Attention, Perception, & Psychophysics
2010, 72 (1), 100-113
doi:10.3758/APP.72.1.100

Visual orienting in dynamic broadband (1/*f*) noise sequences

Christoph Rasche and Karl R. Gegenfurtner
Justus-Liebig-Universität, Giessen, Germany

Visual orienting has typically been characterized using simple displays—for example, displays with a static target placed on a homogeneous background. In the present study, visual orienting was investigated using a dynamic broadband (1/*f*) noise display that should mimic a more naturalistic setting and that should allow saccadic orienting experiments to be performed with fewer constraints. In Experiment 1, it was shown that the noise movie contains gaze-attracting features that are almost as distinct as the ones measured for (static) real-word scenes. The movie can therefore serve as a strong distractor. In Experiment 2, observers carried out a luminance target search that showed that saccadic amplitude errors were substantially higher (18%) than the ones measured in simple displays. That error is certainly one of the primary factors making gaze-fixation prediction in complex scenes difficult. Supplemental figures for this study may be downloaded from http://app.psychonomic-journals.org/content/supplemental.

Visual orienting has predominantly been investigated on homogeneous backgrounds with well-defined targets (see, e.g., Findlay & Gilchrist, 2003; Roos, Calandrini, & Carpenter, 2008); for example, a saccade is made from a central fixation cross toward a peripheral target of sufficiently high contrast. Such simple, highly controlled conditions are suitable to accurately determine parameters and dependencies such as saccadic latency toward a target, the dependence of latency on target parameters, target detectability in the periphery, and so on. How those parameters and dependencies change when more complex visual input is presented has been investigated only marginally. Analogously, Roos et al. even made a distinction between evoked saccades, the ones that are triggered in laboratory conditions, and spontaneous saccades, the ones that are triggered when freely viewing real images. Their analysis concentrated primarily on intersaccadic intervals. The orienting experiments mentioned so far were carried out with a "static" observer. Visual orienting is also investigated with "dynamic" observers to determine visual selection during daily activities such as tennis playing or walking through the forest (e.g., Einhäuser, Rutishauser, & Koch, 2008; Land, Mennie, & Rusted, 1999). Such experiments are beneficial for gaining a more detailed understanding of (spontaneous) saccadic target selection during actions, but they do not allow for the control of the visual environment to investigate precise orienting behavior in response to targets—for instance, how accurately a target is foveated during search. Our approach to the analysis of visual orienting aimed between the static and dynamic conditions: On the one hand, we desired control over the stimulus, and for that, the observer was preferably static; on the other hand, we wanted a stimulus that could possibly mimic the same degree of attraction (or distraction, depending on task) as that experienced during "free moving"—a simulation of the dynamic condition. For this purpose, we generated a movie based on dynamic broadband noise, which was displayed on a typical experimental CRT monitor. The broadband noise had an inverse frequency amplitude spectrum (1/*f*), similar to the amplitude spectrum of natural images (Field, 1987). This allowed us to investigate saccadic orienting properties under close-to-naturalistic conditions.

Our first goal was to characterize visual selection in those noise movies—specifically, to elucidate whether, and to what degree, the dynamic 1/*f* movie stimulus provided preferred features of fixation selection. From scene-fixation studies, it is known that there exists a preference for high contrasts (see, e.g., Parkhurst & Niebur, 2003; Reinagel & Zador, 1999; Tatler, Baddeley, & Vincent, 2006). One may therefore wonder whether this contrast preference also exists in a mere 1/*f* display (in our case, a movie). To investigate this question, observers merely performed free viewing in Experiment 1. In analyzing the image properties of the fixation locations, we found that some of them were indeed similar to the image properties of fixation locations of natural scenes. From this finding, we concluded that the movie provided a good "natural" attractor. It made sense, therefore, to investigate orienting behavior in those movies by carrying out a luminance search task, which we did in a second experiment. It turns out that saccadic landing precision decreased, as compared with measurements that were made with evoked saccades and reported in other studies.

C. Rasche, rasche15@gmail.com

The use of 1/*f* noise has become increasingly popular over the years because natural images show the same amplitude spectrum (Field, 1987; Simoncelli & Olshausen, 2001). For instance, Geisler, Perry, and Najemnik (2006) investigated the search behavior for Gabor targets placed in a static 1/*f* noise background using gaze-contingent methodology. They found that both the search time and the number of fixations increase with spatial frequency and noise contrast. Our present experiment, rather, addressed saccadic orienting parameters, such as saccadic latency, constant (mean) error, and variable error.

Cormack and collaborators investigated the detection and identification of shapes. For instance, using the method of classification images, Rajashekar, Bovik, and Cormack (2006) found that saccadic target selection was based on structural cues, despite the noisiness of the 1/*f* displays. Tavassoli, van der Linde, Bovik, and Cormack (2007) continued that line of research, but additionally proposed a variant of the classification image paradigm, with which one can obtain useful classification images in as few as 200 trials. Their method consisted of dividing the display into tiles. Although we used the method of classification images as well, their tiling method was not straightforward to apply, since we used targets that could appear anywhere in the display and not only at selected locations. 1/*f* displays were also used by White, Stritzke, and Gegenfurtner (2008) to determine the latencies of evoked saccades. Employing a gap paradigm, their measurements showed that latencies toward Gabor patches were smaller in 1/*f* displays than on a homogeneous background lacking any structure.

In the present study, we extended the use of 1/*f* noise to dynamic displays in order to create a moving input that potentially was as attracting as that perceived during natural activities. The noise movie was created from a two-dimensional 1/*f* image (Figure 1)—the same type of image as has been used in other studies (e.g., Geisler et al., 2006; Rajashekar et al., 2006). In those studies, the image was presented statically, but in our present study, it was used as the source for a one-dimensional movie by iterating row-wise through the source image: Each row was stretched vertically in space and presented as a frame (Figure 2); the movie thus appeared as a flickering bar code. The choice of spatial one-dimensionality was done for technical simplicity, but is justifiable by the fact that a large number of saccades are made along the cardinal axes (Einhäuser et al., 2007).

EXPERIMENT 1
Free Viewing

Observers were asked to freely view the noise movie—that is, to view it without any specific instruction. They did so naturally on the horizontal axis only (because this is where the dynamic stimulus appeared). An example of a scan path of the horizontal eye position was plotted into the source image (Figure 1); vertical lines represented fixation periods, horizontal lines described saccadic jumps. To determine whether gaze was drawn toward specific locations/features of the stimulus, we compared the fixated movie patches with a set of randomly collected movie patches. This analysis was done analogously to the analysis carried out for scene-fixation studies, in which a gaze prediction for real-world scenes is sought (see, e.g., Carmi & Itti, 2006; Parkhurst & Niebur, 2003; Reinagel & Zador, 1999; Tatler et al., 2006; Vig, Dorr, & Barth, 2009). They typically used (static) images of real-world scenes while observers performed free viewing. The fixated image patches were then compared with a set of randomly collected image patches in an effort to characterize the bottom-up component of saccadic target selection in natural scenes. In our present study, the fixated image patches corresponded to spatiotemporal image patches before saccadic landing. Specifically, for a given fixated spatial location, its preceding space–time patch was extracted (see the white rectangles in Figure 1). This space–time patch presumably has triggered the visual system to place its gaze on it. It is therefore called a "trigger patch."

To investigate whether there exists a feature in a trigger patch, a first simple step was to look at the average of the collected trigger patches—a method known as the *classification image* (Ahumada, 2002; Shimozaki, Chen, Abbey, & Eckstein, 2007; Tavassoli et al., 2007). The method is typically applied to image patches of "false alarms" in experimental tasks containing the four response outcomes—hits, misses, false alarms, and correct rejections—or by linear combination over classification images for all response outcomes. The classification images typically reveal aspects of stimulus representation and recognition mechanisms when Gaussian noise is used, but recently the methodology has been applied to 1/*f* noise (see, e.g., Geisler et al., 2006; Rajashekar et al., 2006). In our present study, saccadic landing positions can be loosely regarded as false alarms, because no actual target has been placed at those locations, and the classification image may therefore contain hints about the presence of a possible trigger feature for saccadic target selection in 1/*f* displays. In our present study, the classification image corresponded to the average of the space–time trigger patches, whose temporal cross section could also give us clues about the saccadic decision dynamics.

