A Photon Accurate Model of the Human Eye
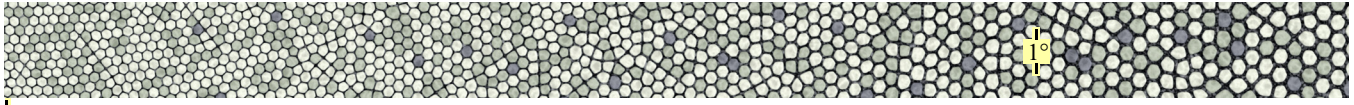
Michael F. Deering

deeringm@acm.org

Figure 1: Starting at the foveal center on the left, the first 1.3° of our synthesized retinal cones, continuing outward on the next 9 pages.

Abstract

A photon accurate model of individual cones in the human eye perceiving images on digital display devices is presented. Playback of streams of pixel video data is modeled as individual photon emission events from within the physical substructure of each display pixel. The thus generated electromagnetic wavefronts are refracted through a four surface model of the human cornea and lens, and diffracted at the pupil. The position, size, shape, and orientation of each of the five million photoreceptor cones in the retina are individually modeled by a new synthetic retina model. Photon absorption events map the collapsing wavefront to photon detection events in a particular cone, resulting in images of the photon counts in the retinal cone array. The custom rendering systems used to generate sequences of these images takes a number of optical and physical properties of the image formation into account, including wavelength dependent absorption in the tissues of the eye, and the motion blur caused by slight movement of the eye during a frame of viewing. The creation of this new model is part of a larger framework for understanding how changes to computer graphics rendering algorithms and changes in image display devices are related to artifacts visible to human viewers.

CR Categories: I.3.1 [Computer Graphics]: Hardware Architecture - Raster Display Devices, I.3.3 [Computer Graphics]: Picture/ Image Generation - Display Algorithms, I.3.7 [Computer Graphics]: Three Dimensional Graphics and Realism.

Additional Keywords and Phrases: eye models, schematic eyes, synthesized retina, human eye cone models, display devices.

1 Motivation

All applications of computer graphics and displays have a single ultimate end consumer: the human eye. But while enormous progress has been made on our models for rendering graphics, much less corresponding progress has been made on models for showing what the eye actually perceives in a given complex situation. This paper describes a system using a combination of computer graphics rendering techniques and known anatomical and optical properties of the human eye to produce a much finer grain simulation of the image forming process: a photon accurate model of the human eye.

This new eye model is a key part of a larger graphics, display, and perceptual simulation effort. Many would say that new display technologies, 9 megapixel panels and projectors for example, are coming ever closer to "matching the resolution of the human eye", but how does one measure this, and in what areas are current displays and rendering techniques still lacking? Having a much more accurate model of the interaction of modern rendering and display technology

NOTE: *This is an expanded version of a paper and an implementation sketch that were published in ACM SIGGRAPH 2005. This long form should be viewed as the "technical report" version. Some of the changes were just to make some figures larger, but other text and figures have been added.*

with the fine details of the human visual system is a necessary component for answering these questions. The scope of the current paper is the role of photons in this process: from their creation by a display device, to their eventual absorption by cones in the human eye.

The resolution perceived by the eye involves both spatial *and* temporal derivatives of the scene; even if the image is not moving, the eye *is* ("drifts"), but previous attempts to characterize the resolution requirements of the human eye generally have not taken this into account. Thus our photon model explicitly simulates the image effects of drifts via motion blur techniques; we believe that this effect when combined with the spatial derivatives of receptive fields is a necessary component of building a deep quantitative model of the eye's ability to perceive resolution in display devices.

The system described in this paper can be used to understand through simulation the retinal image quality of images generated by new hardware before it is actually built. Such hardware includes projectors, foveal-peripheral displays, rendering hardware, video compression/decompression hardware, etc. Defects in retinal image quality can point to weakness in proposed designs; but lack of defects can show where hardware might be simplified. A general benefit of perceptually based rendering is discovering what things the visual system *doesn't* notice, leading to possible savings in computer graphics computations; many such experiments are possible with the system as presented. The current system is intended as a solid base upon which incremental features can be added to address specific needs.

There has been some previous work in this area. [Deering 1998] tried to characterize the resolution limits of the human eye as when the display pixel density matches the local cone density. Unfortunately, this simple approximation can understate the resolution requirements in the fovea, where more pixels than cones may be needed, and overstate the resolution limits in the periphery, where large receptive fields rather than cones are the limit. Looking at this another way, there are five million cones in the human eye, but only half a million receptive field pairs outputting to the optic nerve. In [Barsky 2004] a system was described in which the composite optical transform function for all the optical elements of a particular person's eye was used to produce retinal images, though chromatic effects are not included.

2 Overview

The paper describes our combined model of displays and the human eye. After a description of our display model, most of the paper focuses on a description of the anatomical and optical properties of the human eye, and an explanation of which are or are not included in our system, and how they are simulated. While some attention must be paid to all of the portions of the optical pipeline, we go into implementation detail only for the two major new contributions of this work: the retinal cone synthesizer (which produced Figure 1 above), and the rasterization of individual photon events into individual photoreceptors within this synthesized retina. Because of the

focus on color displays, our eye model is a photopic model, and only simulates retinal cones; rods are not included.

3 The Literature

[Glassner 1995] provides a good introduction to the human eye from a computer graphics perspective. [Oyster 1999] is the best single book on in-depth information about nearly all aspects of the human eye. [Rodieck 1998] covers a little less ground, focusing more on the early stages of vision (up through a little beyond how the cones work), but covers the material in a superb way. [Atchison and Smith 2000] is the best single book on the optical aspects of the human eye. The book "Seeing" edited by Karen De Valois (of which [Thibos 2000] is one chapter), is contains an excellent series of chapters traversing the entire visual pipeline. Many of the other papers referenced in this paper are from the Journal Vision Research, and excellent (if expensive) source of thousands of primary papers on many of the topics described here, including several papers on the early development of the human eye that validate some aspects of our retinal synthesizer.

4 Conventions and Units

Screen sizes are stated in units of centimeters (cm), all gross eye anatomical features are stated in units of millimeters (mm), fine anatomical detail are given in micrometers (microns) (μ), light wavelengths are expressed in nanometers (nm). Angles will be expressed in units of degrees (°), arc minutes, and arc seconds. Most angles are **visual angles** or **visual eccentricity**: the angle that a ray from an external point in space to the first nodal point of the eye makes with the optical axis of the eye. However some angles will be explicitly expressed in terms of **retinal eccentricity**: on a spherical coordinate system on the surface of the retina, the angle of a point on the retina (colatitude from the center of the fovea). The other coordinate in this system is the angle of retinal meridian (or longitude). On a right eye the zero meridian is the horizontal meridian running from the center of the fovea through the optic disc. Note that due to optical magnification, each degree of retinal eccentricity cover 1.5× to 1.3× less of the retina than each degree of visual eccentricity. Also note that the region of the retina up to an eccentricity θ covers an angular field of view of 2θ. Distances or areas on the surface of the retina expressed in units of mm or μ always assume a nominal 24 mm diameter retina. Retinal eccentricity is sometimes expressed in these mm. Individual eyes are either left or right, but by using relative terminology (nasal, temporal) the eye is generally described independent of its side. Our model produces either, the one simulated in this paper is a right eye. (Pairs of simulated eyes are useful in understanding stereo effects.)

5 Point Spread vs. Optical Transfer Functions

The optical transfer function (**OTF**) is a powerful technique for expressing the effects of optical systems in terms of Fourier series. The OTF works best (or at least needs fewer terms) when the details being modeled are also well described by Fourier series, as is the case for most analog and continuous inputs. But the sharp discontinuous sides and inter-pixel gaps that characterize both emissive pixels in modern displays and the polygonal cone optical apertures of the receptive pixels of the human eye do not fit this formalism well. So for our system, we use the mathematically equivalent point spread function (**PSF**). Since we are modeling the emission of each photon from a display surface pixel element as a discrete event this is a fairly nat-

ural formulization. At the retinal surface of cones, the PSF becomes the probability density that the photon will appear at a given point.

Both formulizations apply only to light of a specific wavelength; the PSF of a broadband source of photons is the sum of the PSF at each wavelength within the broadband spectrum, weighted by the relative number of the photons at that wavelength. While resolution is often thought of as a grey scale phenomena, many times chromatic aberration can be the limiting factor. Thus in our system all optical properties and optical effects are computed over many different spectral channels. Specifically, all of our spectral functions cover the range from 390 to 830 nm; in inner loops, we use 45 channels at 10 nm increments, elsewhere we use 4401 0.1 nm channels.

6 Photon Counts of Displays

In deciding how to simulate the effects of display devices on the cones in the human eye, a natural question is how many quanta events (photons) are we talking about?

Consider the example of a 2,000 lumen projector with a native pixel resolution of 1280×1024 @60 Hz, projected onto a (lambertian) screen 240 cm wide with a gain of 1, viewed from 240 cm away by a viewer with a 40 mm^2 entrance pupil. By definition, a lumen is a radiant flux of $4.09*10^{15}$ photons per second at an optical wavelength of 555.5 nm. In 1/60th of a second frame time, a single pixel of that display will emit $1.04*10^{11}$ photons, spread over a 2π steradian hemisphere from the screen. At our viewer's 240 cm distance, this hemisphere is ~36 meters^2 in area, and a 40 mm^2 entrance pupil will capture only 114,960 photons from that pixel. Only 21.5% of these photons will make it through all the tissue of the cornea, lens, macula, and into a cone to **photoisomerize** and cause a chemical cascade resulting in a change in the electrical charge of that cone, or about 24,716 perceived photons.
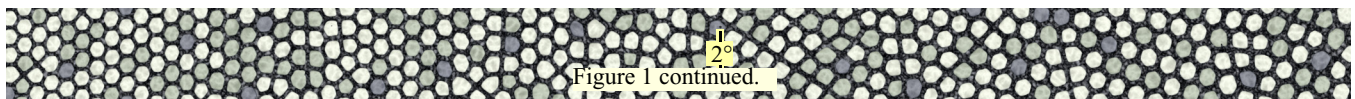
Not counting any optical aberrations, this single pixel will cover an angular region of 2.5×2.5 minutes of arc, or about 5×5 cones (in the center of the fovea). Thus each cone will receive ~1/25th of the photon count, or one pixel will generate 996 perceived photons per cone per 1/60 second. Note that this is for a full bright maximum value white pixel. Dimmer colored pixels will produce corresponding less photons.

