Biomimetic Eye Modeling & Deep Neuromuscular Oculomotor Control

MASAKI NAKADA*, ARJUN LAKSHMIPATHY*, HONGLIN CHEN, NINA LING, TAO ZHOU, and DEMETRI TERZOPOULOS, University of California, Los Angeles, USA
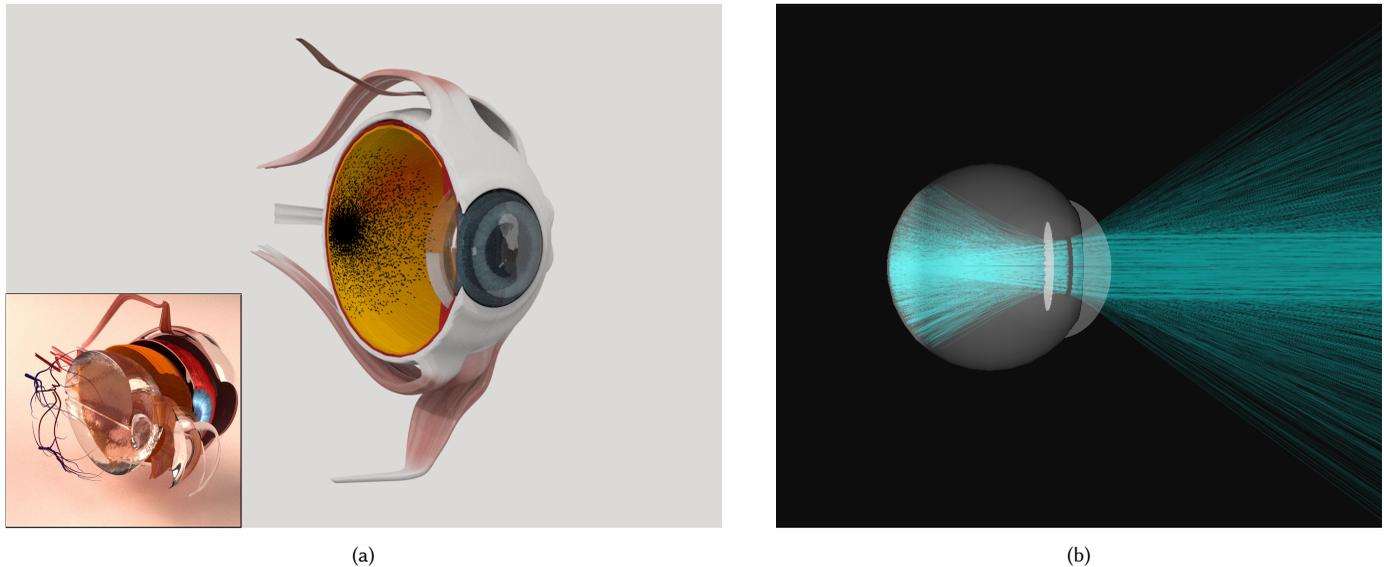
(a)

(b)

Fig. 1. (a) Cross section of a detailed geometric model of the left eye with an inset rendering of the "dissected" model. The black dots on the surface of the retina indicate the positions of our retinal photoreceptors. (b) Rays are cast from the positions of the photoreceptors through the finite-aperture pupil and out into the 3D environment; the irradiance responses of the photoreceptors are computed by ray tracing.

We present a novel, biomimetic model of the eye for realistic virtual human animation. We also introduce a deep learning approach to oculomotor control that is compatible with our biomechanical eye model. Our eye model consists of the following functional components: (i) submodels of the 6 extraocular muscles that actuate realistic eye movements, (ii) an iris submodel, actuated by pupillary muscles, that accommodates to incoming light intensity, (iii) a corneal submodel and a deformable, ciliary-muscle-actuated lens submodel, which refract incoming light rays for focal accommodation, and (iv) a retina with a multitude of photoreceptors arranged in a biomimetic, foveated distribution. The light intensity captured by the photoreceptors is computed using ray tracing from the photoreceptor positions through the finite aperture pupil into the 3D virtual environment, and the visual information from the retina is output via an optic nerve vector. Our oculomotor control system includes a foveation controller implemented as a locally-connected, irregular Deep Neural Network (DNN), or "LiNet", that conforms to the nonuniform retinal photoreceptor distribution, and a neuromuscular motor controller implemented as a fully-connected DNN, plus auxiliary Shallow Neural Networks (SNNs) that control the accommodation of the pupil and lens. The DNNs are trained offline through deep learning from data synthesized by the eye model itself. Once trained, the oculomotor control system operates robustly and efficiently online. It innervates the intraocular muscles to perform illumination and focal accommodation and the extraocular muscles to produce natural eye movements in order to foveate and pursue moving visual targets. We additionally demonstrate the operation of our eye model (binocularly) within our recently introduced sensorimotor control framework involving an anatomically-accurate biomechanical human musculoskeletal model.

CCS Concepts: • **Computing methodologies → Computer graphics**; **Animation**; **Bio-inspired approaches**; **Neural networks**; **Artificial life**; **Physical simulation**.

Additional Key Words and Phrases: Human eye modeling and animation; Oculomotor control; Biomimetic modeling and simulation; Neural network learning; Deep learning; Biomechanical human animation.

ACM Reference Format:
Masaki Nakada, Arjun Lakshmipathy, Honglin Chen, Nina Ling, Tao Zhou, and Demetri Terzopoulos. 2019. Biomimetic Eye Modeling & Deep Neuromuscular Oculomotor Control. *ACM Trans. Graph.* 38, 6, Article 221 (November 2019), 14 pages. https://doi.org/10.1145/3355089.3356558

---

*Co-primary author.

Authors' Address: Computer Science Department, University of California, Los Angeles, CA 90095, USA..

---

1 INTRODUCTION

Animating the eyes, metaphorically the "windows to the soul", is a critical yet insufficiently explored aspect of the realistic animation of humans and other animals. Although some pioneering work has been done in developing computational models of the eye, the research community has not yet developed a realistic eye model for highly dynamic, autonomous, sensorimotor human simulation. Rather than attempting to adjust the parameters of ad hoc procedural models in the hope of generating plausible kinematic eye movements, our work takes a strongly biomimetic approach to tackling this challenge. Our conviction is that increasingly detailed functional modeling and simulation of the various parts of the human body, particularly the eyes, will naturally produce realistic, dynamic movements without the need for substantial quantities of external data from eye trackers and other motion capture systems. To this end, the present paper makes two principal contributions.

First, we introduce a biomimetic model of the eye, which consists of the following functional components:

(1) Submodels of the 6 extraocular muscles, innervated by muscle activations, which actuate realistic eye movements.
(2) An iris submodel, actuated by sphincter/dilator muscles, that can accommodate to incoming light intensity by constricting and dilating the pupil.
(3) Corneal and physics-based deformable lens submodels that refract incoming light rays and adjust the focal distance.
(4) A retina with a multitude of photoreceptors (the black dots in Fig. 1a) arranged in a biomimetic, foveated distribution.

Retinal photoreceptor responses to 3D visual stimuli are computed using ray tracing from photoreceptor positions through the finite-aperture pupil (as opposed to a simplistic pinhole camera) into the 3D virtual environment (Fig. 1b), and the visual information from the retina is output by the model eye via an Optic Nerve Vector (ONV).

Second, to control our eye model, we introduce a compatibly biomimetic oculomotor control system, which includes the following functional components:

(1) Two Shallow Neural Networks (SNNs) that control the accommodations of the iris and lens.
(2) A fully-connected neuromuscular motor Deep Neural Network (DNN) that controls the extraocular muscles to produce natural eye movements.
(3) A new, locally-connected, irregular foveation DNN, or "LiNet", that conforms to the nonuniform retinal photoreceptor distribution and drives eye movements.

Of course, we take a machine learning approach to training the four neural networks in our oculomotor control system. This is automatically carried out offline from visual data systematically synthesized by the eye model itself. The properly trained neural networks then operate efficiently and robustly online, innervating both the intraocular muscles to achieve accommodation and the extraocular muscles to produce natural eye movements.

Although there is some variance among oculomotor theories, three principal types of human eye movements and fixation have been identified [Bahill et al. 1975; Becker 1989; Leigh and Zee 2015], as follows: *Saccade movements:* Quick eye movements that bring visual targets from peripheral vision to the center of the retina—the

fovea—where visual acuity is greatest.[1] *Smooth pursuit movements:* Given a moving visual target, the eye pursues the fixated target as it moves; the trajectory of the eye movement is not entirely smooth, but exhibits random small-amplitude oscillations. *Vestibulo-Ocular Reflex (VOR) movements:* When the head rotates, the eye makes a reflex compensatory rotation in order to stabilize the visual target in the foveal region while the head is in motion.

Our neuromuscular motor DNN outputs the activation signals to the 6 extraocular muscles to synthesize all these types of eye movements, and it does so entirely naturally, without trajectory planning or the imposition of any contrived constraints on velocity and acceleration. Rather, it is driven by the output of the foveation DNN, which is fed directly by our eye model's ONV.

In a series of experiments, we evaluate the performance of our biomechanical eye model and its associated neuromuscular motor control system in foveating visual targets, fixating them, and pursuing their motions, and we compare it to human eye movement data. Furthermore, we demonstrate the life-like operation of our eye model (binocularly) as an important component of a complete sensorimotor control system for an anatomically-accurate biomechanical human musculoskeletal model.

2 RELATED WORK

In the context of human animation, the eyes are arguably the most influential facial features [Looser and Wheatley 2010], and the emotions of characters can often be conveyed by means of gaze and eye movements. Various research efforts have focused on realistic eye animation. They can be characterized as physics-based, statistical, or data-driven modeling [Ruhland et al. 2014].

In their pioneering work, Lee et al. [2002] took a statistical approach to the modeling of saccadic movements of the eye using eye-tracking data. Lee and Terzopoulos [2006] demonstrated a procedural kinematic eye movement synthesis method in the context of a biomechanical neck-head-face system. Komogortsev et al. [2013] employed a 2D linear oculomotor model with a simplified pulse-step neuronal control signal to account for horizontal and vertical movements of the eyeball.

In robotics, Lesmana and Pai [2011] and Lesmana et al. [2014] implemented a robotic oculomotor pulse-step controller that learns an internal model of the eye plant from measurements and produces realistic eye movements. Other robotic eye models and controllers are cited in their papers.

