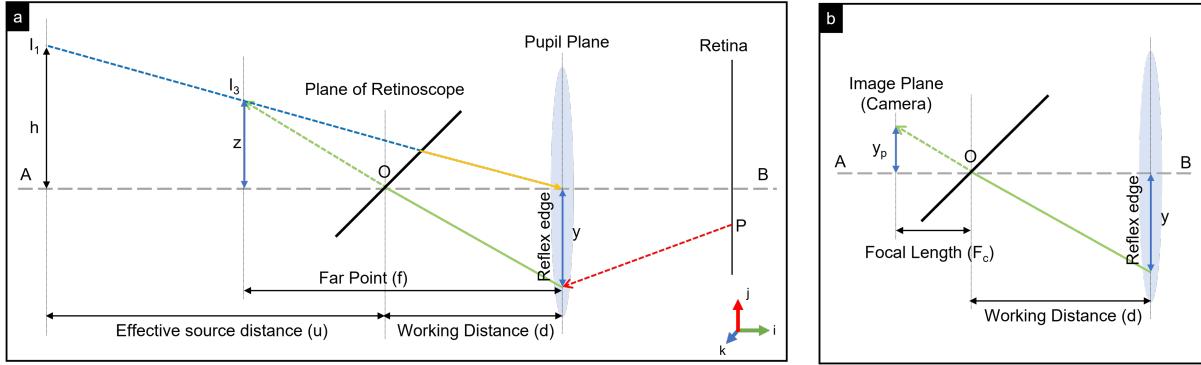


Fig. 5. Mathematical modelling of the retinoscopy principle to estimate the refractive power of a given eye  $P_{net}$  as a function of effective light sources distance  $u$ , working distance  $d$ , and the ratio  $y_p/x_p$  ( $= r$ ), where  $y_p$  is the image location of reflex and  $x_p$  is the image location of beam on the camera sensor.

Substituting  $y/x$  from equation 5 in equation 4:

$$\frac{y_p}{x_p} = \frac{u * f}{(f - d) * (u + d)} \quad (6)$$

$$y_p * (f - d) * (u + d) = u * f * x_p \quad (7)$$

Note that in this formulation  $x_p$  and  $y_p$  are measured from the optical axis (AB) of the pupil, which is not known. To get around this, we compute  $y_p$  and  $x_p$  at two different timestamps  $t_1$  and  $t_2$ , and take their difference. In equation 7,  $u$  and  $d$  are fixed for a patient during a retinoscopy session, and so is the far point distance  $f$ , as it is a property of the patient's eye. Thus, computing equation 7 at times  $t_1$  and  $t_2$ , and take the difference yields:

$$[(y_p)_{t_2} - (y_p)_{t_1}] * (f - d) * (u + d) = u * f * [(x_p)_{t_2} - (x_p)_{t_1}] \quad (8)$$

$$\Delta y_p * (f - d) * (u + d) = u * f * \Delta x_p, \quad (9)$$

where  $\Delta y_p = (y_p)_{t_2} - (y_p)_{t_1}$  and  $\Delta x_p = (x_p)_{t_2} - (x_p)_{t_1}$  are referred to as the displacement of the reflex and the beam between timestamps  $t_1$  and  $t_2$ .

Next, we estimate the refractive error in terms of the far point distance  $f$ . Let the eye's optical system be represented by a single lens at the pupil plane with the retina at a distance  $R$  from the pupil plane. For an emmetropic eye (i.e., eye without any refraction error), the eye should have a refractive power  $P_e$ , such that rays from infinity focus on the retina. Applying lens equation for this condition yields:

$$P_e = \frac{1}{\infty} - \frac{1}{R} = -\frac{1}{R} \quad (10)$$

Now, if the retinoscopic reflex is formed at a far point distance  $f$ , it means that the eye has a refractive power,  $P_u$ , such that the image of point P at distance  $R$  is formed at  $f$  (Figure 5a). Applying the lens equation for this condition now yields:

$$P_u = \frac{1}{f} - \frac{1}{R} \quad (11)$$

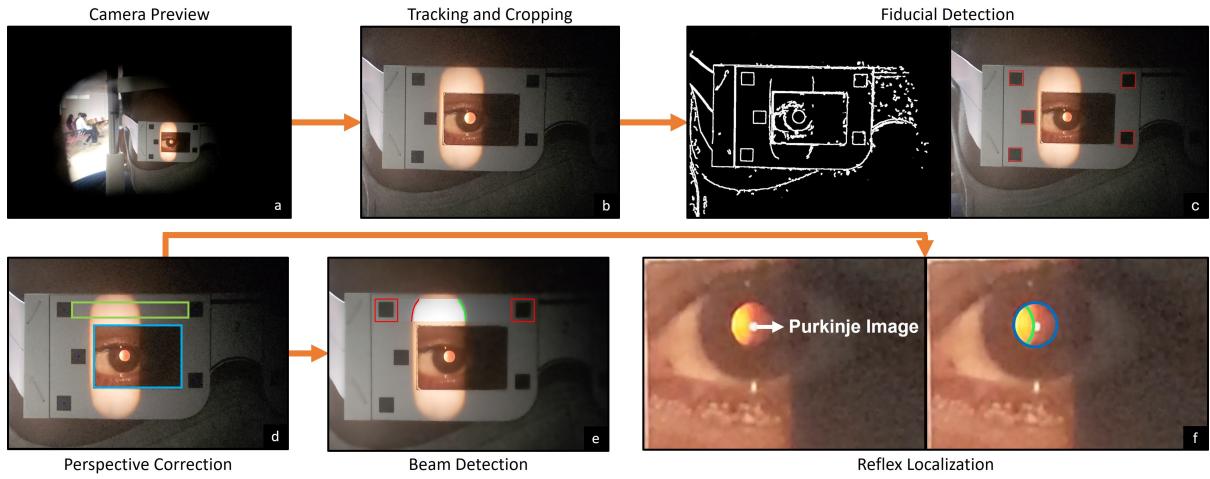


Fig. 6. Our proposed video processing pipeline: (a) captured video frame using smartphone camera, (b) frame is initialized with a tracker and cropped to get region of interest, (c) fiducial detection on the edge map, (d) perspective correction via homography estimation, (e) detection of retinoscopic beam edges, and (f) pupil detection and reflex edge detection.

thus the power of the corrective lens,  $P_{net}$ , needed to fix such an eye should be such that  $P_e = P_u + P_{net}$ . Thus  $P_{net} = P_e - P_u = -\frac{1}{f}$ . Finally, substituting  $f$  from equation 9 and replacing  $\Delta y_p / \Delta x_p$  as  $r$ , we obtain  $P_{net}$  as:

$$P_{net} = -\frac{1}{f} = \frac{u - (u + d) * r}{r * d * (u + d)} \quad (12)$$

At this point  $P_{net}$  represents the net refractive power of the eye along the vertical meridian  $\hat{j}$ . To obtain spherical, cylindrical power and cylindrical axis of the lens, we can repeat this process along two more predetermined axis and use the results from [9].

## 6 PROPOSED SOLUTION: VIDEO ANALYSIS PIPELINE

Our video analysis pipeline takes as input the recorded video and the frame size (small / medium / large) used to capture the video, and outputs the net refractive power of the eye along the scoped meridian (i.e., the meridian along which the retinoscopic beam was scoped). For estimating refractive power, four data points are needed: (1) camera parameters, (2) pixel location of the fiducial centers, (3) pixels moved by the retinoscopic beam between timestamps  $t_1$  and  $t_2$ , and (4) pixels moved by the retinal reflex between timestamps  $t_1$  and  $t_2$ . Only the camera parameters are known from the smartphone specifications. Below we describe the different components of our proposed pipeline to compute the remaining variables. Note: As the video frames are very noisy, we use Otsu's thresholding on different regions of the frame—beam search space, reflex search space, etc.—as Otsu is dynamic and adaptive, and hence is generalizable across video quality and lighting conditions.

