Biol Cybern
DOI 10.1007/s00422-007-0199-4

ORIGINAL PAPER

Biological Cybernetics

Through a barn owl's eyes: interactions between scene content and visual attention

Shay Ohayon · Wolf Harmening · Hermann Wagner · Ehud Rivlin

Received: 31 August 2007 / Accepted: 22 October 2007

Abstract In this study we investigated visual attention properties of freely behaving barn owls, using a miniature wireless camera attached to their heads. The tubular eye structure of barn owls makes them ideal subjects for this research since it limits their eye movements. Video sequences recorded from the owl's point of view capture part of the visual scene as seen by the owl. Automated analysis of video sequences revealed that during an active search task, owls repeatedly and consistently direct their gaze in a way that brings objects of interest to a specific retinal location (retinal fixation area). Using a projective model that captures the geometry between the eye and the camera, we recovered the corresponding location in the recorded images (image fixation area). Recording in various types of environments (aviary, office, outdoors) revealed significant statistical differences of low level image properties at the image fixation area compared to values extracted at random image patches. These differences are in agreement with results obtained in primates in similar studies. To investigate the role of saliency and its contribution to drawing the owl's attention, we used a popular bottom-up computational model. Saliency values at the image fixation area were typically greater than at random patches, yet were only 20% out of the maximal saliency value, suggesting a top-down modulation of gaze control.

S. Ohayon (✉) · E. Rivlin
Israel Institute of Technology (Technion),
Haifa 32000, Israel
e-mail: shay.ohayon@gmail.com

W. Harmening · H. Wagner
Institute of Biology II, RWTH University,
Kopernikusstrasse 16, 52074 Aachen, Germany

1 Introduction

The non-uniform distribution of photoreceptor cells in the primate retina is one of the factors contributing to the alternate fixate-saccade behavior, where the latter is used to redirect the highest acuity retinal area (fovea) toward objects of interest. The early works of Buswell and Yarbus have shown that fixations are not random, but fall on semantically informative parts of the scene (Buswell 1935; Yarbus 1967). Follow up investigations revealed important aspects of the mechanisms that underlie overt and covert visual attention (Treisman and Gelade 1980).

Little is known, however, about similar mechanisms in avian species, partly due to anatomical dissimilarities. The retinal structure of birds exhibits a different spatial organization compared to the primate retina. In barn owls, for example, the ratio of photoreceptor cell density at the area centralis to cell density at the periphery is several folds smaller compared to the ratio in humans (Wathey and Pettigrew 1989). In addition, the distribution of photoreceptor cells is not along circular iso-contour lines, but along an elongated horizontal visual streak. Other raptors, such as falcons, have multiple distinct areas that are specialized for acute vision (Tucker 2000).

These anatomical differences pose interesting questions with respect to the nature of fixation and its role in visual attention. In this paper we first address the problem of fixation consistency in barn owls. That is, investigate whether the object of interest repeatedly falls on a specific retinal point, or whether owls attend to objects while they appear anywhere in their visual field. We then address the problem of what draws attention in barn owls by exploring the statistical nature of image properties that are fixated, and compare our findings to similar studies done in humans (Henderson et al.

2007; Parkhurst and Niebur 2003; Reinagel and Zador 1999; Krieger et al. 2000).

Primates are able to control eye position by using a set of eye muscles, causing the eye to rotate in its socket. In contrast, barn owls possess a tubular eye structure that limits eye movements, and nearly immobilizes their eyes (Steinbach and Money 1973; Du Lac and Knudsen 1990). To compensate for the limited eye movement owls have a flexible neck that permits head rotations to extreme angles (Masino and Knudsen 1990). These rotations are extremely fast (up to 700°/s), ballistic in nature and resemble the saccadic eye movements of primates (Payne 1971; Masino and Knudsen 1993; Knudsen and Konishi 1979; Ohayon et al. 2006; Du Lac and Knudsen 1990; Knudsen et al. 1979) Likewise, fixation in barn owls refers to a time interval of stable head position.

In our experiments, we have used a miniature wireless camera attached to the owl's head to derive quantitative image measurements. Betsch used a similar approach with cats and obtained interesting statistical properties of natural images, as seen from the cat's perspective (Betsch et al. 2004). However, only global analysis of images was possible since cats' ability to move their eyes is much greater compared to owls (Vanni-Mercier et al. 1994). The nearly fixed position of owls' eyes, in contrast, permits a finer, localized analysis. In Sect. 2.1.4 we introduce a projective model that explains how a retinal point can be mapped back to the captured images, enabling the measure of image statistics at specific retinal areas.

A heated debate regarding the role of low level image features in attracting fixation is currently taking place. On the one hand, there is evidence that specific image features, such as luminance contrast or edges, attract fixation and are statistically different at fixation positions (Reinagel and Zador 1999; Mannan et al. 1996; Parkhurst et al. 2002; Baddeley and Tatler 2006). On the other hand, models that are based only on low level image features fail to predict fixation positions during an active search task (Henderson et al. 2007; Navalpakkam and Itti 2005; Itti 2005), suggesting that the process is governed by high level, top down mechanisms (Henderson and Hollingworth 1999; Hayhoe and Ballard 2005). To evaluate how these concepts extend to avian species, we give a detailed description on correlations between scene content and image features at fixation positions in freely behaving barn owls.

One of the most influential work on modeling the shifts in visual attention was proposed by Koch and Ullman (Koch and Ullman 1985). They proposed a model that combines several features onto a single map, from which the next fixation position is selected. Several variations of this model have appeared since (Rao et al. 2002; Zaharescu et al. 2004; Itti et al. 1998), and have claimed to predict fixation positions of humans under free viewing conditions (Zaharescu et al. 2004; Parkhurst et al. 2002). The common ground for most of these biologically plausible models is the saliency hypothesis, stating that fixation is determined according to conspicuous regions in the visual field. Saliency is typically computed across several image scales by applying a local difference operator to various feature maps, such as luminance, color, edge density, orientation, depth, motion, etc. The next fixation is selected in a winner-take-all strategy by selecting the position with the highest saliency.

Can a simplified saliency model account for a barn owl's head movements and correctly predict fixation positions? From the neuronal point of view, there is enough evidence to suggest the notion of feature extraction. Evidence for orientation selectivity (Pettigrew and Konishi 1976; Liu and Pettigrew 2003), binocular disparity (Nieder 1999), as well as neurons responding to subjective contours (Nieder and Wagner 1999), all suggest that owls have a highly developed visual system capable of feature extraction and integration. Integration also occurs across different modalities, such as the visual and auditory pathways, supporting the notion that attention plays a major role in the modulation of sensory input (Johnen et al. 2001; Whitchurch and Takahashi 2006). The last topic we address in this paper, therefore, is the extent to which owl fixation correlates with a popular bottom-up saliency model (Itti et al. 1998).

2 Fixation area in barn owls

Our experiments were designed to determine whether owls consistently fixate the object of interest on a specific retinal point. The basic paradigm used was an active search task in which owls had to find a food item among a set of distracters that shared a similar appearance to the food item.

2.1 Materials and methods

2.1.1 Experimental setup and protocol

Two adult barn owls (*Tyto alba pratinocla*), YA and SC, were tamed by hand rearing and could be easily handled. Owls weight was maintained at approximately 85% of their ad-libitum weight. They received the necessary food, one dead chick in the course of each experiment. No positive reinforcement was given. Both owls carried a small metal plate that had been fixed to their skull under anaesthesia (Wagner 1993). All experimental procedures were approved by the Regierungspräsidium Köln.

A miniature wireless camera (subsequently abbreviated as owl-cam) was attached to the metal plate on the owl's head (Fig. 1). Video from the owl-cam was acquired using a 1/4" CMOS type sensor with a resolution of 380 vertical lines. The sensor could operate at light levels above 2 LUX and

Fig. 1 A barn owl with the miniature wireless camera attached to its head

contained automatic gain control. A 900 MHZ radio frequency transmitter was connected to the image sensor and transmitted images in NTSC standard (480 × 720) to a nearby receiver. Both camera and transmitter were powered using a single 3 V lithium battery (CR2 type). The battery output was connected to a push-up voltage circuit (MAXIM 756 chip), producing stabilized 5 V. Power consumption of both camera and transmitter was measured at 80 mA, resulting in a continuous operation of approximately 10 h. The owl-cam total weight was 28 g. We used a standard 10 mm diameter lens, covering a field of view of 42.84° in the horizontal axis and 27.12° in the vertical axis. The camera's focal length was estimated using a method described in (Zhang 2000), and was found to be approximately 800 pixels.

Owls are known for their frontal eye position, which provides a large binocular overlap (Martin 1982). However, exact visual coverage of both eyes in barn owls is unknown (van der Willigen 2000). We assume that the lens angle of view intersects with the owl central visual field. We aligned the camera on the owl's head iteratively in a series of pre-experiments until food items caught by the owl appeared in the captured images. Note that the center of these images does not necessarily correspond to the owl's center of visual field.

Video sequences were recorded in the birds' aviary. The aviary measured 165 cm in width, 425 cm in length and 250 cm in height. Owls were free to move and fly in this environment. Prior to an experiment, a dead chick (target) and several dummy distracters (3–6) were randomly positioned across the floor. Distracters were either a small yellow paper, cut in the form of a chick, or a small rectangular shape with yellow feathers on top. Each recording session started by inserting the owl into the aviary. The owl typically flew towards one of the perches and observed its environment by scanning the targets and distracters that were randomly positioned on the ground. At times, it flew from one perch to the other to observe the items from a different angle. Experiments ended after approximately 20 min elapsed, or if the owl flew toward the target or one of the distracters.

Table 1 Statistics of aviary video sequences

Data statistic	Owl YA	Owl SC
Number of video sequences	19	21
Average movie length (min)	35.32	21.89
Total recording time (min)	671	459

2.1.2 Data collection

Table 1 summarizes the number of video sequences recorded, as well as their average duration. Each video sequence represents an experiment in which the owl was introduced to the aviary. Three typical images acquired in the aviary, showing the room from the owl's point of view are depicted in Fig. 2a–c. Dummy targets appear as yellow blobs. Notice that in Fig. 2a and b, two different yellow targets appear in the same image position (highlighted by a red circle).

2.1.3 Scene-cut detection

A common method to obtain fixation positions in behavioral experiments is to use an eye tracker. Typically, a constrained subject sits in front of a computer display, while a high resolution IR camera tracks the pupil and determines the fixation point on the computer display. Eye movement raw data are segmented to intervals of fixations and saccades, according to eye velocity (Itti 2005).