To characterize the degree of distinctness of the trigger patches, we applied Tatler, Baddeley, and Gilchrist's (2005) and Tatler et al.'s (2006) methods that were developed for a discrimination of fixation and nonfixation patches. The methods determine the receiver-operating characteristic (ROC) area value in a very simple manner, and we refer to this as the "fixation prediction."

Method

Observers. Two male and 3 female students (age = 23–30 years) served as observers. All of the observers had normal or corrected-to-normal vision. All of the observers were naive about the aim of the experiment.

Equipment. Observers were seated in a dimly lit room and faced a 21-in. CRT monitor (ELO Touchsystems, Fremont, CA) that was driven by an ASUS V8170 (Geforce 4MX 440) graphics board with a refresh rate of 100 Hz, noninterlaced. At a viewing distance of 47 cm, the active screen area subtended 45º × 36º of visual angle in the horizontal and vertical directions, respectively. With a spatial resolution of 1,280 × 1,024 pixels, this resulted in 28 pixels/deg.

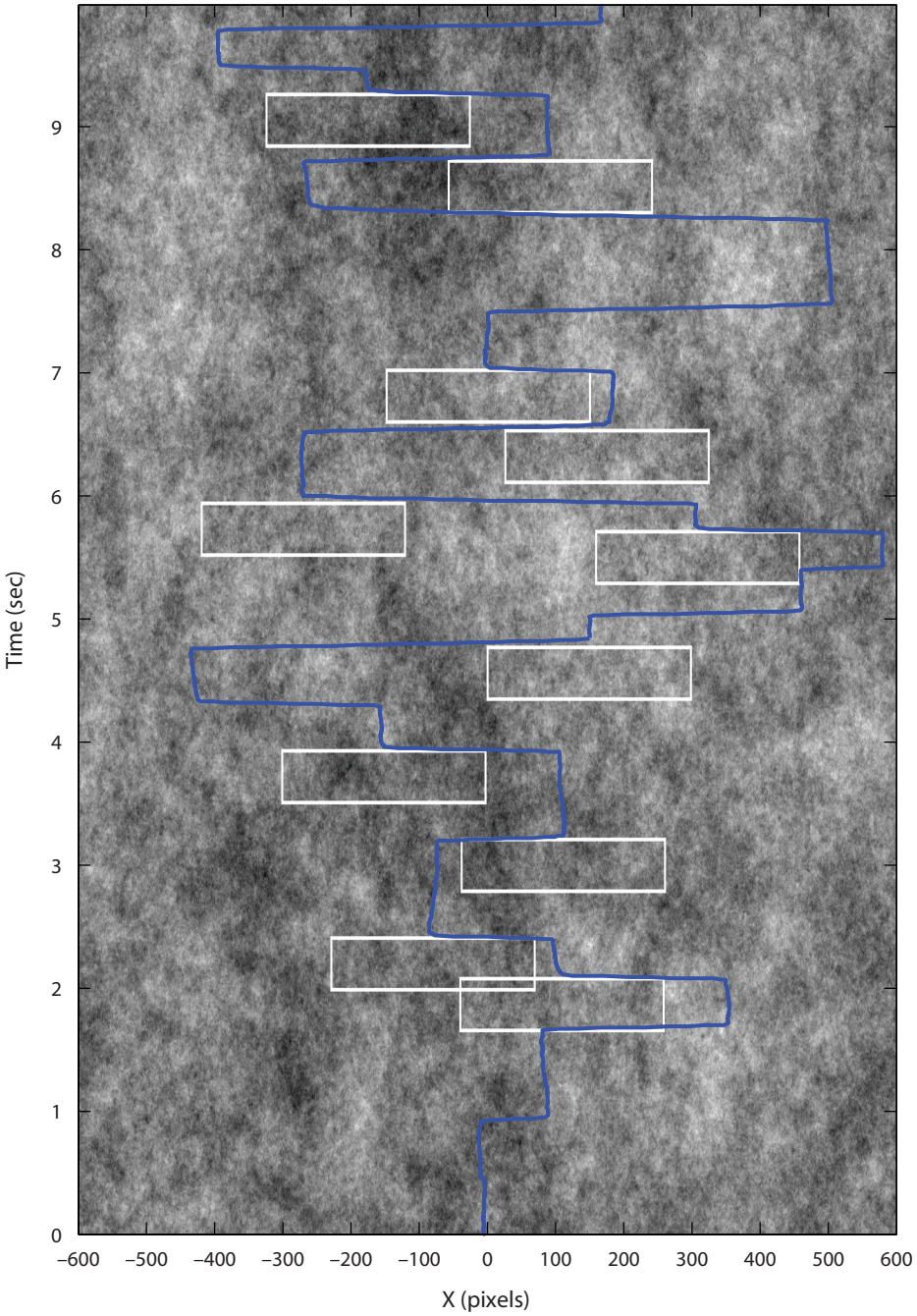

Figure 1. Image source for a noise movie, with an example of a horizontal eye-position scan path (black; online version, blue) overlaid (fixations run vertically; saccades are the horizontal jumps). The source is a two-dimensional 1/*f* image, $I(x,t)$, and represents a space–time plot (width = 1,200 pixels = 42.3°; 1° = 28.4 pixels). Each row is used as a source for a movie frame (10 msec) and is stretched vertically (see Figure 2). The white rectangles outline the selected trigger patches [$P_i(x,t)$; shown schematically only] centered around saccadic landing position (trigger patches close to the image border were omitted).

The observer's head was stabilized in place using a chinrest. Eye-position signals were recorded with a head-mounted, video-based eyetracker (EyeLink II; SR Research Ltd., Osgoode, Ontario, Canada) and were sampled at 250 Hz. Observers viewed the display binocularly through natural pupils. The stimulus display and data collection were controlled by a PC using modified C routines from SR Research Ltd. and SDL routines.

Stimulus. The two-dimensional 1/*f* image, $I(x,t)$, was generated using a 2-D image of normally distributed random pixel-intensity values, whose frequency spectrum was then transformed to describe

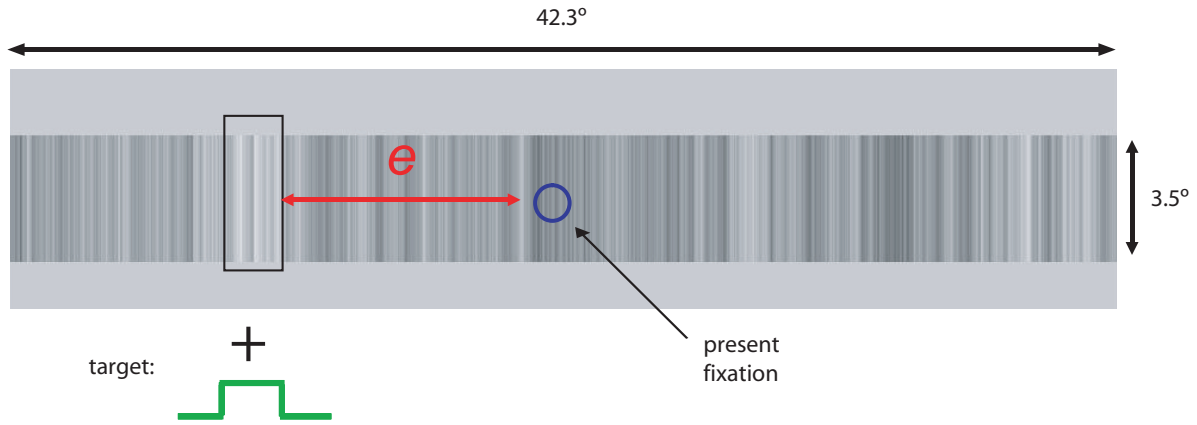

Figure 2. A sample frame of the noise movie and target generation. The "bar code" represents one frame (duration = 10 msec) whose luminance distribution is taken row-wise from the two-dimensional 1/*f* image, as in Figure 1. In a luminance-target search (Experiment 2), a target was generated by adding a rectangular function (gray; online version, green) to the luminance profile of the bar code. This was done for a fixed pulse amplitude (fixed-amplitude condition), as well as for a scaled pulse amplitude, which depended on fixation eccentricity (scaled-amplitude condition) (circle: eccentricity at present point in time).

a 1/*f* frequency decline. The image size was 1,000 × 1,200 pixels, which defined the temporal and spatial characteristics of the stimuli, respectively. Each row was the source for a single frame: The row was stretched vertically to a height of 100 pixels and was placed into a gray background that was presented as 8-bit luminance resolution (40 cd/m² luminance). A frame was shown for 10 msec; a movie thus lasted 10 msec × 1,000 (pixels) = 10 sec and constituted one trial. Each movie, $I(x,t)$, was different to avoid potential learning effects.

