Disparity sensitivity in man and owl: Psychophysical evidence for equivalent perception of shape-from-stereo

Robert F. van der Willigen   Department of Biophysics, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, The Netherlands

Wolf M. Harmening   Department of Zoology and Animal Physiology, RWTH Aachen, Aachen, Germany

Sabine Vossen   Department of Zoology and Animal Physiology, RWTH Aachen, Aachen, Germany

Hermann Wagner   Department of Zoology and Animal Physiology, RWTH Aachen, Aachen, Germany

The perception of shape-from-stereo is best characterized by the spatial disparity-contrast sensitivity function (DSF). This is the stereo analogue of the well-known luminance-contrast sensitivity function (CSF). In principle, the DSF and CSF portray a visual system's ability to detect spatial modulation as specified by changes in binocular disparity and luminance, respectively. In humans, less fine detail is visible in the stereo domain than is possible in the luminance domain. Here, we characterize for the first time the DSF in a non-human species, viz. the barn owl. At the same time, we re-examined the human DSF with identical apparatus and methods to directly compare between two vertebrate species that evolved stereovision independently. We discovered a close relationship between the owl and human ability to detect shape-from-stereo. In particular, the shift in absolute position between the human and owl DSF, as measured here, closely corresponds to the shift in absolute position between their respective CSFs, as known from the literature. In conclusion, our study establishes unprecedented experimental proof of a striking similarity in the prowess of humans and owls to achieve shape-from-stereo.

Introduction

Although comparative physiology has incorporated behavioral measurements of stereopsis in a wide range of animals (for review, see Howard and Rogers, 2002), the perceptual judgment of 3D shape has so far not been assessed. A more complete characterization of a visual system's ability to represent disparity-defined form (hereafter, shape-from-stereo) can be obtained by measuring the disparity sensitivity function (DSF). DSF measurements yield information about an individual's ability to detect the shape of objects defined by small-disparities over an extended range of object sizes and orientations. Typically, the DSF is a plot of the reciprocal of stereo acuity (i.e., stereo sensitivity to small disparities) as a function of the spatial frequency of depth modulations in cycles per degree of visual angle. It has the form of an inverted-U envelope, or band-pass characteristic. To be more precise, the DSF provides a disparity-based transfer characteristic at threshold, showing a distinct peak and decreasing sensitivity on either side of this peak (i.e., a high and low frequency fall-off in sensitivity, respectively). Thus, the high frequency fall-off in sensitivity provides a measure of stereo resolution: the finest spatial resolution of changes in depth that can be seen (Tyler, 1973, 1974). Most notably, besides the above described upper limit of stereo resolution, the DSF also has a lower limit: the largest possible modulations in disparity beyond which stereopsis collapses and fusion is not possible. Thus, no disparity modulation can produce a sensation of 3D shape outside these limits (Filippini & Banks, 2009; Tyler, 1975, 1991).

The DSF has been measured in humans alone (Bradshaw & Rogers, 1999; Rogers & Graham, 1982; Tyler, 1974), using sinusoidally modulated disparity gratings as can be provided by pairs of random-dot patterns (Figure 1). In such corrugated stereograms (hereafter, corrugated-RDSs), the disparity information is concealed by random matrices of thousands of minute dots. Under stereo viewing, however,
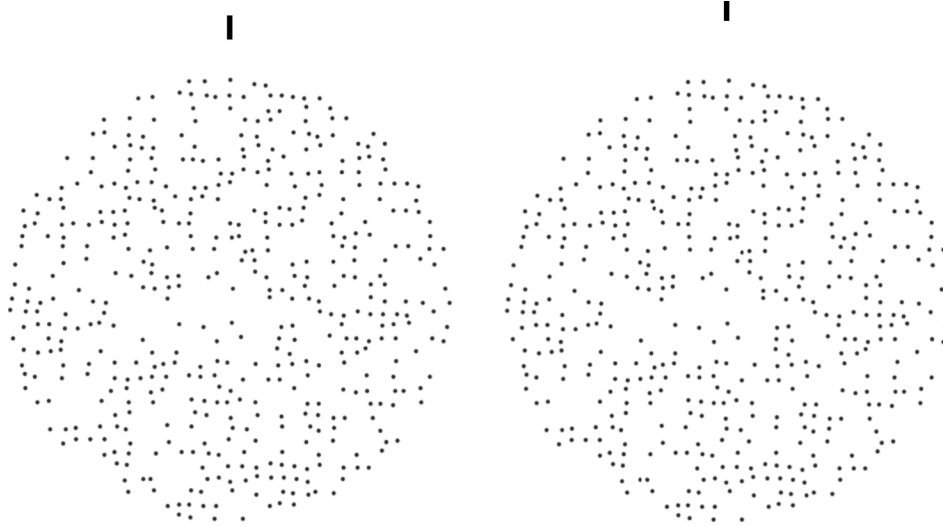
Figure 1. Example of a horizontal corrugated RDS stereo pair. When cross-fused (use vernier lines to fuse), the stereogram reveals a grating in depth, defining three complete cycles with a concave central corrugation. Divergent fusion reverses the sign of apparent depth and would evoke a convex central corrugation. A pictorial representation of the induced 3D sensation is illustrated in Figure 2. In the experiments, subjects were required to indicate the depth sign of the central corrugation, i.e. concave vs. convex (compare inset Figure 3).

corrugated-RDSs with high dot-densities (Banks, Gepshtein, & Landy, 2004) can produce a compelling percept of 3D shape: a corrugation in depth. Thus, corrugated-RDSs constitute an analytical technique of relating the sensitivity for perceiving corrugations in depth to their disparity modulation in its purest form.