### 6.1 Image Cropping and Tracking

The first step in our processing pipeline is to define the region of interest (ROI) in each frame of the video. Since the recorded videos are at a  $3180 \times 2160$  pixels (width  $\times$  height) resolution, a significant portion of the frame comprises of the background (Figure 6a), which needs to be removed for subsequent processing. Using the fiducials, we detect the paper frame and form a bounding box around it as our ROI (Figure 6b). The ROI is

detected on a single frame, and tracked on successive frames using CSRT tracker [39]. In the following steps, we use this cropped ROI ( $I_{ROI}$ ) for each frame as input.

## 6.2 Fiducial Detection

In the cropped region,  $I_{ROI}$ , the location and size of fiducial squares need to be detected accurately to estimate the working distance. Moreover, location of fiducials also define the search space for beam and reflex detection. As the videos were collected in a dark room, there is uneven distribution of light on the paper frame (e.g., the region of the paper frame closer to the retinoscopic beam is more illuminated than other regions) which makes accurate detection of fiducials non-trivial. We perform three preprocessing steps: (1) convert  $I_{ROI}$  to gray-scale based on CCIR-601 standards to obtain  $I_G$ , (2) noise reduction on  $I_G$  using median filtering with a  $5 \times 5$  kernel, and (3) adaptive histogram equalization [64]. Next, we obtain an edge map of  $I_G$  using Canny edge detector [11]. As the square shaped fiducials were captured with a camera panning horizontally, in every frame, a few fiducials appeared as quadrilaterals. So, we identify contours and detect shapes having 4 corners in the edge map. As the fiducial markers would be captured as roughly of same size, we remove quadrilaterals whose area is significantly greater/lesser than median area of the detected quadrilaterals.

To further remove any falsely detected fiducial(s), we leverage the unique fiducial pattern and its dimension to find a fixed ordering among the fiducials. We start with the top-left detected quadrilateral in the frame, and try to fit a set of other detected quadrilaterals using the known fiducial pattern dimension. This step is repeated, starting with each of the identified quadrilaterals, to find the best fit. If the set of identified fiducials is less than five, the locations of the unidentified fiducial(s) is interpolated/extrapolated by tracking their location in consecutive frames. Due to the low resolution of fiducials ( $\approx 40 \times 40\text{px}$ ), we found aliasing artifacts on the fiducial edges. Hence, we estimate the center of detected fiducials using intensity-based weighted averaging on the negative of  $I_G$  (i.e.,  $255 - I_G$ ).

## 6.3 Perspective Correction

Note that during retinoscopy, participants are asked to look at a logMAR chart placed at a distance of 3m to relax the eye and minimize accommodation. We followed a similar guideline for our data collection. To avoid blocking patient's view of the logMAR chart, we captured our data by keeping the retinoscope at an angle of  $\sim 10\text{-}15^\circ$  from the patient's line of sight. This resulted in a perspective distortion in the recorded frames. In this step, we remove this distortion in each frame by homography correction using the detected fiducials. The fiducial markers on the paper frame lie on a 2D plane in the 3D world. Hence we define a  $3 \times 3$  homography matrix ( $H$ ) which maps the points from the captured image ( $I_{ROI}$ ) to a fixed known 2D plane [29]. We compute this  $H$  matrix by mapping the centers of the five detected fiducials in  $I_{ROI}$  to the centers of the known fiducial pattern dimension on the paper frame, using a least square method. Finally, we apply the transformation  $H$  to the image  $I_{ROI}$  to generate a distortion free image with the correct perspective.

## 6.4 Beam Detection

We detect the retinoscopic beam edges in the beam search space (Figure 2) between the top two fiducials. Since the videos are recorded in a low-light setting, we used bilateral filtering on  $I_{ROI}$  to reduce noise while preserving the beam edges. The beam is orange in color, hence after cropping the beam search space, we perform non-local mean denoising [10] on the red channel of  $I_{ROI}$ .

There are two possible beam positions. First, the edge is not present in the beam search space. We remove such frames from analysis by comparing the threshold value obtained using Otsu thresholding algorithm [45] with the empirically found mean intensity of the beam. Second, the beam is partially/fully present in the beam search space. In that case, the search space can be divided into two intensity-based clusters—beam and background. To

identify the beam edges, we use Canny edge detector [11] with two threshold values—Otsu threshold and median intensity of the image as threshold—and compute a bitwise OR operation between their outputs. Finally, we label the beam edges as left/right based on the sign of the horizontal gradients at these edges. The gradient value is positive for the left edge as we are moving from low to high pixel intensity and negative for the right edge.

## 6.5 Pupil Detection

The search space for pupil detection is the empty rectangular area in the interior of the paper frame. We use Hough transform [31] on the edge map of the cropped reflex search space ( $I_{eye}$ ) to detect the circular pupil. The reflected light from the retina illuminates eye's blood vessels making it appear reddish-orange in color, thus we use the red channel in the RGB color space for pupil detection in the Hough transformed image.

In some frames, Hough transform fails and outputs false positives. Since the patient focuses on the LogMAR chart during data collection, the position of the pupil with respect to the fiducial pattern remains approximately fixed. Therefore, we create a 2D histogram plot of the x,y coordinates of the Hough-detected circle centers with the top-left fiducial as the origin. The peak in this histogram is selected as the pupil coordinates for that frame.

## 6.6 Reflex Edge Localization

The region of interest for reflex detection is the pupil. Finding the accurate location of the reflex edges (in  $I_{eye}$ ) in each frame is the very challenging, because of the small resolution of the pupil (typically  $\sim 15$  pixels). As a result, even a single pixel error in reflex edge localization significantly impacts our results. Therefore, we super-resolve  $I_{eye}$  to 4X using FSRCNN (Fast Super-Resolution CNN) [20], which helps us to obtain a sub-pixel resolution. Next, we perform Otsu adaptive thresholding [45] on the super-resolved  $I_{eye}$  to localize one/two boundaries of the retinal reflex in each frame. One of the challenges in the reflex edge localization is the presence of reflections from different surfaces of the patient's eye, referred to as Purkinje images [17] (Figure 6f). Our video data contains Purkinje image formed by the reflections from the outer surface of cornea, which appears as a bright region, and looks similar to the retinal reflex. To remove Purkinje reflections in  $I_{eye}$ , we traverse the binary mask indicating the reflex / not-reflex pixels outputted by the Otsu thresholding, and for each column count the percentage of pixels identified as reflex. Since the Purkinje reflections are typically much smaller than the reflex, we can discard columns with have only a small percentage of pixels identified as reflex (we used a threshold of 40% obtained empirically).

Finally, we label the boundaries as left/right edge based on the sign of the gradients along horizontal axis. Furthermore, we use a gradient-based weighted averaging, where the kernel is a gradient along horizontal axis at that pixel, thus increasing the accuracy of reflex edge localization.

## 6.7 Refractive Error Estimation

After processing all the frames in a video, we obtained five fiducial centers, left and right edge of the beam, pupil center and radius, and left and right edge of the reflex, for each frame. To estimate the net refractive error of the eye according to our formulation 12, we calculate the distances moved by retinal reflex and the retinoscopic beam between a given time interval. However, since the smartphone camera is moving across frames, we need a static object as the point of reference for calculating the distance moved. We use the center of the fiducial markers as the origin since its position does not change in the video.