While the more broadband emissions of a real projector will generate more (but less photo-effective) photons, the total number of quanta events at the cone level will remain the same. This is a small enough number that we decided to model each individual quanta emission and absorption. With modern computing power, we can simulate every photon that effects a portion of the retina for small numbers of display video frames in a few hours of CPU time.

7 Display Pixel Model

Unlike the Gaussian spots of CRT's (described in [Glassner 1995]), modern digital pixel displays employ relatively crisp, usually rectangular, light emitting regions. In direct view displays, each of the three (or more) color primaries have separate non-overlapping regions. In projection displays, the primaries generally overlap in space, though not in time for field sequential color display devices. At the screen, projection based displays have less crisp color primaries regions and more miss-alignment, due to distortions caused by optics, temperature, and manufacturing miss-alignment.

Figure 2 a-d shows for several different devices their light emitting regions. Our system implements a general parmeterized model of

2°

Figure 1 continued.

(a) Direct view CRT   (b) Direct view LCD

(c) Projection 3 chip DLP™   (d) Projection 1 chip DLP™
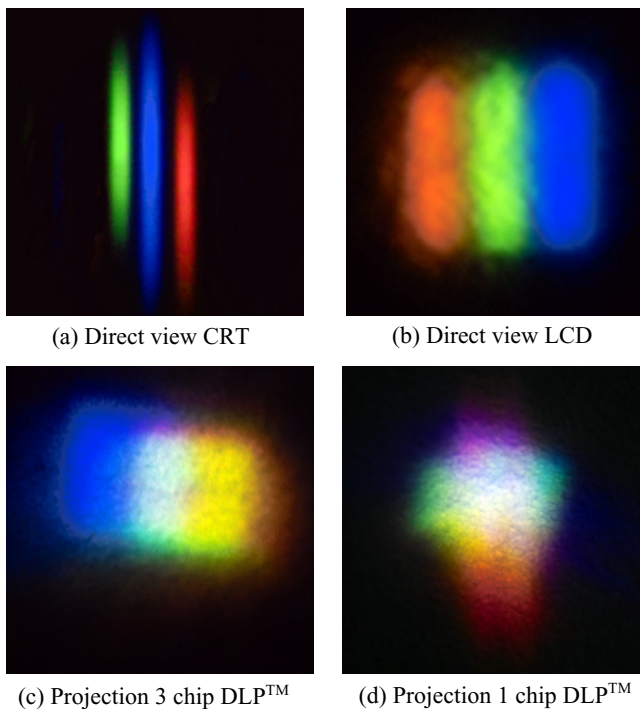
Figure 2: 2 direct view & 2 projection display device single pixels.

this sub-pixel structure. Each color primary also has its own spectral emission function.

The temporal structure of the sub-pixels varies wildly between different types of displays, and can have great effect on the eye's perception of the display. CRT displays (direct view or projected) have a fast primary intensity flash, which decays to less than 10% of peak within a few hundred microseconds. Direct view LCD devices have considerable lag in pixels settlings to new values, leading to ghosting, though this is beginning to improve. LCDs generally also use spatial and temporal dithering to make up the last few bits of grey scale. DLP™ projection devices are single bit intensity pixels dithered at extremely high temporal rates (60,000 Hz+); they also use several forms of spatial dithering. LCOS projection devices use true grey scale, and some are fast enough to use field sequential color.

All of these temporal and spatial features must be emulated for each different display simulated in our system. The example shown in Figure 12 used a simulation of a simplified three chip LCOS projector, with 80% pixel fill factor, and highly accurate registration of the red, green, and blue pixel primaries. The example shown in Figures 13 and 15 used a simulation of a different type of display device: room light reflected off a 1200 dpi laser printer printout.

The spectral properties also varies with display type. Direct view LCD and projection LCD, LCOS, and DLP™ devices effectively use fairly broadband color filters for each of their primaries. CRT green and blue phosphors resemble Gaussians, through red phosphors are notorious for narrow spectral spikes. (Appendix G4 of [Glassner 1995] shows an example of CRT spectra, as does Figure 22.6 of [Poynton 2003].) Laser based displays by definition have narrow spectral envelopes.

8 Schematic Eyes

Schematic eyes [Atchison and Smith 2000] are greatly simplified optical models of the human eye. **Paraxial schematic eyes** are only for use within the paraxial region where sin[x] ≈ x, e.g. within 1° of the optical axis. In many cases the goal is to model only certain simple optical properties, and thus in a **reduced** or **simplified** schematic eye the shape, position, and number of optical surfaces are completely anatomically incorrect.

Having optical accuracy only within a 1° field is a problem for an anatomically accurate eye, as the **fovea** (highest resolution portion of the retina) is located 5° off the optical axis. **Finite** or **wide angle** schematic eyes are more detailed models that are accurate across wider fields of view.

9 Variance of Real Eyes

Even finite schematic eyes generally come in one size and with one fixed set of optical element shapes. The idea is to have a single fixed mathematical model that represents an "average" human eye. The problem is that real human eyes not only come in a range of sizes (a Gaussian distribution with a standard deviation of ±1 mm about 24 mm), but many other anatomical features (such as the center of the pupil) vary in complementary ways with other anatomical features such that they cannot be simply averaged. Because our goal is to simulate the interaction of light with fine anatomical details of the eye, we had to construct a **parameterized eye**, in which many anatomical features are not fixed, but parameters. Which features are parameters will be discussed in later sections. There is a beginning of a trend for such more anatomically individualizable eye models in other application areas, such as design simulations for the replacement or addition of a lens within an eye.

10 Photon Count at Cones

Schematic eyes are generally not used to produce images, but to allow various optical properties, such as quantified image aberrations, to be measured. One alternative to using a schematic eye is to employ a composite optical transfer function measured from a real eye. In [Barsky 2004] this was used to form an image on the surface of the retina. But this image is not what the eye sees, because it has not taken into account the interaction of light with the photoreceptor cones, nor the discrete sampling by the cone array.

The human retinal cones generally form a triangular lattice of hexagonal elements, with irregular perturbations and breaks. Sampling theory in computer graphics [Cook 1986; Cook et al. 1987; Dobkin et al. 1996] has demonstrated the advantages of perturbed regular sampling over regular sampling in image formation.

Some previous simulations of the human eye have taken the effects of sampling on a completely regular triangular grid into account [Geisler 1989]. In [Ahumada and Poirson 1987] the locations of cone centers are perturbed, but the underlying grid is still perfectly triangular. In [Hennig et al. 2002] the cones are placed on a completely regular triangular grid, but the receptive fields that sample these cones are perturbed with respect to one another. In [?], a small irregular cone sample pattern was constructed by taking digitized centers from an image of a number of actual cones, and then this pattern was stepped and flipped multiple times to produce a larger retinal cone sampling pattern.

We wanted to model the specific sampling pattern of the eye in detail and over the entire retina as part of our system. Thus we had to construct a retina synthesizer; a program that would produce an an-

3°

Figure 1 continued.
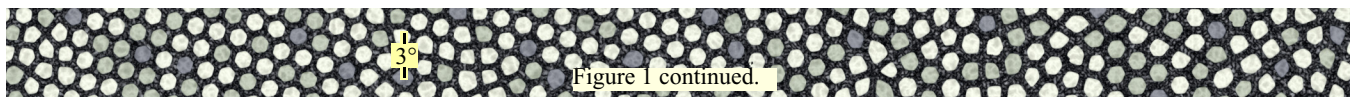
atomically correct model of the position, size, shape, orientation, and type distribution (L M S) of each of the five million photoreceptor cones in the human retina.

11 Eye Rotation During Viewing

Even while running and looking at another moving object, the visual and vestibular systems coordinate head and eye movements to attempt to stabilize the retinal image of the target. Errors in this stabilization of up to 2° per second slip are not consciously perceivable, though measurements of visual acuity show some degradation at such high angular slips [Rodieck 1998; Steinman 1996].

At the opposite extreme, for fixation of a non-moving object (such as a still image on a display device), three types of small eye motions remain: tremor (physiological nystagmus), drifts, and microsaccades [Martinez-Conde et al. 2004]. Microsaccades are brief (~25 ms) jerks in eye orientation (10 minutes to a degree of arc) to re-stimulate or re-center the target. **Drifts** are brief (0.2 to 1 second) slow changes in orientation (6 minutes to half a degree of arc per second) whose purpose may be to ensure that edges of the target move over different cones. Tremors are 30 to 100 Hz oscillations of the eye with an amplitude of 0.3 to 0.5 minutes of arc. These small orientation changes are important in the simulation of the eye's perception of display devices, because so many of them now use some form of temporal dithering. There is also evidence that orientation changes are critical to how the visual system detects edges.

Our system allows a unique orientation of the eye to be set for *each* photon being simulated, in order to support motion blur [Cook 1986]. While the orientation of the eye could be set to a complex combination of tremor, drifts, and microsaccades as a function of time, because there is some evidence that cone photon integration is suppressed during saccades, we have concentrated on a single drift between microsaccades as the orientation function of time to simulate. We assume that drifts follow Listing's law [Haslwanter 1995]; the drift is a linear interpolation of the quaternions representing the orientation of the eye relative to Listing's plane at beginning and end of the drift; our default drift is 6 minutes of arc per second at a 30° to the right and up. Our neutral vergence Listing's plane is vertical and slightly towed in corresponding to the 5° off-center fovea.

The rotational center of the eye is generally given as 13.5 mm behind the corneal apex, and 0.5 mm nasal [Oyster 1999]. Our model uses this value; we do not simulate the few hundred microns shift in this location reported for large (±20°) rotations.