The use of corrugated-RDSs in the stereo domain (i.e., shape-from-stereo) is closely related to the use of sinusoidally modulated luminance gratings to characterize the spatial contrast sensitivity function (CSF) in the luminance domain (i.e., shape-from-luminance) (Campbell & Robson, 1968; Watson, Barlow, & Robson, 1983). When considering the resolution for luminance gratings, however, stereo resolution is surprisingly poor (Banks et al., 2004; Bradshaw & Rogers, 1999; Regan, 2000). The DSF peaks around 0.5 cycles/deg, whereas the CSF peaks around 5 cycles/deg. In the luminance domain, therefore, more fine detail is visible then is possible in the stereo domain. A second, major difference between the DSF and the CSF is their orientation dependency, or lack of it (Bradshaw & Rogers, 1999; Rogers & Graham, 1983). Typically, at low frequencies, thresholds for detecting vertically oriented corrugated-RDSs are higher than those for detecting horizontally oriented corrugated-RDSs. The anisotropy disappears or may reverse for high frequencies and large individual differences in its extent are often observed (Hibbard, Bradshaw, Langley, & Rogers, 2002). Although sensitivity for luminance defined gratings is subject to an 'oblique effect', no anisotropy for the principal orientations (vertical vs. horizontal) has been observed in CSF testing (Campbell & Robson, 1968; Watson et al., 1983).

Conceivably, low spatial stereo resolution and stereoscopic anisotropy reflect fundamental constraints of optimal neural networks that make possible the process of shape-from-stereo in biological systems (Banks et al., 2004; Burt & Julesz, 1980; Tyler, 1991). At the same time, it is most plausible to assume that luminance sensitivity inadvertently plays a role in shape-from-stereo because the texture elements of real world objects are most often defined by differences in their luminosity (Regan, 2000). Thus, although the perceived 3D shape of a natural image is inherently ambiguous and likely to be influenced by luminance sensitivity, the brain could capitalize on statistical regularities of natural images to single out the 3D interpretation that is statistically most likely (Todd, 2004). By reversing this argument, we postulate that: (1) if both low spatial resolution and orientation selective mechanisms in the stereo domain are due to brain's strategy to maximize the probability of correctly interpreting a given object's shape, and (2) if both the extent and magnitude of the best possible stereo resolution critically depends on shape-from-luminance; then all species with functional stereopsis should have a DSF with: (1) a band-pass characteristic that displays anisotropic sensitivity for the principal orientations, and (2) a peak sensitivity that is shifted to the left relative to the peak sensitivity of the CSF. Here, we determine the DSF in both humans and barn owls, using identical displays and psychophysical procedures to make possible a legitimate comparison between these two evolutionary distinct visual systems. Since the CSF has been determined for both humans and owls, we also aim to establish if the CSF has predictive power. That is, if the CSF is known: Is it possible to estimate the absolute position and peak sensitivity of the DSF? To our best knowledge, this study is the first to examine the DSF in a non-primate that is known to possess functional stereopsis (Howard &

Rogers, 2002; Medina & Reiner, 2000; Nadler, Angelaki, & DeAngelis, 2008; van der Willigen, Frost, & Wagner, 1998).

Materials and methods

Subjects and apparatus

Two adult barn owls (Tyto alba pratincola) were cared for and treated under a permit from the Regierungspräsidium Köln (Germany). The principles of laboratory animal care (National Institute of Health publication No. 86-23, revised 1985) were followed. A detailed description of the animal preparation and care has been given elsewhere (van der Willigen, Frost, & Wagner, 2002, 2003; Wagner, 1993). During experimentation, the birds wore a spectacle frame containing polarized filters plus a small head-tracking sensor. The apparatus, stimuli, and procedures used to test two naive adult human subjects were identical to those used to test the owls, unless specified otherwise. Both human subjects had normal visual acuity and stereo acuity of at least 40 arc sec (Randot stereotests).

A detailed description of the experimental setup and calibration methods, has been given elsewhere (van der Willigen et al., 2002, 2003). Briefly, the stimulus sequences, reinforcing contingencies and data processing were controlled by a Silicon Graphics workstation (Mountain View, CA). Two response bars, within easy reach of the owls, were symmetrically placed to the left and right of a remotely operated food dispenser that delivered small pieces of meat. Head position and orientation was monitored online by tracking the owls' gaze direction with aid of a magnetic tracking device (model: MiniBird; Ascension, Burlington, VT) under infrared illumination. A color (P22-phosphor) cathode ray tube (CRT) functioned as the stimulus presenting panel. Gamma-correction was applied to produce a linear relationship between luminance and the color level specified by the workstation. Spatial calibration involved the creation of a look-up table that converted the desired visual directions into CRT coordinates. Experiments took place in a darkened room after 10 minutes of dark adaptation. The viewing distance approximated 110 cm, making one pixel of the screen subtend 1.24 arcmin of visual angle.

Stimulus presentation and configuration

To create stereoscopic displays, one eye's image was written to the even scan lines, and the other eye's image to the odd scan lines of the CRT, each with a frame rate of 60 Hz. RDSs were polarized by use of a liquid crystal modulator (LCM, model: SGS310; Tektronix, Beaverton,

OR) placed directly in front of the CRT. The LCM transmitted and blocked the left and right eye's images (hereafter, half-images) alternately in synchrony with the CRT frame rate (120 Hz) when viewed through a set of differently polarized filters. All stereograms were displayed in red because the liquid crystal shutter works best at long wavelengths (interocular crosstalk <2%).

Anti-aliasing was obtained by use of Gaussian blobs. This allowed dot displacement at virtually arbitrary positions between integral pixel locations by placing the center of maximum luminance at the desired location using a Gaussian function. That is, all stimuli comprised sparse blobs in the form of circular Gaussian luminance profiles (standard deviation, i.e. half-width at half-height equaled 2 arcmin). The individual blobs had a maximum luminance of 4.0 cd*m^{-2}.

The stimuli proper were red stereoscopic surfaces presented against a dark background (<0.01 cd*m^{-2}). These surfaces consisted of two half-images each covered with a fresh matrix of randomly distributed Gaussian blobs. Blob density was homogeneous, approximating 10 dots/deg^2. All blobs were contained within a circular aperture (diameter size: 15 degree), and thus around 1780 individual blobs were visible in the stimulus. The relative disparity, r, of the individual blobs was calculated as follows:

$$r(x, y) = \pm(A/2)\cos[2 \cdot \pi \cdot CF(y\cos(\rho) - x\sin(\rho)) + \theta], \quad (1)$$

where x and y provide the blob coordinates within the circular aperture, and A, CF, ρ and θ represent the peak-to-trough amplitude, spatial frequency, orientation, and phase of the corrugated-RDSs. During each trial, the subject's absolute viewing distance was monitored by means of a magnetic tracking device. Trials during which the viewing distance deviated more than 1 cm form the desired 110 cm were aborted and repeated at a later time.

