Vision Research 86 (2013) 35–42

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Scene categorization at large visual eccentricities

Muriel Boucart [a,*], Christine Moroni [a], Miguel Thibaut [a], Sebastien Szaffarczyk [a], Michelle Greene [b]

[a] Lab. Neurosciences Fonctionnelles & Pathologies, Université Lille-Nord de France, CHU Lille, CNRS, France
[b] Department of Computer Science, Stanford University, United States

ARTICLE INFO

Article history:
Received 7 June 2012
Received in revised form 13 March 2013
Available online 15 April 2013

Keywords:
Scene perception
Peripheral vision
Coarse-to-fine

ABSTRACT

Studies of scene perception have shown that the visual system is particularly sensitive to global properties such as the overall layout of a scene. Such global properties cannot be computed locally, but rather require relational analysis over multiple regions. To what extent is observers' perception of scenes impaired in the far periphery? We examined the perception of global scene properties (Experiment 1) and basic-level categories (Experiment 2) presented in the periphery from 10° to 70°. Pairs of scene photographs were simultaneously presented left and right of fixation for 80 ms on a panoramic screen (5 m diameter) covering the whole visual field while central fixation was controlled. Observers were instructed to press a key corresponding to the spatial location left/right of a pre-defined target property or category. The results show that classification of global scene properties (e.g., naturalness, openness) as well as basic-level categorization (e.g., forests, highways), while better near the center, were accomplished with a performance highly above chance (around 70% correct) in the far periphery even at 70° eccentricity. The perception of some global properties (e.g., naturalness) was more robust in peripheral vision than others (e.g., indoor/outdoor) that required a more local analysis. The results are consistent with studies suggesting that scene gist recognition can be accomplished by the low resolution of peripheral vision.

1. Introduction

Many insights have been made in vision research from the study of natural scenes. Human perception of real-world scenes is surprisingly robust, even in the face of limited presentation time (Greene & Oliva, 2009a; Joubert et al., 2007; Potter, 1976; Rousselet, Joubert, & Fabre-Thorpe, 2005; Thorpe, Fize, & Marlot, 1996), limited spatial frequency information (Schyns & Oliva, 1994), and limited attentional allocation (Li et al., 2002; but see Cohen, Alvarez, & Nakayama, 2011; Evans & Treisman, 2005). These studies typically use relatively small stimuli (<20° of visual angle) shown on a computer monitor. However, the real-world vision that we are trying to simulate is immersive, encompassing the entirety of the visual field. Additionally, it is also known that many aspects of form perception are substantially impaired in the periphery, relative to central vision. Indeed, visual acuity is not uniform across the visual field. The density of cone photoreceptors, responsible for high-resolution perception, decreases considerably as eccentricity increases from the fovea. Their number drops to about 50% at 1.75° from the fovea and to less than 5% at 20° from the fovea (Curcio, Allen, et al., 1991). Receptor density is also reduced in peripheral retina. A study from Chui, Song, and Burns (2008) showed that cone photoreceptor packing density decreases from 28,000 cells/mm^2 to 7000 cells/mm2 from a retinal eccentricity of 0.3–3.4 mm along the superior meridian. Receptive fields are larger in periphery, leading to loss of spatial resolution, and retinotopic projection to cortex prioritizes foveal inputs, resulting in a disproportionately large representation of central retinal locations in the visual cortex whilst peripheral parts of the retina are less and less represented on the cortical surface as eccentricity increases (Azzopardi & Cowey, 1996; Duncan & Boynton, 2003; Popovic & Sjostrand, 2001). In fact, over 30% of primary visual cortex (V1) represents the central 5° of the visual field (Horton & Hoyt, 1991). As a consequence, peripheral vision is far less capable of fine discrimination, even after its low spatial resolution has been compensated for by increasing size (M-Scaling, Näsänen & O'Leary 1998; Saarinen, Rovamo, & Virsu, 1987), by contrast enhancement (Mäkelä et al., 2001) or by increasing temporal integration (Swanson, Pan, & Lee, 2008). Moreover, crowding, a phenomenon whereby target visibility decreases in the presence of nearby objects, is more pronounced in the periphery (Leat, Li, & Epp, 1999; Levi, 2008; Pelli, 2008). This has been demonstrated with letters, digits, bars and gabor stimuli (Bouma, 1970; Kooi et al., 1994; Strasburger, Harvey, & Rentschler, 1991; see Strasburger, Rentschler, & Jüttner, 2011 for a review). Contour perception of gabor elements is impaired at an eccentricity of 6° (May & Hess, 2007) and figure/ground segregation in impaired in peripheral vision (Thompson et al., 2007).