In the context of ophthalmological research, computational models of the human eye and its extraocular muscles may be traced back to pioneering work by Robinson [1964]. Several models have been developed to support the planning of strabismus corrective surgeries. For example, Buchberger [2004] implemented an anatomically accurate model of the eye, including a muscle-force prediction model and a quasi-static model that utilized optimization to balance the muscle forces. Wei et al. [2010] developed the first fully dynamic biomechanical model of human eye movement for clinical applications, which accounted for the nonlinear kinematics of the oculomotor plant's geometry and extraocular muscle mechanics using strands.

---

[1] *Fixation:* After the eyes saccade to foveate the visual target, fixation maintains visual attention on the target; the eye seems still and stable, but small oscillations and drifts are typically observed.

With regard to modeling the optics of the eye for use in ophthalmology, Greivenkamp et al. [1995] modeled visual acuity using ray tracing, and Fink and Micol [2006] developed an optical model of the eye that employed ray tracing to visualize retinal images that account for various eye conditions such as myopia, hyperopia, astigmatism, cataract, etc. Wei et al. [2014] accelerated the ray tracing through a static geometric eye model by parallelization on the GPU.

The primary motivation for our work is the importance of eye models in binocular visuomotor systems for autonomous, physics-based virtual characters. Lee and Terzopoulos [2006] demonstrated a basic visuomotor system for a fully dynamic, biomechanical, muscle-actuated human neck-head-face model. Yeo et al. [2012] presented a more sophisticated visuomotor system, albeit for a kinematic anthropomorphic virtual character. Both these systems lacked biologically-motivated visual processing.

Our approach to eye modeling and simulation is inspired mainly by our recent work [Nakada et al. 2018b] that introduced an elaborate, biomimetic sensorimotor system for a full-body, biomechanical human musculoskeletal model whose neural control mechanisms are based on deep learning. Unlike the uniform, Cartesian grid arrangement of most artificial imaging sensors, visual sampling in the primate retina is known to be strongly nonuniform [Schwartz 1977]. Accordingly, our human model in [Nakada et al. 2018b] had biologically inspired eyes with foveated retinas (see also [Nakada et al. 2018a]). However, they were overly simplistic, purely kinematic eyes modeled as pinhole cameras. Given the nonuniform distribution of photoreceptors, the retinas optically sensed the 3D scene through ray tracing, which best emulates how the human retina samples scene radiance from the incidence of light on its photoreceptors. However, this and all earlier eye models for use in computer graphics largely ignored the organs of biological eyes that are essential to human vision, namely the cornea, iris/pupil, and lens. The substantially more realistic, muscle-actuated biomechanical eye model we introduce in the present paper includes functional simulations of these optical organs. Advancing beyond the simple pinhole-camera eye model to a realistic model—one that incorporates a finite-aperture pupil between a transparent cornea and lens that refract incoming light—necessitates the use of more advanced ray tracing techniques. Furthermore, we equip the retina in our eye model with a substantially greater number of photoreceptors, and this necessitates the use of a new type of locally-connected, irregular DNN in the sensory system of our eye model. Last but not least, through the use of machine learning, we tackle the more difficult neuromuscular visuomotor control problems that our elaborate eye model presents as part of the sensorimotor system of an advanced biomechanical human musculoskeletal model.

In summary, our novel, fully functional model of the human eye is unprecedented in its unique combination of features.

3 BIOMIMETIC EYE MODEL

Our geometric eye model (Fig. 1a), which is commercially available from TurboSquid.com, consists of the cornea, iris, lens, sclera, retina, optic nerve, ciliary muscle, and 6 extraocular muscles—the medial rectus, lateral rectus, superior rectus, inferior rectus, superior oblique, and inferior oblique (Fig. 2). We model the virtual eye as a sphere of radius 12 mm that can be rotated with respect to its center around its
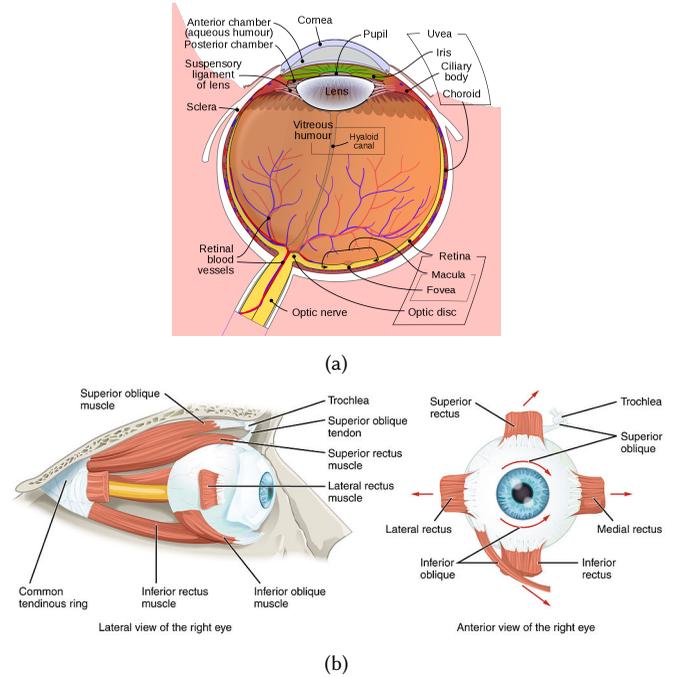

(a)

(b)

Fig. 2. (a) Schematic diagram of the right eye (superior view of a transverse section). (b) Lateral and anterior views of the eye showing the extraocular muscles and (right) eye movements they induce (red arrows). (Image sources: (a) https://commons.wikimedia.org/wiki/File:Schematic_diagram_of_the_human_eye_en.svg; (b) https://commons.wikimedia.org/wiki/File:1412_Extraocular_Muscles.jpg.)

vertical y axis by a horizontal angle θ, around its horizontal x axis by a vertical angle ϕ, and around its view axis z by a torsional angle ψ [Haslwanter 1995]. The eyes are in their neutral positions looking straight ahead when $\theta = \phi = \psi = 0°$. The horizontal and vertical fields of view are $167.5°$.[2]

Mechanically, the eyeball is treated as a rigid body of mass 7.5 grams with a 3 rotational degree-of-freedom ball joint, which is actuated by the extraocular muscles. Appendix A provides the equations of motion of the eyeball.

3.1 Retina

The retina is the light-sensitive innermost layer of shell tissue of the eye. It is populated by a nonuniform distribution of photoreceptors (cones and rods). The density of cones decreases radially from the fovea toward the periphery. To emulate foveated perception in our model, we use the noisy log-polar photoreceptor distribution [Nakada et al. 2018a,b]. We include a larger number of photoreceptors—14,400

---

[2] According to human physiological data [Bekerman et al. 2014; Dagnelie 2011; Riordan-Eva and Cunningham 2011], the transverse size of an average eye is 24.2 mm and its sagittal size is 23.7 mm. Its average mass is 7.5 g. The approximate field of view of an eye is 30° to superior, 45° to nasal, 70° to inferior, and 100° to temporal. When the two eyes are combined, the binocular field of view becomes approximately 135° vertically and 200° horizontally.

in total—situated at:

$$\boldsymbol{d}_k = e^{\rho_j} \begin{bmatrix} \cos \alpha_i \\ \sin \alpha_i \end{bmatrix} + \begin{bmatrix} \mathcal{N}(\mu, \sigma^2) \\ \mathcal{N}(\mu, \sigma^2) \end{bmatrix}, \quad \text{for } 1 \le k \le 14{,}400, \quad (1)$$

where $0 < \rho_j \le 40$, incremented in unit steps, and $0 \le \alpha_i < 360°$, incremented in $1°$ steps, and where $\mathcal{N}$ denotes additive IID Gaussian noise. We set mean $\mu = 0$ and variance $\sigma^2 = 0.0025$. The black dots on the retinal surface in Fig. 1a depict the placement of photoreceptors according to this noisy log-polar model. Other placement patterns are readily implementable, including more elaborate procedural models [Deering 2005] or photoreceptor distributions empirically measured from biological eyes [Grady 2004].

3.1.1 Computing Irradiance at the Photoreceptors. The irradiance at any point on the hemispherical retinal surface at the back of the eye may be computed using ray tracing. Multiple sample rays from the position of each photoreceptor are cast through the finite-aperture pupil and out into the 3D virtual world where they recursively intersect with the visible surfaces of virtual objects and query the virtual light sources in accordance with the Phong local illumination model. The irradiance values returned by these rays determine the light impinging upon the retina that is captured by that photoreceptor. The contribution of each sampling ray is combined via a weighted sum to determine the preliminary irradiance value at each photoreceptor. Note that the weighted sum does not correspond to the average, as the sum of the weights is allowed to exceed unity. Following this computation, we approximate the Stiles-Crawford Effect [Stiles et al. 1933] by applying a Guassian multiplier that is maximized for photoreceptors at the center of the fovea. The final irradiance value is thus the product of the preliminary weighted sum and the Stiles-Crawford approximation multiplier, clamped to a maximum value of 1.0 for full activation. Fig. 1b illustrates the retinal imaging process.[3]

3.2 Iris and Pupil

The iris controls how much light enters the eye through the pupil. The normal size of the human pupil is 2–4 mm diameter under bright light, and 4–8 mm in the dark. We model the pupil as a circle with diameter ranging from 1 mm to 4 mm, and the iris as an annular region with a 6 mm external radius.

The annular iris sphincter muscle is modeled as a composition of eight individual Hill-type muscles [Lee et al. 2009] conjoined in an octagonal pattern with a manipulatable force point at each vertex (Fig. 3a). Each vertex is connected to its corresponding closest point on the iris perimeter via a 1D prismatic joint, with the screw axis always oriented toward the center of the complex. A passive spring element, whose natural length is achieved when the pupil is maximally dilated, is also situated along each joint. The pupil diameter at any given time step is computed as the average distance between each pair of opposing vertices of the sphincter complex.