On the CRT, the cosinusoidal test corrugation frequency (test-CF) varied between half a cycle and 17 complete cycles on the display: 1/2, 1, 3, 5, 9, 17 cycles. Accordingly, stereoscopic sensitivity (see section Data analysis) was measured for six CFs across a 3.5 octave range: 0.0439, 0.0877, 0.2632, 0.4386, 0.7895 or, 1.4912 cycles per degree of visual angle (cpd). The corrugations were centered on the midline of the observer. The central corrugation was set to be "concave" (pointing away from the observer) or "convex" (pointing towards observer). Both vertical and horizontal corrugations were tested. To create surface curvature in two dimensions either a vertical (in case of a horizontal test-CF) or, a horizontal (in case of a vertical test-CF) baseline corrugation of exactly 1/2 a cycle on the display was added. The test-CF to baseline-corrugation amplitude ratio equaled 2.0. These baseline curvatures were used to create stimulus surfaces that looked more 'natural.' Moreover, they prevented the

subjects making use of the disparity of only a few central blobs while performing the concave/convex judgement (see also section Control experiments). A gray cross (1° by 4°, 0.18 cd*m^{-2}) at the center of a completely darkened stimulus aperture functioned as the observation stimulus.

Peak-to-trough disparities were created by introducing equal but opposite shifts to corresponding blobs in the half-images of a single RDS. Surfaces with horizontally oriented corrugations were generated by applying cosinusoidal shifts to each row of blobs. Vertically oriented corrugations were generated by applying cosinusoidal shifts to each column of blobs. Calculation of binocular retinal disparity was based on the assumption that the observers maintained fixation at a distance of 110 cm. The interpupillary distance of owl O1 was 40 mm (±4 mm, *SD*); that of O2 was 38 mm (±4 mm, *SD*). The sign of the depth displacement (or disparity) relative to the fixation plane was made positive when the central corrugation was convex, and made negative in case of a concave central corrugation. In this way, the surfaces contained either negative or, alternatively, positive disparities relative to the plane of fixation.

Animal training procedures and transfer to the discrimination paradigm

A two alternative forced-choice (2AFC) task was used to shape the discriminative performance of the owls (van der Willigen et al., 2003). Initially, the animals learned to avoid head movements and were required to fixate the center of the stimulus aperture from a predefined primary position when presented with the observation stimulus; head orientation deviations of less than 2 degree of visual angle were allowed.

Subsequently, the owls were trained to discriminate depth corrugations. These corrugations could contain either a horizontal or a vertical corrugation of exactly three complete cycles on the display, as shown in Figure 2. The peak-to-trough disparity amplitude was 12 arcmin. By pecking at one of the two response bars the owls reported whether the central corrugation was concave, or alternatively, convex. A trial was initiated after the owls attended (motionless) to the observation stimulus for 3 up to 6 seconds. Prior to stimulus onset the stimulus display was made dark for approximately 500 ms. In addition, the animals learned to avoid head movements just before stimulus onset and were required to gaze in the direction of the fixation-cross of the observation stimulus. To this end the vertical and horizontal orientation directions of the owls' head were detected automatically (see section Subjects and apparatus). Trials in which the head-orientation deviations were greater or equal to 1.5° were aborted and repeated later, which occurred in <10% of all trials. After the owls attained reliable performance again, sensitivity was measured as described below. Note
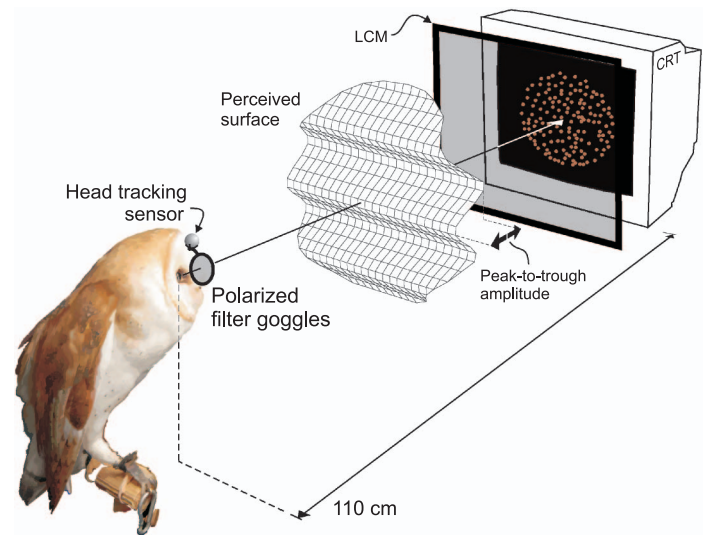

Figure 2. Simplified schematics of the experimental setup. A liquid crystal modulator (LCM) grants successive viewing of alternating random dot stereo pairs, shown on the stimulus display (CRT). When viewed through polarized filter goggles, a corrugation in-depth is perceived. The corrugated surface shown here resembles the percept (a convex central corrugation that points towards the observer) arising from Figure 1 when viewed under divergent fusion. For clarity, the two response keys and the feeder apparatus are not shown, and individual parts of the setup are not drawn to scale.

the central corrugation of each surface was aligned symmetrically, both vertically and horizontally, to the owls' viewing geometry using the position information as provided by the tracking device. Trial progression was self-paced. However, the actual exposure time was approximately 800 ms (owl 1: mean 731 ms, *SD*: 51 ms; owl 2: mean 815 ms, *SD* 42 ms).

The two human subjects were required to perform the same task as the owls under identical conditions. However, presentation times and inter-stimulus interval were fixed at 1000 and 700 ms, respectively. Care was taken to ensure that the angle of convergence was appropriate under these viewing conditions by use of dichoptic vernier fixation lines (one for each eye) that was made visible prior to the stimulus proper. The observers had to signify that the vernier lines were aligned by pressing the middle button of a computer mouse. The left and right button signified the presence of a 'concave' and 'convex' corrugated-RDS, respectively.