Our method of using a head mounted camera, on the other hand, outputs images as raw data and requires a different method to segment behavior into fixation and saccade. A typical video sequence obtained from the owl-cam is characterized by intervals of stabilized images and intervals of blurry images. Conceptually, "image speed" corresponds to head velocity. Any frame of such video may be categorized into two classes—fixation (stabilized image) or saccade (large observed image motion).[1] Adjacent fixation frames may be concatenated into a fixation interval, which we refer to as a scene. Scenes are separated by saccade intervals and are represented by a single image (hereafter—the fixation image).

[1] In principle, two adjacent frames are needed to determine the class, since motion cannot be extracted from a single image frame.

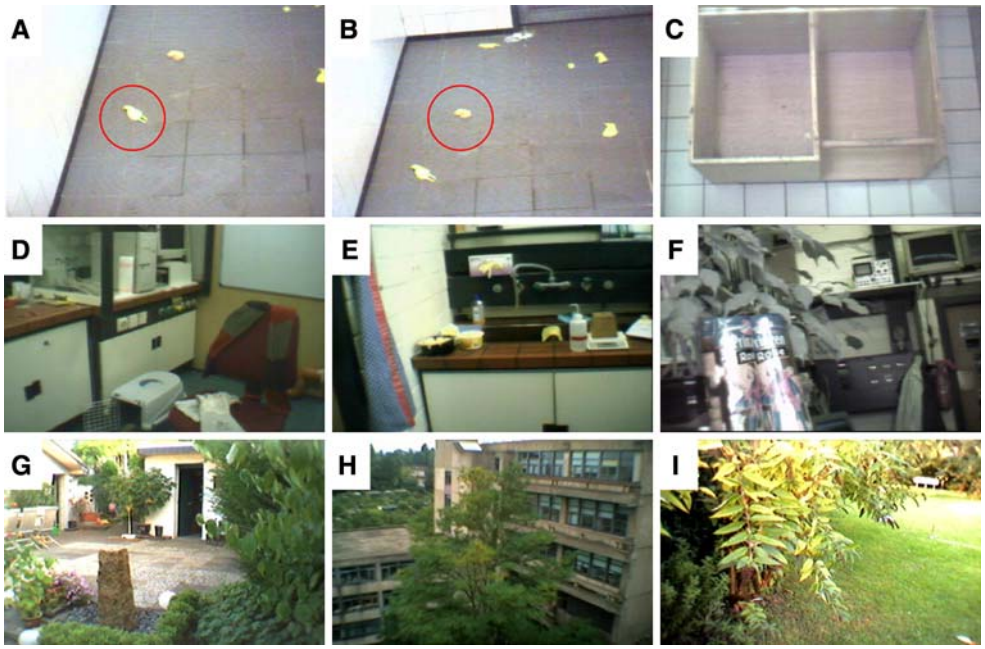

Fig. 2 Video snapshots acquired using the owl-cam. The aviary sequences, shown in the first tow contained yellow targets with a similar appearance to the true food item. Owl's fixation spot is marked with a red circle in subplot **a** and **b**. **c** snapshot during a flight towards the perch. **d, e, f** Snapshots from our office environment. **g, h, i** Snapshots from the outdoors environment

The fixation image of a fixation interval is defined as the center frame of the interval.

Automatic segmentation of a video sequence into scenes is named scene-cut detection and is a known problem in video archiving.[2] Most scene-cut detection algorithms are based on adjacent frame comparison. Typically, features are extracted from each image and are compared to yield a similarity measure. Low similarity values represent a scene-cut (Dailianas et al. 1995).

For our analysis, we have developed a novel scene-cut algorithm based on a local edge density correlation measure. The algorithm segments our raw video data into fixation and saccade intervals and can be summarized to the following operations:

1. Compute the edge map of frame K and frame $K-1$ (Fig. 3b).
2. Divide the frames into non-overlapping blocks of width B (Fig. 3b).
3. Compute the frame signature, which is the mean edge response in each block. The result is a vector of length $\frac{MxN}{B^2}$, where image size is MxN (Fig. 3c).
4. Compute the correlation between edge response of frame K and frame $K-1$ (Fig. 3d).
5. If correlation is above a fixed discrimination threshold f, label frame K as fixation, otherwise—label it as saccade.

Free parameters, such as the edge detection method or block size, may be fine-tuned according to the data set. In our implementation, we selected a block size of 60×60 pixels, yielding a signature of length 120. We used Laplacian of Gaussian as the edge detection method, mainly due to its fast Matlab implementation (Marr and Hildreth 1980).

To quantify the performance of this algorithm and compare it to competitive algorithms, a receiver operating characteristic (ROC) curve may be constructed. First, a test sequence is prepared and each of its frames is labeled either as a fixation or as a saccade. Then, a fixed discrimination threshold f is selected. The performance can then be measured in terms of sensitivity and specificity. Furthermore, in certain cases, one can find an optimal discrimination threshold by computing the correlation between algorithm output and ground truth manual classification (Baldi and Brunak 2001). The correlation function is defined as:

$$\operatorname{argmax} \frac{\langle TP, TN\rangle - \langle FP, FN\rangle}{\sqrt{\langle TP+FN\rangle \langle TP+FP\rangle \langle TN+FP\rangle \langle TN+FN\rangle}}, \tag{1}$$

where TP, TN, FP, FN correspond to true positive, true negative, false positive and false negative, accordingly.

[2] Scene-cut detection should not be confused with scene change detection. The latter is usually referred to the detection of a change within a single scene.

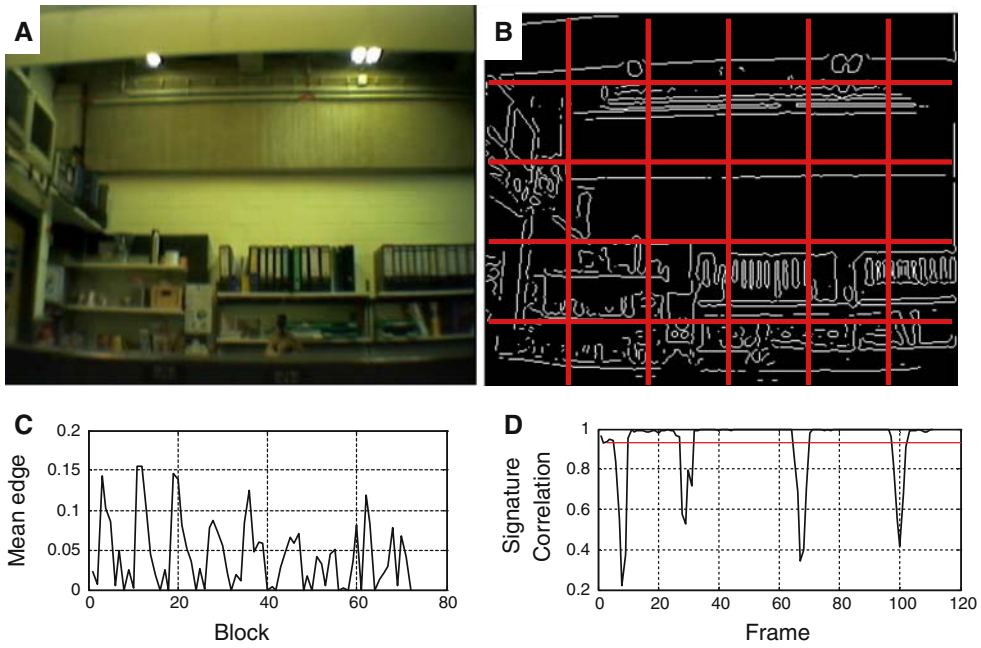

Fig. 3 Scene-cut algorithm. **a** Original image frame. **b** Edges and segmentation into image blocks. **c** Frame signature (mean edge response). **d** Correlation between adjacent frame signatures

2.1.4 Camera-retina model

The methods described so far can only evaluate the temporal characteristics of fixations. Spatial properties, such as fixation positions, require a finer analysis of the process that generates the given images. Consider a simplified projection model which correlates an owl's retina to a camera's focal plane (Fig. 4). Any given point r on the owl's retinal surface forms a ray $\tilde{l}$ when passing through the eye optical center $O = (O_X, O_Y, O_Z)$. This ray is projected onto the camera focal plane as line $\tilde{k}$. In computer vision literature, a line that relates two projective devices is commonly known as an epi-polar line (Hartley and Zisserman 2000). The intersection of all epi-polar lines is at the epi-pole e, the corresponding point of O. Points along line $\tilde{l}$ are represented as $P(\alpha) = O + \alpha d$, where d is a normalized direction vector and α is a distance measure from O. The projection of these points onto the image plane is obtained via a perspective transformation:

$$p(\alpha) = \left[f \frac{O_X + \alpha d_x}{O_Z + \alpha d_z}, f \frac{O_Y + \alpha d_y}{O_Z + \alpha d_z} \right], \tag{2}$$

where f denotes camera focal length.

This simplified model implies that if objects of interest are repeatedly being fixated by the owl on the same retinal position, they will be imaged along the same epi-polar line. Furthermore, if objects of interest are far enough from the owl ($\alpha > \alpha_{\text{Near}}$), they will be mapped to a very short line segment. As distance increases, these points will eventually converge to the point at infinity, defined by the following:

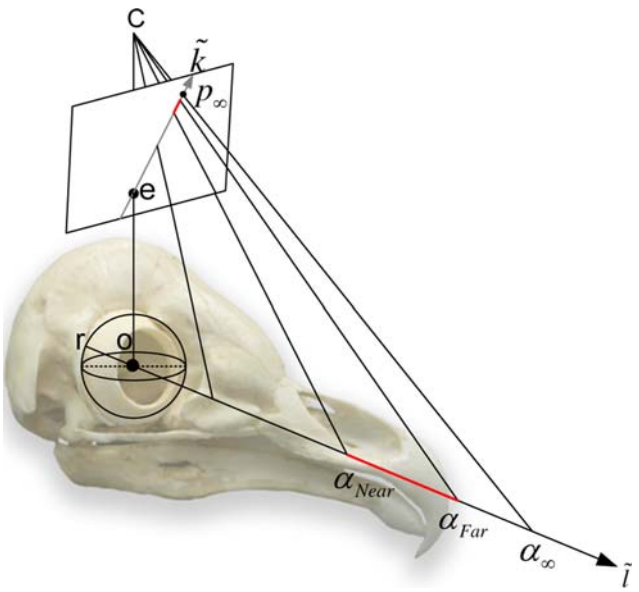

Fig. 4 Projection model. A retinal point r passes through the eye optical center O to form a ray. This ray is projected onto the camera focal plane to form the epi-polar line $\tilde{k}$. As the distance of an object increases (larger α values), its projection will converge to p_∞

$$\lim_{\alpha \to \infty} p(\alpha) = \left[f \frac{d_x}{d_z}, f \frac{d_y}{d_z} \right]. \tag{3}$$

Although the transformation parameters between the camera and the retina are unknown, we can still approximate the length of the projected line segment using Eq. 2. A reasonable assumption is that the camera lies approximately 5 cm above the owl's retina. To simplify the computations, let $O = (0, 5, 0_{\text{cm}})$. Furthermore, in our setup, the range of possible distances an object may appear is between approximately 200 and 400 cm. When these numbers are applied, along with the focal length of the owl-cam, we obtained an upper bound of 10 pixels on the maximal projected line segment.