12 The Optical Surface Model

Most of the traditional finite schematic eye models were too anatomically inaccurate for our application. The most anatomically correct and accurate image forming simple model we found was [Escudero-Sanz and Navarro 1999]. It is a four optical surface model using conic surfaces for the front surface of the cornea (conic constant -0.26, radius 7.72 mm) and both surfaces of the lens (conic constants -3.1316 and -1.0, radii 10.2 and -6.0 mm respectively), and using portions of a sphere for the back surface of the cornea (radius 6.5 mm) and the front surface of the retina (radius 12.0 mm). The optics and pupil are assumed centered. Indices of refraction of the mediums vary as a four point polyline of optical wavelength. We used their model as a starting point for our optical elements. The first thing we had to change when focusing on a fovea 5° off the corneal optical axis was to decenter the pupil by 0.4 mm, which is consistent with decenter measurements on real eyes (see the iris section
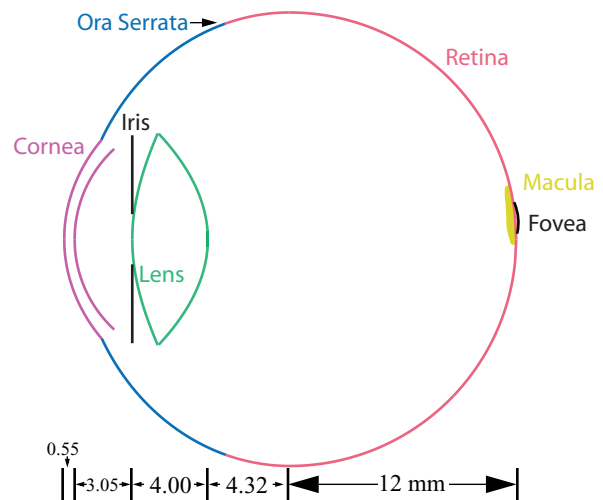

Figure 3: Optical model. All dimensions are in mm.

below). We also change the parameters of the front surface of the lens and the position of the pupil to model accommodation to different depths and different wavelengths of light. Our modified version of the Escudero-Sanz schematic eye is shown in Figure 3.

While it has been known for over a hundred years that the human eye lens employs a variable index of refraction, until recently there has been very little empirical data on how the index varies, and the recent data is still quite tentative [Smith 2003]. Nevertheless, in our search for anatomical accuracy, we simulated a number of published models of variable index lenses [Atchison and Smith 2000], including that of [Liou and Brennan 1997], who's schematic eye did include a decentered pupil. When layered shells of constant refractive index are used, we came to the same conclusion as several other authors: modeling the lens using less than 400 shells produces too many quantization effects. Even with this many shells, we could not get any of the models to produce acceptable levels of focus (to be fair, most did not claim high focus accuracy). Most of the models let the index of refraction vary as a quadric function of position within the lens; an analysis of our focus errors showed that this may be too simple to model the lens well. Because the primary emphasis of our system is the retina and its interaction with diffracted light, we chose for our initial version to stay with the simple non-variable index conic surface lens model as shown in Figure 3.

New measurement devices mean that we now also have more accurate data on the exact shape of the front surface of the cornea [Halstead et al. 1996]. However there are accuracy issues in the critical central section of a normal cornea. So while our goal was to create a framework where we can insert more anatomical elements into our parameterized eye as needed, for the front surface of the cornea again we decided for our initial version to use a conic model as shown in Figure 3.

13 The Human Retina & Photoreceptor Topography

The term **retina** refers to the interior surface of the eye, containing the photoreceptors and associated neural processing circuitry. Geometrically, the retina is a sub-portion of a sphere: a sphere with a large hole in it where the retina starts at the **ora serrata** (see Figure 3). The ora serrata boundary is not a perfect circle; while it lies in the vicinity of retinal eccentricity = 120° circle, the eccentricity to points on the ora serrata varies by 6% (a slight dip nasal). Because
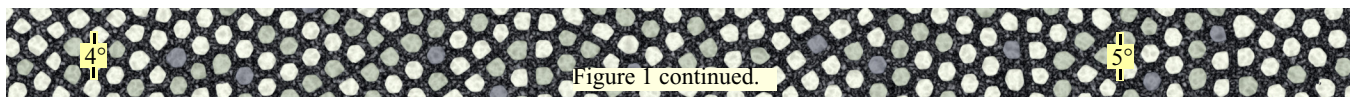

Figure 1 continued.

of the magnification of the eye's optics, in *visual eccentricity* the ora serrata lies at about 95°.

Most real retinas are actually slightly ellipsoidal; if the length of the eye does not match the refractive power of the eye, the result is myopia or hyperopia; inequality in width to height can produce astigmatism. The actual shape deviates further from a sphere as you look at fine details: flattening both near the optic nerve and the cornea, and deepening at the foveal pit. We wanted a simulation that could be easily extended to these cases, but our initial version uses a perfectly spherical model.

The human retina contains two types of **photoreceptors**: approximately 5 million **cones** and 80 million **rods**. The center of the retina is a rod-free area where the cones are very tightly packed out to 0.5°. [Curcio et al. 1990] is the major work describing the variation in density of the cones from the crowed center to the far periphery, where it turns out that the density is not just a function of eccentricity, it is also a function of meridian. There is a slightly higher cone density toward the horizontal meridian, and also to the nasal side. We refer the reader to the reference for the fine details of the distribution; our model uses these data points as cone density input targets for different regions of the retina.

The distribution of the three different cone types (**L**, **M**, and **S**, for long, medium, and short wavelength, roughly corresponding to peek sensitivity to red, green, and blue light) is further analyzed by [Roorda et al. 2001]. The L and M cone types vary completely randomly, but the less frequent S cone type tends to stay well away from other S cones. There are a range of estimates of the ratios of S to M to L cones in the literature, and there certainly is individual variation. For our system this is an input parameter; we use 0.08 to 0.3 to 0.62 as a default. Out to 0.175° in the fovea there are no S cones; outside that their percentage rapidly rises to their normal amount by 1.5°.

At the center of the fovea, the peak density of cones (measured in cones/mm^2) varies wildly between individuals; [Curcio et al. 1990] reports a range of values from 98,200 to 324,100 cones/mm^2 for different eyes. Outside the fovea, there is much less variance in cone density. [Tyler 1997] argues that the density of cones as a function of visual eccentricity outside the central 1° is just enough to keep the photon flux per cone constant, given that cones at higher eccentricities receive less light due to the smaller exit pupil they see, against the slightly larger amount of flux they receive for being closer to the exit pupil. Tyler's model for density from 1° to 20° of visual eccentricity θ is:

$$\frac{\text{cones}}{\text{mm}^2}[\theta] \;=\; 50000 \cdot \left(\frac{\theta}{300}\right)^{-2/3} \qquad (1)$$

We use this model in our system, modulated by the cone density variations due to meridian from [Curcio et al. 1990]. Beyond 20°, we follow Tyler's suggested linear fall-off to 4,000 cones/mm^2 at the ora serrata.

14 What Does a Cone Look Like?

Before describing our retina synthesizer, some additional three dimensional information about the shape of the optically relevant portions of photoreceptor cones needs to be given. In the fovea, the cone cell's terminal, axon, and nucleus are relatively transparent, and pulled away from the optically active inner and outer segment of the cone. (All other retinal processing cells are also pushed to the side of the retina.) Figure 4 shows three neighboring cone cells. In-
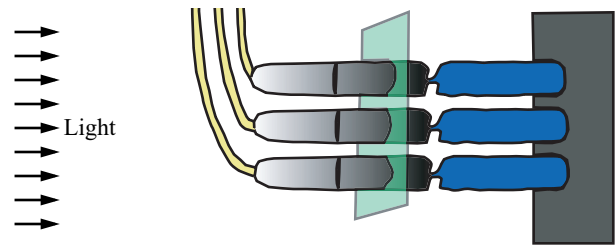

Figure 4: Three neighboring foveal cones. The inner segments are shown as grey; the left and right halves are the myoid and the ellipsoid portions, respectively. The outer segments are shown as blue. The connection to the nucleus heads off the top in yellow. The transparent green plane is the section we model as polygons. *(This figure is a placeholder; a more correct version needs to be drawn!)*

coming light first hits the inner segment, which due to having a higher optical index than the surrounding extracellular medium acts like a fiber optics pipe to capture and guide light into the outer segment. The outer segment contains the photoreceptor molecules whose capture of a photon leads to the perception of light. In the fovea, these portions of the cone cells are packed tightly together, and the combined length of the inner and outer segment is on the order of 50 microns, while the width of the inter segment may be less than 2 microns across. A section through the ellipsoid portion of the inner segment, shown in transparent green in Figure 4, is the optical aperture that is seen in photomicrographs of retinal cones, and is the element simulated by our retina synthesizer. Outside the fovea, the cone cells are more loosely packed, shorter (20 microns), wider (5-10 microns), and interspersed with rod cells. In addition, the rest of the cone cell and all of the other (mostly transparent) retinal processing cells and (not so transparent) blood supply lie on top of the cones and rods. Photomicrographs of foveal cones may not always have their limited depth of field focused precisely on the ellipsoid portion of the inner segments; S cones look either larger or smaller than L and M cones depending on focus depth. Optically, another diffraction takes place at the entrance aperture of the cone inner segment; thus especially in the fovea where the separation between the active areas of cones are less than the wavelength of light, it is not entirely clear where within the first few microns of depth of the inner segment the aperture actually forms. In our system, we use the polygonal cell walls as created.

15 The Retina Synthesizer

A major new contribution of this work is the retina synthesizer: given parameterized statistics of a retina, as described in the previous section, it "grows" a population of five million packed tiled cones on the inside of a spherical retina. The description of each cone as a polygonal aperture for light capture is passed on as data for later stages of the system. The rest of this section will describe how the retina synthesizer works.

A retina is started with a seed of seven cones: a hexagon of six cones around the center-most cone. The retina is then built by several thousand successive growth cycles in which a new ring of cones is placed in a circle just outside the boundary of the current retina, and then allowed to migrate inward and merge in with the exiting cones. Each new cone is created with an individual "nominal" target radius: the anatomical radius predicted for the location within the retina that the cone is created at.

Figure 1 continued.

6°
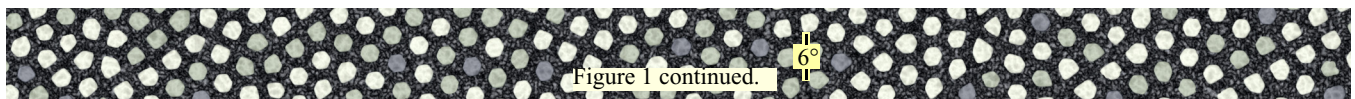
Each cone is modeled as one center point plus a set of points that define the cone's (convex) polygonal cell boundary. The inward movement and jostling for position of cones is modeled by a two phase sub-cycle: the first phase considers only the cell center points and uses simulated forces between them to migrate these points; the second phase re-forms the (polygonal) cell boundaries based on the updated location of the cell centers. After each new ring of cones is added, 25 to 41 of these sub-cycles are run to seamlessly merge them into the existing cone mosaic.