Procedure. Observers performed blocks of 50 trials, generally three blocks per day and six blocks per experiment. Each block was preceded by a nine-point calibration and validation. Before each trial, observers fixated a central spot, and a drift correction was carried out to minimize errors of headband slippage or other factors. The average spatial measurement accuracy was 0.28°±0.55°, which was as accurate as that in other studies (e.g., 0.4°±0.1° in Tatler et al., 2005). Asking observers to perform free viewing is not as easy as it seems, because some observers immediately ask what they are supposed to look for. To obtain "unbiased" viewing, observers were simply asked to "be inspired." This unusual instruction was given to avoid any mentioning of the overall goal of the experiments. Observers who participated in multiple experiments started with the free-viewing task.

The classification image was generated for trigger patches of sufficiently large size to understand the dynamics. A trigger patch was centered spatially on the point of saccadic landing and temporally at the time of saccadic onset. Trigger patches exceeding the image border were omitted in addition to the very first one after a trial was started. Trigger patches were aligned according to the direction of saccadic flight: Patches resulting from a leftward flight were mirrored along the spatial dimension. An equal number of random patches were collected for statistical comparison, which was chosen as follows: Given the spatial location of a trigger patch, a corresponding random patch was taken whose spatial location was mirrored along the spatial axis with the half image width as the line of reflection. In other words, the random patches were chosen from the scan path, which was flipped along the spatial axis. In total, an average of 2,990 patches was collected per person. In addition to the mean, the variance of the trigger patches was computed.

Fixation prediction. To determine the degree of distinctness of the trigger patches, the average luminance values of the individual patches were compared with the average values of the random patches. Specifically, a decision threshold was slid through the two luminance distributions, and true and false positives were determined; from these, the ROC area value was calculated (see Tatler et al.'s [2005] and Tatler et al.'s [2006] methods for an elaborate justification). The average luminance value for an individual patch was taken from a small area that was centered around the minimum of the classification image—for instance, 1° wide and 60 msec long.

Results

Because the use of our movie sequences was novel, we ensured that saccadic search behavior roughly corresponded to regular search behavior. Histograms for fixation durations showed a peak at around 320 msec (on average) and a 1/time function decay (see Supplemental Figure 1); the ones for saccadic amplitudes showed a peak between 2° and 8°, with long-tailed distributions for the majority of observers.

The classification image was generated for patches starting from several hundreds of milliseconds before saccadic onset and extending some hundreds of milliseconds after onset in order to observe the entire dynamics (the patches in Figure 1 are smaller for purposes of illustration). Figure 3 is an example for 1 observer. It shows a depression (trough) in the luminance level. To highlight the area that was significantly different from chance, those values were set to the mean grayscale value (0.5) and were within the range of values given by the classification image for the random patches (see Figure 4 for a cross section). This highlighted area is now called a salient area. The salient area's spatial minimum is located 0.61° on average beyond the landing point, which was set at 0° eccentricity (SD = 0.67°; 1 person showed overshoot, not shown); the minimum occurs approximately 130 msec before saccadic onset. The width of the salient area extends approximately 5°–7° at the time of the minimum; its duration lasts more than 1 sec before saccadic onset. This large size is primarily due to the 1/*f* correlation along both dimensions (which emphasizes slowly changing low frequencies). The classification image of 4 other observers is shown in Figure 4, left column: They all clearly show a

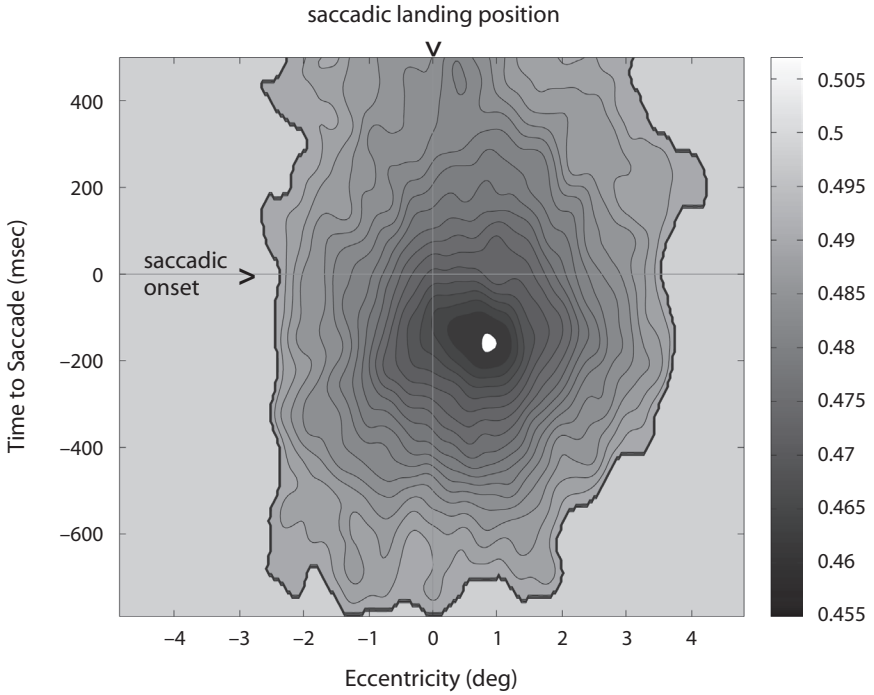

Figure 3. The classification image for 1 observer for the free-viewing experiment (3,823 fixations), the average of the white-outlined trigger patches in Figure 1. The white dot represents the minimum luminance value of the entire field. *y*-axis, countdown to saccadic onset, in milliseconds (0 = saccadic onset); *x*-axis, eccentricity relative to saccadic landing position, in degrees (0 = saccadic landing position; minus and positive values represent less and more eccentricity, respectively). The total luminance level ranges from 0 to 1. Values that were within the range of the values from the random classification image (nontrigger) patches were set to 0.5.

large depression of similar size with a minimum at about the same spatial and temporal locations. The right column displays the variance of the trigger patches. Only small salient areas could be found, but they were not consistent across observers.

The analysis of fixation prediction yielded ROC area values ranging from 0.54 to 0.60 for different observers ($M = 0.56$).

Discussion

The classification image of trigger patches clearly shows the presence of a large depression for all observers tested and distinguishes itself clearly from the random background. The depression indicates that the visual system preferred to direct its gaze toward a static, looming dark spot in this type of movie stimulus.

The classification image reveals two more aspects of the spot. (1) The location of the spatial minimum—1º beyond saccadic landing—can be interpreted in two ways: The visual system selected a location of large contrast (e.g., the edge of a dark vertical bar), or it selects the center of a dark spot and lands just short of it. Thus, there is the possibility that the bare $1/f$ "background" structure of a visual image may contain gaze-triggering features. (2) The temporal minimum occurred at approximately 130 msec before saccadic onset, which indicates that the saccadic system decided to jump when the dark spot started to disappear, given a saccadic decision time of 75 to 100 msec (see, e.g., Caspi, Beutter, & Eckstein, 2004; Nazir & Jacobs, 1991).

With the fixation prediction analysis, we determined that this gaze-attracting feature can be distinguished from a random background stimulus by an ROC area value of up to 0.60 ($M = 0.56$). To ensure that the values were approximately comparable with those from other methods, we used the reverse-correlation method that was developed for the analysis of visual neurons (Simoncelli, Paninski, Pillow, & Schwartz, 2004), as well as the support-vector machine classification method by Kienzle, Franz, Schölkopf, and Wichmann (2009). With all of these methods, similar ROC area values were obtained that hinted that the values were independent of the method. In the next section, it will be reported that the "distinctness" of false alarms for a luminance target was of similar magnitude. Thus, even if the movie consisted of only $1/f$ noise, it may have provided the same degree of attraction experienced during natural input.