To conclude, the projective model provides a mathematical description how a single point on the retina may be mapped, as a function of distance, to the camera's image plane. We do not imply that the given image perfectly matches the image on the retina, but rather that we can correlate a specific retinal point with a well defined small region in the image.

2.1.5 Automatic target detection

During experiments conducted in the birds aviary, a target (dead chick) and several yellow distracters were randomly positioned on the floor (Fig. 5 top). Both were automatically detected in fixation images according to the following scheme. First, the image was transformed from RGB into CYM color space. Then, color values were mapped into a single 2D image by enhancing the Y channel using the following transformation:

$$\text{Yellowness} = \frac{Y(x, y)}{\varepsilon + Y(x, y) + C(x, y) + M(x, y)}. \quad (4)$$

A fixed threshold transformed the yellowness measure into a binary mask. Morphological operations (erode, dilate) were applied to discard some of the irrelevant candidates, such as elongated shapes or very small blobs (Gonzalez and Woods 2001). The final binary mask obtained using the above procedure contained the spatial positions of the target and distracters in a scene (Fig. 5 bottom panel).

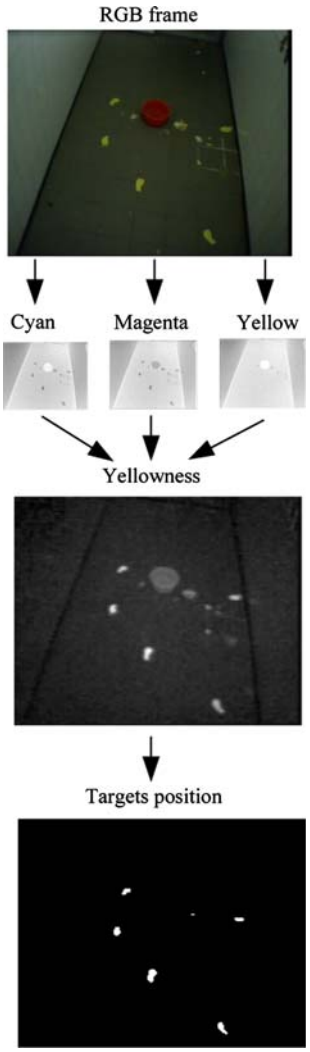

Fig. 5 Automatic target segmentation. A fixation image (*top*) is decomposed into CYM color space and integrated back to form a yellowness measure. The yellowness image is thresholded to obtain potential targets position (*bottom*)

2.2 Results

2.2.1 Scene-cut detection

To evaluate the performance of our algorithm, we manually labeled each frame in a test sequence containing 4,000 frames. The test sequence was extracted from one of the longer sequences in our database. Our scene-cut algorithm was applied on this test sequence and results were compared to manual labeling by measuring hits and misses per frame.

Using different discrimination threshold levels, we constructed a ROC curve accordingly (Fig. 6). Two additional algorithms were implemented to assess the performance level of our novel algorithm. The first was based on measuring the absolute frame difference, and the second on histogram difference (Ahanger and Little 1996).

We found that local edge density similarity measure was substantially superior to the two competing algorithms (Fig. 6a). We selected the best discrimination threshold using Eq. 1. A distinct local maxima in this correlation function was found at approximately 0.9 (Fig. 6b). The overall performance obtained a sensitivity of 90% and a specificity of 95%.

Our scene-cut algorithm was subsequently applied to all video sequences recorded in the aviary. A total of 9,236 fixations were obtained in experiments from owl YA and 6,561 fixations from owl SC. Histograms of fixation and saccade durations was calculated. The results from owl SC are

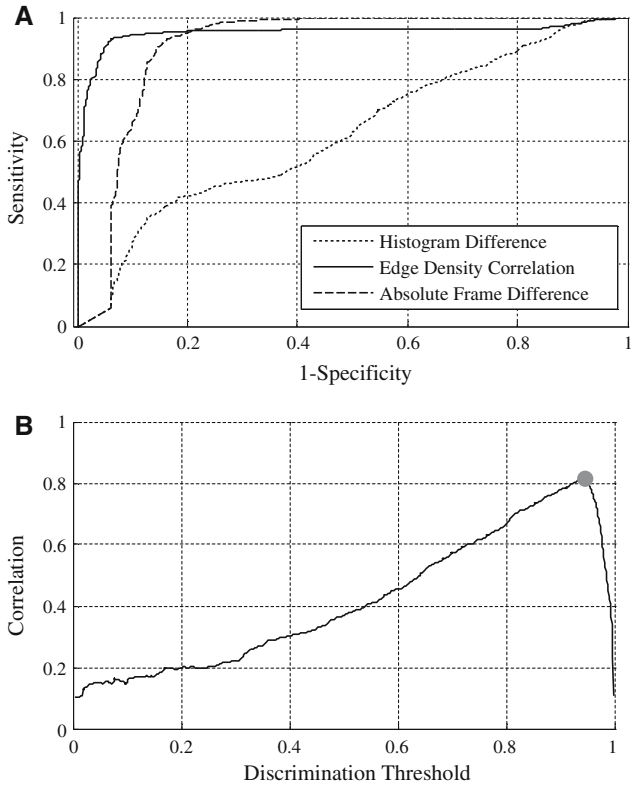

Fig. 6 **a** Quantitative comparison of three scene-cut algorithms with varying sensitivity parameter. Edge density correlation measure outperforms histogram difference and absolute frame difference methods. **b** Correlation between ground truth and results obtained using local edge density correlation on the test sequence

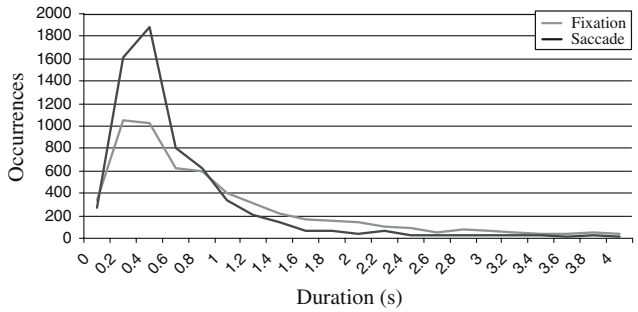

Fig. 7 Fixation and saccade duration histograms in aviary sequences for owl SC

shown in Fig. 7. Saccade durations peaked at 0.5 s while fixation durations peaked at approximately 0.4 s. Notice that the number of fixations is equal to the number of saccades since each fixation always appears between two adjacent saccades. Fixations had a broader distribution and dominated durations longer than 1.0 s, suggesting that the subject fixate objects for longer periods compared to saccade movements. The long saccade durations typically corresponded to intervals in which owls shook their head trying to remove the owl-cam. A manual inspection of video sequences revealed that most saccade movements were precise. At times, a secondary micro saccade would follow to correct the large saccade movement. This was probably due to the unusual weight of the owl-cam on the owl's head.

2.2.2 Image fixation area

The scene-cut detection algorithm reduced the amount of data to be further analyzed considerably by retaining only a small subset of fixation images from the full video sequence. Each fixation image was processed using the automatic target detection algorithm (Sect. 2.1.5). The result was a set of binary masks (one per fixation image), representing the spatial position of target and distracters, relative to the owl's retina.

In each experiment we summed up all binary images to obtain a map which represents the likelihood of a target (or a distracter) appearing somewhere in the owl's visual field. These maps were normalized to form a probability density function (PDF) of object position. A typical PDF represents the integration of approximately 600 fixations, all of which occurred in a single experiment. Figure 8a–d depict a well localized peak in each PDF, suggesting that the owl repeatedly directed its gaze in a way that aligned a specific retinal position with a target or distracter. Similar results were also obtained from shorter sequences containing about 40 fixations.

Upon summation of binary maps from all experiments, a well localized circular cluster was found in the cumulative map (Fig. 8e), similar to the peak found in each single experiment. A multi-variate gaussian with symmetric covariance matrix was fitted at the cluster center-of-mass. Variance was estimated at $\sigma^2 \approx 100$ pixels, roughly corresponding to a circular area with a 30 pixels radius, or to 1.5° in the owl's visual field.

These results indicate that during a search task, owls scan their environment in a serial manner and repeatedly direct their gaze in a way that brings the target (or distracters) to a specific retinal position, which we refer to as the retinal fixation area.

From hereon, we define the owl's *fixation spot* as the computed center-of-mass of the cumulative map (Fig. 8e). In addition, we define the owl's *image fixation area* as a circular disc, centered at the fixation spot, with varying radius *r*.

In contrast to traditional fixation acquisition systems, where subject's fixation position is measured relative to a display screen and changes with each saccade, our video sequences are registered relative to the observer. Put in other words, the fixation spot never changes its position in video sequences regardless of subject's movements. Furthermore, the detection of the fixation spot does not depend on the spatial configuration of target and distracters on the floor, since they may appear anywhere in the owl's visual field.