---

[3]The number of rays emitted by each photoreceptor depends on its position on the retina and the size of the pupil. Under average lighting conditions, a photoreceptor would typically cast from 15 to 20 rays. Therefore, with 14,400 photoreceptors, on average the retina performs a multi-refraction trace of approximately $14{,}400 \times 17 = 244{,}800$ rays. On a 3.4 GHz Intel Core i7-6700 CPU, the ray tracing computation requires from 2.2 s for a constricted pupil to 2.5 s for a dilated pupil. By comparison, the computations for the remainder of the oculomotor system take only about 60ms per time step.
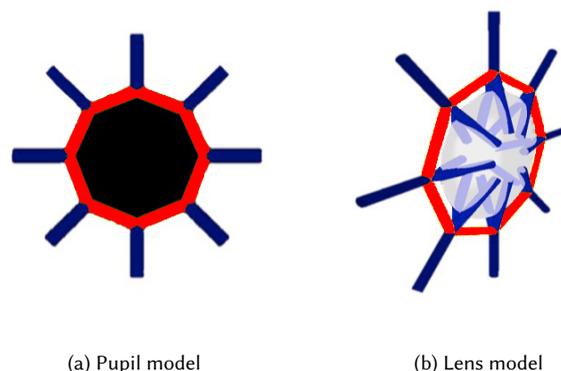
(a) Pupil model   (b) Lens model

Fig. 3. Biomechanical pupil and lens models. (a) The active sphincter muscles (red) are suspended from the perimeter via a series of passive spring elements (blue). (b) The active ciliary muscles (red) are conjoined with the remaining ciliary body via a series of passive spring elements (exterior blue). The passive zonular tendons (interior blue) suspend the lens body to the ciliary muscle.

During pupil constriction, the same activation signal is sent to all muscle actuators comprising the sphincter. The resulting contractile forces pull each of the eight vertices towards the center of the complex, which reduces the pupil diameter. During pupil dilation, the passive spring elements pull the sphincter vertices towards the perimeter of the complex, which increases the pupil diameter. The biological eye has both sphincter and dilator muscles in the iris-pupil complex; however, our model reasonably approximates the reciprocal relationship of the muscle activations with just a single active muscle group.

3.3 Cornea, Aqueous Humor, Lens, and Vitreous Humor

These four components refract light rays, resulting in a focusing of the image through the finite-aperture pupil. In the human eye, the greatest amount of refraction (approximately two-thirds) is caused by the cornea, which has a fixed focus. The incoming light is then refracted by the aqueous humor—a fluid that fills the iris-pupil complex. Finally, the lens makes the necessary deformation to refocus light from nearby objects onto the retina, accounting for the additional refractory properties of the gelatinous vitreous humor inside the eyeball. The aqueous and vitreous humors are modeled as simple block refractive media, while the remaining organs are elaborated upon subsequently. Appendix D tabulates the refractive index of each medium.

The cornea is modeled as a negative meniscus lens with non-uniform radii of curvature. More specifically, its outer face has a smaller radius of curvature than its inner face closer to the iris-pupil complex. A uniform index of refraction is used throughout the entire organ to reduce the number of refraction computations. Spherical approximations for the two surfaces are used for the purposes of intersection computations in order to reduce to $O(1)$ the otherwise $O(\log n)$ complexity of a K-D tree search normally associated with a generic polygon mesh. This approximation is a reasonable tradeoff in our model as only the rays that are admitted by the pupil are needed for computation, and the majority of such rays are highly

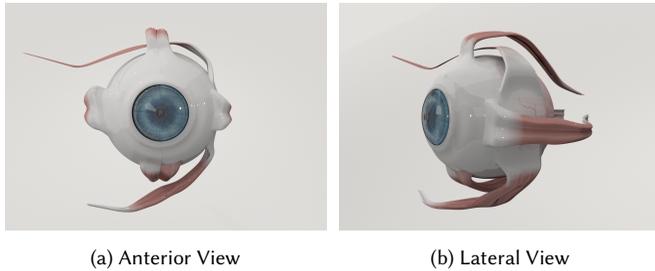(a) Anterior View   (b) Lateral View

Fig. 4. Geometric model of the eye and its extraocular muscles.

concentrated near the central region of the body; therefore, the additional accuracy gained by a more accurate geometry would largely be wasted in the scope of our model. Consistent with the biological eye, the cornea is treated as a non-deformable body.

The lens is modeled as a volume-preserving deformable body with initial conditions determined by the approximate rest length radius of the surrounding ciliary muscle and a baseline minimum thickness (Fig. 3b). The body is connected to the surrounding muscle via a series of passive tendons known as zonules. When the ciliary muscle is completely deactivated, tension in the zonular tendons is maximized. The collective tension across all of the zonules generates a net compression force on the lens body, which in turn causes the lens to compress and achieve its minimum thickness. Conversely, at full activation the lens' restorative internal force causes the body to expand until an equilibrium is reached between the restorative force and the steadily building zonular tension caused by the expansion. The pressure is modeled as a series of damped springs between the center of mass of the lens and the zonule attachment points, with the springs having an approximate rest length just slightly beyond the point at which equilibrium is achieved with the zonular tendons. We model the pressure via a finite element method to exclude any computations that may apply to portions of the surface that are not directly affected by the ciliary muscle activation. Spherical approximations are used to estimate the radius of curvature for both surfaces of the organ; however, as is the case in the biological eye, the radii of curvature of the surfaces typically are not equal. We make this approximation for the lens for reasons similar to the cornea, and similarly use a uniform index of refraction for the entire lens body interior to reduce the number of refraction computations.

The area of the lens actually used for photoreceptor sampling is determined by the radius of the pupil. Once the area has been determined, a set of sampling points is designated on the surface. Each retinal photoreceptor then emits a single ray passing through each of the sampling points. Therefore, as the pupil expands, each photoreceptor collects information from more sampling rays. The total number of rays cast during any simulation time step is therefore dependent on the luminosity of the environment, which determines the size of the pupil.

3.4 Extraocular Muscles

The human eyeball is protected by the orbital bones of the skull. Fig. 2 shows an anterior diagram of the right eye, which reveals the positions of the six extraocular muscles. The ocular nerves that

innervate these muscles pass through an opening in the superior orbital fissure of the eye. The eyeball is tightly suspended in the orbit and its motion is dictated by the direction of the six extraocular muscles that surround it.

The 4 rectus muscles as well as the superior oblique all originate from a localized area towards the back of the eye socket of the skull called the orbital apex, while the inferior oblique originates from the maxilliary bone. The muscles that contribute to the movements of the eye are principally the lateral and medial recti for horizontal movements, the superior and inferior recti for vertical movements and the superior and inferior obliques for torsion movements.

Fig. 4 illustrates the eye model. Note that all 4 of the rectus muscles originate from a similar position towards the orbital apex, while the oblique muscles are shown to originate elsewhere. The extraocular muscles are modeled using the Hill-type muscle model [Lee et al. 2009]. Details about the muscle actuators are presented in Appendix B.

Away from the primary forward position, eye movements may be grouped into 8 tertiary gazes, but each gaze orientation involves strong activations of only a couple extraocular muscles, with the remaining muscles contributing weakly. Fig. 2b illustrates these poses along with the primary muscle drivers of each pose. Direct vertical movements are effected by a co-contraction of a rectus and oblique muscle, whereas horizontal movements are driven solely by the medial and lateral recti.

4 NEUROMUSCULAR OCULOMOTOR CONTROL

Fig. 5 presents an overview of our neuromuscular oculomotor control system for our model of the eye, revealing its sensory and motor subsystems, and the figure caption summarizes the function and connectivity of its 4 neural controllers, which include 2 SNNs and 2 DNNs.

4.1 Optic Nerve Vector

The retinal RGB "image" captured by the eye is output for further processing down the visual pathway, not as a 2D array of pixels, but as a 1D vector of length $14,400 \times 3 = 43,200$, which we call the Optic Nerve Vector (ONV). The raw sensory information encoded in this vector feeds the neural networks that directly control pupil size, lens accommodation, and eye movements. More specifically, as shown in Fig. 5, the ONV from the retina is fed to a pupil control SNN, which outputs a control signal (a change in muscle activation) to the iris that constricts or dilates the pupil. The ONV also feeds a lens control SNN, which outputs a control signal to the ciliary muscle that controls lens deformation, hence optical focus. Finally, the ONV feeds a foveation DNN, whose output drives a neuromuscular motor controller DNN that innervates the 6 extraocular muscles to produce eye movements.

4.2 Motor Subsystem

The extraocular muscle actuators are controlled by specifying for each of them a time-varying, efferent activation signal $a(t)$. Referring to Fig. 6, the output of a neuromuscular motor controller is

$$\boldsymbol{a}(t + \Delta t) = \boldsymbol{a}(t) + \Delta \boldsymbol{a}(t). \tag{2}$$
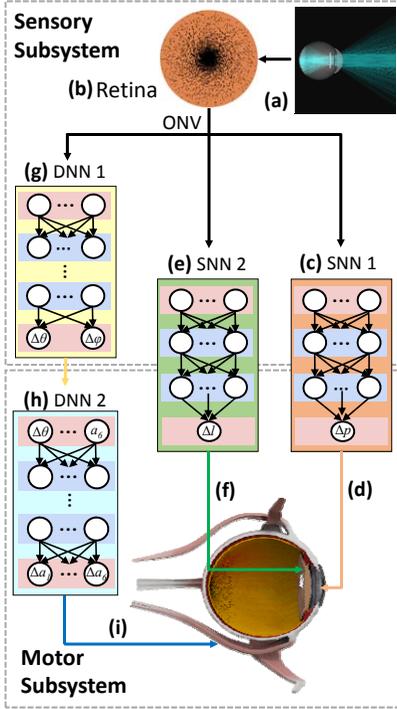
Fig. 5. Oculomotor control system architecture for the (left) eye, showing the modular neural network controllers in the sensory subsystem (top) and motor subsystem (bottom), including two DNNs (1,2) and 2 SNNs (1,2), of 4 types, colored orange, green, yellow, and blue. Within the sensory subsystem, (a) each photoreceptor on the retina casts rays into the virtual world to compute the irradiance at the photoreceptor. (b) The arrangement of the 14,400 photoreceptors (black dots) on the foveated retina. The retina outputs a 43,200-dimensional Optic Nerve Vector (ONV). The (orange) iris SNN (c) inputs the ONV and outputs a constriction/dilation activation to the iris muscles. The (green) lens (e) controller inputs the ONV and outputs a lens muscle activation to control focal accommodation. The (yellow) foveation DNN (g) inputs the ONV and outputs angular discrepancy estimates to the (blue) neuromuscular motor controller DNN (h), which outputs activations to control the 6 extraocular muscles in order to foveate and pursue visual targets.

Note that the feedback loop formed by the vector a of muscle activation signals makes the neuromuscular controller a Recurrent Neural Network (RNN).