15.1 Definition of Neighbor Cone

For two cones p and n the normalized distance between them is:

$$D[p,n] = \frac{|p - n|}{(p.r + n.r)} \qquad (2)$$

where p and n also denote the locations of the center of the cones, and p.r and n.r are their birth radius targets. A cone n is said to be a *neighbor* of a cone p by the neighbor predicate function N:

$$N[p,n] = D[p,n] \geq 1.5 \qquad (3)$$

15.2 The Cone Force Equation

Two forces operate on cell centers during the first phase of the sub-cycle: a drive for each cone to move in the direction of the center of the retina; and a exponentially increasing repulsive force pushing cone centers that are neighbors away from each other. The center driving force includes a random component, and its overall strength K2 diminishes as a growth cycle progresses (effectively simulated annealing) through between 25 to 41 sub-cycles. A simplified version of the summed force equation is:

$$p' = p + K1 \cdot \vec{pn} + K2 \cdot \vec{r} - K3 \cdot \sum_{n}^{N[p,n]} spline[D[p, n]] \cdot \frac{n - p}{|n - p|}$$

$$(4)$$

where $\vec{pn}$ is unit normal from the cone p in the direction of the retinal center, and $\vec{r}$ is a vector with a random magnitude between 0 and 1 in a random direction. Both these vectors are three dimensional, but are tangent to the local retinal surface (a sphere). K1 and K3 are constants, and spline[] is defined by:

$$spline[x] = \begin{cases} 0 & x > 1 \\ (1 - x) \cdot 3.2 & 0.75 < x \leq 1 \\ (3.75 - x) \cdot 0.266 & 0 \leq x \leq 0.75 \end{cases} \qquad (5)$$

Again, this is a simplified version; the real implementation is a 200 line Java method, not a single equation. The "real" force equation actually operates in a metric space in the plane of the local cone, which at large eccentricities can be tilted 20° or more from the local tangent to the spherical retinal surface because cones point at the exit pupil of the eye.

15.3 The Re-Forming of Cone Cell Borders

The process starts off with a just the set of bare cone cell center points, all previous cell wall information having been discarded. The re-forming of cell walls is a topological and connectivity problem similar to that solved by Vornonoi cell construction, but with additional cell size constraints. This sub-section will describe the details of the algorithm.
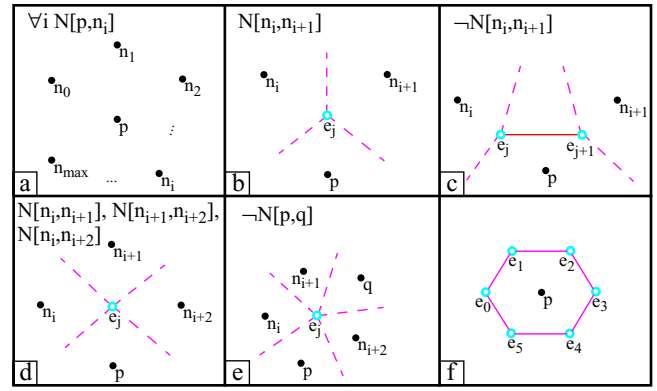

Figure 5: Patterns for new edge cell vertex e_j creation.

The standard Vornonoi cell solution has several biologically undesirable features. First, it allows cell walls to form arbitrary far from cell center points, ignoring the constraint of maximum cone cell size. Second, its rule for forming corners in cell walls does not balance the internal pressure of the cells. And finally, it forms cell wall corners based only on three points, while real cell wall corners are frequently the point of mutual contact of four or five cells.

In order to address these limitations, a new algorithm for forming cell walls given a set of cell center points was designed.

For each cone p, a spatially indexed data base of cones is searched to find all others cones n_j such that $N[p,n_j]$. The neighbors are sorted into clockwise order about p, as shown in Figure 5a. Then new cell wall edge vertex creation pattern matching rules are applied all around the circle of neighbors, resulting in a polygonal cell wall, as shown for the six sided case in Figure 5f.

For each successive neighbor n_i around p, rules are tried involving 5, 4, 3, or 2 consecutive neighbors (most to least complex rules). Figures 5b-e graphically show some of these rules.

In the simple case of Figure 5b, cones p, n_i, and n_{i+1} are all mutually neighbors. The new cell wall edge vertex e_j is created by simply averaging the center points of the three cones. The dashed magenta lines show where cell walls will connect to e_j; the particular direction of the walls will depend on the positions of other cell wall edge vertices. (The same convention is used in Figures 5c-e).

Figure 5c shows the contrasting case in which n_i and n_{i+1} are both (by definition) neighbors of p, but unlike in Figure 5b, not of each other. p will share a new cell wall edge vertex e_j with n_i that is created by averaging the center points of just these two cones; e_{j+1} is created similarly just between p and n_{i+1}. The cell wall edge between e_j and e_{j+1} (shown in red) will be the edge of a (usually temporary) void in the retinal cone mosaic.

Figure 5d shows what happens when three consecutive neighbors of p are mutually neighbors *and* n_i or n_{i+2} is closer to p than n_{i+1}; all four cones will share a new cell wall edge vertex e_j, the average of all four cone center points. The key here is that n_i and n_{i+2} are neighbors, and n_{i+1} is further away from p. By symmetry, all four cone cell walls will push out together to meet at a single point. The same logic applies with five mutual neighbors (not shown). In rare cases (Figure 5e), five cones will not be mutual neighbors, but a complex test shows they should share only one new cell wall edge vertex.

After forming new cell walls for each cone, cell walls marked as having voids on one side can be connected together to discover po-

Figure 1 continued.
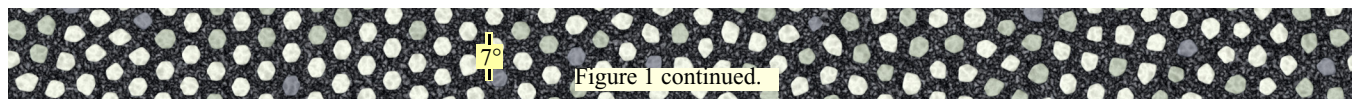
lygonal voids in the mosaic. Any voids encountered that are similar in or of larger area than that of the local cones are seeded with a new cone. In practice the vast majority of voids are just areas where the (code that estimated the) local birth density of cones was too low, and the void filling is just back-filling for this. Specifically, the interesting "breaks" in the regular hexagonal pattern of the cones are cause by the packing process, not by void filling. The last stable cell wall created for a cone becomes its final output polygon.

Just as with the force equation, the real implementation of the cone cell boarder re-forming algorithm operates in a metric space in the plane of the local cone. As an overall result, our synthesized polygonal cone boundaries generated become elliptical (in the local retinal tangent plane) at high eccentricities, just as real human eye cones do.

15.4 Tuning the Algorithm

The number of relaxation sub-cycles used has an effect on the regularity of the resulting pattern. A large number of cycles (80) is enough for great swaths of cones to arrange themselves into completely regular hexagonal tiles, with major fault walls only occasionally. A small number of cycles (20) doesn't allow enough time for the cones to get very organized, and the hexagonal pattern is broken quite frequently. The "just right" number of cycles (41) produced a mixture of regular regions with breaks at about the same scale as imagery from real retinas. After we had set this parameter empirically, we discovered that real central retinal patterns have been characterized by the average number of neighbors that each cone cell has - about 6.25. Our simulated retinas have the same number of average neighbors with our parameterization; different parameterizations generate different average neighbor counts. We let the number of sub-cycles drop outside the fovea to simulate the less hexagonally regular patterns that occur once rod cells start appearing between cone cells in the periphery. Our retina synthesizer does not simulate rods explicitly, but it does reduce the optical aperture of cones (as opposed to their separation radii) in the periphery to simulate the presence of rods.

The algorithm as described does not always produce complete tilings of the retina, even discounting small voids. Sometimes a bunch of cones will all try to crowd through the same gap in the existing retina edge, generating enough repulsive force to keep any of them from filling the gap; the result is a void larger than a single cone. Other times a crowd of cones will push two cones far too close together, resulting in two degenerate cones next to each other. Such faults are endemic to this class of discrete dynamic simulators, and while a magic "correct" set of strength curves for forces might allow such cases to never occur, we found it more expedient to seed new cones in large voids, and delete one of any degenerate pair. We have grown retinas as large as 5.2 million cones with very few voids larger than a cone.

It is not possible to dynamically simulate forces on such a large number of cones simultaneously. Instead, we mark cones by their pathlength (number of cone hops) to the currently growing edge. Cones deep enough are first "frozen": capable of exerting repulsive force, and changing their cell walls, but no longer capable of moving their centers; and then "deep frozen": when even their cell walls are fixed, and their only active role is to share these walls with frozen cells. Once a cone has only deep frozen cones as neighbors, it no longer participates in the growth cycle, and it can be output to a file, and its in-core representation can be deleted and space reclaimed. The result is a fairly shallow (~10 deep) ring of live cones

expanding from the central start point. Thus the algorithm's space requirement is proportional to the square root of the number of cones being produced. Still, the program takes about an hour of computation for every 100,000 cones generated, and unlike other stages of our system, cannot be broken up and executed in parallel. (But retinas are used much more often then they are generated.)

The optic disc (where the optic nerve exits the eye) is modeled in our system as a post process that deletes cones in its region: 15° nasal and 2° up from the foveal center, an ellipse 5° wide and 7° tall.

It must be emphasized that each cone is modeled individually, and that the initial target cone radius is just used to parameterize the forces generated by and on the cone. The final radius and polygonal shape of each cone is unique (though statistically related to the target), and even in areas where the cone tiling is completely hexagonal the individual cones are not perfect equal edge length hexagons, but slightly squashed and lining up on curved rows. *It is these nonperfect optical apertures that is the desired input to the later stage of rasterizing diffracted defocused motion blurred photons.*

The resulting patterns are very similar to photomicrographs of real retinas. Figure 8 is a rendering of a portion of our synthetic retina, and is very similar to the photomicrographs of portions of retinas near the fovea in [Curcio et al. 1990]. Figure 7a is an image from [Roorda and Williams1999] of living retina using adaptive optics taken at 1° visual eccentricity, and about 0.5° across. Figure 7b is the same region of the retina in our system, rendered using light reflected from the bottom of each simulated cone, and then blurred. Our system doesn't directly simulate retinal blood vessels in front of the cones (the dark river running down 7a), but we do simulate projecting images on our retina (see Section 21 below); so 7b borrows a low pass blood vessel shadow map from 7a. Figure 1, running along the bottom length of this paper, is a small portion (0.005%) of one of our synthesized retinas, running from the center of the fovea on the left (0°) to 13° visual eccentricity on the right. (*Note*: the relative cone area model used in rendering this image is too low at large eccentricities; there should be less space between cones. The model has since been fixed, but replacement images have not yet been rendered) (The optic disc was not removed from Figure 1; no cones would have shown in the last portion of the last image.) Figures 8 and 9 give two more views of our data. In Figures 1 and 9, the L cones are white, but we tint the M cones a little darker and slightly greenish, and the S cones a little more darker and slightly bluish.