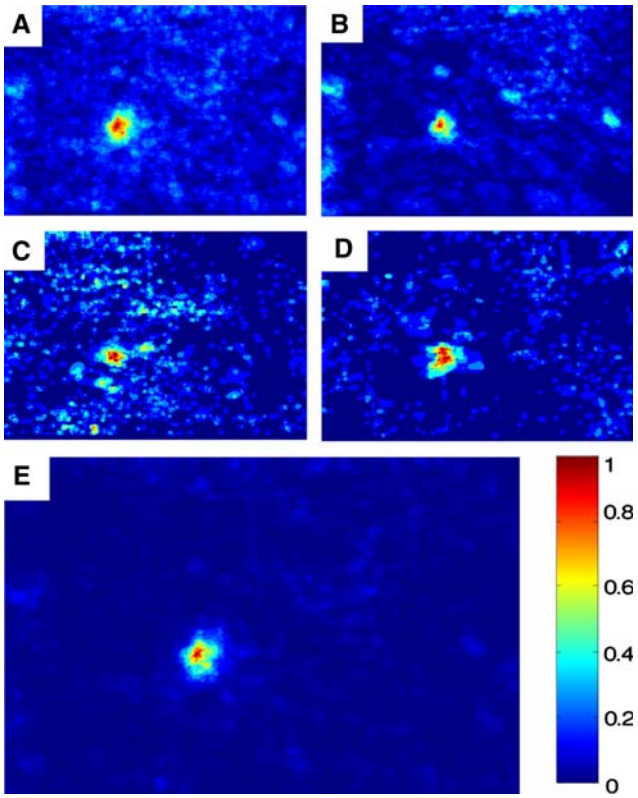

Fig. 8 **a–d** Superimposed images of target and distracter positions in the visual field. Each image depicts the result of a different experiment. **e** summation of all target positions in all experiments. The **color bar** indicates the normalized number of occurrences

We validated our assumption that the object appearing in the image fixation area was indeed the object of interest by observing the last video frames in each experiment. In all experiments that ended with the owl flying toward an object and catching it, the object caught by the owl was the one which appeared in the image fixation area. Furthermore, during the short flight the object remained at the image fixation area, growing in size and moving downward along the epi-polar line, until it was caught.

3 Image content and fixations

The results described in Sect. 2 indicate that owls have a specific retinal fixation area which is repeatedly directed towards yellow objects during fixation, which could be the target or a distracter. But, what happens during a free viewing task? Is the owl, and hence, the retinal fixation area drawn to parts of the scene that are different in some sense (semantically or statistically)? What qualities in the visual scene draw the attention of an owl? In this section, we address these questions by analyzing images that were recorded in three different environments while owls freely viewed a scene. Statistical analysis of image patches is described and results are compared to results of similar studies in primates.

Table 2 Recorded video

Data statistic	Owl YA	Owl SC
Aviary		
Total frame count	966, 400	661, 860
Total length (min)	671	459
Fixation images	9, 236	6, 561
Office		
Total frame count	136, 269	334, 479
Length (min)	94	232
Fixation images	1, 865	5, 913
Outdoors		
Total frame count	106, 577	122, 134
Length (min)	74	84
Fixation images	1, 093	1, 601

3.1 Materials and methods

3.1.1 Experimental setup and protocol

To examine local properties at the image fixation area, the same two owls were used in another series of experiments. Prior to experiments, the fixation spot of each owl was determined over several trials using the method described in Sect. 2.1.5. Experiments were conducted in three types of environments. The first was the aviary. We used the same video sequences as in Sect. 2.1.1. Since owls were not always engaged searching for the food item, but also observed their surroundings, we find it still valid to question what local image features in the aviary drew owl's attention. The second was our laboratory office environment which mainly contained objects, such as the ones seen in Fig. 2 (Sub plots d,e,f). The owl-cam was mounted on the owl's head and the owl was tied to a perch in various locations around the room. Even though the owl was constrained, it could still freely move its head and explore the full 360° of its visual environment. The owl was left alone in the lab while images were recorded for approximately 10–20 min (Table 2). The third type of environment was outdoors. We used two locations, one in a home garden and the other on the roof of our laboratory. Both locations contained trees, shrubs, and other objects, as can be seen from images recorded using the owl-cam in Fig. 2 G,H,I. Short video sequences obtained using the owl-cam are available online at http://brainstory.info/Research/OwlAttention.

3.1.2 Luminance at the image fixation area

We define a circular region around the fixation spot (x_0, y_0) as:

$$Q_{(x_0,y_0,r)} = \left\{(x, y) \,|\, (x - x_0)^2 + (y - y_0)^2 \leq r^2\right\} \qquad (5)$$

Mean luminance, sampled at a circular area of radius r, centered at position (x_o, y_o) in the kth fixation image I_k is defined as:

$$L_k(r, x_0, y_0) = N(r)^{-1} \sum_{(x,y)\in Q_{(x_0,y_0,r)}} I_k(x, y) \quad (6)$$

where $N(r)$ denotes the number of (discrete) pixels inside the circular area.

Luminance values are computed from the color images by eliminating the hue and saturation channels using the standard Matlab function rgb2gray, resulting in a possible range of values between [0,1].

3.1.3 Luminance contrast at the image fixation area

Several recent studies have shown that luminance contrast levels at the fixation point in humans are greater compared to contrast levels sampled at random areas (Reinagel and Zador 1999; Parkhurst et al. 2002; Kayser and Logothetis 2006). Following the definitions of Reinagel and Zador (Reinagel and Zador 1999), we define contrast, sampled at a circular area of radius r, centered at position (x_o, y_o) in fixation image I_k as:

$$C_k(r, x_0, y_0) = \bar{I}_k^{-1} \sqrt{N(r)^{-1} \sum_{(x,y)\in Q_{(x_0,y_0,r)}} (I_k(x, y) - L_k(x_0, y_0, r))^2} \quad (7)$$

where $\bar{I}$ denotes average luminance of the entire kth fixation image. Contrast is measured as the standard deviation of luminance in a circular region, normalized by the mean image luminance. This normalization provides a unitless measure and is less affected by global illumination of the entire scene.

3.1.4 Edge density at the image fixation area

Recently, Baddeley and Tatler have suggested that contrast plays a minor role in drawing human fixations, while high frequency edges are the dominant factor (Baddeley and Tatler 2006). To examine whether edges plays a role in drawing owl's attention, we have measured local edge density at the image fixation area using the following definition:

$$E_k(r, x_0, y_0) = N(r)^{-1} \sum_{(x,y)\in Q_{(x_0,y_0,r)}} Edge_k(x, y), \quad (8)$$

where $Edge_k(x, y)$ denotes a binary image that is obtained by applying canny edge detector on the kth fixation image.

3.1.5 Two-point correlation at the image fixation area

To evaluate how gray scale values are correlated within the small region patch, we used a two-point correlation function, defined as:

$$R(x, y, x_0, y_0) = \sum_k \sqrt{\left(I_k(x_0, y_0) - \bar{I}(x_0, y_0)\right)\left(I_k(x, y) - \bar{I}(x, y)\right)} \quad (9)$$

where $\bar{I}(x, y)$ denotes the average luminance at pixel (x, y) across all fixation images. Note that we measure the correlation relative to point (x_0, y_0), which represents the owl's fixation spot. The normalized two-point correlation function is defined as

$$\hat{R}(x, y, x_0, y_0) = R(x, y, x_0, y_0) / R(x_0, y_0, x_0, y_0). \quad (10)$$

We discard the orientation component in the two-point correlation function, by transforming the coordinates to polar and averaging over all orientations:

$$R(\rho) = \frac{1}{2\pi} \sum_{\Theta=0}^{2\pi} \hat{R}(\rho\cos\Theta, \rho\sin\Theta, x_0, y_0). \quad (11)$$

Thus, $R(\rho)$ represents the extent of pixels at a distance of ρ which are correlated with the central pixel of the fixation spot.

3.1.6 Saliency ratio at the image fixation area

What is the role of saliency in attracting owls' attention? To answer this question, we have tested a common computational model of attention on the obtained fixation images (Itti et al. 1998; Koch and Ullman 1985). If the visual saliency hypothesis is true for owls, high saliency values are expected at the image fixation area. Saliency was computed using the Saliency Toolbox for Matlab (Walther and Koch 2006). This toolbox computes saliency maps from images as well as predictions of fixation positions according to the winner-take-all and inhibition of return schemes (Itti et al. 1998). In our implementation, we have used four feature maps, namely: contrast, color, edge orientation, and hue that were computed across several scales. We define the saliency ratio measure at the image fixation area as

$$S_k(r, x_0, y_0) = (\max S_k)^{-1} \max_{(x,y)\in Q_{(x_0,y_0,r)}} Sal_k(x, y), \quad (12)$$

where $Sal(x, y)$ denotes the saliency image obtained using the Saliency Matlab Toolbox. The maximal value at the image fixation area was selected as the highest possible response using this model over the given radius. The computed saliency

value are normalized by the maximal saliency value across the entire fixation image, mapping the possible range between 0 and 1.

3.1.7 Random observer(s)

The quantitative measurements described above are meaningless unless they are compared to measurements extracted at other areas of the visual field. Therefore, in our experiments we have used the notion of a random observer. In each fixation image, the fixation spot of a random observer is generated by sampling a uniform distribution. Note that this method does not guarantee that the random observer never fixates at a position that was fixated by the owl since the randomly selected position appears in the owl's visual field and could have been fixated at a different frame. This issue is discussed in more detail in Sect. 4.

The measures taken for owl YA and owl SC do not always exhibit the same range, since they depend on environmental factors that were out of our control. For example, global illumination levels depend on weather and affect both luminance and luminance contrast (even in indoor scenes, such as our office room). Taking this into consideration, we compared the normalized measures, which are derived by dividing the owl's measure by the corresponding random observer's measure.

An alternative method of comparison is to test the obtained measures at the fixation spot against other spatially consistent areas in the owl's visual field. In this approach, the visual field is divided into small non-overlapping areas and low-level image properties are evaluated in each area. This method allows us to make quantitative evaluations between values at the fixation spot, and values at other parts of the scene.

3.1.8 Spectral analysis of fixation images

Within the set of all possible images lies a small subset of images, which represent natural stimuli to our visual system. Although natural images may vary in terms of objects, scenes, textures and luminance levels, several studies have reported that they all share a common property; the average power spectrum has a typical fall-off at $1/f^2$, where f is spatial frequency (Field 1987; Parkhurst and Niebur 2003). Here, we investigate whether images from the owl's point of view have a similar spectral slope as that reported for random natural images. We define the power spectral density of a fixation image $I_k(x, y)$ of size $[MxN]$ as:

$$PSD\,(u, v) = 10 \log_{10} \left(\frac{F_k\,(u, v)\, F_k\,(u, v)^*}{MN} \right), \tag{13}$$

where $F_k\,(u, v)$ represents the fourier transform of a filtered fixation image:

$$F_k\,(u, v) = \sum_x \sum_y H_k\,(x, y)\, e^{-2\pi i (ux+vy)}. \tag{14}$$

$H_k(x, y)$ denotes a filtered fixation image which is obtained by point-wise multiplication of $I_k(x, y)$ with a smoothing Hanning window. This removes boundary effects that are caused by the discrete fourier transform assumption about a periodic signal. The spectral slope of an image was estimated by transforming $PSD\,(u, v)$ into polar coordinates $PSD\,(\rho, \theta)$, averaging all θ and robust fitting a straight line.