Next, we want to select the timestamps  $t_1$  and  $t_2$  between which to calculate the distance of retinal reflex and beam. We aim to choose frames with reflex edges clearly visible and which are farthest apart in time. Choosing frames farthest apart in time gives more robust estimates, as any localization errors in reflex / beam edge become smaller *relative* to the overall motion. Note, however that since our formulation assumes standard optics assumptions (perfect lens, small angles, etc.), we want to stay near the center of the pupil. Thus, we restrict

Table 2. Sensitivity and Specificity for the 4-class refractive error screening

Class	Power Range	# of eyes	Sensitivity	Specificity
High Hyperopia	Above 4D	3	—	—
Moderate Hyperopia	1D to 4D	15	60.0 %	98.2 %
Normal	-1D to 1D	96	74.0 %	91.0 %
Moderate Myopia	-4D to -1D	51	82.4 %	80.6 %
High Myopia	Below -4D	23	82.6 %	96.3 %

our search space to only the central 50% of the pupil. We proceed by fitting a linear curve [46] on the position of reflex with respect to time. Then we select the start timestamp  $t_1$  closest to the frame where the left/right edge of the retinal reflex covers at least 25% of the pupil diameter, and the end timestamp  $t_2$  closest to the frame where the same edge covers just beyond 75% of the pupil diameter.

Next, to calculate the distance travelled by a reflex edge between the selected frames, we draw a line along the scoped meridian and measure the movement along this line. For robustness, we do this computation along multiple such lines and take the median distance. We do a similar computation for the retinoscopic beam. Finally, this is substituted into equation 12 to obtain the refractive error.

## 7 RESULTS

To evaluate the effectiveness of the proposed video analysis pipeline and the mathematical model, we conducted a clinical study where we collected retinoscopy videos at a local eye hospital. The retinoscopy videos were captured with the vertical streak beam scoping the horizontal meridian of the eye. Our video analysis pipeline predicts the net refractive error along the horizontal meridian. We compare our estimation with refractive error measurements from autorefractor (Topcon KR 8900), retinoscopy, and subjective refraction along the same horizontal meridian (using  $\text{net refractive error} = \text{spherical} + (\text{cylindrical} \times \sin(\text{axis})^2)$ , where  $\text{axis} \in \{0^\circ - 179^\circ\}$  [9]). We first examine the efficacy of the proposed video processing pipeline as a screening tool by classifying if the person has a refractive error or not, followed by a 4-class classification to understand the severity and type of refractive error. Next, we evaluate our estimated net refractive power with the ground truth subjective refraction measurements, and compare our results with prior work.

### 7.1 Clinical Study

The study included 185 distinct eyes with a mean refractive error of  $-1.13 \pm 2.03\text{D}$  along horizontal meridian, ranging from  $-6.25\text{D}$  to  $3.23\text{D}$ . A person should refer to an eye doctor for further tests if their refractive power is above 1D or below -1D [23]. We use the same cutoff values with the net refractive power for the binary classification—if the patient has a refractive error or not. Among the 185 eyes, 96 eyes were rated as normal (net refractive error between -1D and 1D), while the remaining 89 eyes had a refractive error, according to the subjective refraction measurements. Our proposed method achieves a sensitivity of 91.0% and specificity of 74.0% for this binary classification. Since we had an approximately uniform distribution for normal and refractive error cases, the high sensitivity value indicates that the proposed solution effectively recognizes patients with refractive power. However, our solution achieves a slightly low specificity of 74.0 %, meaning normal eyes were marked as eyes with refractive error. In a medical setting, it is more critical not to let refractive error undiagnosed, hence our proposed solution has the potential to be used as a screening tool.

Next, we examine the type (myopia or hyperopia) and severity (moderate or high) of refractive error. For this we categorize our dataset into five classes: (a) high hyperopia ( $\geq 4\text{D}$ ), (b) moderate hyperopia (1D to 4D), (c) normal (-1D to 1D), (d) moderate myopia (-4D to -1D), and (e) high myopia ( $\leq -4\text{D}$ ). Due to the low prevalence of high

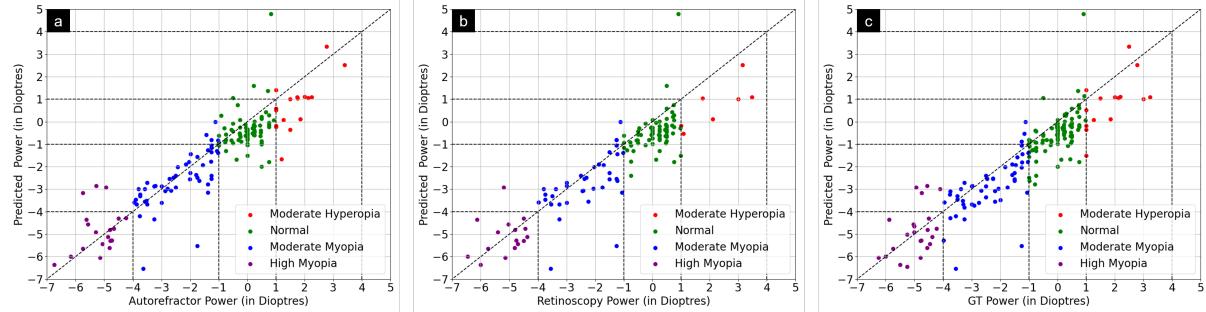


Fig. 7. Correlation plots for refracted power predicted by our proposed system compared to (a) autorefractor, (b) manual retinoscopy, and (c) gold-standard subjective refraction.

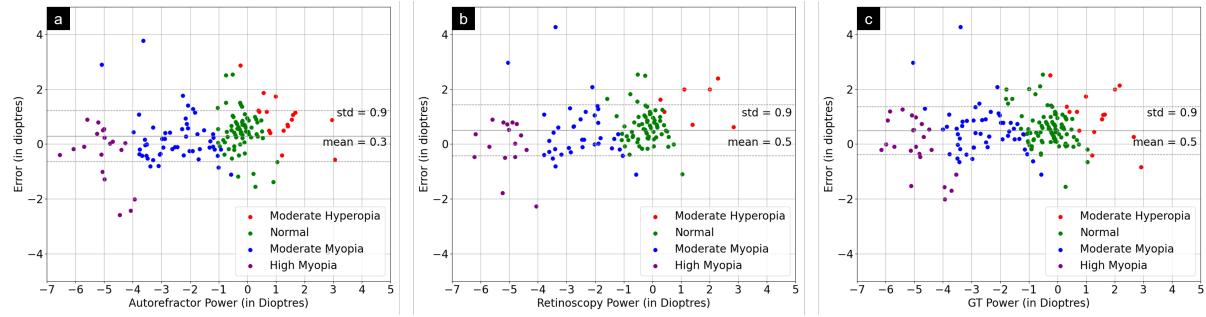


Fig. 8. Bland-Altman plots for refracted power predicted by our proposed system compared to (a) autorefractor, (b) manual retinoscopy, and (c) gold-standard subjective refraction.

hyperopia cases in the general population (<1%, Figure 3a) and the small sample (3 eyes) in our dataset, we ignore them for this classification. For evaluating the 4-classes, we use one-vs-all approach, i.e., if the predicted class matches the ground truth class, we classify it as true positive, otherwise we consider it to be a false prediction. We achieve high sensitivity and specificity ( $\geq 80\%$ ) for myopic eyes (Table 2). This further supports our claim for the proposed system to be used as a screening tool because of the high prevalence of myopic cases in the general population (Figure 3a). For moderate hyperopia cases, we have a small sample size (of 15 eyes), and our approach achieves a low sensitivity of 60.0%, which may be due to the steep gradient of ratio of distances moved by reflex and beam between two timestamps for +1D to +4D refractive power (Figure 9). Existing approaches based on subjective refraction also report low sensitivity (69.2%) and specificity (58.1%) with hyperopic eyes, as evident in the prior work with 4497 eyes [38].