While our algorithm is not intended to be an exact model of the biology of retinal cone cell growth, it does share many features with the known processes of human retinal cone formation, where foveal cones are still migrating towards the center of the retina several years *after* birth.

While this paper focus on cones, the retinal synthesizer has all the connectivity information it needs to also generate receptive fields of cones (and does so). Small receptive fields are created using a single cone as the receptive field center, and all of that cone's immediate neighbors (ones that it shares cell edge boundaries with) as the surround. Larger receptive fields are created by using a cone, and one or more recursive generations of immediate neighbors as the center, and then two or more recursive generations of immediate neighbors outside the center as the surround. Separate algorithms are used to set the relative strength of the center and its antagonistic surround, and do perform the processing of inputs to these receptive fields. The results of this processing also generate images, this time
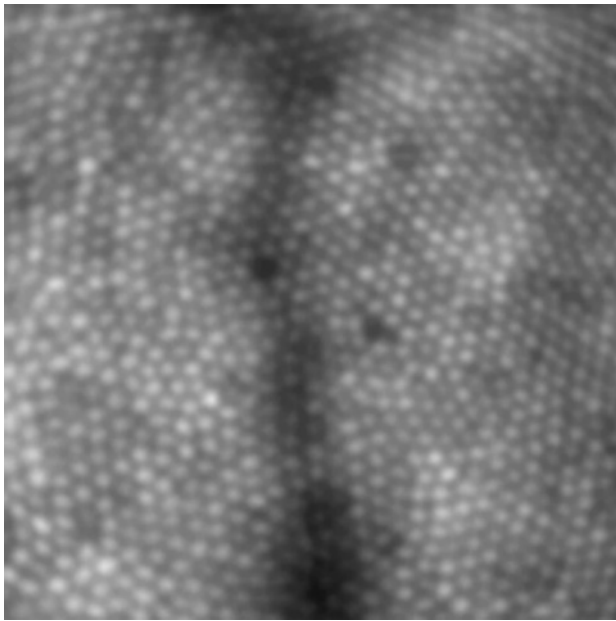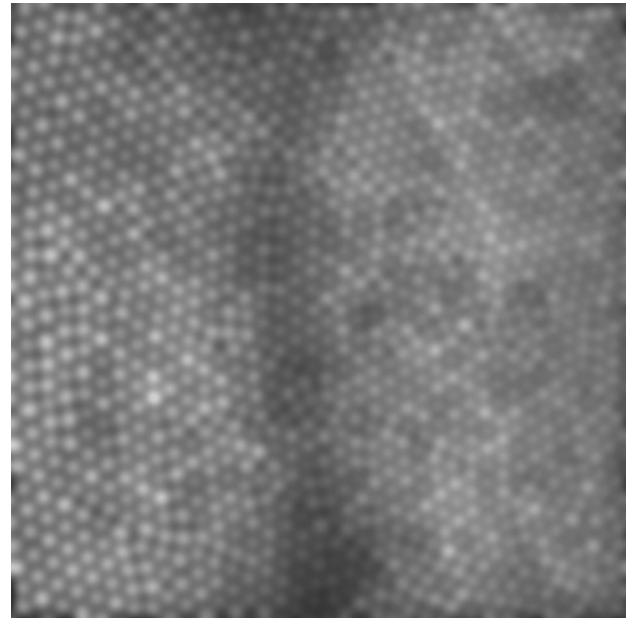

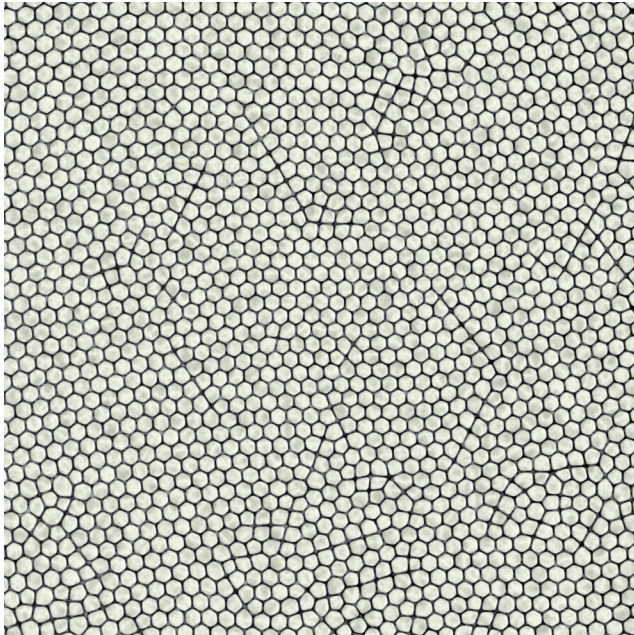Figure 1 continued.
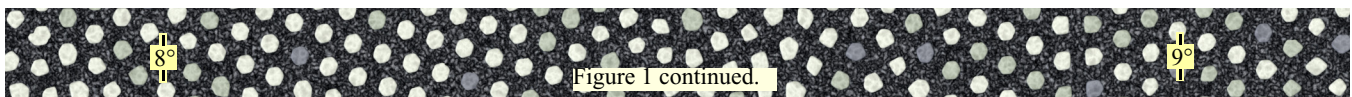
Figure 7: (a) Roorda Image

(b) Our synthetic

Figure 8: 0.5° FOV centered on our synthesized fovea.

Figure 9: Close up of our synthesized cones, ~3 μ diameter each.

of retinal receptive fields; the values are passed onto the parts of the simulator that emulates the LGN and beyond.

16 The Iris

The eye's physical pupil is the hole in the iris. When the iris dilates, the pupil changes in size. When the lens accommodates (changes focus), it does so by bulging outward, and since the iris rests on the front surface of the lens, the iris moves forward in the eye with

changes in accommodation (up to 0.4 mm). Our system includes this effect.

The physical pupils of real eyes are decentered relative to the optical axis established by the cornea. The reason for the decentering is generally believed to be to compensate for the center of the fovea being 5° away from the center of this corneal optical axis. As with most other anatomical features, the amount of decentering varies with the individual. [Oyster 1999; Atchison and Smith 2000] state that most physical pupils are decentered by ~0.5 mm, while [Wyatt
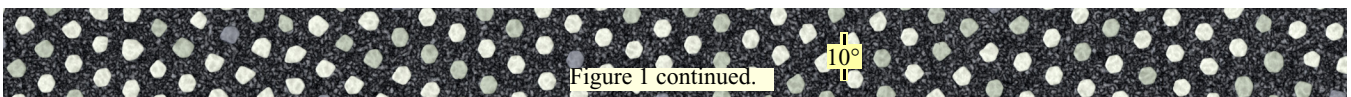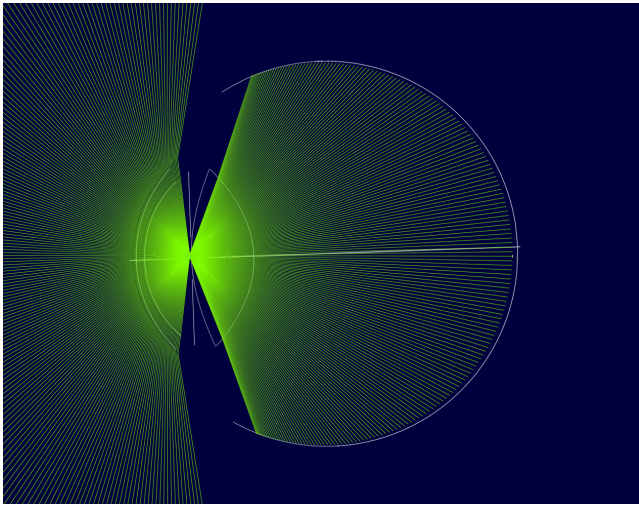

Figure 1 continued.

10°

Figure 10: 190° of rays into schematic eye. Note that at high eccentricities rays do *not* hit the retina at angles normal to it.

1995] measures decentering values for several eyes, and gives an average value of 0.25 mm. As described above, we found physical pupil decentering to be necessary for our model, where a default value of 0.4 mm is used.

The center of the physical pupil actually moves by a small amount laterally as its size changes. [Wyatt 1995] measured an average shift of 0.1 mm; extremes as large as 0.4 mm have been reported. Real physical pupils are not only slightly elliptical in shape (~6%), but have further irregular structure [Wyatt 1995]. The physical pupil is also not infinitely thin; high incident angle rays will see an even more elliptically shaped pupil due to its finite thickness (~0.5 mm). In building our system we considered these additional pupil shape details. However, at the density that our system samples rays through the pupil, none of these details other than the decentering make a significant difference in our computation results, so they are not currently model parameters. ([Wyatt 1995] comes to a similar conclusion.)

The entrance pupil of the eye is the physical pupil as viewed from outside the eye (through the cornea); in the literature, usually the term "pupil" refers to the entrance pupil. The range of its size changes is generally given as 2 to 8 mm. A number of physiological factors can effect pupil size, but there are simple models of predicted pupil diameter as a function illumination levels. To minimize optical aberrations, we generally used a somewhat smaller pupil size than these formulas would predict for the illumination levels of the video display devices being simulated (usually a 2 mm or 3 mm entrance pupil).

The actual anatomical physical pupil size (as simulated) is ~1.13 time smaller than the entrance pupil. The size and position of the pupil that the cones see (through the lens) changes again: the exit pupil. The relative direction to the center of the exit pupil from a given point on the surface of the retina is an important value; this is the maximal local light direction that the cones point in, and is involved in the Stiles-Crawford Effect I below. (This direction is sometimes just described as towards the second nodal point of the eye, but that point is ill-defined for wide field optics and decentered pupils.) Figure 10 shows 190° of external rays through the schematic eye onto the retina. The retinal synthesizer models this tilt in cone cross-sections; within the plane of the retinal sphere cones are elongated and more spread out density-wise by the reciprocal of the cosine of the angle between the direction to the center of the retinal sphere and the direction to the center of the exit pupil.