3.2 Results

3.2.1 Luminance

We define a set of variables $L_{\text{owl}}\,(r)$, each representing the mean luminance levels at owls' image fixation area for a given radius. We measured L_{owl} for radius intervals ranging between 2 and 40 pixels, corresponding to 0.1 and 3° of the owl's visual field. These values were selected within a consideration to the image fixation area size, as calculated in Sect. 2.2.2. Samples of these variables were collected from all fixation images by substituting (x_o, y_o) in Eq. 6 with the owl's known fixation spot coordinates.

In the same manner, we define a set of variables $L_{\text{rand}}\,(r)$, representing the mean luminance levels at random image patches as a function of radius. Samples of these variables were collected by replacing (x_o, y_o) in Eq. 6 with random coordinates drawn from a uniform distribution, scaled to image resolution, at each fixation image.

The mean values of variables L_{owl} and L_{rand} are plotted in the first row of Fig. 9 as a function of their radius, at three different environments (aviary, office, outdoors). A non parametric Mann–Whitney U test indicated a significant difference between the two populations. The most prominent difference was found for owl YA in the aviary environment (p value $< 10^{-20}$).

The main characteristic of this measure was the inconsistency relative to the random observer. For example, in the aviary sequence, both owls typically fixated at patches with lower luminance levels, but in the outdoors one owl's luminance levels was approximately 10% greater compared to its corresponding random observer, while the second's was about 5% lower.

The variance across subjects was not found to be significant. For example, in the aviary sequences the standard deviation was in the range of 0.22–0.2 for owl SC, and between 0.25 and 0.23 for owl YA. The standard deviation typically decreased monotonically as a function of the radius. In outdoors environment, the difference was even less significant

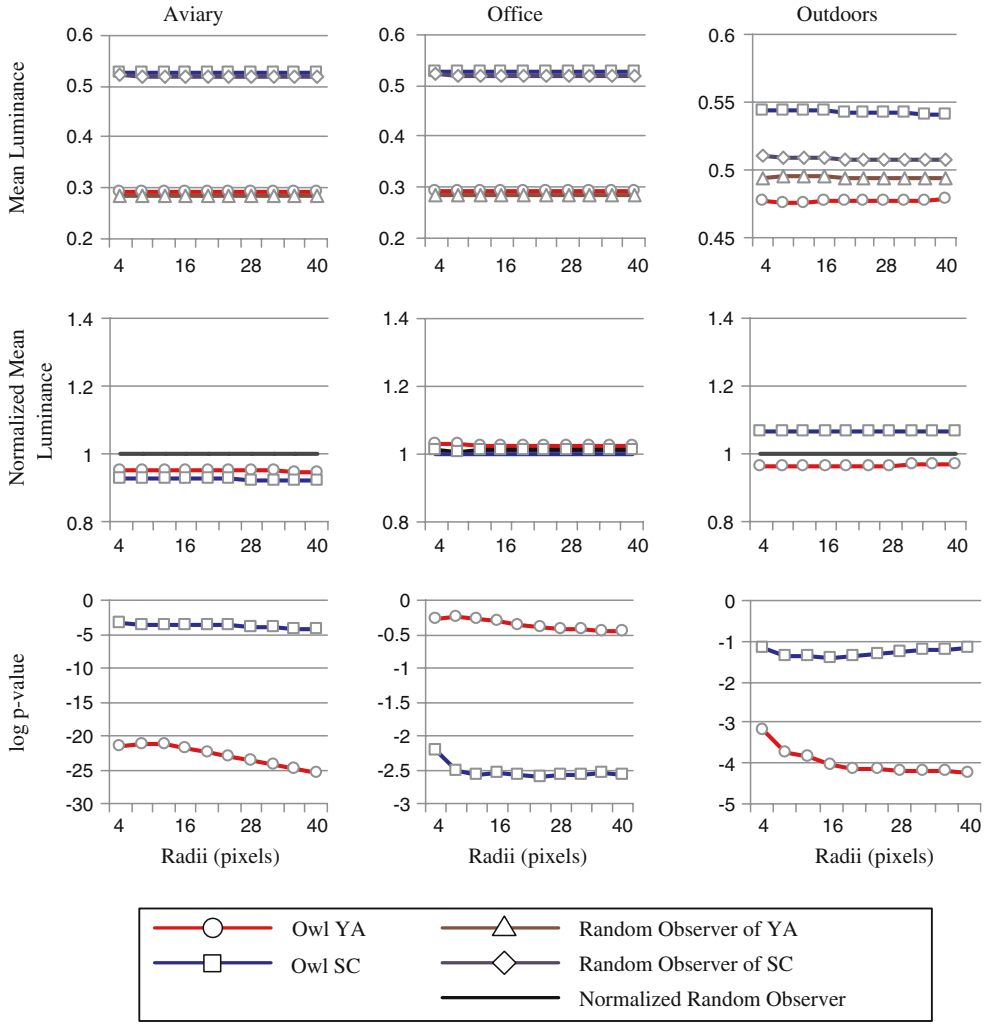

Fig. 9 Comparison of luminance levels of owls and their corresponding random observers. Raw values are depicted in the first row. These values were divided by their corresponding random observer values to obtain the normalized values, depicted in the second row. Third row displays p values of a Mann–Whitney U test between $L_{\text{owl}}(r)$ and $L_{\text{rand}}(r)$

as standard deviation was in the rage of 0.26–0.21 for owl SC, and 0.27–0.22 for owl YA.

3.2.2 Luminance contrast

Similar to the estimation of luminance, we define a set of variables $C_{\text{owl}}(r)$ and $C_{\text{rand}}(r)$, representing local image contrast. Samples were collected from fixation images using Eq. 7.

In all environments, the contrast had a typical curve which increased with radius. Furthermore, both owls consistently fixated at areas that contained greater contrast, compared to their corresponding random observer. Contrast levels were approximately 20–30% greater at the fixation area, and were statistically significant, according to Mann–Whitney U test. Most p values that were computed between the owl measurements and their corresponding random observer were below 10^{-5}.

Similar results were obtained when alternative definitions of contrast were used, such as the following min-max definition: $C = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$. In addition, we verified that contrast levels of C_{owl} and C_{rand} converge at large radius values. Thus, at radius of 60 pixels and above p values were well above -0.5 (Fig. 10).

3.2.3 Edge density

Local edge density, as defined by Eq. 8, was sampled at owls' fixation area $E_{\text{owl}}(r)$ and at random observer fixation spots $E_{\text{rand}}(r)$. In all recorded sequences the mean edge density at the owl's fixation area was greater compared to the values of the corresponding random observer. Furthermore, the

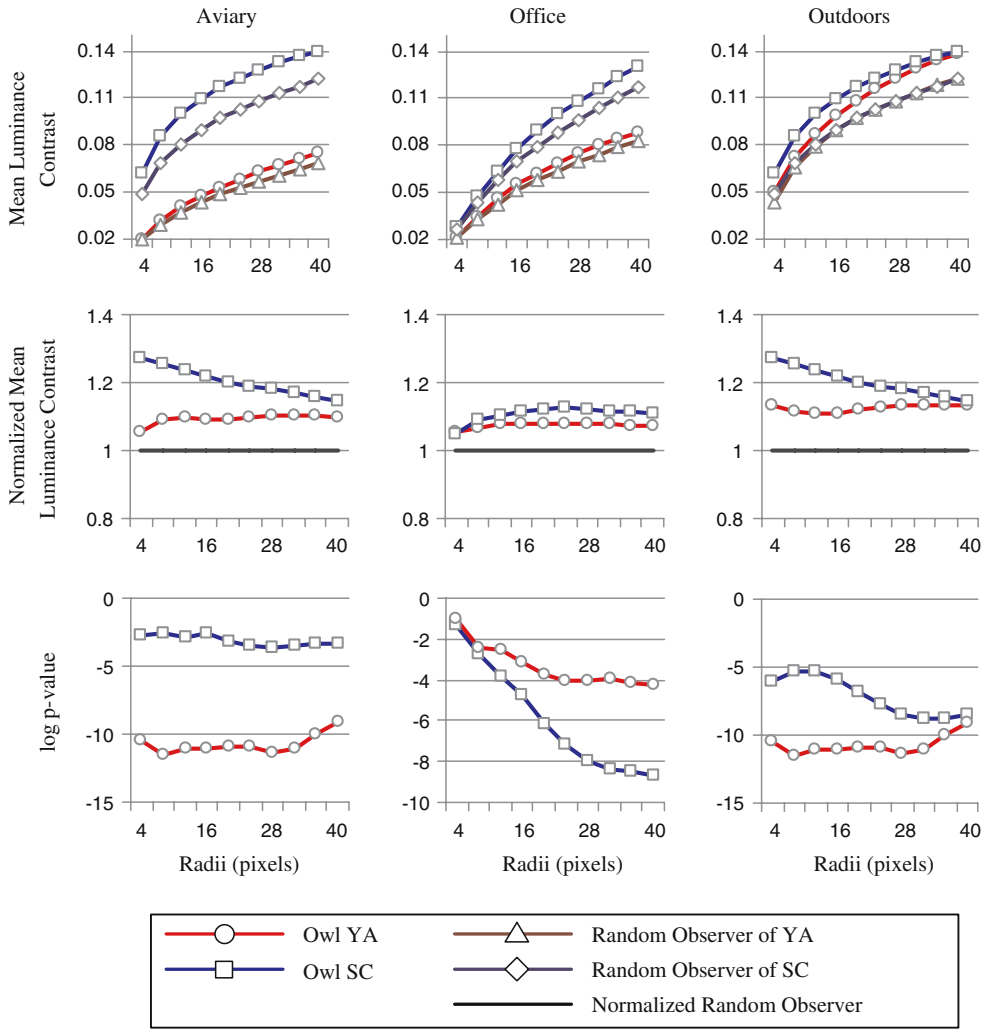

Fig. 10 Luminance contrast levels of owls and their corresponding random observers. Raw values are depicted in the first row. Normalized values are shown on the second row, and the p-values between $C_{\text{owl}}(r)$ and $C_{\text{rand}}(r)$ are displayed on the third row

normalized values of owl SC were typically greater compared to the normalized values of owl YA. The highest difference was recorded in the outdoors environment, where edge density of owl SC was typically 60% greater compared to its corresponding random observer. Most p values were well below 10^{-5} indicating that they are statistically different (Fig. 11).