We compare our estimated net refractive error (along the horizontal meridian) against the subjective refraction (ground truth), manual retinoscopy, and autorefractor measurements collected during the data collection (Table 3). Among the 185 eyes in our dataset, we have autorefractor measurements for 161 eyes, and manual retinoscopy measurements for 130 eyes. Figure 7c shows a strong correlation between our predictions and the ground truth subjective refraction measurements, with the mean absolute error of  $0.75 \pm 0.67$ D. Similarly, Bland-Altman plots (Figure 8c) of the same measurements show a good agreement with a mean difference of  $0.5 \pm 0.9$ D. The Bland-Altman plot also reveals the limits of agreement (95%) as -1.26D to 2.26D. We also achieve a strong Pearson

Table 3. Comparison of mean absolute error (MAE)  $\pm$  standard deviation for the different measurements in our dataset.

Method	Compared against	# of eyes	MAE $\pm$ Std dev
Manual Retinoscopy Autorefractor	Subjective Refraction	130	0.18 $\pm$ 0.27D
	Subjective Refraction	161	0.33 $\pm$ 0.33D
Ours	Manual Retinoscopy	130	0.78 $\pm$ 0.72D
Ours	Autorefractor	161	0.71 $\pm$ 0.67D
Ours	Subjective Refraction	185	0.75 $\pm$ 0.67D

Table 4. Comparison of different approaches to estimate refractive error using portable devices. (\*: Only myopic eyes)

Principle	Method	# of eyes	Mean Age	Power range	MAE $\pm$ Std dev
Self-subjective Refraction	Eye Netra [32]	152	51.9 $\pm$ 18.3 years	-15.25D to 4.25D	0.69 $\pm$ 0.53D
	USee [4]	120	53.1 $\pm$ 18.6 years	-6.00D to 5.00D	-
Eccentric Photorefraction	Plusoptix A09 [50]	64	58 months	-	0.52 $\pm$ 1.20D
	SPOT [33]	134	29.7 years	-	0.66 $\pm$ 0.56D
	Fu et al. [25]	172	11.34 $\pm$ 0.96 years	-4.00D to 0.50D	0.65D
	Yang et al. [61]	165	10 years	<=0D	0.80D
Wavefront Aberrometry	QuickSee [21]	82	26.4 $\pm$ 9.7 years	-6.00D to 4.00D	0.41 $\pm$ 0.53D
	SVOne AR* [15]	92	24.8 years	-6.63D to 0D	0.43D
Retinoscopy	Ours	185	30.2 $\pm$ 11.1 years	-6.25D to 3.23D	0.75 $\pm$ 0.67D

correlation coefficient,  $r(185) = 0.90$ ,  $p < 0.01$ . Figure 7a,b and Figure 8a,b show similar results, comparing our estimation with measurements from autorefractor and manual retinoscopy. Interestingly, 43.2% of our predictions were within  $\pm 0.5$ D, and 74.0% were within  $\pm 1$ D of the prescription given by an experienced optometrist using subjective refraction. These results are at par with the self-subjective refraction methods, like Eye Netra [32] (38.0% predictions within 0.5D) and USee [4] (77.5% predictions within 1D). Though these results are promising, the mean absolute error of manual retinoscopy ( $0.18 \pm 0.27$ D) and autorefractor ( $0.33 \pm 0.33$ D) highlight a scope for improvement (refer to Table 3).

Furthermore, we compare our evaluation dataset statistics and performance of our retinoscopy-based approach with prior work (Table 4). Subjective refraction is used as ground truth measurement for reporting mean absolute error in all the methods. Although smartphone-based approaches using eccentric photorefraction [25, 61] report similar performance to our retinoscopy-based approach, their evaluation is limited to the pediatric population, mostly with myopic eyes. On the flip side, self-subjective refraction-based methods (like Eye Netra [32] and USee [4]) evaluate only with the adult population highlighting their reliance on patient's cooperation. Meanwhile, our proposed automatic retinoscopy-based approach achieves comparable result on a diverse range of refractive power and age groups, in a low-cost manner. Expensive devices based on wavefront aberrometry (like QuickSee [21] and SVOne AR [15]) demonstrate the strongest agreement (< 0.5D) with the ground truth subjective refraction.

## 7.2 Mathematical Model Analysis

Here we examine results from the mathematical model of retinoscopy (Section 5) to understand the effects of data collection parameters (working distance  $d$  and effective source distance  $u$ ). As the working distance decreases from 66 cm (used by optometrist for manual retinoscopy) to 40-30cm (used by our pipeline) with fixed  $u$ , the variation of  $\Delta y_p / \Delta x_p$  increases for a fixed range of refractive error (Figure 9 plots the curve representing

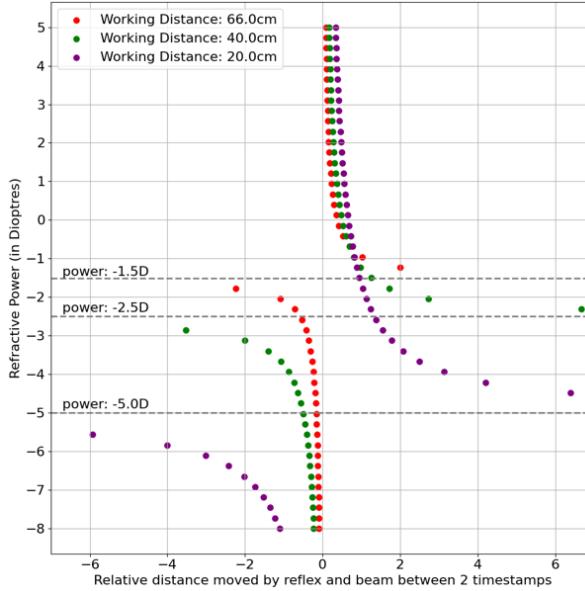


Fig. 9. Plot of refractive power vs ratio of distance moved by retinal reflex and beam, between two timestamps, for three different working distances: 20cm, 40cm, and 66cm. Note: the optimal operating range decreases as the working distance increases.

equation 12), thus enabling higher precision for the video analysis pipeline. Decreasing the working distance below 30cm should theoretically improve our results (Figure 9). However, there are two reasons why we decided against a working distance of 20cm: (1) the patient starts accommodating as their view of the logMAR chart gets obstructed, and (2) the patient's pupil size decreases due to increased beam intensity from retinoscope on the patient's pupil. Hence we strike a trade-off at ~30-40cm. For the effective source distance  $u$ , we selected the lowest vergence sleeve position as it was easiest to keep it consistent across the participants. For any other vergence sleeve position, an extra unnecessary step for calibrating  $u$  would have been required.

## 8 DISCUSSION

In this paper, our goal is to automate the process of retinoscopy for refractive error estimation by measuring the retinal reflex and retinoscopic beam movements. By doing so, we reduce the knowledge barrier to perform retinoscopy, eliminate the dependency on the expensive lens kit, and provide a means to preserve a digital record of retinoscopic videos for referring to it in the future. Our video processing pipeline tracks the retinoscopic beam and retinal reflex in the captured video and outputs net refractive error of the eye based on our extended mathematical formulation of retinoscopy. In our evaluation, we found that the proposed pipeline achieved a sensitivity and specificity of 91.0% and 74.0%, respectively, in diagnosing a patient's eye with a refractive error above 1D or below -1D. We also achieved an acceptable MAE ( $0.75 \pm 0.67$ D) on a range of refractive powers as well as across different age groups. In comparison, prior portable devices for refractive error estimation are either very expensive, require a minimum skill level to operate or have limited evaluation. For example, low-cost devices based on eccentric photorefraction were evaluated on mainly pediatric population, while devices based on self-subjective refraction were evaluated on only adult population (Table 4). Given the versatility of retinoscopy,

we believe that capturing retinoscopic videos using a smartphone has the potential to be an important tool in a variety of eye screenings, and our work is the first step toward this goal. We discuss a few such applications. Moreover, though our evaluation is limited to estimating net refractive error along the horizontal meridian, we discuss how this work can be extended to estimate spherical error, cylindrical error and its axis.