---

[*] Corresponding author. Address: CHRU Lille, Hôpital Roger Salengro, service EFV, Lab. Neurosciences fonctionnelles & pathologies, 59037 Lille, France.
E-mail address: m-boucart@chru-lille.fr (M. Boucart).

0042-6989/$ - see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.visres.2013.04.006

However, in everyday life, we are not aware of the limitations in spatial resolution in peripheral vision because we move our eyes. Furthermore, certain visual discriminations remain robust in the periphery: perception of isolated objects (Boucart et al., 2010; Naili, Despretz, & Boucart, 2006), and even emotion processing (detection of fearful faces (Bayle et al., 2011; Rigoulot et al., 2011) and judgement of facial attractiveness (Guo, Liu, & Roebuck, 2011)), can be relatively good in peripheral vision.

The question of the relative contributions of central and peripheral vision to natural scene perception has been addressed both in normally sighted observers and in clinical patients exhibiting pathologies inducing a visual field loss. Thorpe et al. (2001) examined performances of young, normally sighted people in object categorization (detecting an animal in a natural scene) at large eccentricities. Photographs of natural scenes were randomly presented on a hemispheric screen from 0° (central) to 75° eccentricity. Surprisingly, they found that object detection accuracy was above 70% at 60° eccentricity even though participants claimed to perform the task by guessing. This non-conscious perception at large eccentricities has been confirmed and extended by Boucart et al. (2010) who demonstrated implicit recognition (measured by priming effects) but no explicit recognition (measured by recognition of previously seen pictures) of colored photographs of objects at 50° eccentricity whilst both explicit and implicit recognition occurred at 30° eccentricity in normally sighted young observers, as well as in 4 patients with Stargardt disease (a juvenile maculopathy inducing central vision loss). Larson and Loschky (2009) examined the contribution of central versus peripheral vision on scene gist recognition in a verification task (a matching between a word and a photograph). They presented participants with central photographs of real world scenes ($27 \times 27°$ of visual angle) for 106 ms each. Performance was compared in two conditions: a window condition showing the central portion of the scene and blocking peripheral information, and a scotoma condition blocking out the central portion and showing only the periphery. The radii of the window and scotoma were 1°, 5°, 10.8° and 13.6°. Performance was barely above chance in the 1° window condition suggesting that foveal vision alone is not sufficient for recognizing scene gist. Accuracy increased as the radius of the window increased, and when the overall image areas were equated, peripheral information lead to better categorization performance. The authors suggested that peripheral (and parafoveal vision) is more useful than high-resolution foveal vision for scene gist recognition. Velisavljevic and Elder (2008) examined the contribution of low-level factors (luminance and color statistics) and high level factors (configural properties) in determining the variation in recognition across the visual field. A photograph of a scene ($31 \times 31°$) was centrally displayed for 1s and followed by a probe display containing two smaller image blocks (3.9°): a target derived from the central or peripheral part of the previous photograph and a distractor from a random image never seen by the observer. The task was to decide which of the two image blocks was part of the previous scene. In this study, a "configural property" was defined as a coherent versus a scrambled version of each scene. They found that while spatial coherence increased recognition rates within the central visual field, the benefit of coherence vanished in the periphery, consistent with Hess and Dakin's (1997) account that curvilinear contour binding mechanisms might operate in the central part of the visual field, within 10° of the fovea.

Altogether, these results suggest that some scene categorization tasks can be performed in peripheral vision. However, the spatial resolution limit of scene perception has had little investigation. Studies using both natural scenes and laboratory stimuli, in central vision, have shown that the visual system is particularly sensitive to the global structure, or overall layout of a stimulus (Greene & Oliva, 2009a; Kimchi, 1992; Navon, 1977; Oliva & Torralba, 2001). For natural scenes, the global structure involves the overall space and shape of the scene, such as its degree of openness, the mean depth of the space, and its degree of expansion. (Greene & Oliva, 2009a; Oliva & Torralba, 2001). Other types of global properties describe the type of scene at the superordinate level (e.g., natural versus urban; Greene & Oliva, 2009b; Joubert et al., 2007). Such global properties cannot be computed locally, but rather require relational analysis over multiple regions. To what extent will observers' perception of these global properties suffer in the far periphery?