3.2.4 Two-point correlation

To determine the extent to which nearby pixels of the fixation spot are correlated, we have used the two-point correlation measure, as defined in Eq. 11. For perfect correlated points, such as the center pixel of the fixation spot ($\rho = 0$), $R(\rho)$ has the value of one. As ρ increases, correlation drops. As one can see in Fig. 12, two-point correlation at owls fixation spot drops faster than their corresponding random observers and is consistent in all environments. The lowest correlation (given the radii interval we used) was measured in the outdoors environment, and was 0.78.

3.2.5 Saliency

Measurements of saliency were sampled at owls' image fixation area $S_{\text{owl}}(r)$ and at random patches $S_{\text{rand}}(r)$. In all environments saliency increased with radius, suggesting a salient region in a nearby area to the sampling location (Fig. 13). However, the saliency values measured were at most 20% out of the maximal observed saliency in each fixation image, indicating that the most salient object was not at the owl's fixation area.

Interestingly, although the mean saliency value of owl SC in the aviary was almost twofold greater compared to its corresponding random observer, p values showed no statistical significance between the two populations. Nevertheless,

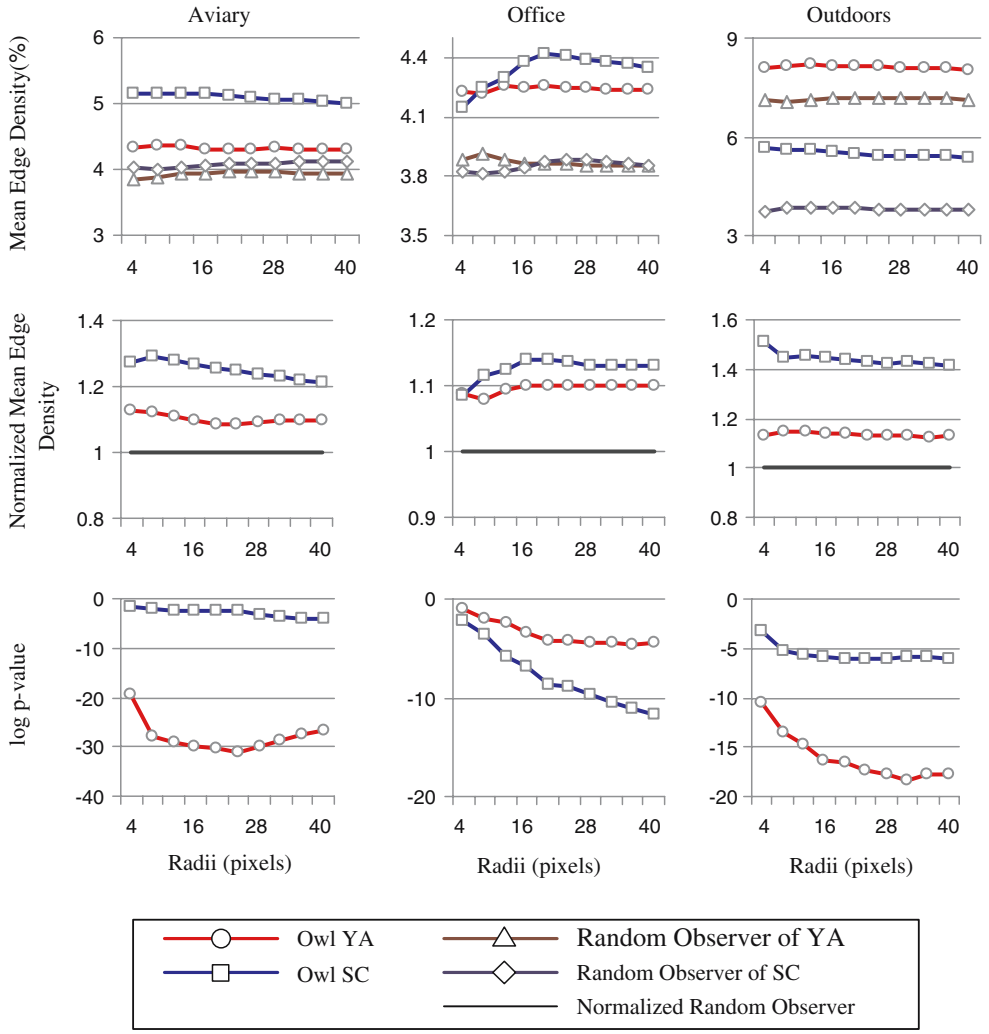

Fig. 11 Edge density values of owls and their corresponding random observers. Raw values are depicted in the first row. Normalized values are shown on the second row, and the p values between $E_{\text{owl}}(r)$ and $E_{\text{rand}}(r)$ are displayed on the third row

in all other measurements, owl values were about 20–60% greater with p-values well below 10^{-3}.

3.2.6 Global analysis of fixation images

The analysis presented so far has focused at a small retinal area, which corresponds to a narrow angle in the owl's visual field. However, global analysis of the entire fixation image provides additional insight to the type of input owls' visual system is operating on. First, we investigated the spectral properties of images from different environments using standard fourier analysis (Sect. 3.1.8). We found a larger concentration of power density spectrum at low frequencies, The average spectral slope of fixation images from the outdoors environment was estimated at 2.0. In both aviary and office environments (indoors), the slope was greater and was estimated at 2.5 (Fig. 14).

In addition, we looked at average values of luminance, contrast, edge density and saliency across the entire recorded visual field by dividing the image into small non-overlapping regions (19 × 10 blocks). We summed the response of each measure to test whether the values at the fixation spot were not only different compared to a random observer, but also to other consistent areas in the visual field.

We found that the average luminance typically decreases along the vertical axis, probably due to the fact that light source in all sequences came from above (Fig. 15, top row). In addition, a non-isotropic decrease of values along radial lines were found in both average contrast and average edge density. Average saliency was less structured, but displayed greater values near the fixation spot as well.

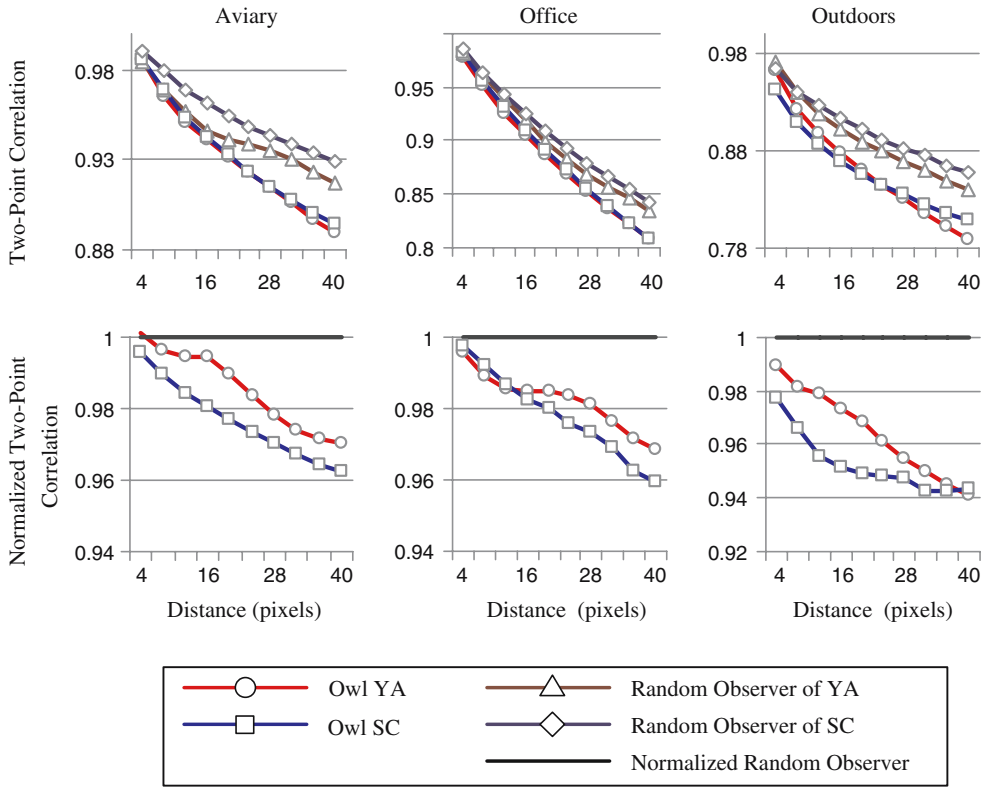

Fig. 12 Two-point correlation as a function of the distance from the center fixation spot. Columns correspond to the three different environments. *First row* two-point correlation values of owls and random observers. *Second row* two-point correlation normalized according to the corresponding random observer

4 Discussion

The primary objective of this study was to examine the interactions between scene content and owls' visual attention. With this in mind, we initially searched for an area that is attended more in the visual field, and then continued to study whether scene content selected by active head movements is correlated to low-level image statistics at the attended location.

4.1 Owls' fixation area

In a recent analysis of barn owls' head movements, Ohayon et al. have proposed that owls' head movements may be categorized to three main classes, namely: fixation, saccade and peering movements (Ohayon et al. 2006). Although the kinematic aspects of saccade and peering movements have been thoroughly investigated, fixations, being static in nature, lacked a proper spatial characterization.

Here, we describe, to the best of our knowledge, the first study which attempts to rigorously define the spatial characteristics of fixations in barn owls. We have addressed the question of fixation consistency and have shown a strong correlation between a specific region in the owl's retina and objects appearing in the visual field. Despite the fact that objects of interest may appear anywhere in the owl's visual field, we found that owls serially scan their environment and direct their gaze in a way that brings the object of interest to a specific retinal area which we have referred to as the retinal fixation area.

Global analysis of the entire visual field showed a strong correlation between low-level image features and the image fixation area. Luminance contrast at the fixation area was not only greater compared to random areas, but was also among the highest in the entire visual field. This observation suggests that integration of luminance contrast, edge density or other low-level image features over a long sequence can be used for automatic detection of the fixation spot without any need for pre-experiment calibrations using yellow targets.