Sensitivity measurements and data analysis

On a daily basis, stimulus levels were presented in a randomly intermixed sequence to form a single psychometric session. That is, sensitivity was determined separately for each of the stimulus conditions using blocks

of nine possible peak-to-trough disparity amplitudes corresponding to $-4, -3, -2, -1, 0, 1, 2, 3$ or 4 times the step-size. A successful daily session consisted of six blocks containing a total of 6×9 trials. In pilot trials, the step-size was determined for each of the test-CFs.

Data were collected in a successive viewing of a balanced quasi-random sequence of surfaces that could be typified as either concave, or alternatively, convex. Reliable performance was defined as: $P(X \geq 83\%) < 0.0001$ when $N = 70$ trials, where P represents the two-sided, independent binomial probability calculated from the number of correct and incorrect responses with an expectation of 0.5 of being correct by chance alone.

Psychometric functions were constructed, and inspected visually (compare Figure 3). The proportion of trials in which the subject indicated the presence of a convex central corrugation was calculated for each stimulus level. Likelihood maximization (based on 10,000 Monte Carlo simulations) (Wichmann & Hill, 2001a) was used for parameter estimation of an expected 2AFC performance function that combined the data from ten successive, successful daily-sessions in which their individual slopes

deviated less than 20% from the mean slope. The maximized performance function for a given test-CF (comprising $10 \times 6 \times 9$ trials in total) was a cumulative Gaussian probability curve defined as:

$$P(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} e \frac{-(x-\mu)^2}{2\sigma^2} dx. \qquad (2)$$

Here, x is disparity, μ and σ are the mean position and the standard deviation (SD), respectively. In particular, the function $P(x)$ corresponds to the probability of indicating a convex shaped central corrugation. The μ parameter represents the bias towards either negative ($\mu > 0$) or positive ($\mu < 0$) disparities. The stereo acuity parameter, σ, represents a direct measure of the observer's ability to perform the discrimination task. The log-likelihood ratio, based on 10,000 Monte-Carlo simulations, allowed verification of the goodness-of-fit: two-sided $\chi^2_{deviance}$ (7) > 12, $p < 0.05$ (Wichmann & Hill, 2001b). In other words, the likelihood of finding a deviance greater than 12 (with 7 degrees of freedom) by chance alone for all fitted functions was less than 5%.

The 95% Confidence intervals (95%-CI) reported throughout this study were estimated using Efron's non-parametric, bias-corrected and accelerated bootstrapping algorithm (Efron, 1987). Spatial frequency bandwidth estimations (i.e., full-width at half-height) of the DFSs were obtained by use of a shape preserving interpolation algorithm (Wolberg & Alfy, 2002). Note that goodness-of-fit statistics are not defined for interpolants. The fit residuals are always zero (within computer precision) because interpolants pass through the data points.
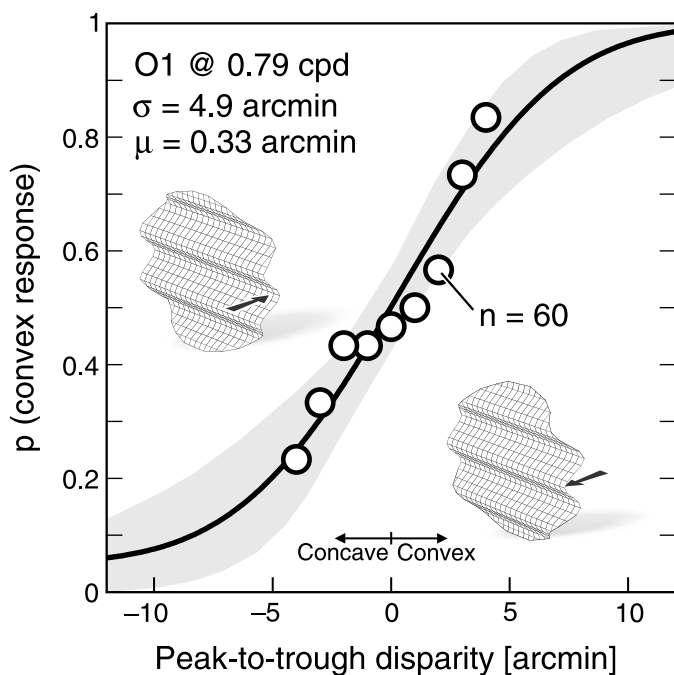

Figure 3. Psychometric function for the forced-choice concave/convex judgments of owl subject 1. The example shown here was recorded on ten days with a horizontal grating at 0.79 cpd corrugation frequency. Data points denote the proportion of trials in which the subject indicated the presence of a convex central corrugation. The gray area represents the surface covered by 10,000 cumulative Gaussian fit functions from which the psychometric function (black) was derived by likelihood maximization. Stereo acuity threshold was defined as the standard deviation, σ, of that function. A measure of response bias towards either convex or concave corrugations is the x-value at which the function inflects, μ.

Results

Human performance

First, we wished to confirm the basic findings on the stereoscopic bandwidth limitation in human stereopsis as reported by Bradshaw and Rogers (1999): (i) the DSF is inverted U-shaped, implying that sensitivity is highest at some intermediate spatial frequency, (ii) thresholds for detecting vertically oriented RDS corrugations are higher than those for detecting horizontally oriented corrugations. This orientation anisotropy is especially evident with low corrugation frequencies.

The DSFs derived from the forced-choice concave/convex judgments for two non-experienced human observers, H1 and H2, are shown in Figure 4, top row. The data represent stereoscopic sensitivity (i.e., the inverse of σ; Equation 2) as a function of spatial frequency, and are plotted for horizontally (black lines) and vertically (gray lines) oriented corrugations. Both human subjects

showed small response biases (<30 arcsec), μ (see Equation 2), with 95%-CIs overlapping with zero disparity, indicating a symmetry in the detection of convex and concave corrugations.