EXPERIMENT 2
Target Search

In Experiment 2, a temporarily appearing luminance target was inserted into the dynamic noise movie. The target search allows one to determine orienting properties,

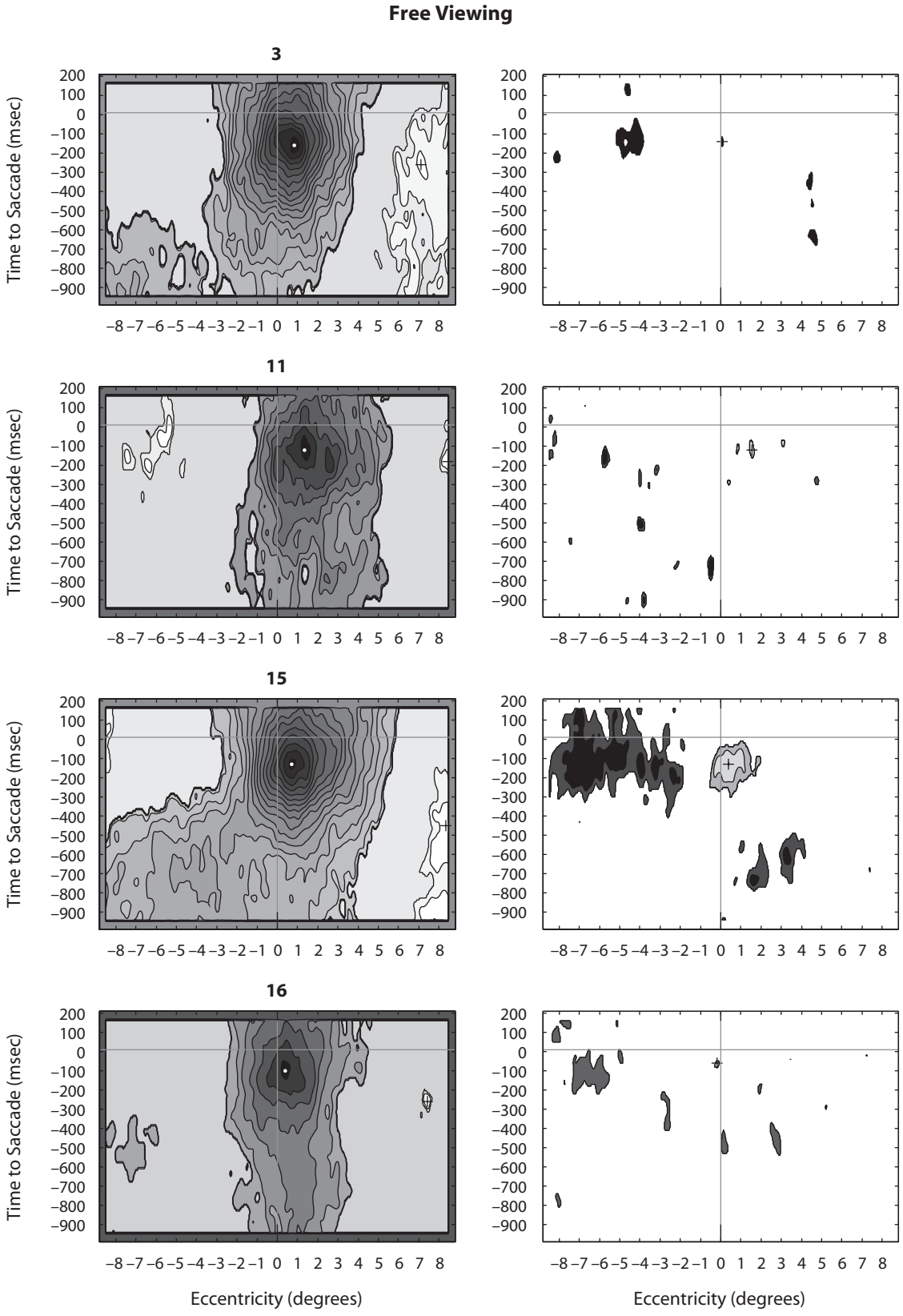

Figure 4. Classification images of trigger patches for the free-viewing experiment for 4 observers (N = 2,839, 2,823, 4,527, and 5,037 fixations per observer, respectively). Left column: Mean of trigger patches. The white dot represents the minimum (observer identification number given above each graph on the left). Right column: The variance of trigger patches. Axes are the same as those in Figure 3.

such as saccadic latency and saccadic precision, which then are compared with measurements that are reported in studies with evoked saccades under single static stimulus conditions. It was assumed that, in particular, saccadic variability would increase because of the presence of the dynamic background noise, which was shown to be a strong gaze attractor in Experiment 1 and which now served as a strong distractor.

The target was a rectangular function added to the luminance profile and was therefore not constant in amplitude (see Figure 2). The target was shown for a few hundred milliseconds and so appears as a bright bar in the display. It appeared only occasionally—approximately three times a trial. The target was of a very low contrast and was initially difficult to detect; therefore, it could be considered "just noticeable."

In preparation for a cuing experiment (to be reported elsewhere), we attempted to create a luminance target that was equally visible in the periphery. To compensate for the peripheral decline in visual acuity, we increased the target amplitude with increasing target eccentricity in a gaze-contingent manner. This is called the "scaled-amplitude" target condition, whereas the measurements made with unmodified target amplitude are called the "fixed-target amplitude" condition ("fixed" means that the added rectangular function was fixed).

Method

Observers. Two male and 4 female students (age = 23–30 years) served as observers. All of the observers had normal or corrected-to-normal vision. All of the observers were naive about the aim of the experiment.

Target. The movie stimuli were the same as those for the free-viewing condition. Targets were added as a rectangular function to the luminance profile of the source image with an increment of value $a_{trg} = 0.2$ (total intensity range from 0 to 1), for a duration of $d =$ 300 msec (30 frames), and a spatial width of 1º at the same location (see Figure 2). Targets were presented spatially and temporally randomly, with an average frequency of 0.333 Hz. This constituted the fixed-amplitude condition.

To generate the scale-amplitude target, it required a function that would compensate for the decline in visual acuity. The decline has been described by various functions—for example, a logarithmic decay (Schwartz, 1980), or a one-over eccentricity function (Wilson, Levi, Maffei, Rovamo, & Devalois, 1990); see also Rovamo and Virsu (1979). We therefore chose an exponentially saturating function in which the target amplitude, a_{trg}, depended on eccentricity e with reference to the present eye position by $a_{trg} = a_{min} + a_{max} - \exp(-e)a_{max}$, whereby a_{min} is a minimal amplitude and a_{max} is a maximal amplitude; the function starts at a_{min} and saturates at $a_{min} + a_{max}$. The parameter values were $a_{min} = 0.2$, matching the amplitude of the fixed-amplitude condition, and $a_{max} = 0.5$, chosen heuristically after a few initial tests.

Analysis of saccadic orienting properties. Observers were instructed to move their gazes toward the targets and to press a button when seeing a conspicuity. Target detection was defined as the temporal coincidence of a "saccadic hit" and a buttonpress. A saccadic hit required a saccadic flight toward the target and a spatial landing within 5º of target eccentricity. The temporal tolerances for saccadic latency and button response were 400 and 1,200 msec, respectively. Each search condition was carried out by 4 to 5 persons. An observer typically performed six blocks of one experiment, during which the target was presented 850–900 times (approximatley three target presentations/trial).

Classification image. Trigger patches resulting from saccades made toward a target were omitted. Although the number of targets was small (~900) in comparision with the number of saccades, there were more than 2,000 patches, on average, for each person.

Results

Observers were not given any specific details about the target and were asked to detect the target by themselves. It would take the observer a few trials (max. ~20) until he or she had discovered the conspicuity on the first block. Over time, the observer developed a clear understanding about the properties of the target, meaning that all of the observers realized that it was a temporarily appearing, vertically oriented, bright bar, as indicated by the classification image.

Fixation durations and saccadic amplitudes were slightly lower than in Experiment 1 (decrease of 40 msec and 1º, respectively; see Supplemental Figure 2). This difference is equally as minor as that measured in another study for static images, which compared search versus free-viewing behavior (Tatler et al., 2006). (Because of the same small difference, we considered the free-viewing behavior in our noise movies—as was carried out in Experiment 1—as analogous to free-viewing behavior in static scenes.)