Experiment 1 examined the perception of a set of global scene properties such as naturalness, openness, expansion, indoor/outdoor for eccentricities ranging from 10° to 70°. Oliva and Torralba (2001) showed that global scene properties could be determined with a high reliability from a diagnostic set of low-level image features without invoking grouping operations. If the perception of global scene properties can be accomplished on a coarse version of the image, then basic-level scene categorization should also remain robust in the low resolution of large eccentricity. For the same reason, the low resolution of peripheral vision at large eccentricities might be sufficient to categorize some of a scene's global properties, allowing basic-level categorization as basic-level scene categories share global property values (Greene & Oliva, 2009a). On the other hand, basic-level categorization (rivers/mountains or cities/skyscrapers), might require more detailed processing of specific diagnostic features, and therefore be more difficult at large eccentricities. This was assessed in Experiment 2.

2. Experiment 1: Global scene property classification in the periphery

2.1. Method

2.1.1. Subjects

12 participants (9 females), all right handed, ranging in age from 22 to 45 took part in the experiment. Participants were members of the medical staff of the department of ophthalmology of the Lille's university hospital and students in neuroscience. All participants had normal or corrected-to-normal vision. A written consent was obtained from all participants.

2.1.2. Stimuli

The stimuli consisted of 400 colored photographs of natural scenes selected from a large laboratory database that had been previously ranked by human observers along each of the global scene properties (Greene & Oliva, 2009a). Four global properties were examined here: expansion, naturalness, openness, and indoor versus outdoor, using 100 photographs each (50 for each pole). Example images are shown in Fig. 1. The size of each photograph was $20 \times 20°$ of visual angle at a viewing distance of 2.1 m. As contrast sensitivity rapidly decreases at large eccentricities, especially for high spatial frequencies (Cannon, 1985), the images were presented at full contrast. However, no attempts were made to rescale the images to compensate for acuity loss in the periphery.

2.1.3. Apparatus

The stimuli were displayed by means of three projectors (SONY CS5) fixed on the ceiling, and connected to a PC computer. Participants were seated 2.1 meters from a hemispheric rigid light grey ($68\ cd/m^2$) screen covering 90° eccentricity on each side of the central fixation (see Fig. 1). Eye movements were recorded by means of an infrared camera located on the table in front of the observer (see Fig. 2), while head position was maintained by a chin rest. The presentation software (Vision 180) was written by the laboratory engineer. Participants made responses using a two-button box.

Fig. 1. Examples of the four global scene properties used in the experiment. Each property has two poles (shown side by side).

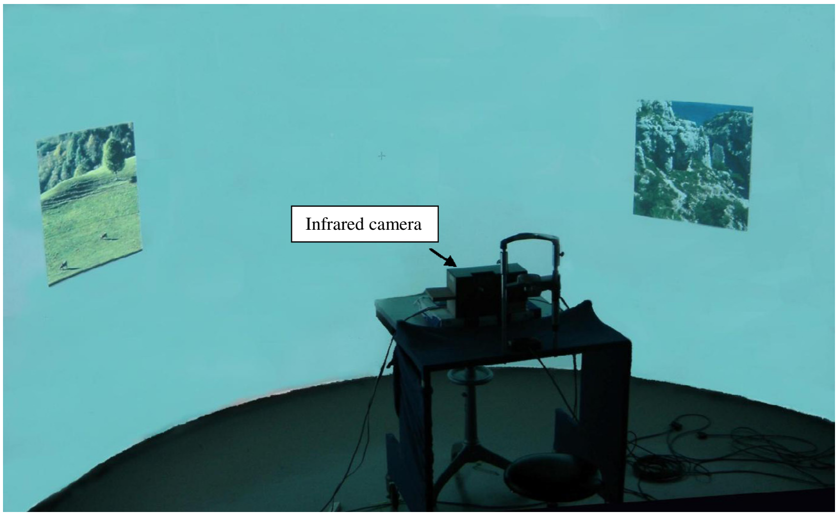

Fig. 2. The panoramic screen (5 m in diameter) covered 180° of visual field. For each categorization task, two scenes were simultaneously displayed to the left and right of fixation for 80 ms. Central fixation was controlled by an infrared camera placed in front of the participant.