## 8.1 Applications

Uncorrected refractive error is a widespread vision problem, which can lead to permanent vision impairment if not addressed timely. Hence many portable medical-grade autorefractors are available to democratize eye care services. However, these devices are still expensive for the Global South, and they only serve a single purpose. In contrast, the videos that we captured during our retinoscopy procedure can be used for multiple applications, although we focused on estimating refractive error in this paper.

**8.1.1 Amblyopia Screening.** Undiagnosed refractive error can also cause amblyopia, a serious eye disorder. Amblyopia, also called lazy eye, mainly occurs due to poor vision in one eye, which can cause the brain to ignore the blurred image coming from the weaker eye. Over time this can lead to permanent vision loss in the weaker eye. Amblyopia affects around 3 out of every 100 children [42]. If diagnosed early (children of age 7 years or less), it is treatable by eye patching and wearing corrective lenses.

According to the AAPOS 2013 guidelines (for age  $\geq 4$  years) [19], amblyopia risk factors are: (a) difference in refractive power between the two eyes (anisometropia)  $> 1.5\text{D}$ , (b) hyperopia  $> 3.5\text{D}$  (along any meridian), (c) myopia  $< -1.5\text{D}$  (along any meridian), (d) astigmatism  $> 1.5\text{D}$ , (e) crossed eyes, and (f) media opacity  $> 1\text{ mm}$  in size. The first four risk factors are dependent on refractive error estimation, which can be measured by our proposed pipeline. Moreover, our approach can be extended to detect crossed eyes from the retinoscopy videos, similar to Huang et al. [30]. Hence, our proposed low-cost portable solution has the potential to be useful in remote and resource-constrained settings [6] for amblyopia screening.

**8.1.2 Keratoconus Detection.** Keratoconus is a severe eye disease that leads to deformation of the cornea, developing a conical bulge. It affects people in the age group of 10-25 years and progresses slowly for several years. If diagnosed early, it can be treated with corrective lenses; however, in advanced stages, it requires a corneal transplant or can lead to partial or complete blindness. Thus timely diagnosis and treatment are essential to cure keratoconus. The diagnosis of keratoconus is performed by expensive and bulky medical devices called corneal topographers, which are not accessible to the masses, especially to people living in low- and middle-income countries. In a recent work [26, 27], a low-cost smartphone-based corneal topographer, SmartKC, was proposed to diagnose keratoconus. However, it requires additional hardware like a 3D printed placido head, USB-powered LEDs, and a paper-based diffuser. Although it achieves high accuracy in keratoconus diagnosis, the additional hardware makes it difficult to assemble by non-experts and requires additional training to use effectively.

On the other hand, retinoscopy has been shown to be very sensitive and reliable for detecting keratoconus even in the early stages of the disease [2, 36]. In keratoconus eyes, the retinal reflex, instead of being evenly distributed and moving in a single direction, shows two edges that move towards each other. This creates the visual effect of a scissor and is known as the *scissors reflex*. A simple extension to our video processing pipeline can be made to identify such scissors reflex accurately. With that, our proposed system has the potential to be used for mass screening of keratoconus.

**8.1.3 Generate Complete Eye Prescription.** For evaluating our current system, we computed the net refractive power along the horizontal meridian ( $0^\circ$ ). However, our approach is not limited to a single meridian. We can perform the same retinoscopy process (*i.e.*, collect retinoscopy videos from the smartphone camera) along different meridians to estimate the net refractive power along those meridians. Leveraging the formulation from meridional

refractometry [9], given the net refractive power along three predetermined meridians of the eye, the complete prescription of the eye (comprising of spherical power, cylindrical power, and cylindrical axis) can be calculated.

**8.1.4 Training Device.** Retinoscopy is dependent on the examiner's perception of neutrality and has been found to be prone to inter-observer variability [41, 53]. Moreover, prior work showed a direct relation between accuracy of refractive error estimation using retinoscopy with years of experience [5], thus highlighting the requirement of highly-trained examiners with a certain level of experience. At the same time, learning to use a retinoscope device is known to be hard, e.g., American Association of Ophthalmology states "*Retinoscopy is not an easy skill to learn. It takes patience and a lot of practice.*" [37].

Looking beyond identifying eye disorders, our system can act as a guide for teaching retinoscopy to medical students. The ability to digitally zoom, pause/play/replay a recorded video, and track progress over time, can be valuable to explain, learn, document, and demonstrate abnormal retinal reflex movements and retinal disorders. Moreover, it can help bridge the communication gap between a teacher and students. There has been prior work [12] supporting the use of digital retinoscope as a training device.

## 8.2 Real-time Assistance

During the pilot stage of data collection, we iteratively experimented with different parameters such as working distance, room brightness and beam width. Despite fixing these data collection parameters, we had to discard ~27% of collected videos due to small pupil size, noisy reflex/fiducial, blurry eyes, and out-of-bound working distance. To make the setup more user-friendly, we plan to implement on-device image processing based quality checks to provide real-time feedback to the operator. First, working distance should be calculated on the preview screen to guide the operator to move forward or backward while collecting data. Second, computing the sharpness (similar to [26, 48]) of the eye region and fiducial markers while capturing the video can prompt the operator to recapture video if either of the regions is out-of-focus/blurry. Third, another problem that we identified in our data was the speed at which the retinoscope was rotated by the data collector. As our video processing pipeline tracks retinal reflex throughout the video, it requires a slow movement of the retinoscope. In future deployments, the smartphone app should provide feedback to the operator regarding the speed of movement of the retinoscopic beam based on the readings from the smartphone's in-built IMU sensors. Finally, thresholds like small pupil size (below 3mm) can be computed in real-time. In such a case, the app can suggest that the operator use cycloplegia drops, to reduce accommodation and increase pupil size.

## 8.3 Failure Cases

Despite a robust and reliable video processing pipeline, we identified four types of cases (Figure 10) where our proposed solution failed to estimate refractive error correctly. These cases can be classified into two categories: outside operating range (Figure 10a,b) and data collection issues (Figure 10c,d).

First, in high hyperopia cases (> 4D), the pupil size of the patient is significantly smaller than the average pupil size [28] (Figure 10a). In such cases, our pipeline failed to track the retinal reflex movement within the pupil. Second, in high myopia cases (< -6D), the retinal reflex tends to be dull [16] (Figure 10b), again resulting in incorrect refractive error estimate. These problems are even faced by trained examiners while performing manual retinoscopy. However, since manual retinoscopy is a trial and error process, the retinal reflex becomes magnified and brighter as the examiner reaches neutrality by adding combinations of lenses in the trial frame. These two scenarios restrict our operating range from -6.0D to +3.0D, comprising ~95% of the clinical cases (Figure 3a).