The classification images of the trigger patches show a salient area, which can be described as a mound or elevation (Figure 5). The salient area was of similar size to the salient area of the free-viewing experiment (compare left columns of Figures 4 and 5). The spatial location of the elevation's maximum was at 1.34º, on average ($SD = 1.07$º), and it roughly corresponded to the minimum of the free viewing's salient area: The spatial minima and maxima were statistically not different. Some salient areas also contained large, flanking depressions (e.g., Observers 9 and 12), which probably reflect individual differences in target representation and were partly caused by the spatial $1/f$ correlation or luminance values.

We can discern more differences between the two types of salient areas by looking at the temporal cross section of the classification images. Figure 6 shows the one for free viewing and the one for the luminance target search, averaged across observers. The one for free viewing was inverted for comparison. The cross section for the target search shows that integration occurred later and at a higher amplitude, reflecting that observers searched for rather abrupt onsets. In comparison, the cross section for free viewing shows that a long-lasting, gradual integration took place: Observers apparently followed the slow dynamics of the dark spots.

The fixation prediction analysis for trigger and random patches (omitting the few patches with targets) yielded ROC values ranging from 0.53 to 0.62 ($M = 0.57$), which were approximately equal to those measured for the free-viewing trials ($M = 0.56$).

Orienting properties. Figure 7 shows the manual reaction times (RTs), saccadic latencies, and detection rates as a function of eccentricity, averaged across observers for the fixed-amplitude and scaled-amplitude target conditions (dashed–dotted and solid lines, respectively). The manual RT for the fixed-amplitude condition increased

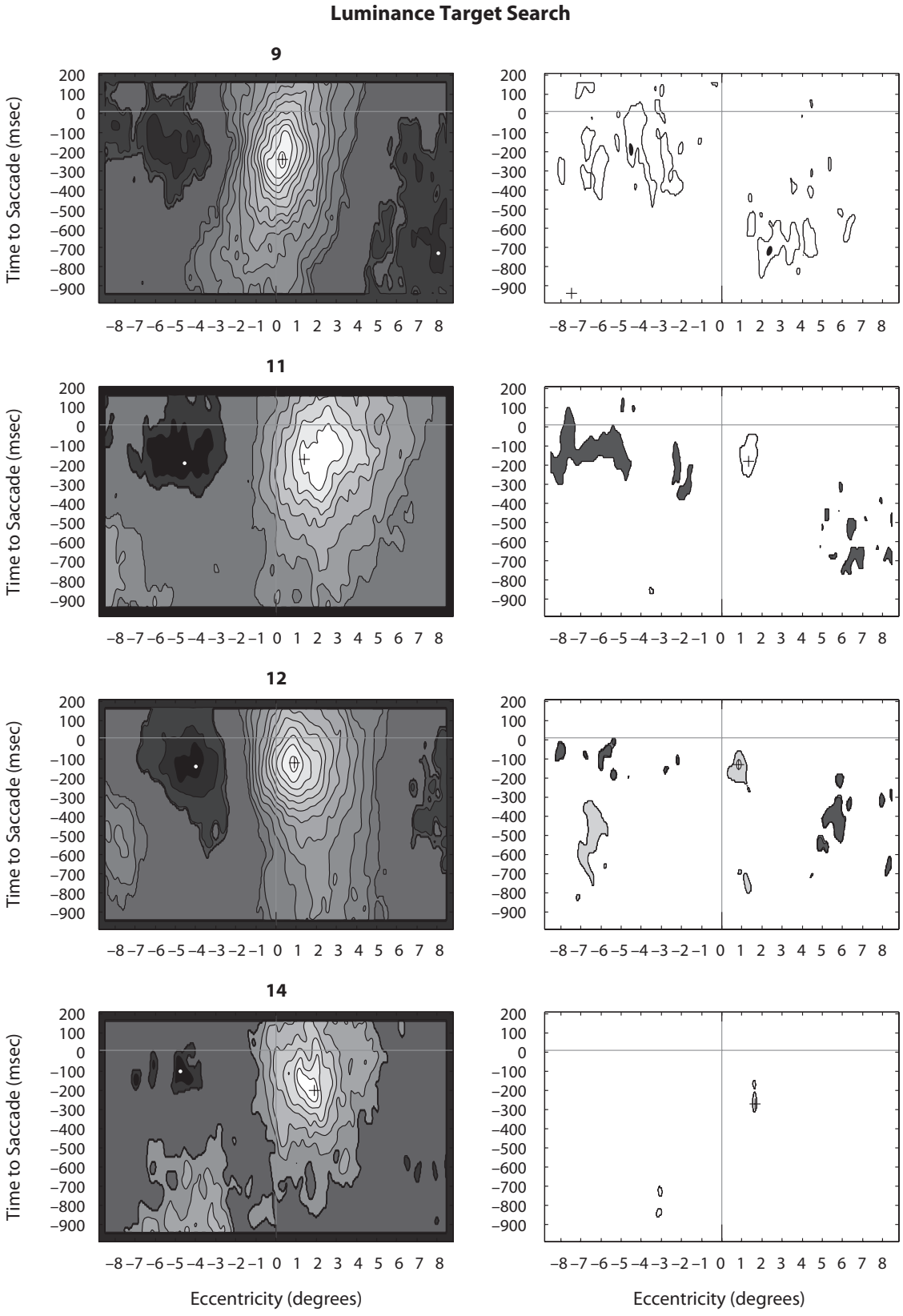

Figure 5. Classification images of trigger patches for a luminance target search. Left column: Luminance average. The plus sign denotes the location of the maximum. Right column: The variance of the trigger patches. Labeling is the same as in previous figures.

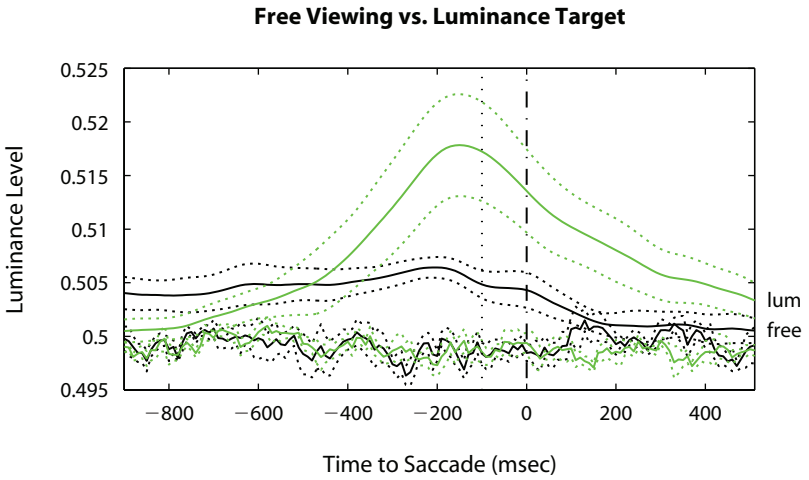

Figure 6. Temporal cross sections of the observer-averaged classification image (4–5 observers). *x*-axis, countdown to saccade, in milliseconds (0 = saccadic onset); *y*-axis, luminance level (total range [0,1]); free, free viewing (black); lum, luminance target search (gray; online version, green). The free-viewing values were inverted for reason of comparison. The functions near 0.5 are from the random patches. The dotted curves denote standard errors of interobserver performance. The vertically dotted straight line is placed at −100 msec.

from approximately 500 msec to approximately 550 msec for 1º to 15º eccentricity (top graph, black; $p < .05$ for a t test that compared ranges from 1º to 10º and from 11º to 20º). For the scaled-amplitude condition—the gaze-contingent increment in target amplitude—RTs remained roughly constant across eccentricity. Saccadic latencies, in contrast, decreased initially for eccentricities up to about 12º, from approximately 240 msec down to 200 msec (top graph); then, they remained constant at around 210 msec. For the scaled-amplitude condition, latencies were slightly lower by about 10–30 msec.

The detection rate for fixed-amplitude targets (lower graph, Figure 7) clearly decreased for large eccentricities as was expected, given the peripheral decline in visual acuity, from a value of approximately 0.5 at 4º to less than 0.3 for 20º and more. The low detection rate for small saccadic amplitudes (1º–3º) was due to the difficulty to determine a saccadic jump toward the target, given the noisiness of the background, and was probably also due to the observer's low propensity to saccade near targets. The detection rate for buttonpresses as a function of eccentricity was roughly constant at a level of approximately 0.6.