4.2.1 Extraocular Muscle Control.
The extraocular muscle motor control DNN (Fig. 6) takes as input the desired change in gaze angles, $\Delta\theta$ and $\Delta\phi$, and produces muscle activation adjustment signals Δa. Note that the torsional gaze angle ψ is not explicitly fed as input on the premise that the torsional angle is determined as a result of a neural, involuntary constraint rather than through voluntary control. This is justified by Donders law and Listing's law [Tweed et al. 1990], which together state that the orientation of the eye in a particular gaze position is always the same and, moreover, that it can be explicitly computed; therefore, there is a unique torsion angle for each voluntary gaze position.
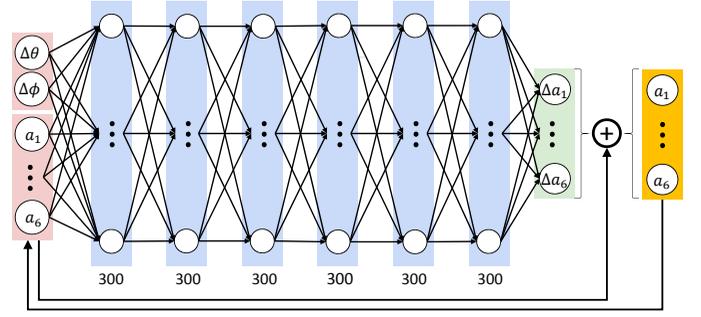
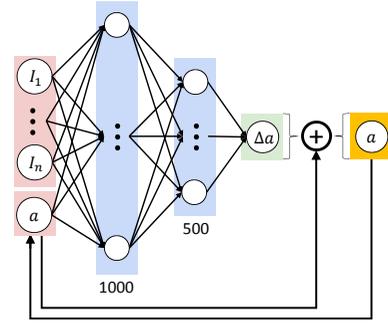Fig. 6. Fully-connected, recurrent neuromuscular motor control DNN architecture for the extraocular muscles.

Fig. 7. Fully-connected, recurrent neuromuscular motor control SNN architecture for the intraocular sphincter muscle of the iris.

DNN Architecture: As shown in Fig. 6, the input layer of the eye's voluntary motor DNN consists of 8 units that include 2 units for the eye rotation target angular discrepancies, $\Delta\theta$ and $\Delta\phi$, and 6 units for the activations a_i, $1 \le i \le 6$, of the six extraocular muscles. The output layer consists of 6 units providing the muscle activation adjustments Δa_i. There are 6 hidden layers, each containing 300 units. The DNN is constructed using Rectified Linear Units (ReLUs).[4]

Offline Synthesis of Training Data: To train the voluntary motor DNN, we employ our biomechanical eye model to synthesize training data as follows: Specifying a target orientation for the eye yields angular discrepancies $\Delta\theta$, $\Delta\phi$, between the current eye orientation and the target eye orientation. With these angular discrepancies and the current muscle activations a_i as the input, we compute for the biomechanical eye model inverse kinematics followed by inverse dynamics with muscle optimization (Appendix C provides additional details). This determines muscle activation adjustments Δa_i, $1 \le i \le 6$, which incrementally reduce the angular discrepancies. These adjustments serve as the desired output of the DNN.

The *offline* training data synthesis requires approximately 0.1 s per input-output pair on a 3.4 GHz Intel Core i7-6700 CPU with 8 GB Ram. Repeatedly specifying random target orientations and repeating the

---

[4]The architecture of our extraocular muscle motor DNN is consistent with that of the neuromuscular motor controllers of the biomechanical human musculoskeletal model that we developed in [Nakada et al. 2018b].
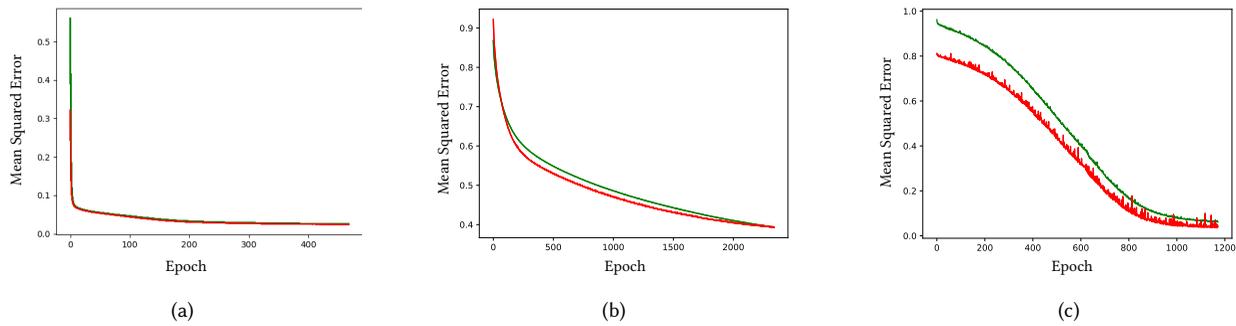
Fig. 8. Progress of the backpropagation training process for (a) the eye motor DNN, (b) the iris SNN, and (c) the lens SNN on the training (green plot) and validation (red plot) datasets.

above synthesis process, we synthesized a large training dataset of 1M input-output pairs in approximately 28 hours.

The motor subsystem datasets were all randomly divided, with 90% of the points comprising the training set and the other 10% comprising the validation set.

DNN Training: The initial weights of the DNN are sampled from the zero-mean normal distribution with standard deviation $\sqrt{2/fan\_in}$, where *fan_in* is the number of input units in the weight tensor [He et al. 2015]. We employ the mean squared error loss function and apply Adaptive Moment (Adam) estimation as the stochastic optimization method, using the following parameters: $lr = 10^{-6}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$, $\alpha = 0.001$, where β_1 and β_2 represent the exponential decay rate for momentum estimates taking an average and an average of squared gradients, respectively, ϵ prevents divide-by-zero errors, lr is the learning rate, and α is the step size. An early stopping condition—negligible improvement for 10 successive epochs—is enforced to avoid overfitting.

The backpropagation DNN training process converged to a small error in 482 epochs. Fig. 8a plots the progress of the training process. After the DNN is trained, it serves as the *online* neuromuscular motor controller for the extraocular muscles.

4.2.2 Iris Sphincter Muscle Control. Fig. 7 shows the architecture of the recurrent neuromuscular motor controller of the iris.

SNN Architecture: The SNN takes as input the ONV—i.e., the photoreceptor irradiance values I_i, $i \leq 1 \leq 43,200$—as well as the current activation a of the iris sphincter. Once trained, the SNN produces the sphincter activation adjustment signal Δa, which is then added to the activation. The 1-unit output layer yields the updated activation a.

The SNN differs from the extraocular muscle motor control DNN in its use of only two hidden layers, but with more units per layer given the high dimensionality of the ONV—1,000 units in the first hidden layer and 500 units in the second. The fully-connected network uses Rectified Linear Units (ReLUs) as the activation function.

Offline Synthesis of Training Data: To train the iris sphincter SNN, we sampled from the space of all potential visual target positions

and environment luminosities. The eye gaze was fixed and only the target position was varied.

First, the target was randomly placed within the field of view and the illumination on both the target and the ambient illumination of the surrounding environment were varied. The eye regulates the amount of incoming light by adjusting the pupil in order to maintain a constant net irradiance summed over all the photoreceptors. We compute the difference between the current and ideal net irradiance, and from that difference derive the corresponding sphincter activation adjustment Δa (refer to Appendix D).

The ONV concatenated with the sphincter activation is the input and Δa is the output of the input-output training pair, which takes approximately 1.9 s to compute. Repeating this process over a variety of visual target positions and illumination intensities, we synthesized a training dataset of roughly 200k input-output pairs in approximately 106 hours.

SNN Training: The SNN is trained *offline* in a manner similar to the motor DNN, making use of both the He initialization and Adam optimizer for the parameter initializations and gradient descent optimization, respectively. The parameters, parameter values, and the enforced stopping condition remain the same. The training data synthesis takes 0.1 s of computational time to solve for 0.02 simulation seconds on a 2.4 GHz Intel Core i5 CPU with 8 GB Ram.

The backpropagation SNN training process converged to a small error after approximately 2,500 epochs. Fig. 8b shows the progress of the training process. After the SNN is trained, it serves as the *online* neuromuscular motor controller for the iris sphincter muscle.

4.2.3 Lens Ciliary Muscle Control. The architecture of the neuromuscular motor controller of the lens is identical to that of the iris (Fig. 7). It too is a fully-connected network driven by the ONV, and the activation a that it produces controls the lens ciliary muscle. The retinal "image" of an object is in focus when the computed focal point (explained subsequently) coincides with the surface of the retina, which we found to be a reasonable assumption based on the observation of ONV patterns (Fig. 9). The trained SNN in the recurrent controller produces Δa, which adjusts the ciliary muscle activation a to induce focal accommodation.
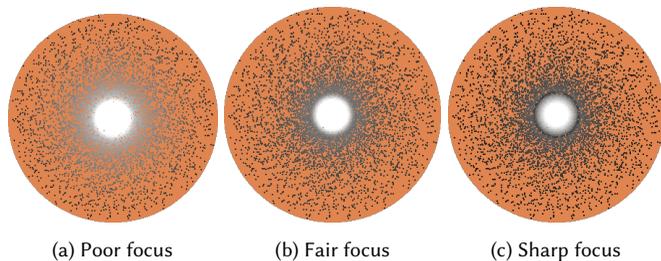
(a) Poor focus   (b) Fair focus   (c) Sharp focus

Fig. 9. Comparison of ONV responses when fixating a white ball under ambient illumination with different lens activations.

Fig. 10. The LiNet architecture. The receptive fields are depicted as circles in the retinal domain.

Offline Synthesis of Training Data: Lens accommodation principally occurs when the gaze is fixated on a visual target; therefore, it is reasonable to sample only the space of target positions near the eye's optical axis. Thus, only the distance of the visual target was varied. Clipping and culling within the eye's view frustum reduces the number of object vertices to consider. Rays are cast along vertex normals with greater bias given to vertices on the forward facing polygons of the object. Those rays that succeed in refracting through the cornea and passing through the pupil converge at some focal point along the optical axis, which is computed via a standard line-ray intersection method, and the signed distance along the optical axis between the computed focal point and the retinal surface is determined.

To increase the robustness of the lens ciliary SNN to changes in illumination, we also sampled the space of potential environment luminosities. The pupil adjusts to the irradiance, which varies the number of rays used to compute the focal point.