The retinal topography of barn owl's includes two areas with increased ganglion cell density. One is a horizontal streak which spans across the entire retina and the other is a temporal area centralis, approximately 2 mm in diameter (Wathey and Pettigrew 1989). Although we have shown the existence of an image fixation area, we could not register it anatomically to prominent retinal features because we lacked a reference point. Put in other words, our analysis has not shown that the retinal fixation area overlap with the known anatomical features. Theoretically, this can be demonstrated by a simultaneous recording of a light source, both using the

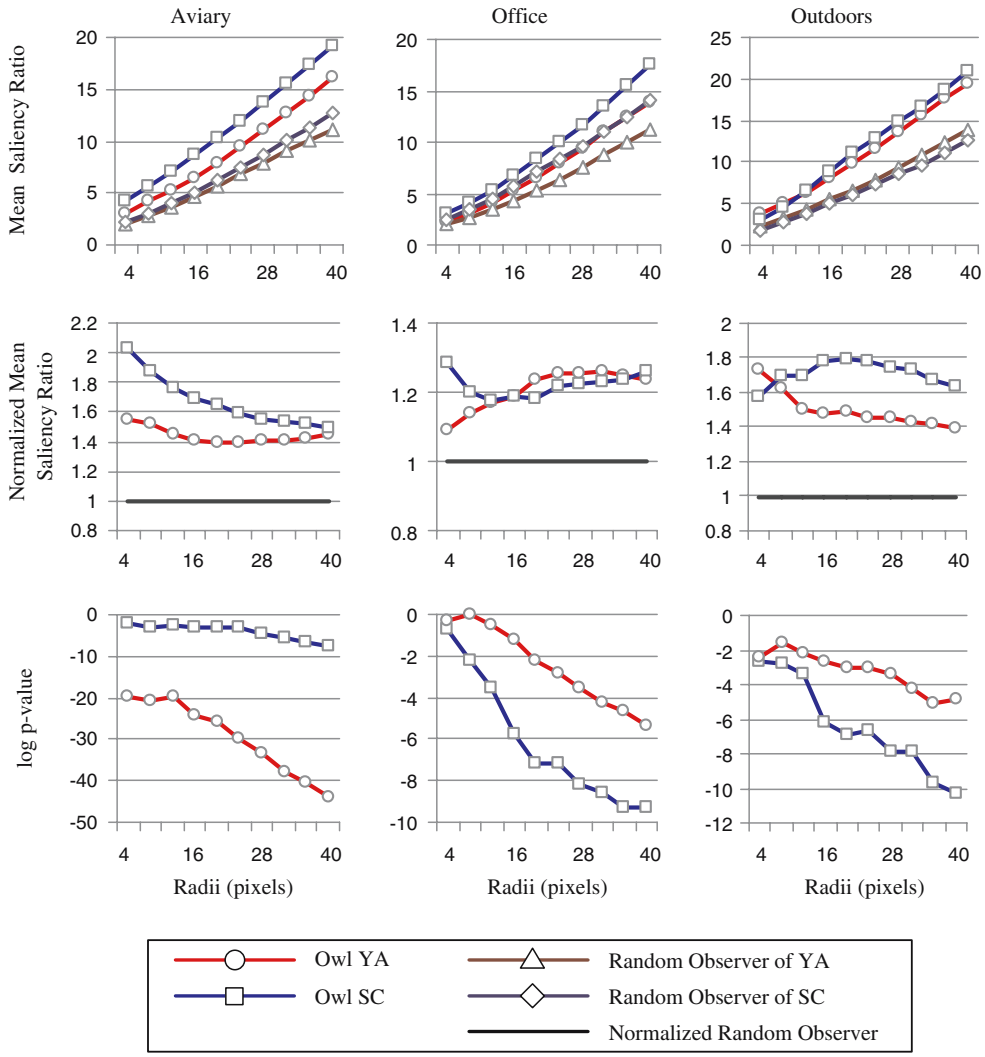

Fig. 13 Saliency values of owls and their corresponding random observers. Raw values are depicted in the first row. Normalized values are shown on the second row, and the p values between $S_{\text{owl}}(r)$ and $S_{\text{rand}}(r)$ are displayed on the third row

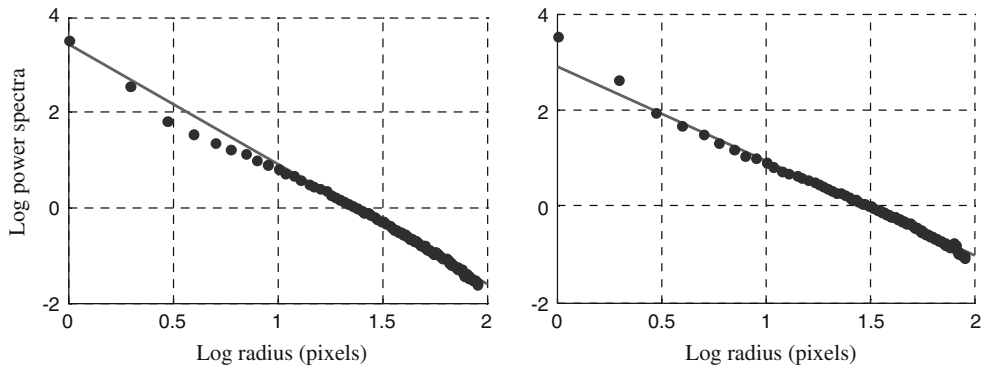

Fig. 14 Spectral slope of natural images obtained at different environments. **a** Indoors scenes (aviary, office), **b** outdoors scenes

owl-cam and using a digital ophthalmoscope. In practice, we hypothesize that the retinal fixation area coincides with the temporal area centralis since it is the highest acuity area in the owl's retina.

4.2 Low level image properties at the fixation area

Several recent studies of low level image properties at fixation positions in humans have indicated a statistical difference of

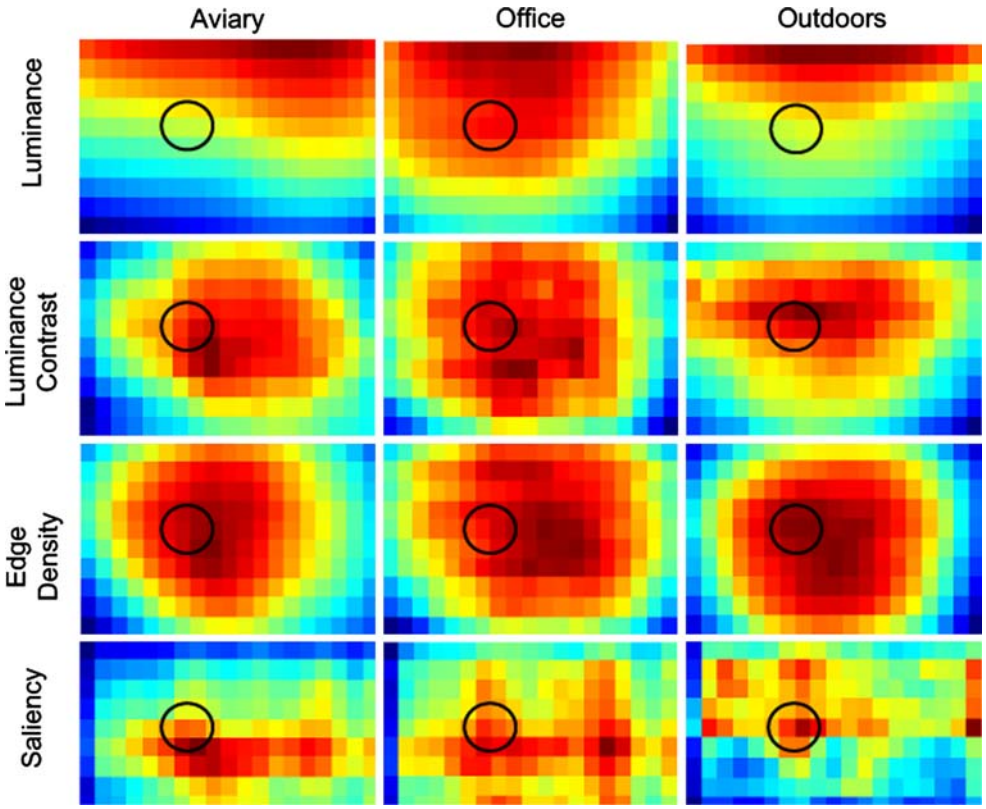

Fig. 15 Global statistics of owl YA, averaged over all fixation images. The fixation area is marked by a *black circle*

various measures at the fixation spot, relative to measurements that were extracted at random. Reinagel and Zador have shown both an increased contrast levels at the fixation spot, as well as a greater decrease in correlation, compared to a random observer (Reinagel and Zador 1999). Similar results were also obtained by Parkhurst and Niebur as they investigated the statistical differences in various image types (Parkhurst and Niebur 2003).

The results we obtained are in agreement with both these studies. Contrast levels were found to be greater by almost 30% near the fixation spot, and two-point correlation was lower by approximately 0.05, which is similar to the decrease found by Parkhurst and Niebur (Parkhurst and Niebur 2003). Furthermore, we found significantly higher values of edge density that were approximately 30–50% greater compared to the random observer. These results are in accord with the recent studies of Krieger et al. (2000) and Henderson et al. (2007) who reported that eye movements in humans are drawn to edges. Interestingly, Henderson also reports lower luminance levels at the fixation spot, which in our experiments we found to be less significant (compared to other measures) and inconsistent across the different environments we recorded at.

A major difference between our experiments and the ones that were done on humans is related to the method in which data were collected. While we have gathered fixations continuously during a period of more than 10 min at which owls could freely look at any direction, fixations in humans were extracted during a short interval (typically, 4 s) after the onset of image stimulus and were restricted to the area of the computer monitor. Closer to our method, the recent studies of Land and Hayhoe (2001) and Hayhoe and Ballard (2005) have managed to extract fixations during natural behavior of humans, but unfortunately did not report low-level image properties and focused on high level cognitive tasks.

Statistical measures obtained from human fixations using eye trackers are commonly compared to a random observer or to a shuffled ensemble. In the shuffled ensemble method, subject's fixation positions are used, but image data are sampled from randomly selected images from the given image set (Parkhurst and Niebur 2003; Henderson et al. 2007). While the first method is similar to the one we used, the latter tries to account for subjects' bias to fixate in the central region of observed images. We could not use the shuffled technique since in our case, fixation position in the image is fixed throughout all experiments.

The statistical comparison to a random observer may introduce a bias and diminish differences between two populations since a random fixation spot may be fixated by the owl at another time instance, thus corrupting the set of random

image patches with samples of owl fixations. The proper comparison should be relative to areas that were never fixated. However, generating the set of positions in images which were never fixated is a complicated computational problem that requires a global alignment of all fixation images by generating a visual field mosaic. Nevertheless, the statistical tests we utilized have shown strong significant differences, supporting the hypothesis that in some cases the actual differences may even be greater than the observed ones.