2.1.4. Procedure and design

Prior to the beginning of each block, participants were presented with a pair of scenes that were not used in the experiment along with a description of the global property target on a piece of paper. On each trial, following a central fixation cross displayed for 500 ms, a pair of photographs representing both poles of the global property (e.g., one indoor and one outdoor scene) was simultaneously displayed left and right of fixation for 80 ms while the participant's gaze was stable on fixation. There was no backward masking. To ensure central fixation, eye position was monitored with an eye tracker. A saccade was detected by the camera if the eye moved by 3° from fixation. A trial was triggered when fixation had been stable for 500 ms. The experiment consisted of four blocks of 96 trials each determined by 2 spatial locations of the target (left/right), 4 eccentricities (10°, 30°, 50° and 70°, measured from the center of each scene) and 12 different images randomly selected for each global property pole. Each block corresponded to one of the four global scene properties, and the order of blocks was counterbalanced across participants. Participants performed a spatial forced choice (left/right) task, indicating whether the left or right image contained the target global property pole. For each of the four global properties tested, half of the participants were given one pole of the global property as a target (e.g. locate the open scene), and the other half were given the opposite pole as a target (locate the closed scene). Odd numbered participants were given the targets *urban*, *indoor*, *open* and *expansion*, while even numbered participants were given the targets *natural*, *outdoor*, *closed* and *no expansion*. Participants responded by pressing the left or the right key depending on the location of the target. There was no specific instruction regarding speed or accuracy. The four eccentricities (10°, 30°, 50° and 70°) and the two spatial locations of the target (left/right) were equally and randomly represented within each block. Each block lasted about 5–7 min, and the entire session lasted about 30 min.

2.2. Results

Analyses of variance, using the software SPSS, were conducted on the reaction times (RTs) and accuracy. As the condition of normality was not filled for accuracy, a logarithmic transformation was applied on the data ($\log_{10}$). The variables were the spatial location of the target (left/right of fixation), eccentricity (10°, 30°, 50° and 70°) the spatial property of the scene (naturalness, openness, expansion and indoor/outdoor) and the property's target pole. The data are presented in Fig. 3.

2.2.1. Accuracy

There was a main effect of the target pole $F(1,10) = 1.33$, $p = 0.27$) (natural: 90.9%, urban: 85.9%, indoor: 68.4%, outdoor: 73.9%, open: 77.3%, closed: 74.1%, expansion: 72.9%, no expansion: 78.1%). Accuracy decreased with increased eccentricity (10°: 88.4%, 30°: 84%, 50°: 73.9%, 70°: 64.6%; $F(3,33) = 17.15$, $p < 0.001$). Accuracy was significantly affected by the target global property ($F(3,33) = 27.73$, $p < .001$). The highest accuracy was observed for naturalness (88.4%). Averaged over spatial location and eccentricity accuracy was higher for naturalness than for openness (75.7% $t(11) = 11.56$, $p < .001$), than for expansion (75.5% $t(11) = 4.84$, $p < .001$) and than indoor/outdoor categorization (71.3% $t(11) = 4.26$, $p < .001$). All scene properties were categorized above chance at 70° eccentricity except for indoor/outdoor for which accuracy did not differ from chance at 70° ($t(11) = 1.76$, ns). There was a small, but reliable effect of spatial location (4.9%), with performance on targets displayed on the right side of fixation higher than on the left ($F(1,11) = 8.8$; $p < 0.013$). There was no statistically significant interaction (global property and eccentricity: $F(9,90) = 1.33$, $p = 0.28$; global property and spatial location: $F(3,30) = 1.02$, $p = 0.37$; eccentricity and spatial location : $F(3,30) = 1.9$, $p = 0.30$; global property, spatial location and eccentricity: $F(9,90) = 0.72$, $p = 0.49$; global property and target pole: $F(3,30) = 0.31$, $p = 0.81$; eccentricity and target pole: $F(3,30) = 0.48$, $p = 0.54$; spatial location and target pole: $F(1,10) = 0.28$, $p = 0.60$; global property, eccentricity and target pole: $F(9,90) = 1.9$, $p = 0.15$; global property, spatial location and target pole: $F(3,30) = 1.19$, $p = 0.32$; eccentricity, spatial location and target pole $F(3,30) = 0.41$, $p = 0.65$; global property, spatial location, eccentricity and target pole: $F(9,90) = 1.7$, $p = 0.20$).