To generate training data, the ONV is concatenated with the ciliary muscle activation a to form the input. To compute the associated desired output, which is the ciliary muscle activation adjustment Δa, we determined after some experimentation that it is beneficial to nonlinearly scale the signed distance using an exponential function (refer to Appendix D and Equation (8)). Input-output training data pairs were thus computed by repeating this process over a variety of visual target positions and illumination intensities. Each training data pair can be synthesized in approximately 2.2 s. In this way we synthesized a training dataset of roughly 200k input-output pairs in approximately 122 hours.

SNN Training: The SNN is trained *offline* in a manner similar to the aforementioned neural networks, making use of both He initialization and the Adam gradient descent optimizer. The parameter values and the enforced stopping condition are the same as before. The training data synthesis process takes 0.1 s of computational time to solve for 0.02 simulation seconds on a 2.4 GHz Intel Core i5 CPU with 8 GB Ram.

The backpropagation SNN training process converged to a small error after approximately 1,500 epochs. Fig. 8c shows the progress of the training process. After the SNN is trained, it serves as the *online* neuromuscular motor controller for the lens ciliary muscle.

4.3 Sensory Subsystem

Since the two previously described SNNs that perform neuromuscular control of the iris and lens directly take as input the ONV from the
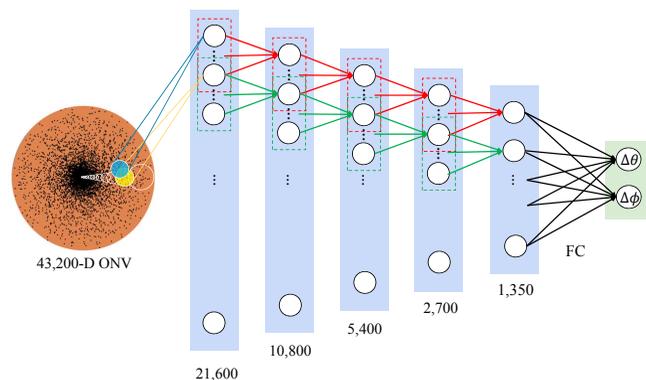
eye, they have both sensory and motor functionality; thus, SNN 1 and SNN 2 are depicted in Fig. 5 as bridging the sensory and motor subsystems of our oculomotor control system. Next, we will describe the sensory subsystem's foveation DNN 1 in Fig. 5.

4.3.1 LiNet Foveation DNN. There are approximately 5 million cones in the human retina [Deering 2005]. With this many photoreceptors in its ONV input, training a fully-connected neural network—as we did in [Nakada et al. 2018b] with only 3,600 retinal photoreceptors—is intractable due to the overwhelming memory requirements. Furthermore, the popular Convolutional Neural Network (CNN), which is designed for conventional, regularly-sampled image arrays, is inappropriate for use with our irregular distribution of retinal photoreceptors. To overcome this obstacle, we propose a novel, locally-connected, irregular DNN architecture, which we call a "LiNet", that effectively reduces memory consumption, thus enabling the incorporation of many more retinal photoreceptors. Fig. 10 illustrates our LiNet architecture.

For a traditional fully-connected neural network, each unit in a hidden layer is connected to all the units in the previous layer. By contrast, for a locally-connected network, each unit in a hidden layer is connected to only a fixed number of neighboring units in the previous layer, which is known as its receptive field. This is similar to a CNN, but unlike a CNN for which all of the receptive fields of units in a hidden layer share the same weights, the receptive fields of a LiNet can have different weights, as is necessary for irregular distributions of photoreceptors.

Offline Synthesis of Training Data: To train the foveation LiNet, we sample from the space of all potential environment luminosities, pupil states, lens states, and visual target positions, where the eye gaze is fixed and only the target position is varied.

First, the target is placed randomly within the field of view and the local illumination of both the target and the ambient illumination of the environment are altered. Next, the iris sphincter and lens ciliary muscles are assigned normalized random activations. Due to the substantially higher dimensionality of the sampling space, the target location, illumination, sphincter activation, and ciliary activation
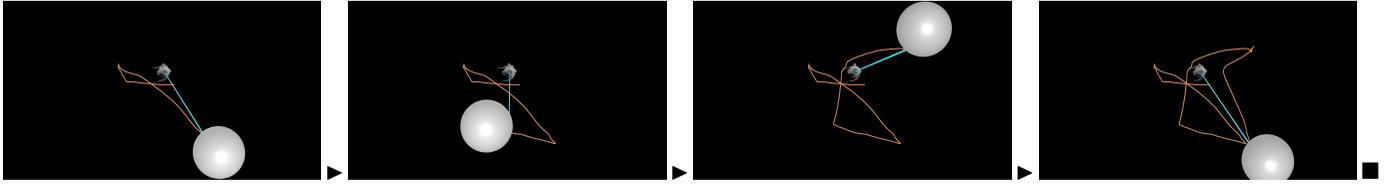
Fig. 11. Sequence of frames from a simulation of saccadic movement dynamics with integrated motor and sensory DNN & SNNs. The blue trace indicates eye movement during the transition from one position of the visual target (white sphere) to the next. The red line indicates the gaze direction.

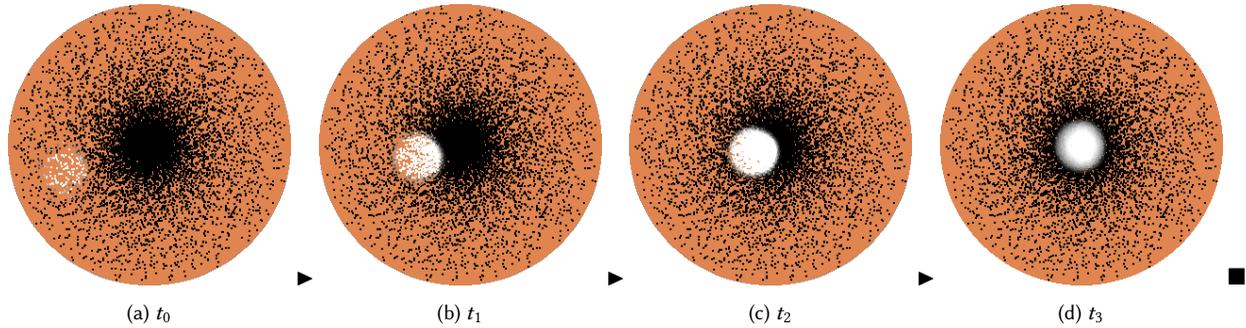(a) t_0 (b) t_1 (c) t_2 (d) t_3

Fig. 12. Time sequence (a)–(d) of photoreceptor responses in the left retina during a saccadic eye movement that foveates and tracks a moving white ball. At time t_0 the ball becomes visible in the visual periphery, at t_1 the eye movement is bringing the ball towards the fovea, the moving ball is viewed within the fovea at time t_2, albeit slightly off center, and at time t_3 the ball is fully centered on the foveal region.
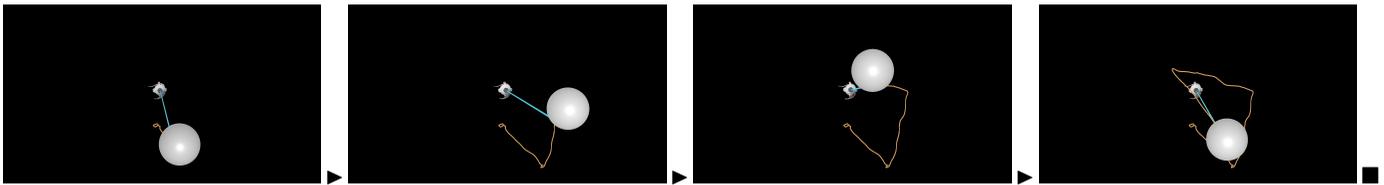

Fig. 13. Sequence of frames from an animation of smooth pursuit movement dynamics with motor and sensory DNN & SNNs. The blue trace indicates eye movement during the transition of the visual target (white sphere) from one position to the next. The red line indicates the gaze orientation.

were changed in every frame in order to collect a broader variety of samples.

To synthesize an input-output training data pair, the ONV serves as the input and the associated desired eye rotation angular adjustments $\Delta\theta$, $\Delta\phi$ serve as the output. Each input-output training pair can be synthesized in approximately 1.2 s. Repeating this process over a variety of visual target positions, illumination intensities, sphincter activations, and ciliary activations, we synthesized a training dataset of roughly 200k highly varied input-output pairs in approximately 67 hours.

5 EXPERIMENTS AND RESULTS

5.1 Dynamic Sensorimotor Control

The following sections present experimental evidence that our neuromuscular oculomotor controller produces realistic saccadic and smooth pursuit eye movements in response to visual stimuli processed by its trained sensory DNN and SNNs.

5.1.1 Saccadic Movement. Fig. 11 shows frames from a simulation that demonstrates noise in both saccadic movement and fixation, which results from the foveation DNN's observation of the visual target. This is characteristic of biological eye behavior, against which we compare below.

5.1.2 Fixation. We performed an experiment with our sensorimotor control model to compare behavior during fixation. Fig. 12 illustrates a typical fixation movement. After a white ball visual target appears in the periphery (Fig. 12a), the eye rotates to foveate the target (Fig. 12b–d). The eye exhibits small yet visible corrections (Fig. 12c,d), as the target is imperfectly centered in the fovea during fixation. This produces a cluster of gaze positions for a particular visual target rather than a single fixed gaze, as can be seen in the integrated saccadic movement traces in Fig. 11. We validate this behavior with a comparison to that of actual human subjects below.