17 Cornea and Lens Density

While the cornea and lens are built from nearly optically transparent tissue, they do pass less light at some wavelengths (mostly short) than others. In the literature, the **prereceptoral filter (PRF)** effects of the combination of the cornea and lens are usually modeled as a single lens density spectra. The best current data set is on the web site [Stockman and Sharpe 2004]. (All instances of this reference in this paper implicitly also reference the original work: [Stockman and Sharpe 2000] and [Stockman et al. 1999].)

Some data exists on the individual effects of the cornea [Van Den Berg and Tan 1994]; in our system we use this data to split the Stockman & Sharpe data into a separate cornea and lens spectral transmittance function. The data is given in an average spectral density form; we had to normalize it by the average path length that rays take within our models of the cornea and lens in order to get a true spectral transmittance functions of physical optical path length.

18 Macular Pigment Density

The **macula** (see Figure 3) is an extra filter of short wavelength (blue) light covering approximately the central to 7° to 10° of the retina; there is considerable individual variation. It is generally thought to reduce lateral blue light scatter that confounds red-green signaling in the fovea. Again, we used the best current data set [Stockman and Sharpe 2004]. The effect is strongest at the center of the fovea, and falls off approximately linearly to zero at its edge.

19 Stiles-Crawford Effect

The **Stiles-Crawford effect I (SCE-I)** [Lakshminarayanan and Enoch 2003] is the reduction of perceived intensity of rays of light that enter the eye away from the center of the entrance pupil. It is caused by the waveguide nature of the inner segment of the retinal cones. It is generally thought to reduce the effect of stray (off axis) light due to scattering within the eye, and also to reduce chromatic aberration at large pupil diameters. While our system currently models scattered light by throwing it away, the chromatic effects are of considerable interest, so we felt it was important to include a simulation of SCE-I in our system.

The SCE-I is generally modeled as a parabola (in log space) as a intensity diminishing function η[r] of the radial distance r (in mm) from the center of the pupil that the light enters:

$$\eta[r] \;=\; e^{-pc \cdot r^2} \qquad (6)$$

where pc has the value commonly cited in the literature of 0.05 mm^{-2} [Lakshminarayanan and Enoch 2003].

In most systems, the SCE-I is modeled by an apodization filter: a radial density filter at the pupil. In our system, we can more accurately model the SCE-I effect at the individual cone level. This allows us to also simulate the 1° perturbations in relative orientation direction within the cones that is thought to occur. But we need to convert the standard equation above to a function of the angle φ relative to the orientation of an individual cone. With our optical model, empirically it was found that conversion from physical entrance pupil coordinates in mm to φ in radians is a linear factor of 0.047, ±0.005. After multiplying by the 1.13 physical to entrance pupil scale factor, this gives a simple first order rule of:

11°

Figure 1 continued.
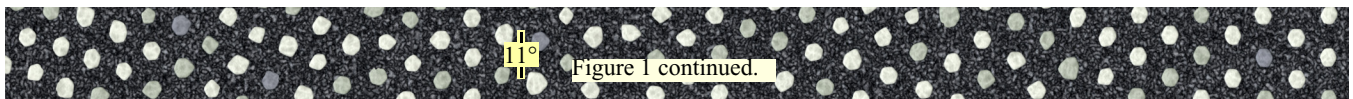
(7)

Some papers argue that the SCE-I is better modeled by a Gaussian (in log space); we have stayed with the parabolic function as the 1° perturbations already change the overall effect.

20 Cone Photopigment Absorptance

Once a photon is known to enter a cone, the probability of it being absorbed depends on its wavelength λ, the type of cone (L, M, or S), and the width and length of the cone. Each cone type has its own absorbance (photopigment optical density) spectra function $A[\lambda]$. Variations in density due to the width and length of cones are modeled as a function $D[\theta]$ of eccentricity θ. Combining these gives us an absorptance function $J[\lambda, \theta]$ of wavelength and eccentricity:

$$J[\lambda, \theta] \ = \ 1 - 10^{-D[\theta] \cdot A[\lambda]} \qquad (8)$$

Again for $A[\lambda]$ we used the spectral data from [Stockman and Sharpe 2004] for the L, M, and S cones. We also used their estimates of $D[0]$ (at the center of the fovea): 0.5 for the L and M cones, and 0.4 for S cones. By 10°, $D[\theta]$ for L and M linearly falls to 0.38; By 13° $D[\theta]$ for S falls to 0.2.

Only 2/3^{rds} of photons that are absorbed by a photopigment **photo-isomerize** the molecule. These photoisomerizations within the cone's outer segment start a chemical cascade that eventually leads to a photocurrent flowing down to the cone cell's axon and changing its output. The effects of this cascade can be fairly accurately modeled by a simple set of differential equations [Hennig et al. 2002], or the output can be even more simply approximated as a logarithmic function of the rate of incoming photons. While under ideal conditions as few as five photoisomerizations within a 50 ms window can be perceived, generally it takes at least 190 photoisomerizations within a 50 ms window to produce a measurable response. The linear response of a cone occurs between 500 and 5,000 photoisomerizations per 50 ms; above this significant numbers (more than 10%) of cone photopigments are in a bleached state, and more complex models are needed. (Mimicking these effects are part of the process of producing high dynamic range images.) However, as was shown in Section 5, sitting right next to a 2,000 lumen digital projector produced only about 1000 photoisomerizations per 16 ms per (foveal) cone, or about 3,000 photoisomerizations per 50 ms. Thus for our purposes of simulating the effects of display devices on the eye, we are generally operating in the simple range, and do not at present have to simulate any non-linear saturation processes. There are many other suspected regional non-linear feed-back mechanisms from other cells on the retina to the cones that may effect the actual output produced by a cone. To separate out these effects, our system currently produces as output a per cone count of the photons that would have been photoisomerized by a population of un-bleached photopigments. Any post-cones retinal circuitry processing is beyond the scope of the present paper.

21 Wavefront Tracing and Modeling Diffraction

The quality of retinal images on the human eye are usually described as optical aberration limited when the pupil is open fairly wide (>3 mm), and as diffraction limited at the smallest pupil diameters (2-3 mm). Some authors (such as [Barton 1999]) approximate both these PSF effects as Gaussians of specific widths, and then use the square-root of the widths squared to obtain a combined distortion Gaussian PSF. Unfortunately, this approach is too simplistic for
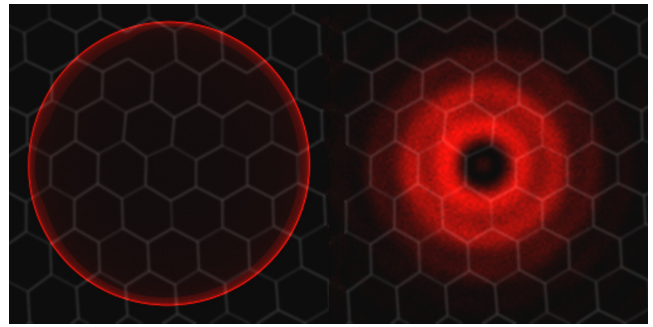

Figure 11: (a) non-diffracted PSF, (b) diffracted PSF (20μ×20μ The faint hexagons are 2.5μ cone apertures to show relative scale.

accurate retinal images, at least for the region of greatest interest (foveal).

For many axial symmetric cases, optics theory provides simple (non-integral) closed form solutions for the PSF (Seidel aberrations, Bessel functions, Zernike polynomials). Unfortunately for the practical case of the living human eye, which is not even close to axial symmetrical, one must solve the integral solutions numerically on a case by case basis. Furthermore, because of loss of shift invariance, different PSFs have to be custom calculated for every different small region of the retina [Mahajan 2001; Thibos 2000].

These PSFs are also different for different wavelengths of light. The PSF produced by defocused optics can produce some surprising diffraction patterns, as is shown in Figure 11b, vs. the non-diffracted PSF of Figure 11a. Note the hole in the center of the diffracted image: a point projects into the *absence* of light. While this strange pattern is reduced somewhat when a wider range of visible wavelengths are summed, it does not go away completely. (For some similar images, see pp. 151 of [Mahajan 2001]). Figure 11b shows why accurate PSFs of the eye cannot be approximated by simple Gaussians.

To numerically compute the diffracted local PSF of an optical system, a wavefront representing all possible photon paths from a given fixed source point through the system must be modeled. When the wavefront re-converges and focuses on a small region of the retina, the different paths taken by different rays in general will have different optical pathlengths, and thus in general the electric fields will have different phases. It is the interference and support of these phases that determine the local PSF, representing the relative probability that a photon emitted by that fixed source point will materialize at a given point within the PSF. At least several thousand rays need their path to the pupil simulated, and then in turn their several thousand *each* possible paths to the surface of the retina need to be simulated, pathlengths and thus relative phases computed, and then phases summed at each possible impact point.

Modern commercial optical packages have started supporting non axial symmetric diffracted optical systems; however for our system we found it more convenient to create our own custom optics code optimized just for the human eye. Our optical code traces the refracted paths of individual rays of a given wavelength through any desired sequence of optical elements: the cornea, the iris, the lens, and to the retina. Along the way, wavelength specific losses due to reflection, scatter, and absorption (prereceptoral filters) are accumulated.

Computing the full diffraction PSF for each photon being simulated would be too much even for modern computing power, so we pre-

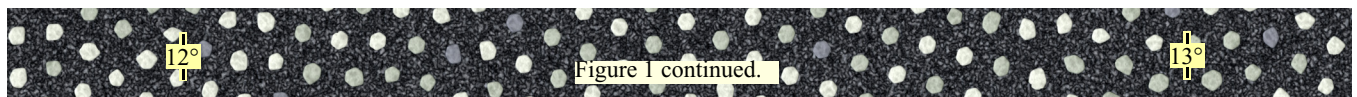12°   Figure 1 continued.   13°

Figure 12: 30×30 pixel display imaged onto synthesized retina. Each 80% fill factor display pixel is visible (broadband light).

compute an array of diffracted PSFs for a given parameterized eye, accommodation, and display screen being viewed. Because the PSF is invariant to the image contents, and to small rotations of the eye, a single pre-computed array can be used for many different frames of video viewing. We also only have to compute an array of PSFs for the particular portion of the retina needed for a given experiment.