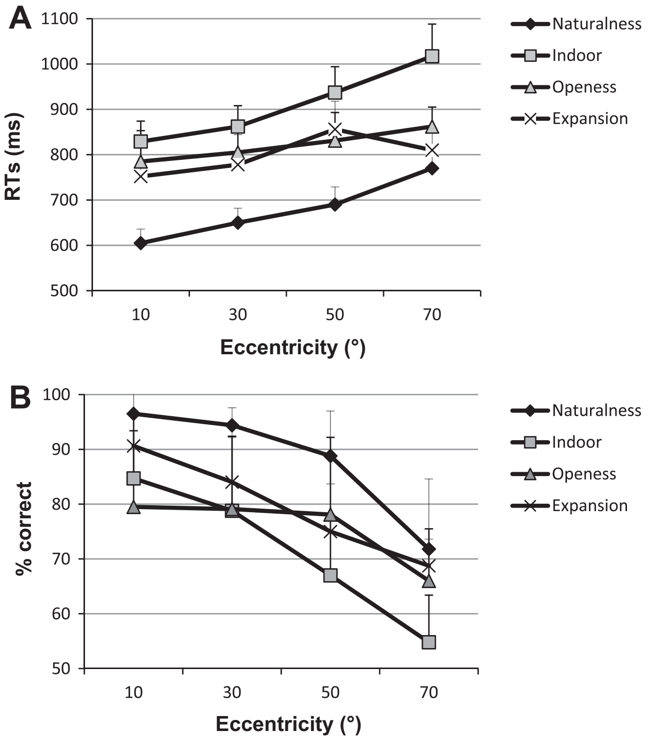

Fig. 3. Mean reaction times (-A-) and mean percent correct responses (-B-) as a function of the global scene property (averaged over the two poles of each property) and eccentricity. The bars represent ±1 standard error of the mean.

2.2.2. Response times

There was no main effect of the target pole for RTs $F(1,10) = 0.58$, $p = 0.46$) (natural: 665 ms, urban: 693 ms, indoor: 910 ms, outdoor: 912 ms, open: 762 ms, closed: 880 ms, expansion: 941 ms, no expansion: 707 ms). RTs increased significantly with increased eccentricity (10°: 743 ms, 30°: 774 ms, 50°: 829 ms, 70°: 890 ms; $F(3,33) = 18.33$; $p < .001$). RTs were significantly affected by the target global property ($F(2,33) = 5.81$, $p = .003$). Averaged over spatial location and eccentricity RTs were significantly shorter for naturalness than for openness (by 142 ms $t(11) = 2.82$, $p < .016$), than perspective (by 145 ms $t(11) = 2.64$, $p < .023$) and than indoor/outdoor categorization (by 232 ms $t(11) = 6.60$, $p < .001$). RTs did not differ significantly for left and right targets (right: 804 ms vs left: 818 ms $F(1,11) = 2.98$, $p = 0.11$). There was no significant interaction (global property and eccentricity: $F(9, 90) = 1.14$, $p = 0.34$, global property and spatial location: $F(3,30) = 0.74$, $p = .53$, eccentricity and spatial location: $F(3,30) = 1.92$, $p = 0.14$, global property, eccentricity and spatial location: $F(9,90) = 1.34$, $p = 0.22$; global property and target pole: $F(3,30) = 2.4$, $p < 0.09$; eccentricity and target pole: $F(3,30) = 1.7$, $p = 0.17$; spatial location and target pole: $F(1,10) = 2.03$, $p = .10$; global property, eccentricity and target pole: $F(9,90) = 0.24$, $p = 0.98$; global property, spatial location and target pole: $F(3,30) = 1.13$, $p = 0.35$; eccentricity, spatial location and target pole $F(3,30) = 1.24$, $p = 0.31$; global property, spatial location, eccentricity and target pole: $F(9,90) = 1.38$, $p = 0.20$).

2.3. Discussion

The results show that classification of global scene properties can be accomplished in the far periphery. The perception of some global scene properties was more robust in the low resolution of peripheral vision than others. For naturalness, accuracy was still around 90% at 50° eccentricity and above 70% at 70°. The ease of perceiving a scene's naturalness has been reported in other studies. Indeed, observers have higher accuracy and faster response times to classifying a scene as a natural than to manmade environment (Greene & Oliva, 2009b; Joubert et al., 2007). In contrast, indoor/outdoor accuracy was above 80% correct at 10° eccentricity and at chance at 70°. If the diagnostic information to perform the task is contained in local features, then performance should be impaired in peripheral vision as the details conveyed by high spatial frequencies are processed centrally. This may be the case with the indoor/outdoor task: wooden tables and chairs are frequently found indoors in dining rooms while plastic tables and chairs can be found outdoors on patios. Interestingly, this is similar to the accuracy fall-off observed by Thorpe et al. (2001) using an object detection task. Furthermore, Tran et al. (2010) reported that patients with central vision loss due to age related macular degeneration were more impaired in categorizing scenes as indoor versus outdoor than as natural versus urban, indicating that the higher resolution of central vision is required to discriminate indoor from outdoor scenes. Moreover, naturalistic real-world scenes are known to self-crowd in the far periphery. For instance, To et al. (2011) examined whether changes in different features (such as color or shape) were differentially compromised in peripheral vision with naturalistic images. They found that, even for isolated peripheral targets, features within an image could crowd one another thus reducing performance. Perhaps indoor/outdoor classification was more affected by internal crowding than the other global scene properties.

Although a number of scene properties have been studied as global scene properties, there is no exhaustive list of such properties. Furthermore, the perceptual availability of global scene properties varies (Greene & Oliva, 2009b). In this work, we examined the indoor/outdoor distinction as a potential global scene property, but unlike the other properties, the chance performance in classification in the far periphery suggests that the perception of a scene as being an indoor or outdoor environment requires local analysis.