5.1.3 Smooth Pursuit. Fig. 13 shows frames from a typical experiment with the smooth pursuit of a moving white ball visual target. As with the saccadic movement comparison, we note here that the

(a) Human (dis) (b) Kinematic model (c) Our DNN model

(d) Human (vel) (e) Kinematic model (f) Our DNN model

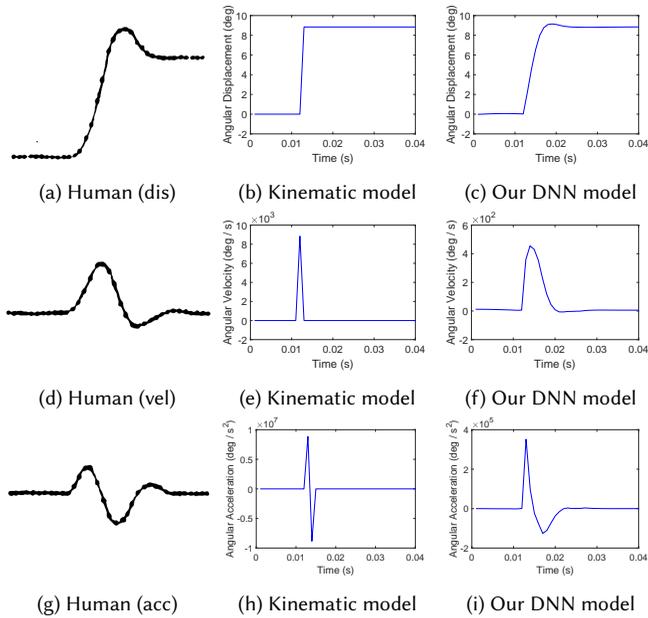(g) Human (acc) (h) Kinematic model (i) Our DNN model

Fig. 14. Comparison of ocular angular (a)–(c) displacement, (d)–(f) velocity, and (g)–(i) acceleration of saccadic movements for a human subject, the baseline kinematic eye model, and our biomechanical eye model with deep neuromuscular oculomotor control. (a) (d) (g) Plots, reproduced from [Thomas 1969], of human eye movement data over approximately 10 ms (since ordinate values were not provided, we can compare only the shapes of the curves).
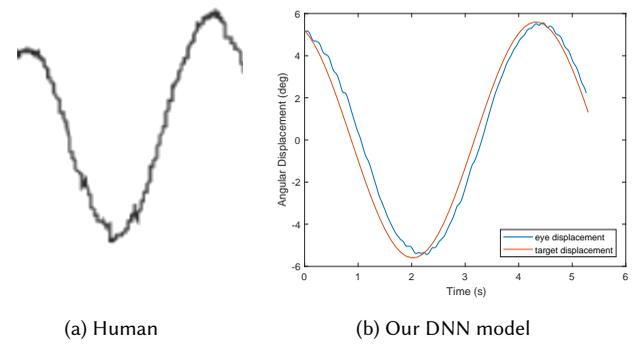

(a) Human (b) Our DNN model

Fig. 15. Plot of eye angular displacement over time in smooth pursuit tracking a visual target, which moves with a sinusoidal curve in the horizontal direction. (a) Human eye angular displacement data reproduced from [Schraa-Tam et al. 2008, Fig. 2, Subfig. 4] in the range from $t \approx 7\,$s to $t \approx 12\,$s. (b) Plot to similar scale showing the angular displacement of our eye model (blue) and of the visual target (red).
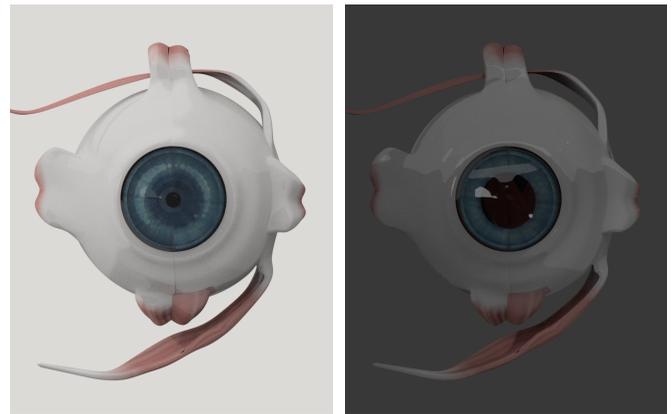

Fig. 16. Close-up view of pupil under different illuminations.

smooth pursuit line trace exhibits more noisy characteristics during transitions between various visual target positions.

5.2 Comparison with Human Eye Movement

We now compare the results of our model both against the baseline kinematic human eye model and against biological data from human subjects [Thomas 1969].

5.2.1 Saccadic Movements. Because it proved difficult to find raw data about actual muscle activations during saccadic movements, we instead focus on three accessible quantities—angular displacement, angular velocity, and angular acceleration of the eye.

Fig. 14 shows a comparison between the eye angular displacement, angular velocity, and angular acceleration curves of a human subject, the baseline kinematic eye model, and our simulated biomechanical eye model for a small saccadic movement of around 10–15 degrees. For angular displacement (Fig. 14a–c), the movement of our eye model resembles that of a human subject's eye movement, whereas the instantaneous shift of the kinematic control is unrealistic. For angular velocity (Fig. 14d–f) and acceleration (Fig. 14g–i), our eye model curves again resemble those of the human subject, whereas the transient spikes generated by the kinematic control system are unrealistic.

Our biomechanical eye model is therefore a clear improvement over the conventional, purely kinematic eye models of Lee and Terzopoulos [2006], Yeo et al. [2012], Nakada et al. [2018b], and other

researchers. We performed additional experiments with various random visual target positions and recorded the corresponding orientations of the eye. Fig. 17a shows results for 6 such random orientations for the θ and ϕ ocular angles.

5.2.2 Smooth Pursuit Movements. Fig. 15 shows a comparison between the eye angular displacement curves of a human subject measured by Schraa-Tam et al. [2008] and our simulated model for the smooth pursuit of a visual target moving in a sinusoidal pattern. As can be seen from the plots, the curve from our biomechanical eye model resembles that of the human subject's eye movements.

5.3 Pupil Accommodation

To test pupil accommodation, we placed a light source at random locations within the field of view of our eye model and changed its luminance.

Fig. 16 illustrates the dilation and constriction of the pupil in response to the varying luminosity of the surrounding environment.
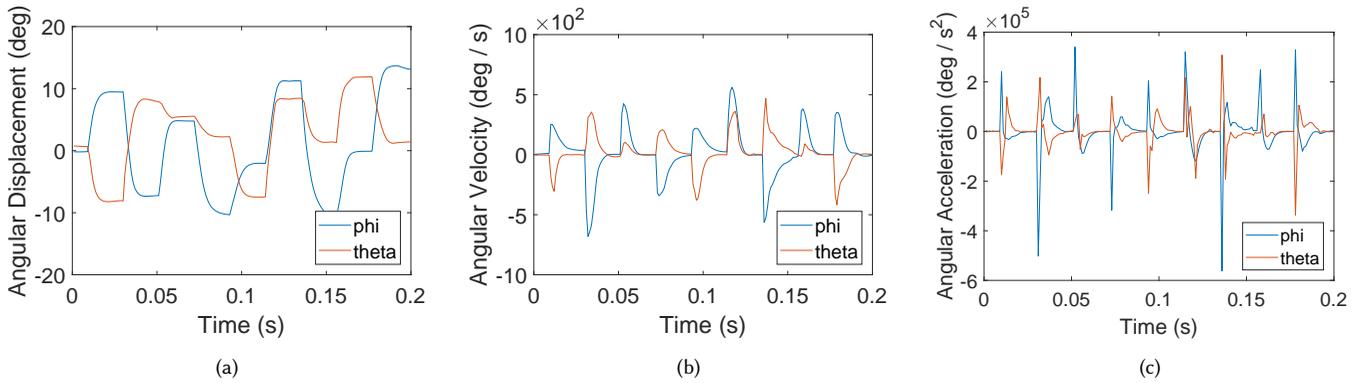
Fig. 17. Plots of horizontal (red) and vertical (blue) ocular angular (a) displacement, (b) velocity, and (c) acceleration in saccadic movement and fixation over 6 random target placements.
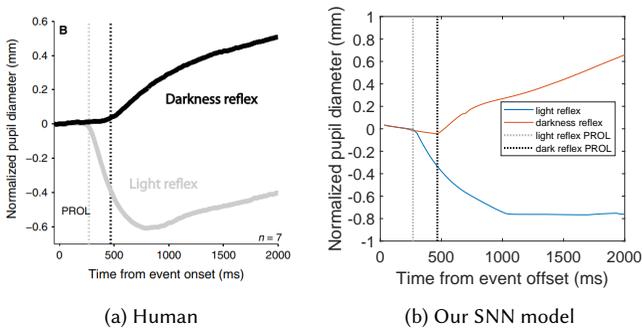

Fig. 18. Plot of pupil diameter over time in response to illumination transitions. (a) Human eye pupil diameter data reproduced from [Wang and Munoz 2014, Fig. 6, Subfig. B]. (b) Plot to similar scale showing the pupil diameter of our eye model under brighter (blue) and darker (red) luminance transitions.
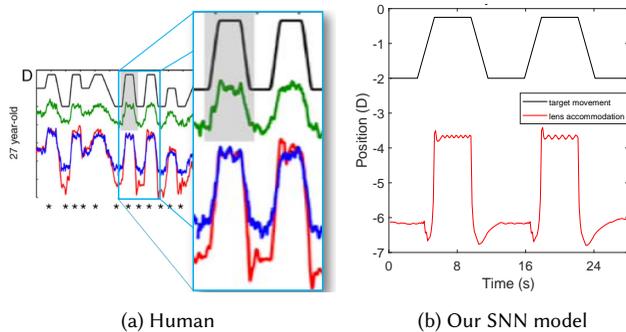

Fig. 19. Plot of diopter data in lens accommodation over time in response to the visual target movements along the visual axis (black). (a) Human eye lens accommodation (red) reproduced from [Bharadwaj and Candy 2008, Fig. 2D] for the duration of approximately 30 s (magnified region). (b) Plot to similar scale showing the lens accommodation (red) of our eye model.

Our experiments revealed that the pupil successfully performed realistic dilations and constrictions in response to the light entering the eye. The constriction and dilation rate varied commensurately with the rate of change of light intensity.

Fig. 18 shows a comparison between the pupil diameter curves of a human subject reported by Wang and Munoz [2014] and our simulated model subjected to background luminance changes. As can be seen from the plots, the curve from our biomechanical eye model resembles that of the human subject's pupil constriction and dilation. Consistent with the human data, we introduced an approximately 300 ms delay in the response of our iris controller to an increase in luminosity and an approximately 500 ms delay in its response to a decrease in luminosity.

5.4 Lens Accommodation

To test lens accommodation, we placed a visual target before our eye model and varied its distance from the eye.

Fig. 20 illustrates a lens accommodation simulation in which the ciliary muscle contracts in order to expand the lens body. The resulting increase in lens thickness caused the computed focal point to shift toward the back of the retina and consequently reduced the blur of the retinal image in a manner similar to Fig. 9. As the target depth was varied, the lens SNN maintained the focal point near the retinal surface. This behavior remained reasonably consistent even with the added complexity of iris-pupil accommodation, and we performed additional experiments to confirm reasonable online operation in varying environment luminosities as well.