It is immediately apparent from Figure 4 that the shape of the DSFs from horizontal and vertical corrugations is very similar: sensitivity decreases for both high and low spatial frequencies, revealing inverted U-shaped envelopes. The estimated spatial frequency bandwidth (mean, full-width at half-height; N = 2) approximated 3.6 and 3.7 octaves for horizontal and vertical oriented corrugations, respectively, as averaged over the data of the two

subjects (horizontal, vertical; H1: 3.759, 4.027; H2: 3.379, 3.311). Maximum sensitivity from horizontal corrugations was found in the range of 6–9 arcsec at optimal frequencies around 0.56 cpd. Maximum sensitivity from vertical corrugations was found in the range of 15–25 arcsec at optimal frequencies around 0.38 cpd.

Expressed another way, at optimal frequencies and a viewing distance of 110 cm a depth difference of 0.6–0.8 mm between the peaks and troughs could be detected from horizontally oriented corrugations, whereas a depth difference of 1.4–2.2 mm could be detected from vertically oriented corrugations. This stereoscopic aniso-
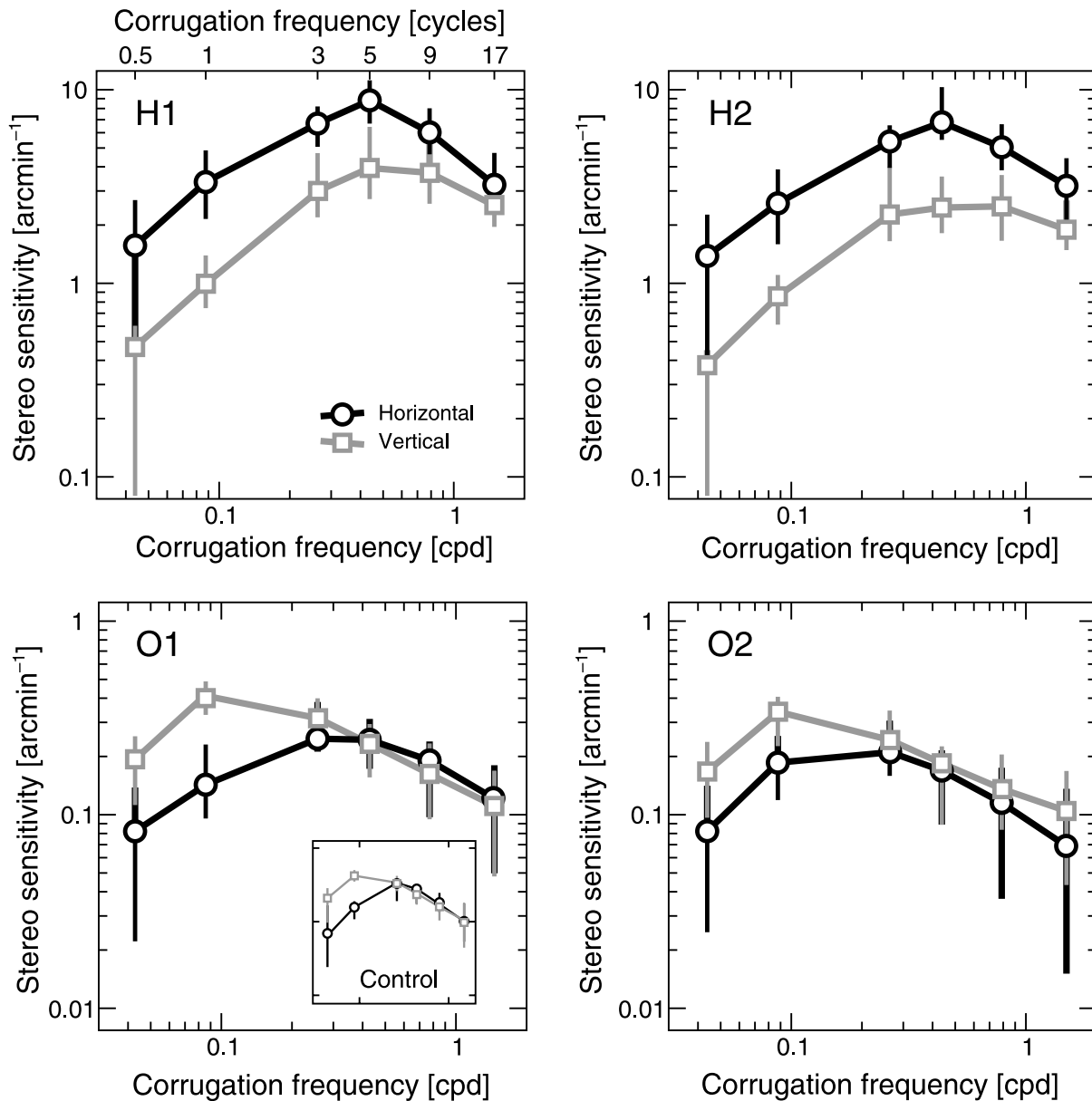

Figure 4. DSFs of two human observers, H1 and H2 (top row), and two barn owls, O1 and O2 (bottom row). Each data point represents 10 × 6 × 9 trials and is derived from the inverse standard deviation parameter, σ, of the estimated cumulative Gaussian performance function (Equation 2), which is a measure of stereo sensitivity. Each curve represents 3240 trials in total. The different markers and line contrast represent data for horizontal (black circles) and vertical oriented corrugations (gray squares). Vertical bars denote bootstrapped 95%-CIs. The inset in the lower left panel shows results from the second control experiment (see section Control experiments for details).

tropy was maximal for the lowest frequency, 0.04 cpd and non-existent for the highest frequency, 1.5 cpd. Moreover, at low frequencies the fall-off in sensitivity with vertical corrugations is more rapid to the fall off found with horizontal corrugations.

In summary, our results confirm those previously reported by Banks et al. (2004) and Bradshaw and Rogers (1999) when taking into account a dot-density of approximately 10 dots/deg^2. We therefore were confident that the present procedure and viewing conditions provide sufficient fidelity and resolution to measure the DSF in owls.

Owl performance

The DSFs derived from the forced-choice concave/convex judgments for the two owls, O1 and O2, are shown in Figure 4, bottom row. The data represent stereoscopic sensitivity (i.e., the inverse of σ; Equation 2) as a function of spatial frequency, and are plotted for horizontally (black lines) and vertically (gray lines) oriented corrugations. Both owls showed small response biases (<50 arcsec), μ (see Equation 2), with 95%-CIs overlapping with zero disparity, indicating a symmetry in the detection of convex and concave corrugations.