The results of Experiment 1 have shown that, with the exception of indoor/outdoor, global scene property classification can be performed accurately in the far periphery. Do human observers utilize global property information to perform basic-level categorization? Oliva and Torralba (2006) suggested that a low resolution "sketch" would probably be sufficient to categorize at the superordinate level but higher resolution analysis could be needed for basic scene categories such as sea, mountain, river, or highway. Consistent with this suggestion, Joubert et al. (2007) and Rousselet, Joubert, and Fabre-Thorpe (2005) showed that observers are faster to classify scenes as being natural or urban (383–393 ms) than to categorize them at the basic-level (405–463 ms, see also Larson & Loschky, 2009) in central vision. We expected that, in the far periphery, the perception of a scene's basic-level category would be limited by the perception of its global properties. Specifically, as rapid scene categorization cannot be predicted from the perception of a single global property, and thus may require the perception of multiple diagnostic properties (Greene & Oliva, 2009a), then we would expect basic-level categorization to be limited by the resolution limits of global property perception. Therefore, we expect basic-level categorization to require at least fine resolution as global property classification. This was assessed in Experiment 2.

3. Experiment 2: Basic level categorization in periphery

3.1. Method

3.1.1. Subjects

12 New observers (8 females) took part in the basic-level categorization task. All participants were right handed, ranging in age from 21 to 33. Participants were all students in medicine or physiology. All had normal or corrected-to-normal vision. A written consent was obtained from all participants.

3.1.2. Stimuli

The stimuli were 1756 colored photographs of natural scenes chosen from a large laboratory scene database. Six basic-level categories were represented: three natural (mountains ($n = 445$ photos), forests ($n = 336$) and rivers ($n = 302$)); and three man-made (cities ($n = 148$), highways ($n = 323$) and skyscrapers ($n = 202$)). Examples are shown in Fig. 4. The angular size of each photograph was $20^\circ \times 20^\circ$ of visual angle at a viewing distance of 2.1 m.

3.1.3. Procedure and design

Participants were tested on six blocks of 96 trials each, consisting of 24 trials (12 with the target on the left and 12 with the target on the right of fixation) and 4 eccentricities (10°, 30°, 50° and 70°). In each block, two basic-level categories were paired as target and distractor. For natural categories, participants were presented with pairs of: rivers/mountains, forests/rivers and mountains/forests in three separate blocks. For man-made categories, they were presented with pairs of skyscrapers/highways, highways/cities and cities/skyscrapers also in three independent blocks. In each block, participants were given a single target category. For instance, odd-numbered participants were given rivers, forests and mountains as targets and even-numbered participants were given mountains, rivers and forests as targets. The same procedure was used for man-made pairs. Following a central fixation cross displayed for 500 ms, a pair of photographs was simultaneously displayed to the left and right of fixation for 80 ms when the gaze was maintained on a central fixation. Central fixation was controlled by the eye tracker. There was no backward masking. The task was to indicate whether the target category was on the right or left side. Responses were given by pressing the left or the right key corresponding to the location of the target.

3.2. Results

Analyses of variance, using the software SPSS, were conducted on RTs and accuracy. When the condition of normality was not filled (on accuracy), a logarithmic transformation ($\log_{10}$) was applied on the data and an ANOVA was conducted. The factors were the spatial location of the target (left/right), the categories of scenes (rivers, mountains, forests, highways, cities and skyscrapers) and the four eccentricities (10°, 30°, 50° and 70°). The results are presented in Fig. 5.

3.2.1. Accuracy

No significant main effect of category was observed ($F(5,55) = 1.43$, $p = 0.22$). There was no significant main effect of the spatial location of the target ($F(1,11) = 1.32$, $p = 0.27$) . Accuracy decreased with the increase in eccentricity: 10°: 85.1%, 30°: 83.5% , 50°: 80.6% , 70°: 73.2% ($F(3,33) = 59.53$, $p < .001$). Eccentricity interacted significantly with category ($F(15,165) = 1.83$, $p < .035$). As can be seen from Fig. 5 the category city was less affected by the increase in eccentricity (chi2 (3) = 3.08, $p = 0.37$) than the others (forest (chi2 (3) = 23.41, $p < 0.001$) mountain (chi2 (3) = 18.34, $p < 0.001$), river (chi2 (3) = 18.42, $p < 0.001$), skyscraper (chi2 (3) = 27.51, $p < 0.001$) and highway (chi2 (3) = 22.82, $p < 0.001$)) for which accuracy decreased with increased eccentricity. No other interaction was statistically significant (category and spatial location: $F(5,55) = 2.15$, $p < .07$; eccentricity and spatial location $F(3,33) = .218$, $p = 0.88$).