The detection rate for scaled-amplitude targets increased from 4º to 8º eccentricity and saturated at a value of 0.6. It then slowly declined for large eccentricities. An eccentricity-dependent saccadic hit tolerance was tested as well (as opposed to the fixed 5º tolerance), but this resulted only in a scaling of the curve's amplitude.

Figure 8 shows the landing precision of saccades for the fixed-amplitude condition. The upper left plot depicts the amount of undershoot (mean saccadic amplitude; constant error) as a function of target eccentricity for the first saccade. The landing variability (variable error) was expressed as the standard deviation of the saccadic amplitudes (upper right plot). All three relationships could be fit with a linear equation and yielded significant results ($p < .05$). Expressed differently, the average saccadic amplitude toward a target undershot by 18%, and its landing position varied with approximately 10% of target eccentricity. A corrective saccade occurred with a probability of 10%, and was preferably triggered for target eccentricities in a range of 10º to 20º (lower left plot). The second (corrective) saccade still showed some undershoot (~5%; see lower right plot).

For the scaled-amplitude condition, the relationships for landing precision showed surprisingly no notable differences (Figure 9). However, the proportion of second saccades seemed a bit altered, with prominently fewer saccades in the range of 1º to 10º and a sharp incline around 11º.

Discussion

At a first glance, the classification image for the false alarms of the luminance target looked like a mere inversion of the classification image for the free-viewing task. But the temporal cross section of the classification images revealed that the visual system followed slower changing dynamics when detecting dark spots during free viewing as compared with detecting the sudden onset of a luminance target.

The fixation prediction analysis returned an ROC value of 0.57 on average, which was equally high as the discrimination for the patches obtained from the free-viewing task. This was surprising, because we expected that they would be higher, given that observers had a specific search target.

The detection of salient targets in noisy one-dimensional movies has been investigated by Neri and Heeger (2002). Their observers were asked to foveally detect and identify a target in a short movie sequence (nine frames, total trial

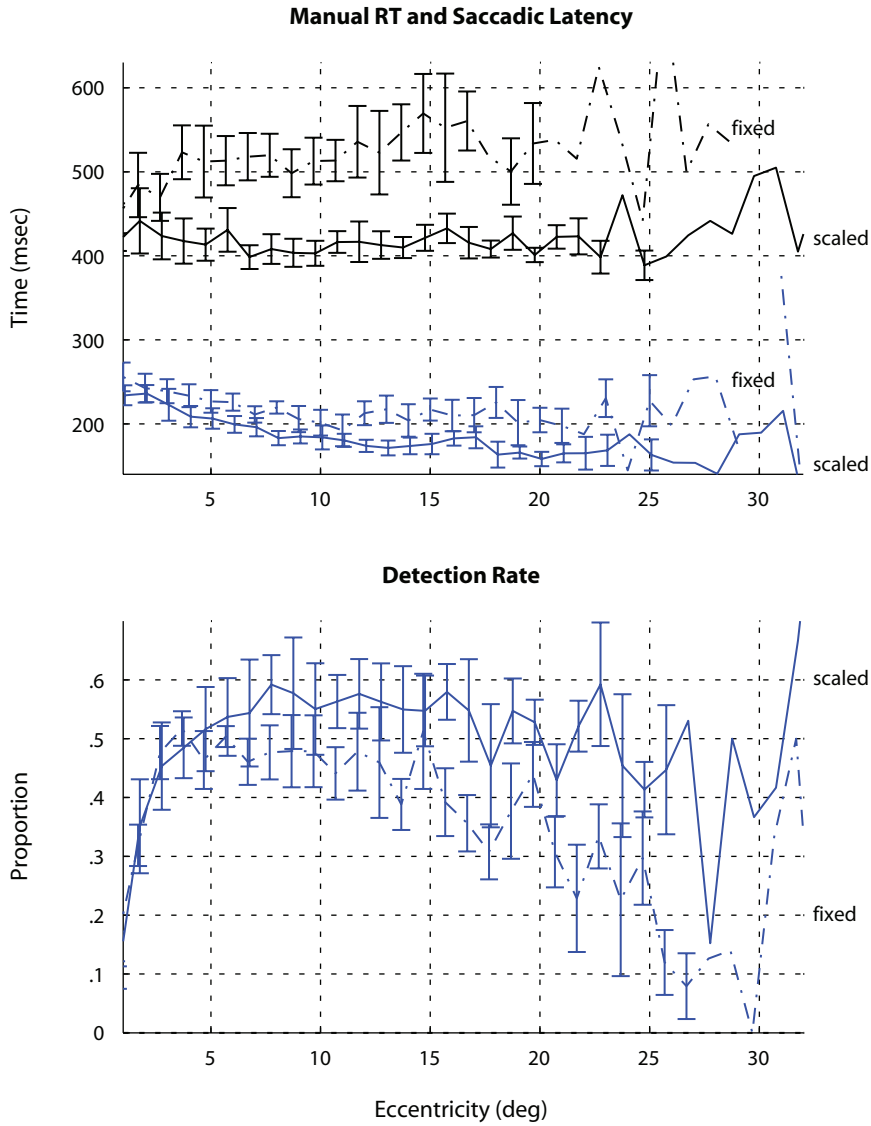

Figure 7. Target search for fixed-amplitude (dashed–dotted) and scaled-amplitude (solid) target conditions. Top plot: Manual reaction times (RTs; upper two traces) and saccadic latencies (lower two traces) in dependence of target eccentricity. Bottom plot: Detection rate for targets that captured gaze and were signaled by a buttonpress. Error bars = standard errors of interobserver performance (lacking error bars = 1 observer only).

duration = 243 msec); their choice of noise was Gaussian. Their primary finding was that this discrimination process consisted of a "detection" and an "identification" stage, in which the former was indicated by a high level of variance preceding the identification stage by approximately 100 msec. Translated to our experiments, this would mean that we should have found a saliency field in the variance plots 100 msec preceding saccadic landing, but no such high levels were detected (right columns in Figures 4 and 5). The lack of such variance may have been due to the many methodological differences between our study and theirs, but the primary reason may have been that our results were based on saccadic decisions and theirs on perceptual decisions.

Orienting properties. The scaled-amplitude condition yielded shorter RTs for both the manual response and the saccadic latency: The manual RTs for the fixed-amplitude condition were slightly increasing across eccentricities, but seemed to decrease for the scaled-amplitude condition. The saccadic latencies for the scaled-amplitude condition seemed to be merely downscaled.

We further compared the saccadic latencies with those measured for evoked saccades in a study by Kalesnykas and Hallett (1994). They determined latencies using gaze shifts starting from the display center and moving to the periphery. Targets simply turned on (step function) and were displayed on a blank background. For eccentricities of 1° to 3°, the latencies slightly decreased, remained constant from