It is immediately apparent from Figure 4 that, also for the owl data, the shape of the DSFs from horizontal and vertical corrugations is very similar: sensitivity decreases for both high and low spatial frequencies, revealing inverted U-shaped envelopes. The estimated spatial frequency bandwidth (mean, full-width at half-height; N = 2) approximated 3.8 and 3.3 octaves for horizontal and vertical oriented corrugations, respectively (horizontal, vertical; O1: 3.955, 3.311; O2: 3.569, 3.228). Maximum sensitivity from horizontal corrugations was found in the range of 3–5 arcmin at optimal frequencies around 0.2 cpd. Maximum sensitivity from vertical corrugations was found in the range of 2–3 arcmin at optimal frequencies around 0.1 cpd.

Expressed another way, at optimal frequencies and a viewing distance of 110 cm, a depth difference of 3–4 cm between the peaks and troughs could be detected from horizontally oriented corrugations, whereas a depth difference of 2–3 cm could be detected from vertically oriented corrugations. This reverse stereoscopic anisotropy (when compared to that observed in the two human subjects) was maximal for the lowest frequencies measured, and non-existent for the frequencies higher than 0.3. Moreover, at low frequencies the fall-off in sensitivity with vertical corrugations is equal to the fall off found with horizontal corrugations.

Control experiments

To ensure contamination-free stimuli and stereopsis driven performance, we worked out two controls. First, to test if our RDS contained monocular cues that could be exploited by the owl to "correctly" distinguish between positive and negative disparities the owls were tested under monocular viewing conditions. That is for each of the tested stimulus conditions they received 30 (suprathreshold: 20 arcmin) trials wherein monocularity was created by placing filters of the same polarization in front of their eyes. The application of this procedure permits only one half-image to stimulate both eyes. Thus, when discriminative behavior is purely based on stereopsis then performance should only be affected by monocular viewing in the sense that performance should be near chance performance (chance of 1 in 2 to be correct). Both owls showed performance at chance level under these viewing conditions. Out of 360 trials, owl subject 1 responded in 191 trials correctly ($p = 0.469$, 95%-CI: 0.417–0.522). Owl subject 2 responded in 173 trials correctly ($p = 0.519$, 95%-CI: 0.467–0.572).

Second, in principle, detection of nonzero disparity of only a few central dots could have been enough to perform our experimental task: reporting whether the central corrugation was concave, or alternatively, convex. This may result in an overestimation of the spatial precision with which observers can discriminate corrugated RDSs. This confound can be tested by use of an orientation identification task (Banks et al., 2004). In case of an identification task where the stimulus as a whole must be interpreted, the observer has to report whether a given corrugation is oriented horizontally, or alternatively, vertically. We therefore trained one owl, O1, to perform this orientation identification task, the data of which are shown in Figure 4, inset. Despite striking differences in the psychophysical tasks, the obtained DSFs do not differ significantly. Note that for this orientation identification task, the phase of a given corrugation was chosen randomly using 7 possible values, ranging from 40/360 to 320/360 cycles separated by steps of 40/360 cycles.

Discussion

The disparity sensitivity function derived from corrugated random-dot stereograms is an assessment of a visual system's ability to form useful impressions of 3D shape over a wide range of corrugation frequencies solely based on modulations in binocular disparity. It is therefore the most direct analytical tool to assess the perception of shape-from-stereo. So far, the DSF has been measured in humans alone. By determining the DSF in both humans and owls, using identical corrugated-RDSs and psychophysical procedures, a direct comparative account about shape-from-stereo is made possible. Here we put forward four novel findings, emphasizing the striking similarity in the aptitude of humans and owls to achieve shape-from-stereo. As such, our set of across-species DSFs discloses

common principles of shape-from-stereo, despite large differences of the visual systems that exist between birds and primates.

First, owls are capable of experiencing shape-from-stereo in much the same way as humans do (Figure 4). In particular, we discovered that the visual systems of both owls and humans are more effective in processing intermediate spatial-frequencies and are less effective at both high and low spatial-frequencies relative to their eyes' resolving power (Harmening, Nikolay, Orlowski, & Wagner, 2009). This pattern of results establishes an equivalent, inverted U-shaped DSF for the processing of shape-from-stereo in these two species.

Second, both species exhibit an orientation-dependent processing of disparity modulation, known as stereoscopic anisotropy (Rogers & Graham, 1983). Most notably, the owl's anisotropy was maximal at low spatial frequencies, but opposite in sign as was the case for our human subjects. That is, our owls were more sensitive to vertical than to horizontal oriented corrugations (Figure 4). This reciprocity in anisotropy between owls and humans cannot be attributed to limitations inherent to the stimulus, because the corrugated-RDSs used here were identical. Although anisotropy in the stereograms of corrugated-RDS is evident and frequently reported in the literature, it is not fully understood, because large individual differences in its extent are commonly observed (for review see (Howard & Rogers, 2002)). It is even possible to reverse the human anisotropy (Serrano-Pedraza & Read, 2009), which makes it equivalent to the anisotropy observed here in our owls. Despite these controversies, the origin of stereoscopic anisotropy may be explained by work emanating from the theoretical analysis of natural images. In recent years, it has been recognized that the number of possible 3D shape interpretations of natural images are highly constrained such that they all are related by a limited class of "generalized bas-relief transformations" (Belhumeur, Kriegman, & Yuille, 1999), or by the statistical properties of optical deformations such as smooth occlusion contours and steep disparity gradients (Geisler & Perry, 2009; Huang & Lee, 2000). Thus, although the perceived 3D shape of a natural image is inherently ambiguous (see also Yang and Purves, 2003) the brain could capitalize on geometric regularities of natural images to single out, or promote, a certain disparity gradient that is behaviorally most relevant. For instance, the need of the owl to strike its preferred pray, mice, in the direction of their movement with its talons placed along the vertical axis of the mouse's body (Martin, 1990; Payne, 1971), may have led to a biased, or heightened, sensitivity to vertical changes in 3D shape.