3.2.2. Response times

There was no significant main effect of category ($F(5,55) = 1.99$, $p < .093$). Neither was there a significant main effect of the spatial location of the target ($F(1,11) = 0.5$, $p = 0.2$). The effect of eccentricity was significant ($F(3,33) = 56.3$, $p < .001$) with an increase in RTs as eccentricity increased: 10°: 546 ms, 30°: 557 ms , 50°: 594 ms, 70°: 626 ms). There was a significant interaction between category, spatial location of the target and eccentricity ($F(15,165) = 2.5$, $p < .002$). No other interaction was significant: category and eccentricity ($F(5,165) = 1.24$, $p = 0.24$); category and spatial location ($F(5,55) = 1.34$, $p = 0.26$); eccentricity and spatial location ($F(3,33) = 2.51$ $p = 0.76$).

4. Discussion

The results of Experiment 2 demonstrate that basic-level categorization can be performed in the far periphery, as accuracy was still at 70% correct at 70° from fixation. Indeed, averaged over the three natural categories (mountains, rivers and forests) and the three urban categories (highways, skyscrapers and cities) RTs were faster (by 100 ms for natural scenes and by 98 ms for urban scenes) in Experiment 2 than in categorization on global properties (in Experiment 1), though the two experiments cannot be compared directly as, with a 80 ms exposure time, discrimination