Fig. 19 compares the diopter data curves of a human subject in lens accommodation as reported by Bharadwaj and Candy [2008] and our model for visual target movements along the visual axis. As can be seen from the plots, the responses of our biomechanical eye model resembles human lens accommodation.

5.5 Oculomotor Control in a Musculoskeletal Model

Finally, we validated the performance of our biomechanical eye model and its neuromuscular oculomotor controller within a full-body biomechanical human musculoskeletal model. Our human model is a completed version of the one in Nakada et al. [2018b] that had
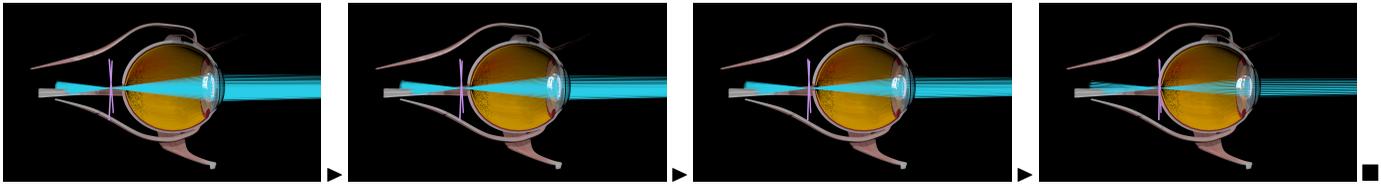
Fig. 20. Sequence of frames from an animation of lens accommodation. The lens deforms in order to align the focal point of the incoming rays (displayed as an 'X') with the retinal surface.

(a) Balanced, ready stance   (b) Reaching to the right   (c) Springing to the left   (d) Leaping to the right
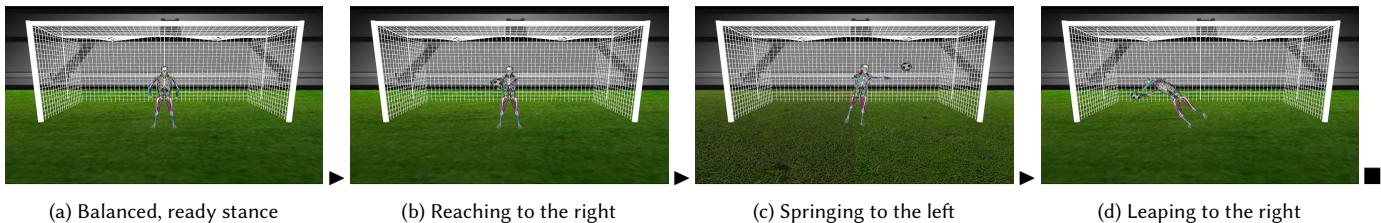
Fig. 21. Sequence of frames from a simulated soccer goaltending scenario involving an autonomous biomechanical human musculoskeletal model incorporating our biomimetic virtual eye models. Under deep neuromuscular oculomotor control, our goalie's eyes observe and persistently track incoming soccer balls by making saccadic foveation and smooth pursuit eye movements. The eye movements furthermore drive cervicocephalic head movements also under deep neuromuscular motor control. With additional deep neuromuscular motor controllers, our goalie controls its torso, arms, and legs to reach out and even leap at approaching balls to deflect them away from the goal.

an immobile spine and simple, kinematic eyes.[5] Afforded binocular visual perception by virtue of our eye model, the virtual human puts it to use in a simulated soccer goaltending scenario, illustrated in Fig. 21. Using its two biomechanical eyes, our virtual human successfully foveates and visually pursues a moving target, the soccer ball, while balancing its body in an upright ready stance. Our virtual human reacts to the ball's trajectory, reaching out with its arms and possibly leaping at the ball in order to deflect it away from the goal, and its two eyes continue to visually track the moving ball while its head and body are in motion, which demonstrates Vestibulo-Ocular Reflex (VOR).

6 DISCUSSION

Although not perfect, the inverse dynamics and neural network controlled models both demonstrated a reasonable degree of fidelity relative to biological findings. However, because we imposed no constraint on the optimization of these activations save for a lower bound on each muscle, there is nothing biasing the computation to favor certain muscles in particular eye orientations to ensure biological consistency. This of course is not critical if the goal is to achieve a model that works, but it is a concern if the goal is a model that accurately matches recorded extraocular muscle activation level data during real human eye movement.

We have stayed true to using a purely voluntary neuromuscular control scheme since there does not seem to exist an underlying muscle reflex control layer in the oculomotor system. This does not cause much difference in the behavior of the inverse dynamics model, but it introduces some noise into the neuromuscular control scheme. This is expected since, in an effort to avoid over-fitting, the

voluntary neural network does not necessarily learn every nuance in muscle activation transitory behavior. Moreover, the noise actually appears to be more natural, because it seems infeasible for the brain to compute inverse dynamics on a millisecond basis in order to obtain exact control solutions. Rather, fast compensation for errors and noise generated by the imperfect motion control is the more natural strategy, which has guided our approach.

7 CONCLUSIONS AND FUTURE WORK

Our primary contributions in this paper have been as follows: We have developed a novel biomechanical model of the eye with unprecedented biomimetic detail. We have also introduced a neuromuscular oculomotor controller for this model. The controller incorporates deep and shallow neural networks, which are trained offline using data synthesized by the eye model itself and, once trained, work efficiently online to control the eye's extraocular and intraocular muscles. The large number of nonuniformly-distributed photoreceptors on the retina of our model eye necessitated the use of a new, locally-connected, irregular deep neural network for foveation. Finally, we demonstrated the consistency of our deep neuromuscular oculomotor control framework relative to human eye movements, and we incorporated our biomechanical eye/oculomotor model into the sensorimotor control system of our advanced biomechanical human musculoskeletal model, which marks a new milestone in human simulation and animation.

Our biomechanical eye model provides an ideal test platform for observing the effects of muscle placements in eye actuation, and for ensuring consistency with the muscle-optimized inverse dynamics control results. In particular, we were able to test that full activation of any particular muscle indeed actuated the eye in accordance with biological principles. However, we employed a simple Hill-type

---

[5]The elaborate structural and operational details of our human model fall outside the scope of the present paper and are presented elsewhere.

muscle model and the pulleys (trochleae) were largely neglected. In future work, we plan to introduce a pulley model and potentially a more sophisticated muscle model, perhaps similar to the strand muscle model proposed by [Wei et al. 2010].

ACKNOWLEDGMENTS

A   EQUATIONS OF MOTION OF THE EYE MODEL

Following the formulation in [Lee et al. 2009], the equations of the motion for the eye can be written as:

$$M(q)\ddot{q} + C(q, \dot{q}) = P(q)f_c + J(q)^T f_e, \tag{3}$$

where q is a state vector comprising the rotational degrees of freedom of the eyeball, $\dot{q}$ and $\ddot{q}$ are the corresponding velocities and accelerations, M is the mass matrix, and C accounts for gravity and Coriolis forces. The Jacobian matrix J transforms the external forces to joint torques, the moment arm matrix P maps the muscle forces to the ball joint space, f_c denotes the contractile muscle forces, and f_e denotes the external forces. Equation (3) may be written compactly as $\ddot{q} = \phi(q, \dot{q}, \tau)$. We solve forward dynamics to compute ϕ by calculating $\ddot{q}$ from the generated torque produced by the muscle forces. Then, we use the implicit Euler time integration method to solve the linearized equations of motion. We compute velocity at the next time step, where Δt is the step size, by solving

$$\dot{q}(t + \Delta t) - \dot{q}(t) = \Delta t\phi(q(t + \Delta t), \dot{q}(t + \Delta t), \tau). \tag{4}$$

Using a first-order approximation, we rewrite this as

$$\delta\dot{q} = \Delta t \left[\phi(q(t), \dot{q}(t), \tau) + \frac{\partial\phi}{\partial q}\Delta t(\dot{q}(t) + \delta\dot{q}) + \frac{\partial\phi}{\partial\dot{q}}\delta\dot{q} \right], \tag{5}$$

where $\delta\dot{q} = \dot{q}(t + \Delta t) - \dot{q}(t)$, thus computing the angular velocities of the ball joint at the next time step. Finally, the joint angles at the next time step are computed using explicit Euler time integration.

B   MUSCLE MODEL

Extraocular muscles exhibit a curved geometry largely as a result of their tangential attachment points to the surface of the eyeball and their passing through a set of elastic pulleys. To reduce computational cost, we approximate each extraocular muscle by connecting several uniaxial Hill-type actuators in series, with one end of the muscle originating at the orbital apex and the other at a tangential insertion on the ocular plant. Increasing the number of actuators yields better approximations to the curvature of the extraocular muscle, albeit at greater computational cost. The muscle activation signal is sent equally to each actuator in the series. Actuator activation contracts the entire muscle thereby applying a tangential force on the eyeball that induces eye rotation. The natural lengths of our extraocular muscle models are adjusted such that in their unactivated state, the eyes assume the "anatomical position of rest", with gaze directions that typically deviate outward by 15–25 degrees [Rosenfield 1997].

For each Hill-type actuator, the muscle force $f_m = f_P + f_C$ combines two components. The passive element f_P, which produces
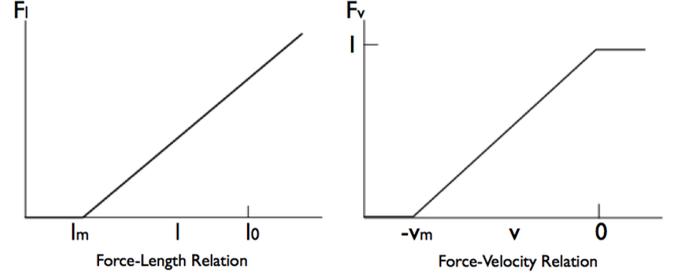

Fig. 22. The force-length and force-velocity relations of the Hill-type muscle model.

a restoring force due to the muscle elasticity, is represented as a uniaxial exponential spring as

$$f_P = \max(0, k_s(\exp(k_c e) - 1) + k_d\dot{e}), \tag{6}$$

where k_s and k_d are the stiffness and damping coefficient, respectively, e is the strain of the muscle and $\dot{e}$ is the strain rate. The contractile element f_C, which actively generates the contractile force of the muscle, is computed as

$$f_C = aF_l(l)F_v(\dot{l}), \tag{7}$$

where F_l is the force-length relation, F_v is the force-velocity relation, and a is the muscle activation level ($0 \le a \le 1$), which serves as the control input to the Hill-type actuator.