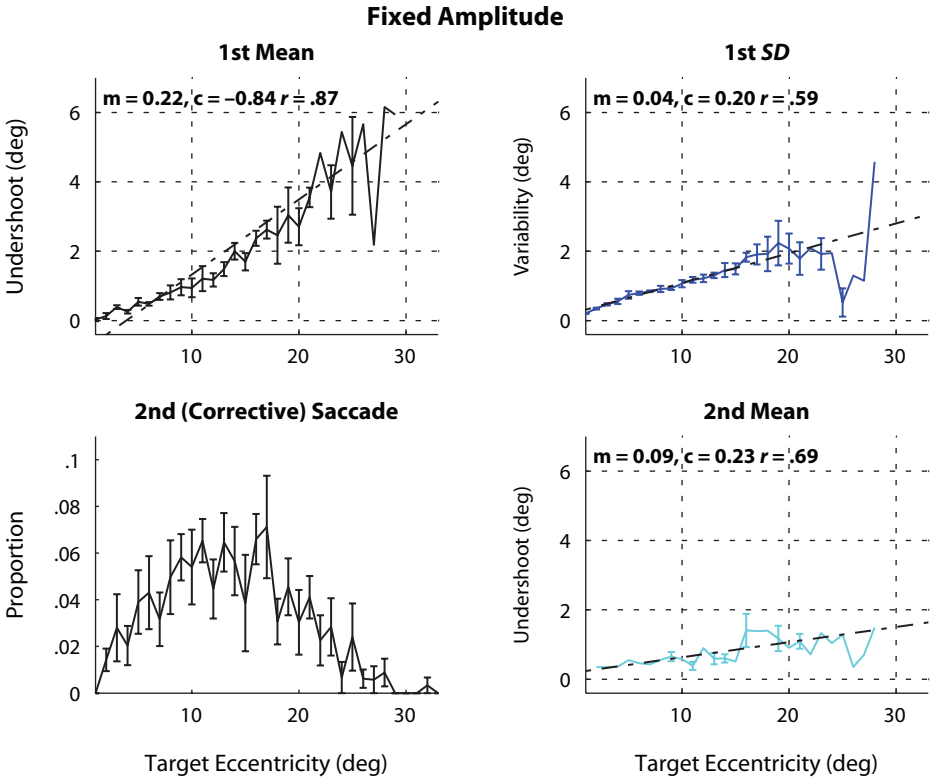

Figure 8. Saccadic landing precision in dependence of target eccentricity (for the fixed-amplitude condition). Upper left: Average undershoot (constant error) for primary saccade [m = slope; c = intercept; *r* = correlation coefficient]. Upper right: Variability (variable error [standard deviation]) for primary saccades. Lower right: Average undershoot for secondary saccades. Lower left: Distribution of corrective (second) saccades (~10% in total). Error bars = standard errors of inter-observer performance (lacking error bars = 1 observer only).

approximately 4º to 12º, and then started to increase for larger eccentricities. The function was therefore given the name "bowl-shaped function." In comparison, in our present study, saccades were made from arbitrary (horizontal) gaze positions on the display with distracting background motion. The latencies for the fixed-scale condition look like a smeared or dampened version of the bowl function (dotted line, lower section of upper graph, Figure 7). Latencies gradually decline for the first 7º to 10º and then remain steady up to approximately 20º. Then, they show signs of increase, although this cannot be statistically tested because of the lack of data points. For the scaled-amplitude targets, the initial decrease seems to extend to even 13º. Thus, the bowl-shaped function seems to hold for more natural viewing conditions, but has broadened. Although there have been attempts to disprove the presence of a constant-latency range—the flat bottom of a bowl function (see, e.g., Hodgson, 2002)—those data were not carried out for a broad range of eccentricities, which makes it difficult to judge about the overall shape of the function.

Kalesnykas and Hallett (1994) also noted a slight deviation for landing precision for eccentricities of up to 6º (see their Figure 10B)—specifically, a (constant) error of 8.3% (undershoot of 0.5º for a target eccentricity of 6º). Earlier studies measured an error of 10% (Becker, 1972; Henson, 1979). The constant error in our experiments was about 18% (undershoot of 3.8º for 20º eccentricity). It is not clear, however, whether the increased variability derives from the search behavior or from the noise background, or from both.

Variability in saccadic landing position (error) was quantified by Aitsebaomo and Bedell (1992), but primarily for different target durations, showing a decrease with increasing target duration. Saccadic landing variability also grows with increasing eccentricity (Figure 8), which was also measured in van Beers's (2007, Figure 2; increasing scatter for landing positions) study. The scaled-amplitude condition yielded the same degree of variability, indicating that although target parameters do affect target detection, and hence saccadic latency, they appeared to have little effect on saccadic amplitude.

The detection rate decreased with increasing eccentricity for fixed-amplitude targets (Figure 7, bottom graph), but for the scaled-amplitude targets, the detection rate was reasonably steady in the range of 4º to 16º, proving that the choice of exponential compensation for the decline in visual acuity was a good first guess of the chosen parameters ($a_{min} = 0.2$, $a_{max} = 0.5$).

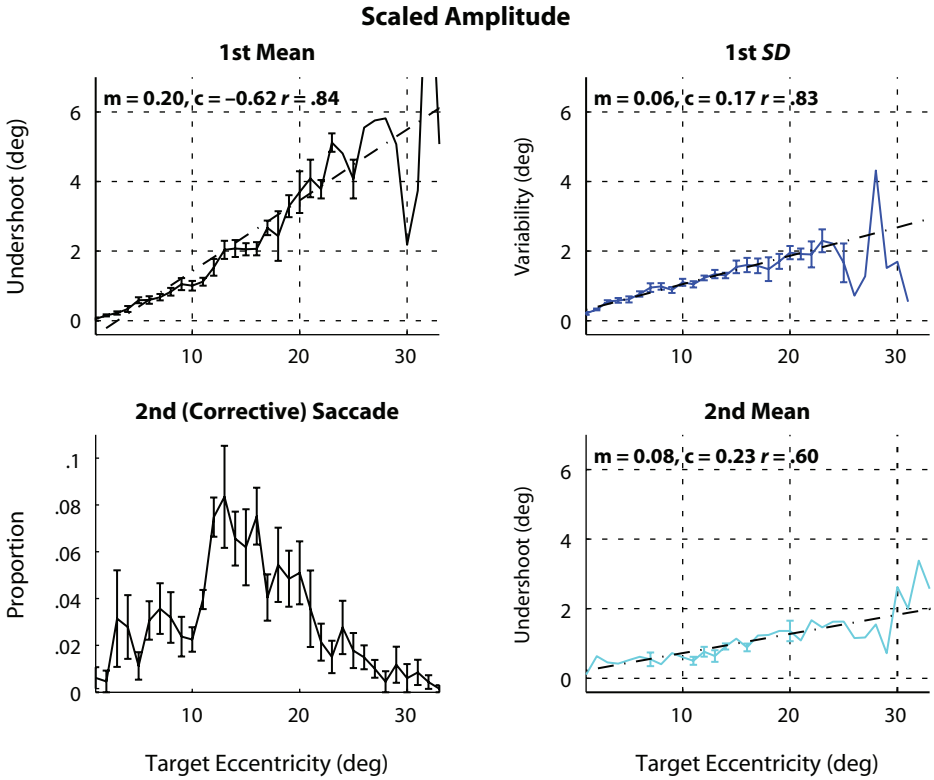

Figure 9. Saccadic landing precision in dependence of target eccentricity for the scaled-amplitude condition. Plots and labeling are the same as in Figure 8.

GENERAL DISCUSSION AND SUMMARY

In the present article, visual orienting was studied with a broadband (1/*f*) noise stimulus, which served as an intermediate form of the typical stimuli used in laboratory experiments (e.g., the gap paradigm), and the stimuli as perceived during natural activities. In Experiment 1, the free-viewing task, we demonstrated that the noise stimulus contained a strong gaze-attracting feature—namely, a static dark, looming spot (Figure 3). This was determined by using the method of classification images (Ahumada, 2002; Shimozaki et al., 2007), whose temporal cross section also revealed that the spot slowly appeared and was foveated when it started disappearing (Figure 6).

To determine the "strength" of this gaze-attracting feature, we employed a trigger/random patch classification analogously to the analysis done for scene fixations (see, e.g., Carmi & Itti, 2006; Parkhurst & Niebur, 2003; Reinagel & Zador, 1999; Tatler et al., 2006). The ROC area values ($M = 0.56$; maximum = 0.60) did not quite reach the same ROC values as did the ones obtained in scene-fixation studies—for example, 0.63 (Tatler et al., 2005) and 0.64 (Kienzle et al., 2009). However, a direct comparison between static and dynamic scenes is difficult because motion attracts gaze exceptionally well (Franconeri & Simons, 2003): It may well be that the dark spot represents just a motion stimulus; after all, it appears to be a looming stimulus.

Concluding Experiment 1, it was shown that our noise display provided a strong degree of visual attraction as possibly experienced during free viewing or moving, yet presenting it in the laboratory allowed us to place well-defined targets and to measure orienting aspects, as we had done in the Experiment 2.

In Experiment 2, a luminance target search was carried out. It revealed that the ROC area values for the target search were not higher, despite the clearer objective of the search task: Although observers had to learn the target by themselves from the first few trials, they all sensed that the target consisted of a bright, vertical bar. Thus, in this noise display, an average value of 0.57 may constitute an upper "distinctness" boundary.