While parameterizable, currently we currently trace 1024 randomly perturbed primary paths from a source point to the pupil, and from each of these we compute the ray, pathlength, and phase to each of 128×128 points on the surface of the retina. Thus PSF$[\bar{p}, \bar{\lambda}]$ is the 128×128 probability density array for a given quantized display surface source point $\bar{p}$ and a given quantized frequency of light $\bar{\lambda}$. The

physical extent of the 128×128 patch on the retina is dynamically determined by the bounds of the non diffracted PSF, but is never allowed to be smaller than 20μ×20μ. This means that at best focus we have probability data at 0.15μ resolution, allowing us to accurately rasterize photon appearance events onto 2.5μ×2.5μ polygonal outline cone optical apertures. Again while parameterized, for the examples shown in this paper $\bar{\lambda}$ was quantized every 10 nm of wavelength, for a total of 45 spectral channels covering wavelengths from 390 to 830 nm. In space, $\bar{p}$ was quantized for physical points on the display surface corresponding to every 300μ on the retina (1°). Photons are snapped to their nearest computed wavelength. The position of the center of the PSF is linearly interpolated between the four nearest
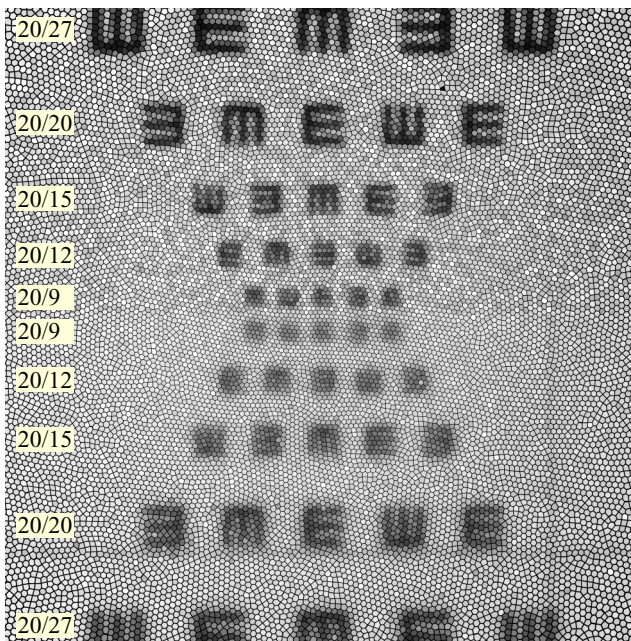
Figure 13: Sloan acuity chart 720×720 pixels imaged on cones. The upper half of the chart is just 543 nm light, lower is broadband

Figure 15: Sloan acuity chart with vertical drift eye motion of 2° pe second. Rendered with 543 nm light.

spatial PSFs; the probability density function itself is snapped to the one of the closest PSFs. The accumulated reflection, scatter, and absorption loss: the prereceptoral filter PRF$[\overline{p}, \overline{\lambda}]$, is stored with each PSF$[\overline{p}, \overline{\lambda}]$, and is also interpolated between them in use.

If we were simulating more general viewing of the three dimensional world, we would also have to generate PSFs from different distances in space as well as level of focus. But these are not needed for simulating the viewing of a flat fixed distance display surface.

22 Putting it all Together: System Results

The first step is to parameterize and synthesize a retina. The same parameterization is then used to interactively adjust the optics (including focus) and the working distance of the simulated display surface; this results in locking down all the optical parameters needed for the next step: computing the array of diffracted PSFs. Those in turn are used as input by the photon simulation. Given the parameters of display surface sub-pixel elements, a frame of video pixels to display, and an eye drift rotation during this frame, we can then simulate the emission of every photon that would occur during this time frame (at least those that might enter the eye). Each simulated photon created is assigned a specific point p in space, t in time, and

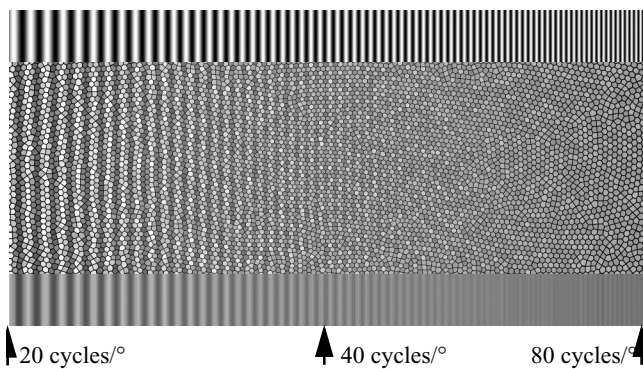| 20 cycles/° | 40 cycles/° | 80 cycles/° |

Figure 14: Spatial frequency ramp test. Top is original sine-wave; middle is sine-wave projected onto synthesized retina; bottom is vertical average of cone signal. Cut-off is 40-60 cycles/degree.
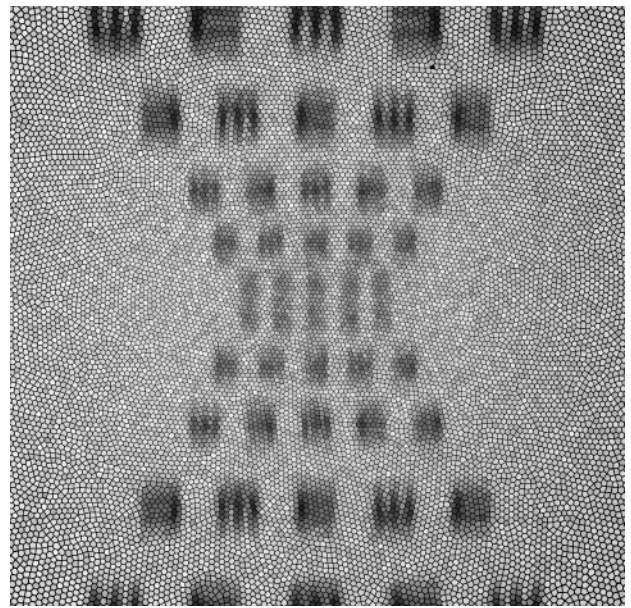
wavelength λ. p is randomly generated from within the sub-pixel display surface element extent. t is randomly selected from within the time interval of the temporal characteristics of the specific display device. λ is randomly generated from within the weighted spectral probability distribution of the display device sub-pixel type.

The quaternions that represent the endpoints of the drift can now be used to interpolate the orientation of the eye given t. This is used to transform p to the point p' on the display surface where the eye would have seen p had no rotation occurred. Using quantized $\overline{\lambda}$ and $\overline{p'}$, PSF$[\overline{p'}, \overline{\lambda}]$ and PRF$[\overline{p'}, \overline{\lambda}]$ can be found, as well as the three closest neighboring values of each. The sum effects of all the prereceptoral filters (cornea, lens, macula) are computed by interpolating between these four PRFs; this gives the probability of the current photon never reaching past the macula. A random number in the range [0 1) is generated, and if it is below this probability this photon is discarded. Otherwise, the center of the landing distribution for the photon is computed by interpolating the centers of the four PSFs by their relative distances from p'. The 128×128 PSFs are actually represented as an accumulated probability array. In this way, a random number is generated and then used to search the 128×128 table until the entry closest to, but not above the random value is found. The associated (x y) location in the table is the location that this photon will materialize at. Using the known retinal center point of the interpolated PSFs, and a 2D scale and orientation transform associated with the PSF, this (x y) location can be transformed to a materialization point on the retinal sphere.

Using spatial indices and polygonal cone aperture outlines imported from the original synthesis of the retina, a list of candidate cones that might contain this point is generated. All candidate cones use plane equations of their polygonal entrance to see if they can claim this photon. If it falls outside all of them (e.g., hit the edge of a cone, or a rod, or a void), the photon is discarded. Otherwise the unique cone owning the photon subjects it to further processing. The individual (perturbed) orientation φ of the cone is used to determine the probability that the SCE-I $\eta[\varphi]$ will cause the photon to be rejected. Otherwise, the photon is subjected to its final test: the probability of absorptance J$[\lambda, \theta]$ by a photopigment in this particular type of cone (L M or S) of a photon of wavelength λ. Now a random number generated with a lower value than this probability will cause this

20/12 line from Figure 15

Motion Frame #1

Motion Frame #2

Motion Frame #3

Motion Frame #4

Motion Frame #5

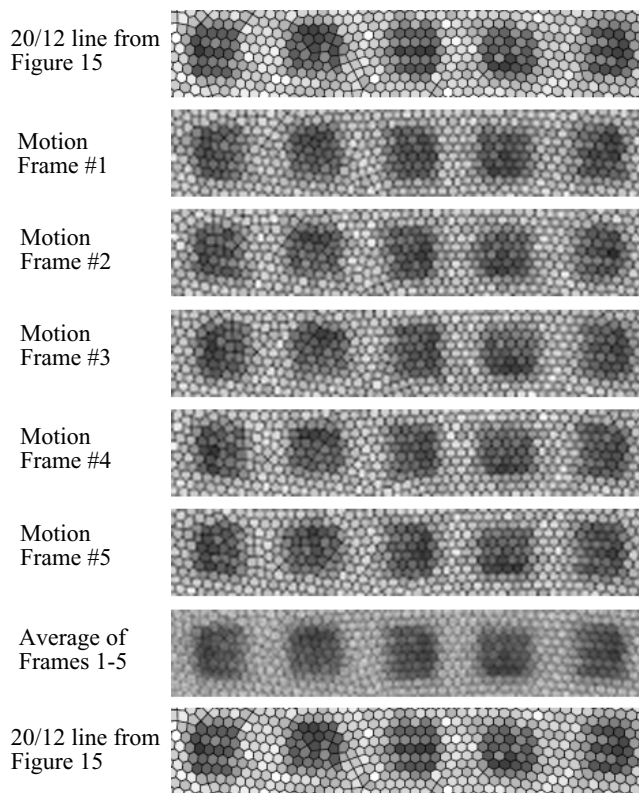**Average of Frames 1-5**

20/12 line from Figure 15

Figure 16: 30 minutes per second drift motion blur of the 20/12 line from Figure 15 (repeated at top and bottom). The average of frames 1-5 is actually more legible than the non-blurred image.

particular cone to increment by one the number of photons that it has absorbed during this frame.