Third, a few recorded videos were out-of-focus and/or noisy (Figure 10c). This resulted in blurred reflex edges, which were not correctly detected by our video processing pipeline. This can be identified and corrected with on-device image processing based quality check (similar to Gairola et al. [26]). Fourth, erroneous fiducial detection (Figure 10d) resulted in a significant deviation from the ground truth. This mainly happened in frames where the

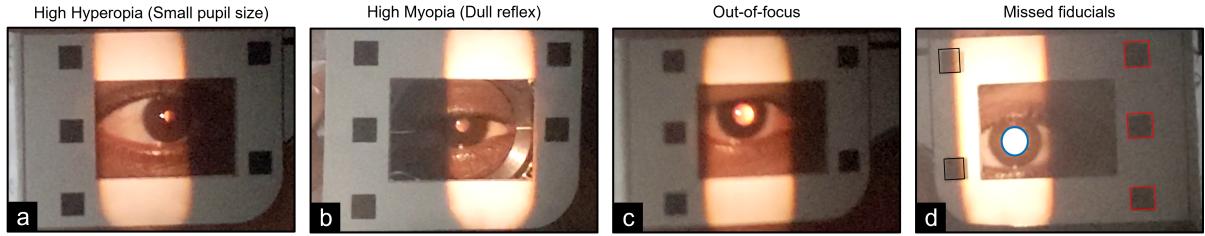


Fig. 10. Cases where our proposed solution failed to estimate refractive error accurately: (a, b) outside operating range, and (c, d) data collection issues.

beam edges are close to the fiducial, resulting in interference. It can also be recognized and corrected during data collection by ensuring that the patient's pupil is approximately positioned at the center of the 'reflex search space' of the paper frame (Figure 2c).

#### 8.4 Limitations

We acknowledge a few limitations of our current work. First, due to COVID-19, we were able to recruit very few patients, hence we were unable to collect a larger dataset (though it is comparable to recent mobile health sensing work [26, 40, 58]). Although our result seems encouraging on the limited range of refractive error, we believe an evaluation with a larger sample size across the range of refractive errors and age group is needed before real-world deployment. Second, our current evaluation is limited to estimating net refractive error along the horizontal meridian. Modern autorefractors output spherical power, cylindrical power, and cylindrical axis. Our setup can also achieve that by computing spherical power along three predetermined axes (using [9]), though the accuracy of that approach needs to be rigorously evaluated. Third, in the current setup we used a medical-grade retinoscope attached to a smartphone camera for data collection. In the future, we plan to use a cheap retinoscope, or a 3D-printed attachment imitating a retinoscope (similar to [26, 27]), thus reducing the setup cost. Finally, for the current evaluation, the data was collected with a single smartphone by two data collectors. For wider deployment, we need to further evaluate our results on other smartphones with more data collectors.

### 9 CONCLUSION

Automating retinoscopy can be an important tool for the early diagnosis of refractive error, a leading cause of vision impairment globally. Along with all the advantages of a retinoscope (like low-cost, portable, and no cooperation required from the patient), a digital retinoscope offers two key advantages—no requirement of a lens kit and no requirement of a trained examiner. In this work, we extended the mathematical formulation of retinoscopy to work on smartphone camera captured retinoscopic videos, and developed a video processing pipeline based on that formulation. We collected 185 eyes video data by attaching a smartphone to a retinoscope with the patient wearing a paper frame. Our approach achieved a sensitivity of 91.0% and specificity of 74.0% for refractive error classification, and a mean absolute error of  $0.75 \pm 0.67$ D for net refractive error estimation. Although more research is needed to develop a functional smartphone-based digital retinoscope at par with commercial medical devices, we believe that this is the first step in showing the feasibility of such a device.

### ACKNOWLEDGMENTS

We would like to thank optometrists Amaravathi P. and Wartson for helping with the data collection, Banukumar Rajendran for helping with retinoscope-smartphone assembly, and all the participants for their time and patience.

## REFERENCES

- [1] Arunika Agarwal, David E Bloom, Vincent P deLuise, Alyssa Lubet, Kaushik Murali, and Srinivas M Sastry. 2019. Comparing low-cost handheld autorefractors: A practical approach to measuring refraction in low-resource settings. *PLoS One* 14, 10 (Oct 2019).
- [2] Haitham Al-Mahrouqi, Saif Oraba, Shihab Al-Habsi, Noufal Mundemkattil, Jithin Babu, Sathiya Panchatcharam, Rashid Al-Saidi, and Abdulatif Al-Raisi. 2019. Retinoscopy as a Screening Tool for Keratoconus. *Cornea* 38 (04 2019), 1. <https://doi.org/10.1097/ICO.0000000000001843>
- [3] American Optometric Association. 2020. Evidence-based clinical practice guideline: Comprehensive pediatric eye and vision examination. *Optometric Clinical Practice* (2020), 2–67.
- [4] Anvesh Annadanam, Varshini Varadaraj, Lucy Mudie, Alice Liu, William Plum, J. White, Megan Collins, and David Friedman. 2018. Comparison of self-refraction using a simple device, USee, with manifest refraction in adults. *PLOS ONE* 13 (02 2018). <https://doi.org/10.1371/journal.pone.0192055>
- [5] Shrikant R Bharadwaj, Menaka Malavita, and Jennifer Jayaraj. 2014. A psychophysical technique for estimating the accuracy and precision of retinoscopy. *Clinical and Experimental Optometry* 97, 2 (2014), 164–170. <https://doi.org/10.1111/cxo.12112>
- [6] Karthik S Bhat, Mohit Jain, and Neha Kumar. 2021. Infrastructuring Telehealth in (In)Formal Patient-Doctor Contexts. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 323 (oct 2021), 28 pages. <https://doi.org/10.1145/3476064>
- [7] Paul Boeder and H. E. Kolder. 1984. Neutralization at Infinity in Streak Retinoscopy. *Archives of Ophthalmology* 102, 9 (09 1984), 1396–1399. <https://doi.org/10.1001/archoph.1984.01040031138042> arXiv:[https://jamanetwork.com/journals/jamaophthalmology/articlepdf/635237/archoph\\_102\\_9\\_042.pdf](https://jamanetwork.com/journals/jamaophthalmology/articlepdf/635237/archoph_102_9_042.pdf)
- [8] Rupert Bourne, Jamie Adelson, Seth Flaxman, Paul Briant, Hugh Taylor, Robert Casson, Mukharram Bikbov, Michele Bottone, Tasanee Braithwaite, Alain Bron, Ching-yu Cheng, Maria Vittoria Cincinelli, Nathan Congdon, Arthur Fernandes, David Friedman, Joao Furtado, Ronnie George, Rim Kahloun, John Kempen, and Theo Vos. 2020. Trends in Prevalence of Blindness and Distance and Near Vision Impairment Over 30 Years and Contribution to the Global Burden of Disease in 2020. *SSRN Electronic Journal* (01 2020). <https://doi.org/10.2139/ssrn.3582742>
- [9] Richard F. Brubaker, Robert D. Reinecke, and Jack C. Copeland. 1969. Meridional Refractometry: I. Derivation of Equations. *Archives of Ophthalmology* 81, 6 (06 1969), 849–852. <https://doi.org/10.1001/archoph.1969.0099010851018>
- [10] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. 2011. Non-Local Means Denoising. *Image Processing On Line* 1 (2011), 208–212. [https://doi.org/10.5201/ipol.2011.bcm\\_nlm](https://doi.org/10.5201/ipol.2011.bcm_nlm).
- [11] John Canny. 1986. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8, 6 (1986), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- [12] Weng Onn Chan, Matthew Crabb, David Sia, and Deepa Ajay Taranath. 2014. Creating a digital retinoscope by combining a mobile smartphone camera and a retinoscope. *Journal of AAPOS : the official publication of the American Association for Pediatric Ophthalmology and Strabismus* 18 4 (2014), 387–8.
- [13] Ying-Ling Chen, bo Tan, and Jwl Lewis. 2003. Simulation of eccentric photorefraction images. *Optics express* 11 (08 2003), 1628–42. <https://doi.org/10.1364/OE.11.001628>
- [14] Yee-Fong Choong, Ai-Hong Chen, and Pik-Pin Goh. 2006. A comparison of autorefraction and subjective refraction with and without cycloplegia in primary school children. *Am J Ophthalmol* 142, 1 (July 2006), 68–74.
- [15] Kenneth Ciuffreda and Mark Rosenfield. 2015. Evaluation of the SVOne: A Handheld, Smartphone-Based Autorefractor. *Optometry and vision science : official publication of the American Academy of Optometry* 92 (11 2015). <https://doi.org/10.1097/OPX.0000000000000726>
- [16] J.M. Corboy. 2003. *The Retinoscopy Book: An Introductory Manual for Eye Care Professionals*. Slack. <https://books.google.co.in/books?id=6I6JeDWonhQC>
- [17] I. Cosme-Cisneros, G. Escamilla-Ruiz, D. Flores-Montoya, G. Hernández-Gómez, and A. Gómez-Vieyra. 2017. *Instrument for Recording Purkinje Images*. 101–109. [https://doi.org/10.1007/978-3-319-28513-9\\_14](https://doi.org/10.1007/978-3-319-28513-9_14)
- [18] L Dandona, R Dandona, M Srinivas, P Giridhar, K Vilas, M N Prasad, R K John, C A McCarty, and G N Rao. 2001. Blindness in the Indian state of Andhra Pradesh. *Invest Ophthalmol Vis Sci* 42, 5 (April 2001), 908–916.
- [19] Sean P. Donahue, Brian Arthur, Daniel E. Neely, Robert W. Arnold, David Silbert, and James B. Ruben. 2013. Guidelines for automated preschool vision screening: A 10-year, evidence-based update. *Journal of American Association for Pediatric Ophthalmology and Strabismus* 17, 1 (2013), 4–8. <https://doi.org/10.1016/j.jaapos.2012.09.012>
- [20] Chao Dong, Chen Change Loy, and Xiaoyu Tang. 2016. Accelerating the Super-Resolution Convolutional Neural Network. In *Computer Vision – ECCV 2016*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 391–407.
- [21] Nicholas J Durr, Shivang R Dave, Fuensanta A Vera-Diaz, Daryl Lim, Carlos Dorronsoro, Susana Marcos, Frank Thorn, and Eduardo Lage. 2015. Design and Clinical Evaluation of a Handheld Wavefront Autorefractor. *Optometry and vision science : official publication of the American Academy of Optometry* 92, 12 (December 2015), 1140–1147. <https://doi.org/10.1097/opx.0000000000000732>
- [22] David B. Elliott. 2016. The good (logMAR), the bad (Snellen) and the ugly (BCVA, number of letters read) of visual acuity measurement. *Ophthalmic and Physiological Optics* 36, 4 (2016), 355–358. <https://doi.org/10.1111/opp.12310>

- arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/opo.12310>
- [23] Daniel Ian Flitcroft, Mingguang He, Jost B. Jonas, Monica Jong, Kovin Naidoo, Kyoko Ohno-Matsui, Jugnoo Rahi, Serge Resnikoff, Susan Vitale, and Lawrence Yannuzzi. 2019. IMI – Defining and Classifying Myopia: A Proposed Set of Standards for Clinical and Epidemiologic Studies. *Investigative Ophthalmology & Visual Science* 60, 3 (02 2019), M20–M30. <https://doi.org/10.1167/iovs.18-25957>  
arXiv:[https://arvojournals.org/arvo/content\\_public/journal/iov/937872/i1552-5783-60-3-m20.pdf](https://arvojournals.org/arvo/content_public/journal/iov/937872/i1552-5783-60-3-m20.pdf)
- [24] LVPEI Center for Innovation. 2018. *FoFo (Folding Foropter) project*. Retrieved January 1,2022 from <https://www.fofo-project.com/>
- [25] Eugene Fu, Zhongqi Yang, Hong Leong, Grace Ngai, Chi Wai Do, and Lily Chan. 2020. Exploiting Active Learning in Novel Refractive Error Detection with Smartphones. 2775–2783. <https://doi.org/10.1145/3394171.3413748>
- [26] Siddhartha Gairola, Murtuza Bohra, Nadeem Shaheer, Navya Jayaprakash, Pallavi Joshi, Anand Balasubramaniam, Kaushik Murali, Nipun Kwatra, and Mohit Jain. 2022. SmartKC: Smartphone-Based Corneal Topographer for Keratoconus Detection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 155 (dec 2022), 27 pages. <https://doi.org/10.1145/3494982>
- [27] Siddhartha Gairola, Pallavi Joshi, Anand Balasubramaniam, Kaushik Murali, Nipun Kwatra, and Mohit Jain. 2022. Keratoconus Classifier for Smartphone-based Corneal Topographer. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE.
- [28] Baran Gencer, Engin Ozgurhan, Tuğba Kurt, Ismail Ersan, Sedat Arıkan, Selçuk Kara, Mediha Coşar, and Ercüment Bozkurt. 2015. Comparision of pupil size in myopic and hyperopic patients. *Pamukkale Medical Journal* 8 (01 2015), 88–91. <https://doi.org/10.5505/ptd.2015.59455>
- [29] R. I. Hartley and A. Zisserman. 2004. *Multiple View Geometry in Computer Vision* (second ed.). Cambridge University Press, ISBN: 0521540518.
- [30] Xilang Huang, Sang Joon Lee, Chang Kim, and Seon Han Choi. 2021. An automatic screening method for strabismus detection based on image processing. *PLOS ONE* 16 (08 2021), e0255643. <https://doi.org/10.1371/journal.pone.0255643>
- [31] J. Illingworth and J. Kittler. 1987. The Adaptive Hough Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9*, 5 (1987), 690–698. <https://doi.org/10.1109/TPAMI.1987.4767964>
- [32] V Swetha E Jeganathan, Nita Valikodath, Leslie M Niziol, Sean Hansen, Hannah Apostolou, and Maria A Woodward. 2018. Accuracy of a Smartphone-based Autorefractor Compared with Criterion-standard Refraction. *Optometry and vision science : official publication of the American Academy of Optometry* 95, 12 (December 2018), 1135–1141. <https://doi.org/10.1097/oxp.0000000000001308>
- [33] Daniela Jesus, Flávio Villela, Luis Orlandin, Fernando Eiji, Daniel Dantas, and Milton Alves. 2016. Comparison between refraction measured by Spot Vision ScreeningTM and subjective clinical refractometry. *Clinics* 70 (02 2016), 69–72. [https://doi.org/10.6061/clinics.2016\(02\)03](https://doi.org/10.6061/clinics.2016(02)03)
- [34] Jorge Jorge, António Queirós, José B Almeida, and Manuel A Parafita. 2005. Retinoscopy/autorefraction: which is the best starting point for a noncycloplegic refraction? *Optometry and vision science : official publication of the American Academy of Optometry* 82, 1 (January 2005), 64–68. <http://europemc.org/abstract/MED/15630406>
- [35] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22, 1 (1951), 79–86.
- [36] Michael Lawless, FRACS FRACO, Anthony DCLP, and Mark FRACS. 2007. Keratoconus: Diagnosis and management. *Australian and New Zealand Journal of Ophthalmology* 17 (11 2007), 33 – 60. <https://doi.org/10.1111/j.1442-9071.1989.tb00487.x>
- [37] Olivia L Lee. 2015. *Retinoscopy 101*. Retrieved February 1,2022 from <https://www.aao.org/young-ophthalmologists/yo-info/article/retinoscopy-101>
- [38] Jody Leone (nee Major), Paul Mitchell, Ian Morgan, Annette Kifley, and Kathryn Rose. 2010. Use of Visual Acuity to Screen for Significant Refractive Errors in Adolescents Is It Reliable? *Archives of ophthalmology* 128 (07 2010), 894–9. <https://doi.org/10.1001/archophthalmol.2010.134>
- [39] Alan Lukežić, Tomáš Vojíř, Luka Čehovin Zajc, Jirí Matas, and Matej Kristan. 2017. Discriminative Correlation Filter with Channel and Spatial Reliability. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 4847–4856. <https://doi.org/10.1109/CVPR.2017.515>
- [40] Alex Mariakakis, Megan A. Banks, Lauren Phillipi, Lei Yu, James Taylor, and Shwetak N. Patel. 2017. BiliScreen: Smartphone-Based Scleral Jaundice Monitoring for Liver and Pancreatic Disorders. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 2, Article 20 (jun 2017), 26 pages. <https://doi.org/10.1145/3090085>
- [41] Sara J McCullough, Lesley Doyle, and Kathryn J Saunders. 2017. Intra- and inter-examiner repeatability of cycloplegic retinoscopy among young children. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)* 37, 1 (January 2017), 16–23. <https://doi.org/10.1111/opo.12341>
- [42] Kaushik Murali, Viswesh Krishna, Vrishab Krishna, B. Kumari, Sowmya Raveendra Murthy, Vidhya C, and Payal Shah. 2021. Effectiveness of Kanna photoscreener in detecting amblyopia risk factors. *Indian Journal of Ophthalmology* 69, 8 (2021). [https://journals.lww.com/ijo/Fulltext/2021/08000/Effectiveness\\_of\\_Kanna\\_photoscreener\\_in\\_detecting\\_19.aspx](https://journals.lww.com/ijo/Fulltext/2021/08000/Effectiveness_of_Kanna_photoscreener_in_detecting_19.aspx)
- [43] Kovin S Naidoo, Janet Leasher, Rupert R Bourne, Seth R Flaxman, Jost B Jonas, Jill Keeffe, Hans Limburg, Konrad Pesudovs, Holly Price, Richard A White, Tien Y Wong, Hugh R Taylor, Serge Resnikoff, and Vision Loss Expert Group of the Global Burden of Disease Study. 2016. Global Vision Impairment and Blindness Due to Uncorrected Refractive Error, 1990-2010. *Optom Vis Sci* 93, 3 (March 2016),