Third, a prominent difference between shape-from-stereo of humans and owls is the shift in frequency at which sensitivity is maximal, as is made explicit in Figure 5. In comparison to the human DSF, the spatial frequency at which the owl DSF peaks is shifted towards the low-frequency end by approximately one log unit. Notice that
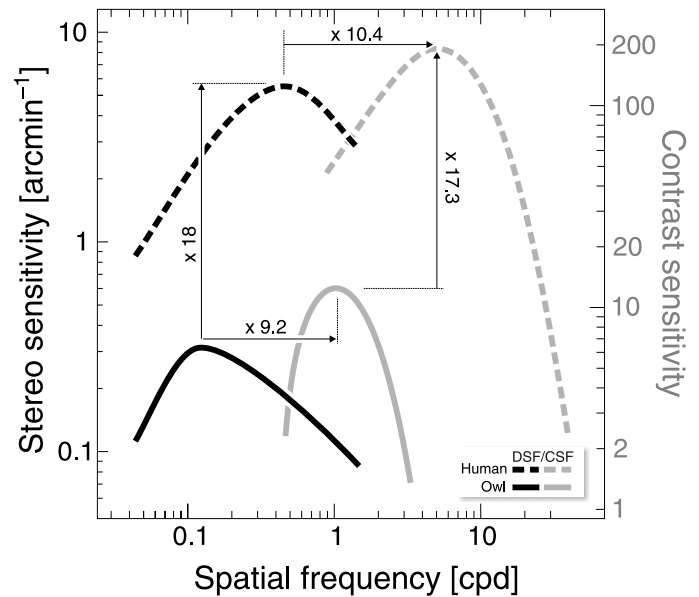

Figure 5. DSF (dark) versus CSF (light) in both humans (dashed) and owls (solid). This straightforward graphical comparison reveals symmetrical relations between the position of peaks of the two functions across species and modalities. DSF data are mean values found in the present study. The human CSF is taken from Campbell et al. (1966) and that of the owl from Harmening et al. (2009).

because of this leftward shift in maximum sensitivity, the high-frequency branch of the owl DSF appears to decline slower than is the case for the human DSF, giving it an overall asymmetrical shape. Unquestionably, this maximum sensitivity shift towards lower frequencies is to be expected when taking into account the differences in the resolving power that exist between the owl and human visual system. In particular, a one-log unit shift can readily be observed in other exclusively neural-driven measurements. These are measures of retinal cell density and sampling capabilities (Ghim & Hodos, 2006; Wathey & Pettigrew, 1989), hyperacuity thresholds, viz. stereo acuity and vernier acuity (Harmening, Göbbels, & Wagner, 2007; van der Willigen et al., 2002), and peak contrast sensitivity (Harmening et al., 2009). In addition, the high optical quality of the owl eye relative to that of the human eye (Harmening, Vobig, Walter, & Wagner, 2007; Schaeffel & Wagner, 1996) may account for the shallower high-frequency roll-off in the barn owl DSF. Because of this remarkable congruency in the data it is most plausible to think that the limitation of shape-from-stereo is primarily of neural rather than optical origin.

Fourth, the second prominent difference between shape-from-stereo of humans and owls is the decrease in their maximum sensitivity, as is made explicit in Figure 5. Sensitivity values of the owls remained below human values across the entire measured spectrum. In other words, less stereoscopic detail was visible for the two barn owls than for the human subjects. A significant reduction

in maximum sensitivity of the owls relative to humans is to be expected, first and foremost from a mere geometrical point of view. The barn owl's inter-ocular distance is on the order of 4 cm, and thus 0.6 times that of man. Moreover, behavioral thresholds for grating acuity are only one tenth of what is reported in humans. Assuming a close-to-linear relationship between hyperacuity and normal visual acuity thresholds, we would have expected that the owl's highest stereo sensitivity values were on the order of $1/(0.6 \times 0.1) = 16.7$ fold lower than that found for our human subjects. This value is close to the decrease in maximal stereoscopic sensitivity reported here. This quantitative accordance may indicate that the weaker ability of the owl to detect shape-from-stereo may have its origins in the monocular processing of the corrugated-RDSs used here. It is most parsimonious to presume that the luminance-defined blobs, that defined our corrugations, have to be detected before binocular disparity can be processed (see also Regan, 2000). Interpreted in this way, the manner in which we structured our corrugated-RDSs affects the process of shape-from-stereo in the sense that it not only requires the detection of disparities modulations but the process of shape-from-luminance as well. A straightforward solution to test this idea, would be to compare the here obtained DSFs with the known CFSs for both humans and owls.

Comparing the sensitivity functions in the disparity and luminance domain (i.e. DSF vs. CSF) of human and owl subjects reveals a quantitative symmetry between those, both across species and across modalities (compare Figure 5) (Campbell, Kulikowski, & Levinson, 1966; Harmening et al., 2009). Relative to the peak of the CSF, the DSF peak is shifted towards lower frequencies in humans and owls. The magnitude of shift is in good accordance in both species, with a factor of 9.2 and 10.4 in owl and human subjects, respectively. Moreover, while maximum stereo sensitivity in the owl is about 18 times lower than that of man, contrast sensitivity is reduced by virtually the same amount. These quantitative consistencies may turn the CSF into a valid predictor for the limits of stereovision in species with similar stereoscopical capabilities (say, species with global stereopsis), at least in the constraints set by the choice of stimuli and methods used here. Based on our results, the following rules of thumb may be derived:

- DSF peak sensitivity is shifted towards lower spatial frequencies by a factor of 10 compared to CSF peak sensitivity.
- Maximum stereoscopic sensitivity is reduced by the same amount relative to the reduction of maximum contrast sensitivity compared with human values.

Because values for maximum contrast sensitivity and maximum disparity sensitivity are well-confirmed in humans (Bradshaw & Rogers, 1999; Campbell et al., 1966), maximum stereoscopic sensitivity may be easily

calculated if the CSF is known. Naturally, these rules remain hypothetical until more DSF/CSF pairs have been identified in further species.