4.3 Attention and saliency

The results presented in Sect. 3.2 indicate that low-level image features are different at the fixation area compared to random image patches. However, what is the contribution of these differences in drawing the owl's attention? To answer this question we quantified the low-level contributions using a popular computational model of attention. The model assumes that low level image features are extracted across multiple scales and are integrated into a unified saliency map. A high score in the saliency map represents a conspicuous object that is most likely to be fixated. Since the images we obtained using the owl-cam are represented relative to the observer, a high saliency value is expected at owl's fixation spot.

The validity of using this model in this research should be mentioned, as this biologically plausible model was originally proposed by Koch and Ullman to predict human fixations (Koch and Ullman 1985). We postulate that from a computational point of view, owls have the necessary neuronal circuits that are needed to compute a saliency map. Owl's visual Wulst is an area which shares similar physiological properties to the mammalian visual cortex. It has precise topographic organization, selectivity for orientation, motion, disparity, and even illusory contours (Pettigrew and Konishi 1976; van der Willigen et al. 2001; Nieder and Wagner 1999). All these support the hypothesis that saliency map computations may take place in owl's visual wulst.

However, as we have found, this model gives poor predictions to owl's fixation spot. Saliency values at the fixation area were typically not the highest in each fixation image and their average was at most 20% out of the maximal saliency value. This comes as no surprise since a pure bottom-up attention model cannot account for all fixations. Objects of interest depend on the task at hand and without a model including top-down processes, peaks in the saliency maps will not correlate well to the correct fixation position.

Furthermore, the underlying reason for the observed statistical differences is still unclear. On the one hand, attention may be drawn to low-level image features, such as the ones we presented. On the other hand, these differences may appear in conjunction to a semantic object that appears in that area. Recent studies have demonstrated that areas with increased localized features, such as contrast, do not attract more attention and that saliency may merely generate a set of locations that are later selected according to their informativeness and whether they are related to the task at hand (Einhäuser and König 2003; Einhäuser et al. 2006; Itti 2005; Henderson et al. 2007).

The search task of finding a real food item among a set of distracters can be considered a conjunction search, because the distracters we used shared the same shape, texture and color to the real food item. It would be interesting to determine whether owls have a pop-out effect and can detect a target without a serial search when it differs only by a single feature.

5 Outlook

The main characteristic of the environments in which we have recorded was their static nature. However, most environments are dynamic and include elements that keep on changing. Information is gathered not only in discrete events, but in a continuous stream of sampling throughout the fixation interval. Our future research will focus on this notion with the goal of obtaining a better understanding of attentional mechanisms that control gaze for selective acquisition of visual information in barn owls. Preliminary results have already shown that owls' attention is drawn toward areas in the visual field with increased motion. It remains to be determined what types of motion attract such attention and under which circumstances.

Acknowledgments We would like to acknowledge the following people who helped us to gather the data and prepare the manuscript: Sandra Brill, Robert van der Willigen and Petra Nikolay. In addition, we would like to thank Peter König for his insightful comments on an earlier version of the manuscript. We would like to thank the coordinator of the Bursary program for Israeli students, at the Center of North Rhine-Westphalia, Arne Claussen, who made Shay Ohayons visit to Aachen possible.

References

Ahanger G, Little T (1996) A survey of technologies for parsing and indexing digital video. J Visual Commun Image Representation 7(1):28–43

Baddeley RJ, Tatler BW (2006) High frequency edges (but not contrast) predict where we fixate: a Bayesian system identification analysis. Vis Res 46:2824–2833

Baldi P, Brunak S (2001) Bioinformatics: the machine learning approach. 2nd edn. MIT, Cambridge

Betsch BY, Einhäuser W, Körding KP, König P (2004) The world from a cats perspective statistics of natural videos. J Biol Cybern 90(1):41–50

Buswell GT (1935) How people look at pictures. University of Chicago Press, Chicago

Dailianas A, Allen R, England P (1995) Comparisons of automatic video segmentation algorithms. In: Proceedings, SPIE photonics

east'95: integration issues in large commercial media delivery systems, Philadelphia
Du Lac S, Knudsen EI (1990) Neural maps of head movement vector and speed in the optic tectum of the barn owl. J Neurophysiol 63:131–146
Einhäuser W, König P (2006) Does luminance-contrast contribute to a saliency map for overt visual attention? Euro J Neurosci 17(5):1089–1097
Einhäuser W, Kruse W, Hoffmann K, König P (2006) Differences of monkey and human overt attention under natural conditions. Vis Res 46(8–9):1194–1209
Field D (1987) Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc Am A 4(12):2379–2394
Gonzalez RC, Woods RE (2001) Digital image processing. Addison-Wesley Longman Publishing Co., Inc., Boston
Hartley R, Zisserman A (2000) Multiple view geometry in computer vision. Cambridge University Press, Cambridge
Hayhoe MM, Ballard DH (2005) Eye movements in natural behavior. Trends Cogn Sci 9(4):188–194
Henderson J, Hollingworth A (1999) High-level scene perception. Ann Rev Psychol 50:243–271
Henderson J, Brockmole J, Castelhano M, Mack M (2007) Eye movements: a window on mind and brain. In: Visual saliency does not account for eye movements during visual search in real-world scenes. Elsevier,Oxford (in prep)
Itti L (2005) Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. Vis Cogn 12(6):1093–1123
Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 20(11):1254–1259
Johnen A, Wagner H, Gaese BH (2001) Spatial attention modulates sound localization in barn owls. J Neurophysiol 85(2):1009–1012
Kayser C, Nielsen KJ, Logothetis N (2006) Fixations in natural scenes: interaction of image structure and image content. Vis Res 46(16):2535–2545
Knudsen E, Konishi M (1979) Mechanisms of sound localization in the barn owl (tyto alba). J Comp Physiol 133:13–21
Knudsen EI, Blasdel GG, Konishi M (1979) Sound localization by the barn owl tyto alba measured with the search coil technique. J Comp Physiol 133:1–11
Koch C, Ullman S (1985) Shifts in selective visual attention: towards the underlying neural circuitry. Human Neurobiol (4):219–227
Krieger G, Rentschler H, Hauske G, Schill K, Zetzsche C (2000) Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. Spatial Vis 13:201–214
Land M, Hayhoe M (2001) In what ways do eye movements contribute to everyday activities? Vis Res 41(25–26):3559–3565
Liu G, Pettigrew J (2003) Orientation mosaic in barn owl's visual wulst revealed by optical imaging: comparison with cat and monkey striate and extra-striate areas. Brain Res 961(1):153–158
Mannan S, Ruddock K, Wooding D (1996) The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. Spatial Vis 10(3):165–188
Marr D, Hildreth E (1980) Theory of edge detection. Proc Roy Soc Lond B207:187–217
Martin G (1982) An owls eye: schematic optics and visual performance in strix aluco. J Comp Physiol 145(341–349)
Masino T, Knudsen EI (1990) Horizontal and vertical components of head movement are controlled by distinct neural circuits in the barn owl. Nature 345:434–437
Masino T, Knudsen EI (1993) Orienting head movements resulting from electrical microstimulation of the brainstem tegmentum in the barn owl. J Neurosci 13:351–370
Navalpakkam V, Itti L (2005) Modeling the influence of task on attention. Vis Res 45(2):205–231
Nieder A (1999) Mechanisms of depth perception in the visual forebrain of the behaving barn owl. PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Fakultät für Mathematik, Informatik und Naturwissenschaften
Nieder A, Wagner H (1999) Perception and neuronal coding of subjective contours in the owl. Nat Neurosci 2:660–663
Ohayon S, van der Willigen RF, Wagner H, Katsman I, Rivlin E (2006) On the barn owls visual pre-attack behavior: I. structure of head movements and motion patterns. J Comp Physiol A Neuroethol Sensory Neural Behav Physiol 192(9):927–940
Parkhurst D, Niebur E (2003) Scene content selected by active vision. Spatial Vis 16(2):125–154
Parkhurst D, Law K, Niebur E (2002) Modeling the role of salience in the allocation of overt visual attention. Vis Res 42(1):107–123
Payne RS (1971) Acoustic location of prey by barn owls (tyto alba). J Exp Biol 54:535–573
Pettigrew J, Konishi M (1976) Neurons selective for orientation and binocular disparity in the visual wulst of the barn owl. Science 193(4254):675–678
Rao R, Zelinsky G, Hayhoe M, Ballard D (2002) Eye movements in iconic visual search. Vis Res 42(11):1447–1463
Reinagel P, Zador A (1999) Natural scene statistics at the center of gaze. Network Comput Neural Syst 10:1–10
Steinbach M, Money K (1973) Eye movements of the owl. Vis Res 13(1):889–891
Treisman A, Gelade G (1980) A feature integration theory of attention. Cogn Psychol 12:97–136
Tucker V (2000) The deep fovea, sideways vision and spiral flight paths in raptors. J Exp Biol 203(24):3745–3754
Vanni-Mercier G, Pelisson D, Sakai K, Jouvet M (1994) Eye saccade dynamics during paradoxical sleep in the cat. Eur J Neurosci 6(8):1298–1306
van der Willigen RF (2000) On the perceptual identity of depth vision in the owl. PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Fakultät für Mathematik, Informatik und Naturwissenschaften
van der Willigen RF, Frost BJ, Wagner H (2001) Encoding of both vertical and horizontal disparity in random-dot stereograms by wulst neurons of awake barn owls. Visual Neurosci 18:541–554
Wagner H (1993) Sound-localization deficits induced by lesions in the barn owl's space map. J Neurosci 13(1):371–386
Walther D, Koch C (2006) Modeling attention to salient proto-objects. Neural Netw 19:1395–1407
Wathey J, Pettigrew J (1989) Quantitative analysis of the retinal ganglion cell layer and optic nerve of the barn owl tyto alba. Brain Behav Evolut 33(5):279–292
Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (tyto alba). J Neurophysiol 96:730–745
Yarbus A (1967) Eye movements and vision. Plenum Press, New York
Zaharescu A, Rothenstein A, Tsotsos J (2004) Towards a biologically plausible active visual search model. In: ECCV Workshop on attention and performance in computer vision
Zhang Z (2000) A flexible new technique for camera calibration. IEEE Trans Pattern Anal Mach Intell 22(11):1330–1334