227–234.

- [44] World Health Organization. 2019. *World report on vision*. World Health Organization. 160 p. pages.
- [45] Nobuyuki Otsu. 1979. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1979), 62–66.
- [46] Art Owen. 2007. A robust hybrid of lasso and ridge regression. *Contemp. Math.* 443 (01 2007). <https://doi.org/10.1090/conm/443/08555>
- [47] Vitor F. Pamplona, Ankit Mohan, Manuel M. Oliveira, and Ramesh Raskar. 2010. NETRA: Interactive Display for Estimating Refractive Errors and Focal Range. *ACM Trans. Graph.* 29, 4, Article 77 (jul 2010), 8 pages. <https://doi.org/10.1145/1778765.1778814>
- [48] Chunjong Park, Hung Ngo, Libby Rose Lavitt, Vincent Karuri, Shiven Bhatt, Peter Lubell-Doughtie, Anuraj H. Shankar, Leonard Ndwiga, Victor Osoti, Juliana K. Wambua, Philip Bejon, Lynette Isabella Ochola-Oyier, Monique Chilver, Nigel Stocks, Victoria Lyon, Barry R. Lutz, Matthew Thompson, Alex Mariakakis, and Shwetak Patel. 2021. The Design and Evaluation of a Mobile System for Rapid Diagnostic Test Interpretation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 29 (mar 2021), 26 pages. <https://doi.org/10.1145/3448106>
- [49] Donatella Pascolini and Silvio Paolo Mariotti. 2012. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology* 96, 5 (2012), 614–618. <https://doi.org/10.1136/bjophthalmol-2011-300539>
- [50] Arnaud Payerols, Clémence Eliaou, Véronique Trezeguet, Max Villain, and Vincent Daien. 2016. Accuracy of PlusOptix A09 distance refraction in pediatric myopia and hyperopia. *BMC ophthalmology* 16 (June 2016), 72. <https://doi.org/10.1186/s12886-016-0247-8>
- [51] J S Rahi, C S Peckham, and P M Cumberland. 2008. Visual impairment due to undiagnosed refractive error in working age adults in Britain. *Br J Ophthalmol* 92, 9 (Sept. 2008), 1190–1194.
- [52] Austin Roorda, Melanie Campbell, and William Bobier. 1997. Slope-based eccentric photorefraction: Theoretical analysis of different light source configurations and effects of ocular aberrations. *Journal of the Optical Society of America. A, Optics, image science, and vision* 14 (11 1997), 2547–56. <https://doi.org/10.1364/JOSAA.14.002547>
- [53] Aran Safir, Lyon Hyams, John Philpot, and Louis S. Jagerman. 1970. Studies in Refraction: I. The Precision of Retinoscopy. *Archives of Ophthalmology* 84, 1 (07 1970), 49–61. <https://doi.org/10.1001/archopht.1970.00990040051013>
- [54] Ashwin Sainani. 2013. Special considerations for prescription of glasses in children. *Journal of Clinical Ophthalmology and Research* 1 (01 2013), 169. <https://doi.org/10.4103/2320-3897.116861>
- [55] Geetha Sravani, Vinay Nilagiri, and Shrikant Bharadwaj. 2015. Photorefraction estimates of refractive power varies with the ethnic origin of human eyes. *Scientific reports* 5 (01 2015), 7976. <https://doi.org/10.1038/srep07976>
- [56] J M Tielsch, A Sommer, K Witt, J Katz, and R M Royall. 1990. Blindness and visual impairment in an American urban population. The Baltimore Eye Survey. *Arch Ophthalmol* 108, 2 (Feb. 1990), 286–290.
- [57] Silay Uğurbas, Numan Küçük, İrem Isik, Atilla Alpay, Cagatay Buyukuslu, and Suat Uğurbas. 2019. Objective vision screening using PlusoptiX for children aged 3–11 years in rural Turkey. *BMC Ophthalmology* 19 (03 2019). <https://doi.org/10.1186/s12886-019-1080-7>
- [58] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N. Patel. 2018. *Seismo: Blood Pressure Monitoring Using Built-in Smartphone Accelerometer and Camera*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3173574.3173999>
- [59] Wolfgang Wesemann, Anthony Norcia, and Dale Allen. 1992. Theory of eccentric photorefraction (photoretinoscopy): astigmatic eyes. *Journal of the Optical Society of America. A, Optics and image science* 8 (01 1992), 2038–47. <https://doi.org/10.1364/JOSAA.8.002038>
- [60] Stephanie West and Cathy Williams. 2016. Amblyopia in children (aged 7 years or less). *BMJ clinical evidence* 2016 (01 2016).
- [61] Zhongqi Yang, Eugene Fu, Grace Ngai, Hong Leong, Chi Wai Do, and Lily Chan. 2020. Screening for refractive error with low-quality smartphone images. 119–128. <https://doi.org/10.1145/3428690.3429175>
- [62] Yaopeng Zhou. 2015. Portable wavefront aberrometer. 9066683 (06 2015). <https://www.freepatentsonline.com/9066683.html>
- [63] Basil J. Zitelli, Sara C. McIntire, and Andrew J. Nowalk. 2012. *Zitelli and Davis' atlas of pediatric physical diagnosis*. Saunders/Elsevier.
- [64] Karel J. Zuiderveld. 1994. Contrast Limited Adaptive Histogram Equalization. In *Graphics Gems*.