We determined the detection performance for fixed-amplitude and scaled-amplitude targets, the latter being an eccentricity-dependent (gaze-contingent) increase in target amplitude to compensate for the peripheral decline in visual acuity (Figure 7, top graph). Peripheral detection performance for scaled-amplitude targets nearly reached the detection performance of fixed-amplitude targets; saccadic latencies as well as manual RTs were even slightly shorter. The amount of saccadic undershoot and variable error was roughly twice as large as that measured for simple displays (see, e.g., Findlay & Gilchrist, 2003; Roos et al., 2008). Only 10% corrective saccades were carried out, which, in turn, were preferably performed for target eccentricities ranging between 10º and 20º. We suspect that during natural activities (free moving), this variability

is even larger and that the number of corrective saccades is higher.

The presence of a constant and variable error can be interpreted as an inaccuracy or insufficiency of the saccadic spatial-orienting mechanism and, hence, a weakness, since it apparently does not allow one to precisely place the gaze on objects with the primary saccade. But for recognition of structure, this primary-saccade precision is not necessary: Scenes and letters can be recognized translation invariant to a large extent (Pelli, Burns, Farell, & Moore-Page, 2006; Rayner, 1998). If any such translation-independent "recognition" took place in our experiments, this may explain why the classification image is relatively blurred. Together with the tracker measurement error of approximately 0.5º, the measured saccadic landing precision can vary by more than 1º for a 5º saccadic jump ($0.5 + 0.18 * 5 = 1.3$), which smears out fine structure in a classification image. This is, in fact, an issue that—to our knowledge—is never really discussed in scene fixation studies. In those studies, it is implicitly assumed that a fixation has landed precisely. Although saccadic orienting is likely more precise in static scenes than in movies, targets in static scenes are typically crowded, and that may add to the uncertainty of saccadic target selection. But even if one assumed a very low error of 8% (the error for single saccades made in simple displays), this adds up to an error of at least 0.9º for a 5º amplitude. This inaccuracy certainly makes it difficult to specify what types of structures are fixated. But we believe that this orienting issue can be further discerned by systematic testing on different image types (line drawings, grayscale scenes, movies) for different types of tasks (detection, recognition).

AUTHOR NOTE

We thank the following members of the Schölkopf department at MPI Tübingen for guidance with some computational methods: Wolf Kienzle for help with the support-vector machines, Jakob Macke for guidance on the reverse correlation technique, and Matthias Franz for an overview of both techniques. We also thank Brian White and Erik Groenewold for comments on an earlier version of the manuscript, and Nadine Hartig for experimental assistance. The present work was funded by the Gaze-Based Communication Project IST-C-033816, European Commission within the Information Society Technologies. Address correspondence to C. Rasche, Abteilung Allgemeine Psychologie, Justus-Liebig-Universität, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany (e-mail: rasche15@gmail.com).

REFERENCES

AHUMADA, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, **2**, 121-131.

AITSEBAOMO, A. P., & BEDELL, H. E. (1992). Psychophysical and saccadic information about direction for briefly presented visual targets. *Vision Research*, **9**, 1729-1737.

BECKER, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca Ophthalmologica*, **82**, 233-243.

CARMI, R., & ITTI, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research*, **46**, 4333-4345.

CASPI, A., BEUTTER, B. R., & ECKSTEIN, M. P. (2004). The time course of visual information accrual guiding eye movement decisions. *Proceedings of the National Academy of Sciences*, **101**, 13086-13090.

EINHÄUSER, W., RUTISHAUSER, U., & KOCH, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision*, **8**(2, Art. 2), 1-19.

EINHÄUSER, W., SCHUMANN, F., BARDINS, S., BARTL, K., BÖNING, G., SCHNEIDER, E., & KÖNIG, P. (2007). Human eye–head co-ordination in natural exploration. *Network: Computation in Neural Systems*, **18**, 267-297.

FIELD, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, **4**, 2379-2394.

FINDLAY, J. M., & GILCHRIST, I. D. (2003). *Active vision*. New York: Oxford University Press.

FRANCONERI, S. L., & SIMONS, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, **65**, 999-1010.

GEISLER, W. S., PERRY, J. S., & NAJEMNIK, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, **6**, 858-873.

HENSON, D. B. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research*, **19**, 57-61.

HODGSON, T. L. (2002). The location marker effect: Saccade latency increases with target eccentricity. *Experimental Brain Research*, **145**, 539-542.

KALESNYKAS, R. P., & HALLETT, P. E. (1994). Retinal eccentricity and the latency of eye saccades. *Vision Research*, **34**, 517-531.

KIENZLE, W., FRANZ, M. O., SCHÖLKOPF, B., & WICHMANN, F. A. (2009). Center-surround patterns emerge as optimal predictors for human saccade targets. *Journal of Vision*, **9**(5, Art. 7), 1-15.

LAND, M. F., MENNIE, N., & RUSTED, J. (1999). The roles of vision and eye movements in the control of activities of everyday living. *Perception*, **28**, 1311-1328.

NAZIR, T. A., & JACOBS, A. M. (1991). The effects of target discriminability and retinal eccentricity on saccade latencies: An analysis in terms of variable-criterion theory. *Psychological Research*, **53**, 281-289.

NERI, P., & HEEGER, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, **5**, 812-816.

PARKHURST, D. J., & NIEBUR, E. (2003). Scene content selected by active vision. *Spatial Vision*, **16**, 125-154.

PELLI, D. G., BURNS, C. W., FARELL, B., & MOORE-PAGE, D. C. (2006). Feature detection and letter identification. *Vision Research*, **46**, 4646-4674.

RAJASHEKAR, U., BOVIK, A. C., & CORMACK, L. K. (2006). Visual search in noise: Revealing the influence of structural cues by gaze-contingent classification image analysis. *Journal of Vision*, **6**, 379-386.

RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.

REINAGEL, P., & ZADOR, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, **10**, 341-350.

ROOS, J. C. P., CALANDRINI, D. M., & CARPENTER, R. H. S. (2008). A single mechanism for the timing of spontaneous and evoked saccades. *Experimental Brain Research*, **187**, 283-293.

ROVAMO, J., & VIRSU, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, **37**, 495-510.

SCHWARTZ, E. L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Research*, **20**, 645-669.

SHIMOZAKI, S. S., CHEN, K. Y., ABBEY, C. K., & ECKSTEIN, M. P. (2007). The temporal dynamics of selective attention of the visual periphery as measured by classification images. *Journal of Vision*, **7**(12, Art. 10), 1-20.

SIMONCELLI, E. P., & OLSHAUSEN, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, **24**, 1193-1216.

SIMONCELLI, E. P., PANINSKI, L., PILLOW, J., & SCHWARTZ, O. (2004). Characterization of neural responses with stochastic stimuli. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences III* (pp. 327-338). Cambridge, MA: MIT Press.

TATLER, B. W., BADDELEY, R. J., & GILCHRIST, I. D. (2005). Visual correlates of fixation selection: Effects of scale and time. *Vision Research*, **45**, 643-659.

TATLER, B. W., BADDELEY, R. J., & VINCENT, B. T. (2006). The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, **46**, 1857-1862.

TAVASSOLI, A., VAN DER LINDE, I., BOVIK, A. C., & CORMACK, L. K. (2007). An efficient technique for revealing visual search strategies with classification images. *Perception & Psychophysics*, **69**, 103-112.

VAN BEERS, R. J. (2007). The sources of variability in saccadic eye movements. *Journal of Neuroscience*, **27**, 8757-8770.

VIG, E., DORR, M., & BARTH, E. (2009). Efficient visual coding and the predictability of eye movements on natural movies. *Spatial Vision*, **22**, 397-408.

WHITE, B. J., STRITZKE, M., & GEGENFURTNER, K. R. (2008). Saccadic facilitation in natural backgrounds. *Current Biology*, **18**, 124-128.

WILSON, H. R., LEVI, D., MAFFEI, L., ROVAMO, J., & DEVALOIS, R. (1990). The perception of form: Retina to striate cortex. In L. Spillman & J. S. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 231-272). San Diego: Academic Press.

SUPPLEMENTAL MATERIALS

Histograms of fixation durations and saccade amplitudes from this study may be downloaded from http://app.psychonomic-journals.org/content/supplemental.

(Manuscript received January 28, 2009;
revision accepted for publication July 31, 2009.)