This process repeats for all photon emission events for each frame of the simulated video sequence. At the end of each frame, the cone photon counts are output to a file. These files are used to generate visualizations of the results by using each cone's photon count to set a constant fill intensity level of its polygonal aperture, normalized by the aperture's area, and the maximum photon count. (And, of course, the image is flipped top-bottom and left-right.) Figures 12-16 were generated this way. Figure 12 shows the result of a 30×30 pixel portion of a display surface being viewed by the central 1.8° of the fovea The video image is of a face, subtending the same visual angle that a person would at a distance of 8 meters. Each cone type has been lightly tinted corresponding to its color, and the relative intensities of each cone type white balanced. The display pixels are 2.5 minutes of arc on a side, and have a fill factor of 80%; the square edges and gaps between the pixels can clearly be seen in the image.

How does one test such a complex simulated eye? The same way you test real eyes: show an eyechart, show it a variable spatial frequency test pattern. Thus in Figure 13 we show a different simulated display device: an acuity chart printed on a 1200 dpi laser printer. Viewed just using light at the chosen focus wavelength, the 20/12 acuity line is mostly readable; with more broad spectrum illumination acuity drops to 20/15. This is consistent with normal human vision of between 20/10 and 20/20. And so also in Figure 14 we describe a variable spatial frequency 100% contrast sine-wave test at 543 nm light; the result is similar to the normal human cut-off of 40-60 cycles/degree.

Figure 15shows what happens when we enable the motion blur due to drifts. The image is the 543 nm illumination of the same Sloan

acuity chart as before, but this time with a vertical drift of 2° per minute. This is 4× faster than the maximum normal drift rate, but is the maximum rate at which your eye can miss-target relative to the fixation point without you being conscious of it. We show this fast drift rate in order to make the motion blur effect maximally visable. Note the strobing horizontal bars in the largest E's in the chart. Figure 16 shows what happens when we enable a slower 30 minute per second motion blur due to drifts. The top image is the 20/12 line from Figure 13.The next five frame down are five consecutive frames from a 30 minute per second drift motion blur rendering of the same 20/12 line. Note how any one of these five frame is blurrier and less legible than the top un-blurred one. But something interesting happens when you average these five frames together, as we did in the next one down (second from the bottom). The average image is actually *more* legible than the un-blurred one (which is duplicated again just below it at the very bottom of the figure for comparison). This is a suspected reason why the human eye almost always has a slow drift: imaging the same external object onto different cone sampling patterns results in the visual system getting a better resolution view of the object.

Doing these sort of averaging over blur and realistic cone sampling patterns was a major goal for the system described in this paper. The actual averaging that the visual system (and eventually our simulator) does is more sophisticated; it involves processing retinal receptive fields of cone outputs, and processing visual cortex spatial/temporal receptive fields of those outputs. The simplistic simple image averaging show here merely illustrates the general concept.

23 Validation

One of the reasons we restricted our lens model to a simple variant of a previous published and validated model was to remove the optics as a validation issue. With non-diffracted rays, we obtained the same scatter plots at various eccentricities as in the original paper [Escudero-Sanz and Navarro 1999]; these did not change appreciably after lens decentering and accommodation to a closer focal distance. Our generalized diffraction calculations generate similar PSFs as other published work [Mahajan 2001].

Our synthesized retinas have the same neighbor fraction ratio (6.25) as studies of human retinas (for the area near the fovea). The density of cones/mm^2 measured *empirically* in the output of our synthesizer matches the desired statistics from [Curcio et al. 1990].

At present, there is no way to physically measure cone photon absorption counts for complex images, this was one of the main reason why we built a simulation. Thus while this section of our simulations were all computed based on known measurements and physical laws, there is no way to directly validate the results. Some past experiments have directly measured optical images formed on the surface of the retina, but these images represent light reflected back *off* the retina, not light selectively absorbed by the cones. Our acuity chart and spatial frequency chart cone photon count images, which are well within the range of physiologically measured values, are at present the best validation of this stage of our results.

24 Future Work

The system at present has most of the mechanisms necessary to also simulate scotopic (rod) vision, the only thing missing are the rods! The retinal synthesizer would need to be extended to deal with mixing the much smaller rod cells with the larger cone cells during the synthetic growth phase. We would like to try more complex surface shapes for the optics and retina. The system already generates receptive fields of cones, in the future we could go beyond simulating only photons and add simulation of current models of some of the rest of the layers of retinal circuitry (such as [Hennig et al. 2002]); beyond that lies the LGN and the simple and complex cells of the

visual cortex. We would like to use such an extended model to produce a more definitive characterization of the resolution limits of the human eye relative to display technology than was done in [Deering 1998]. While color vision theory has its own complications, we do maintain superbly accurate spectral information up to the cone level (of each cone type), and could do something more with it.

25 Conclusions

A new photon accurate model of the interaction of modern display devices and the human eye has been described. The retina synthesizer generates patterns of cones very similar to biological ones; the photon by photon rendering of display images through the diffracted optics of the eye onto these cones produces cone photon count images that match the known spatial response curves of the eye.

Acknowledgements

References

AHUMADA, A, AND POIRSON, A. 1987. Cone Sampling Array Models. *J. Opt. Soc. Am. A 4,8,* 1493-1502.

ATCHISON, A., AND SMITH, G. 2000. *Optics of the Human Eye*. Butterworth-Heinemann.

BARSKY, B. 2004. Vision-Realistic Rendering: Simulation of the Scanned Foveal Image from Wavefront Data of Human Subjects. *First Symposium on Applied Perception in Graphics and Visualization,* 73-81.

BARTEN, P. 1999. *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. SPIE.

VAN DEN BERG, T., AND TAN, K. 1994. Light Transmittance of the Human Cornea from 320 to 700 nm for Different Ages. *Vision Research, 34,* 1453-1456.

Cook, R. 1986. Stochastic Sampling in Computer Graphics. *ACM Transactions on Graphics, 5, 1,* 51-72.

COOK, R, CARPENTER, L, AND CATMULL, E. 1987. The Reyes Image Rendering Architecture. In *Computer Graphics (Proceedings of ACM SIGGRAPH 1987), 21, 4* ACM, 95-102.

CURCIO, C., SLOAN, K., KALINA, R., AND HENDRICKSON, A. 1990. Human Photoreceptor Topography. *J. Comparative Neurology 292,* 497-523.

DEERING, M. 1998. The Limits of Human Vision, in *2nd International Immersive Projection Technology Workshop*.

DOBKIN, D., EPPSTEIN, D., AND MITCHELL, D. 1996. Computing the Discrepancy with Applications to Supersampling Patterns. *ACM Transactions on Graphics, 15, 4,* 354-376.

ESCUDERO-SANZ, I., AND NAVARRO, R. 1999. Off-Axis Aberrations of a Wide-Angle Schematic Eye Model. *J. Opt. Soc. Am. A, 16,* 1881-1891.

GEISLER, W. 1989. Sequential Ideal-Observer Analysis of Visual Discriminations. *Psychological Review, 96, 2,* 267-314.

GLASSNER, A. 1995. *Principles of Digital Image Synthesis*. Morgan Kaufmann.

HASLWANTER, T. 1995. Mathematics of Three-Dimensional Eye Rotations. *Vision Research, 35,* 1727-1739.

HALSTEAD, A., BARSKY, B., KLEIN, S., AND MANDELL, R. 1996. Reconstructing Curved Surfaces from Specular Reflection Patterns Using Spline Surface Fitting of Normals. In *Proceedimgs of ACM SIGGRAPH 1996 23*, Annual Conference Series, 335-342.

HENNIG, M, FUNKE, K., AND WÖRÖGTTER, F. 2002. The Influence of Different Retinal Subcircuits on the Non-linearity of Ganglion Cell Behavior. *The Journal of Neuroscience, 22, 19,* 8726-8738.

LAKSHMINARAYANAN, RAGHURAM, A., AND V, ENOCH, J. 2003. *The Stiles-Crawford Effects*. Optical Society of America.

LIOU, H., AND BRENNAN, N. 1997. Anatomically Accurate, Finite Model Eye for Optical Modeling. *J. Opt. Soc. Am. A 14,8,* 1684-1695.

MAHAJAN, V. 2001. *Optical Imaging and Aberrations: Part II Wave Diffraction Optics*. SPIE.

MARTINEZ-CONDE, S, MACKNIK, S. AND HUBEL, D. 2004. The Role of Fixational Eye Movements in Visual Perception. *Nature Reviews: Neuroscience, 5, March*.

OYSTER, C. 1999. *The Human Eye: Structure and Function*. Sinauer.

POYNTON, C. 2003. *Digital Video and HDTV*. Morgan Kaufman.

RODIECK, R. 1998. *The First Steps in Seeing*. Sinauer.

ROORDA, A., METHA, A., LENNIE, P, AND WILLIAMS, D. 2001. Packing Arrangements of the Three Cone Classes in Primate Retina. *Vision Research, 41,* 1291-1306.

ROORDA, A., AND WILLIAMS, D. 1999. The Arrangement of the Three Cone Classes in the Living Human Eye. *Nature*, 397, 520-522.

SMITH, G. 2003. The Optical Properties of the Crystalline Lens and their Significance. *Clinical and Experimental Optometry, 86,1,* 3-18.

STEINMAN, R. 1996. Moveo Ergo Video: Natural Retinal Image Motion and its Effect on Vision. In *Exploratory Vision: The Active Eye*. M. Landy et al., Ed., Springer Verlag, 3-49.

STOCKMAN, A., SHARPE, L., AND FACH, C. 1999. The spectral sensitivity of the human short-wavelength cones. *Vision Research, 40,* 1711-1737.

STOCKMAN, A., AND SHARPE. 2000. Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research, 39,* 2901-2927.

STOCKMAN, A., AND SHARPE, L. 2004. Colour & Vision Database. http://cvrl.ucl.ac.uk/.

THIBOS, L. 2000. Formulation and Sampling of the Retinal Image. In *Seeing*, K. De Valois, Ed., Academic Press, 1-54.

TYLER, C. 1997. Analysis of Human Receptor Density, in *Basic and Clinical Applications of Vision Science, Ed. V*. Kluwer Academic Publishers, 63-71.

WYATT, H. 1995. The Form of the Human Pupil. *Vision Research, 35, 14,* 2021-2036.