In summary, our results demonstrate both qualitative and quantitative similarities of disparity sensitivity in the visual systems of humans and owls. Owing to the considerable differences in their phylogeny and visual pathways (Cook, 2001; Medina & Reiner, 2000; Zeigler & Bischof, 1993), these findings may reflect fundamental aspects of disparity-based stereo vision in general. These are: (1) the DSF has bandpass, anisotropic characteristics across species, and (2) the DSF is closely related to the corresponding sensitivity function in the luminance domain, the CSF, in the sense that maximum stereo sensitivity and absolute position of the DSF may be readily inferred if the CSF is known.

Acknowledgments

Footnote

Robert F. van der Willegen and Wolf M. Harmening contributed equally to the study.

References

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience, 24,* 2077–2089. [PubMed] [Article]

Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision, 35,* 33–44. [Article]

Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research, 39,* 3049–3056. [PubMed]

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science, 20,* 615–617. [PubMed]

Campbell, F. W., Kulikowski, J. J., & Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *The Journal of Physiology, 187,* 427–436. [PubMed] [Article]

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197,* 551–566. [PubMed] [Article]

Cook, R. G. (2001). Avian visual cognition. [On-line]: www.pigeon.psy.tufts.edu/avc/. [Article]

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of American Statistics Association, 82,* 171–185. [Article]

Filippini, H. R., & Banks, M. S. (2009). Limits of stereopsis explained by local cross-correlation. *Journal of Vision, 9*(1):8, 1–18, http://journalofvision.org/9/1/8/, doi:10.1167/9.1.8. [PubMed] [Article]

Geisler, W. S., & Perry, J. S. (2009). Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience, 26,* 109–121. [PubMed] [Article]

Ghim, M., & Hodos, W. (2006). Spatial contrast sensitivity of birds. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 192,* 523–534. [PubMed]

Harmening, W. M., Göbbels, K., & Wagner, H. (2007). Vernier acuity in barn owls. *Vision Research, 47,* 1020–1026. [PubMed]

Harmening, W. M., Nikolay, P., Orlowski, J., & Wagner, H. (2009). Spatial contrast sensitivity and grating acuity of barn owls. *Journal of Vision, 9*(7):13, 1–12, http://journalofvision.org/9/7/13/, doi:10.1167/9.7.13. [PubMed] [Article]

Harmening, W. M., Vobig, M. A., Walter, P., & Wagner, H. (2007). Ocular aberrations in barn owl eyes. *Vision Research, 47,* 2934–2942. [PubMed]

Hibbard, P. B., Bradshaw, M. F., Langley, K., & Rogers, B. J. (2002). The stereoscopic anisotropy: Individual differences and underlying mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 469–476. [PubMed]

Howard, I., & Rogers, B. (2002). *Seeing in depth: Depth perception.* Toronto, Canada: Porteous.

Huang, J., & Lee, A. (2000). Statistics of range images. *Proceedings of the IEEE Computer Vision and Pattern Recognition, 1,* 324–331.

Martin, G. (1990). *Birds by night.* London, UK: A & C Black Publishers Ltd.

Medina, L., & Reiner, A. (2000). Do birds possess homologues of mammalian primary visual, somatosensory and motor cortices? *Trends in Neuroscience, 23,* 1–12. [PubMed]

Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature, 452,* 642–645. [PubMed] [Article]

Payne, R. S. (1971). Acoustic location of prey by barn owls (*Tyto alba*). *Journal of Experimental Biology, 54,* 535–573. [PubMed] [Article]

Regan, D. (2000). *Human perception of objects.* MA: Sinauer Associates Inc.

Rogers, B., & Graham, M., (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22,* 261–270. [PubMed]

Rogers, B. J., & Graham, M. E., (1983). Anisotropies in the perception of three-dimensional surfaces. *Science, 221,* 1409–1411. [PubMed]

Schaeffel, F., & Wagner, H. (1996). Emmetropization and optical development of the eye of the barn owl (*Tyto alba*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 178,* 491–498.

Serrano-Pedraza, I., & Read, J. C. A. (2009). Horizontal/vertical anisotropy in sensitivity to relative disparity depends on stimulus depth structure. In *Perception ECVP Abstract Supplement* (vol. 38. p. 155).

Todd, J. T. (2004). The visual perception of 3d shape. *Trends in Cognitive Science, 8,* 115–121. [PubMed]

Tyler, C. (1991). *Vision and visual dysfunction: Cyclopean vision.* New York: Macmillan.

Tyler, C. W. (1973). Steroscopic vision: Cortical limitations and a disparity scaling effect. *Science, 181,* 276–278. [PubMed]

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature, 251,* 140–142. [PubMed]

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research, 15,* 583–590. [PubMed]

van der Willigen, R. F., Frost, B. J., & Wagner, H. (1998). Stereoscopic depth perception in the owl. *Neuroreport, 9,* 1233–1237. [PubMed]

van der Willigen, R. F., Frost, B. J., & Wagner, H. (2002). Depth generalization from stereo to motion parallax in the owl. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 187,* 997–1007. [PubMed]

van der Willigen, R. F., Frost, B. J., & Wagner, H. (2003). How owls structure visual information. *Animal Cognition, 6,* 39–55. [PubMed]

Wagner, H. (1993). Sound-localization deficits induced by lesions in the barn owl's auditory space map. *Journal of Neuroscience, 13,* 371–386. [PubMed] [Article]

Wathey, J., & Pettigrew, J. (1989). Quantitative analysis of the retinal ganglion cell layer and optic nerve of the barn owl *Tyto alba. Brain Behavior Evolution, 33,* 279–292. [PubMed]

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature, 302,* 419–422. [PubMed]

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit.

Perception & Psychophysics, 63, 1293–1313. [PubMed] [Article]

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63,* 1314–1329. [PubMed] [Article]

Wolberg, G., & Alfy, I. (2002). An energy-minimization framework for monotonic cubic spline interpolation. *Journal of Computational and Applied Mathematics, 143,* 145–188.

Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience, 6,* 632–640. [PubMed]

Zeigler, H. P., & Bischof, H.-J. (Eds.) (1993). *Vision, brain, and behavior in birds.* Cambridge, Massachusetts: MIT Press.