Fig. 22 plots the force-length and force-velocity relations. The former is represented as $F_l(l) = \max(0, k_{\max}(l - l_m))$, where l_m is the minimum length for the muscle to generate the force, $k_{\max}$ is the maximum stiffness of the activated muscle. The latter is represented as $F_v(\dot{l}) = \max(0, 1 + \min(\dot{l}, 0)/v_m)$, where v_m is the maximum contraction velocity with no load. Coefficient k_c is set to 7 for all the muscles, and I_m is set to $0.5l_0$ and v_m is l_0 s^{-1}. The other coefficients, k_s, k_d, and $k_{\max}$, are scaled to be proportional to the strength of each muscle, which is calculated as roughly proportional to the cross sectional area of the muscle.

In the modified Hill-type model, $F_l(l)$ increases monotonically. This works for our biomechanical model because the extraocular muscles stretch only a modest amount. Thus, we avoid negative stiffness, which may cause instability in the numerical simulation.

C   MUSCLE OPTIMIZATION

Inverse kinematics is first solved to compute the desired rotational joint angles for the eyes to look at a given target object.

With the target kinematic pose computed, the desired acceleration needed to reach the target is computed. Because the eye is able to move at the astonishingly fast rate of $900°/s$, it is not necessary to impose a strict upper bound on the maximal velocity; hence, the acceleration is permitted to be high enough to reach the target orientation in a few time steps. By integrating over successive time steps using these desired accelerations, we are able to compute the torque necessary to actuate the eye to the new orientation.

The desired torque is then used as the basis of a muscle optimizer, which ultimately computes the activations for each muscle, by numerically solving an energy minimization problem, where the

Table 1. Parameter tables.

(a) Refractive indices of the optical organs and mediums used in our model (c.f. Hecht [1987]).

Medium	Refractive Index
Air	1.0
Aqueous Humor	1.336
Cornea	1.376
Lens	1.386
Vitreous Humor	1.337

(b) Parameter values of the exponential function mapping error vector quantities to muscle activations.

Parameter	Iris-Pupil System	Lens System
α	0.01	0.01
β	0.0003	1.1
γ	0.2	0.05

objective is to minimize the total energy output by the combined activations. This may not necessarily be ideal in the case of the eye, as opposed to optimizing the speed of actuation, though neither of these approaches is directly supported by biological evidence.

Additionally, a non-zero constraint is imposed on all of the activations to prevent any of the muscles from being completely deactivated during a movement, since all the extraocular muscles contribute at least minimally to all eye movements. Intuitively, this is expected only to enhance biological accuracy without loss of significant solution space considering that any agonist-antagonist muscle pair could simply be equally co-activated to maintain the eye in a neutral position. We employ the SNOPT optimizer to solve the resulting quadratic objective function with linear constraints.

D PARAMETER VALUES

Our model makes use of several biologically grounded and empirically derived parameter values, which are as follows:

(1) Indices of refraction for each medium in the optical processing pipeline (Table 1a).
(2) Offline SNN training data synthesis parameters for the iris and lens subsystems.

Regarding the latter, to map error vector quantities to muscle activation adjustments, we employ the exponential function

$$\Delta \boldsymbol{a} = \text{sgn}(\Delta U) \max(\gamma, \alpha(e^{\beta|\Delta U|} - 1.0)) \tag{8}$$

as both the iris and lens accommodate to large differences relatively quickly, while only minor adjustments are necessary when the error vector is small. Here, ΔU denotes the error vector, γ represents the maximum possible activation step size per frame, and α and β are scaling factors. The value of γ is empirically determined by observing the smoothness between frame transitions at normal frame rates (between 20–30 fps). The values of α and β are also tuned to adjust the step size reduction in relation to the error vector magnitude. Our empirically determined values of each of the parameters are listed in Table 1b.

REFERENCES

A.T. Bahill, M.R. Clark, and L. Stark. 1975. The main sequence, a tool for studying human eye movements. *Mathematical Biosciences* 24, 3-4 (1975), 191–204.

W. Becker. 1989. The neurobiology of saccadic eye movements: Metrics. *Reviews of Oculomotor Research* 3 (1989), 13.

I. Bekerman, P. Gottlieb, and M. Vaiman. 2014. Variations in eyeball diameters of the healthy adults. *Journal of Ophthalmology* 2014, Article 503645 (2014), 5 pages.

S.R. Bharadwaj and T.R. Candy. 2008. Cues for the control of ocular accommodation and vergence during postnatal human development. *J. Vision* 8, 16 (2008), 14–14.

M. Buchberger. 2004. *Biomechanical Modelling of the Human Eye*. Ph.D. Dissertation. Johannes Kepler University, Linz, Austria.

G. Dagnelie. 2011. *Visual prosthetics: Physiology, bioengineering, rehabilitation*. Springer.

M.F. Deering. 2005. A photon accurate model of the human eye. *ACM Transactions on Graphics (TOG)* 24, 3 (2005), 649–658.

W. Fink and D. Micol. 2006. simEye: Computer-based simulation of visual perception under various eye defects using Zernike polynomials. *Journal of Biomedical Optics* 11, 5 (2006), 054011.

L.J. Grady. 2004. *Space-Variant Computer Vision: A Graph-Theoretic Approach*. Ph.D. Dissertation. Boston University.

J.E. Greivenkamp, J. Schwiegerling, J.M. Miller, and M.D. Mellinger. 1995. Visual acuity modeling using optical raytracing of schematic eyes. *American Journal of Ophthalmology* 120, 2 (1995), 227–240.

T. Haslwanter. 1995. Mathematics of three-dimensional eye rotations. *Vision Research* 35, 12 (1995), 1727–1739.

K. He, X. Zhang, S. Ren, and J. Sun. 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proc. IEEE International Conference on Computer Vision*. Santiago, Chile, 1026–1034.

E. Hecht. 1987. *Optics*. Addison Wesley.

O. Komogortsev, C. Holland, S. Jayarathna, and A. Karpov. 2013. 2D linear oculomotor plant mathematical model: Verification and biometric applications. *ACM Transactions on Applied Perception (TAP)* 10, 4 (2013), 27.

S.P. Lee, J.B. Badler, and N.I. Badler. 2002. Eyes alive. In *Proceedings ACM SIGGRAPH*. ACM, New York, NY, USA, 637–644.

S.-H. Lee, E. Sifakis, and D. Terzopoulos. 2009. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Transactions on Graphics* 28, 4, Article 99 (2009), 17 pages.

S.-H. Lee and D. Terzopoulos. 2006. Heads up! Biomechanical modeling and neuromuscular control of the neck. *ACM Transactions on Graphics* 25, 3 (2006), 1188–1198. Proc. *ACM SIGGRAPH 2006*, Boston, MA, August 2006.

R.J. Leigh and D.S. Zee. 2015. *The Neurology of Eye Movements*. Oxford Univ Press.

M. Lesmana, A. Landgren, P.-E. Forssén, and D.K. Pai. 2014. Active gaze stabilization. In *Proc. Indian Conf. on Computer Vision, Graphics and Image Processing*. 81.

M. Lesmana and D.K. Pai. 2011. A biologically inspired controller for fast eye movements. In *2011 IEEE International Conference on Robotics and Automation*. 3670–3675.

C.E. Looser and T. Wheatley. 2010. The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science* 21, 12 (2010), 1854–1862.

M. Nakada, H. Chen, and D. Terzopoulos. 2018a. Deep learning of biomimetic visual perception for virtual humans. In *ACM Symposium on Applied Perception (SAP 18)*. Vancouver, Canada, 20:1–8.

M. Nakada, T. Zhou, H. Chen, T. Weiss, and D. Terzopoulos. 2018b. Deep learning of biomimetic sensorimotor control for biomechanical human animation. *ACM Transactions on Graphics* 37, 4, Article 56 (August 2018), 15 pages. Proc. *ACM SIGGRAPH 2018*, Vancouver, Canada, August 2018.

P. Riordan-Eva and E.T. Cunningham. 2011. *Vaughan & Asbury's General Ophthalmology*. McGraw Hill Professional.

D.A. Robinson. 1964. The mechanics of human saccadic eye movement. *Journal of Physiology* 174 (1964), 245–264.

M. Rosenfield. 1997. Tonic vergence and vergence adaptation. *Optometry and Vision Science* 74, 5 (1997), 303–328.

K. Ruhland, S. Andrist, J. Badler, C. Peters, N. Badler, M. Gleicher, B. Mutlu, and R. Mcdonnell. 2014. Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In *Eurographics State-of-the-Art Reports*. 69–91.

C.K.L. Schraa-Tam, A. Van Der Lugt, M.A. Frens, M. Smits, P.C.A. Van Broekhoven, and J.N. Van Der Geest. 2008. An fMRI study on smooth pursuit and fixation suppression of the optokinetic reflex using similar visual stimulation. *Experimental Brain Research* 185, 4 (2008), 535–544.

E.L. Schwartz. 1977. Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception. *Biological Cybernetics* 25, 4 (1977), 181–194.

W.S. Stiles, B.H. Crawford, and J.H. Parsons. 1933. The luminous efficiency of rays entering the eye pupil at different points. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character* 112 (1933), 428–450.

J.G. Thomas. 1969. The dynamics of small saccadic eye movements. *Journal of Physiology* 200, 1 (1969), 109–127.

D. Tweed, W. Cadera, and T. Vilis. 1990. Computing 3D eye position quaternions and eye velocity from search coil signals. *Vision Research* 30, 1 (1990), 97–110.

C.-A. Wang and D.P. Munoz. 2014. Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience* 40, 5 (2014), 2822–2832.

Q. Wei, S. Patkar, and D.K. Pai. 2014. Fast ray-tracing of human eye optics on Graphics Processing Units. *Comput. Methods and Programs in Biomed.* 114, 3 (2014), 302–314.

Q. Wei, S. Sueda, and D.K. Pai. 2010. Biomechanical simulation of human eye movement. In *Biomechanical Simulation*. Lecture Notes in Computer Science, Vol. 5985. Springer, Berlin, 108–118.

S.H. Yeo, M. Lesmana, D.R. Neog, and D.K. Pai. 2012. Eyecatch: Simulating visuomotor coordination for object interception. *ACM Transactions on Graphics (TOG)* 31, 4, Article 42 (2012), 10 pages.