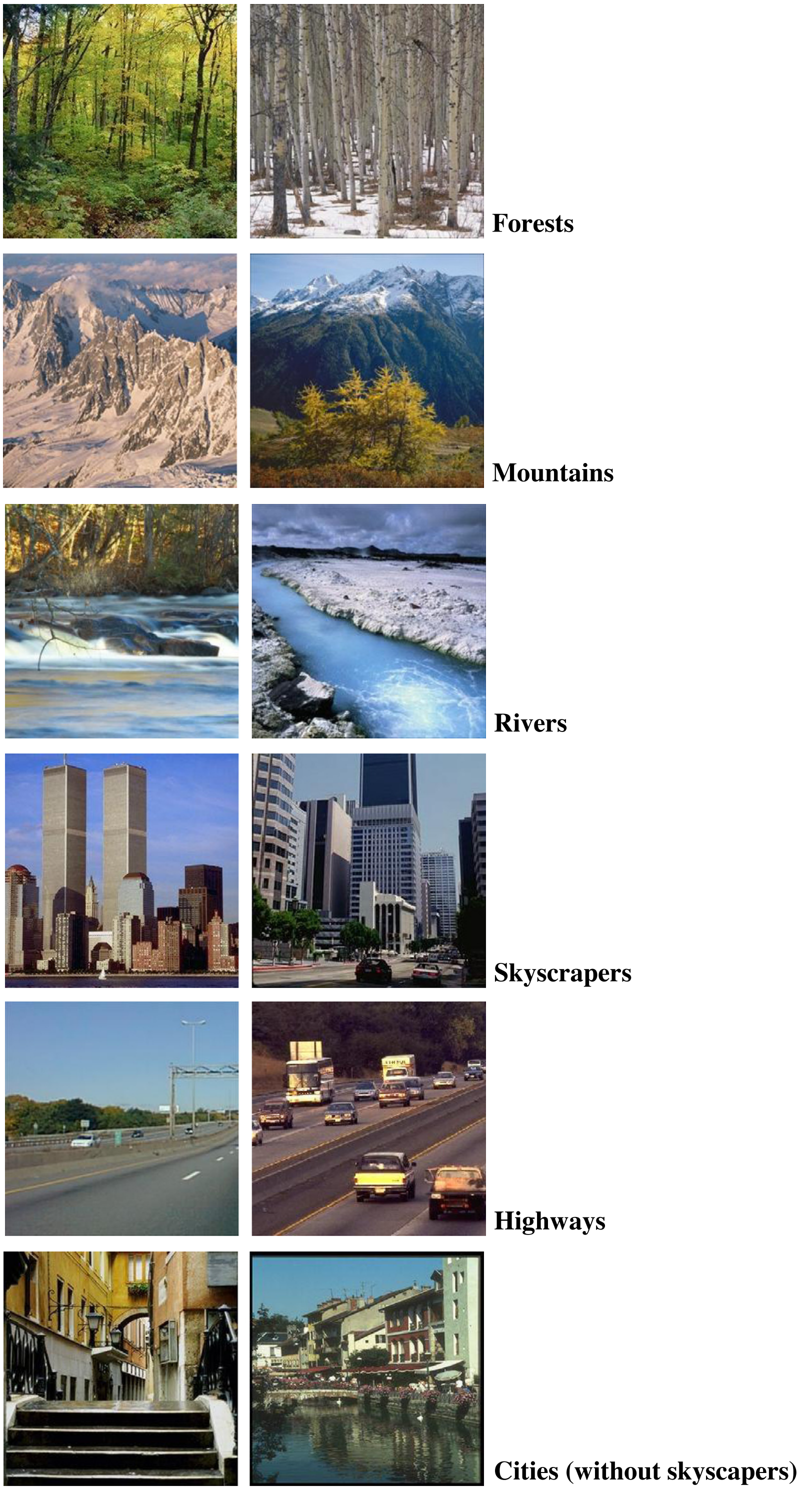

Fig. 4. Examples of scenes used for basic level categorization (mountains, rivers and forests for natural scenes; and highways, skyscrapers and cities for man made scenes).

between expansion/no expansion or open/closed might be more difficult than between forest and river or highway and skyscraper.

The faster RTs for basic level is consistent with previous studies on isolated objects displayed centrally showing a response time advantage for basic level categorization (Palmeri & Gauthier, 2004; Rosch et al., 1976; Tanaka, 2001). This was not necessarily expected, as studies with scenes as stimuli have reported shorter RTs for superordinate-level categorization compared to the basic level (Joubert et al., 2007; Rousselet, Joubert, & Fabre-Thorpe, 2005). Altogether, the results of Experiment 2 indicate that observers can classify scenes at the basic-level rapidly and accurately, even in the far periphery.

5. General discussion

Studies of scene perception have shown that scene recognition is still possible under the impoverished conditions of low spatial

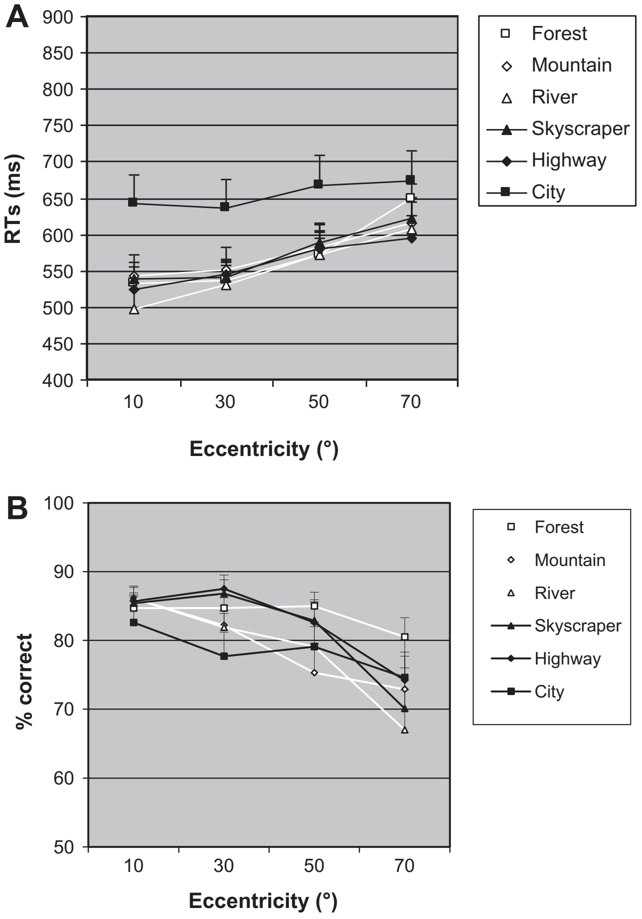

Fig. 5. Mean reaction times (-A-) and mean percent correct responses (-B-) as a function of the scene category and eccentricity in the basic level categorization task. The bars represent ±1 standard error of the mean.

frequency (Peyrin et al., 2003; Schyns & Oliva, 1994) and very short exposure time (Greene & Oliva, 2009b; Rousselet, Joubert, & Fabre-Thorpe, 2005) suggesting that coarse features may capture the diagnostic image information needed for categorization. Spatial resolution decreases dramatically from the fovea to the periphery of the retina. Yet, in two experiments we found that classification of global scene properties (Experiment 1) as well as basic-level scene categorization (Experiment 2) were accomplished with a performance highly above chance (around 70% correct) in the far periphery at 70°. Our results are consistent with Thorpe et al. (2001) who reported that participants were able to detect an animal within a scene above chance at 75° eccentricity. Though our two-alternative forced choice task was more demanding than the go/no-go task on single photographs used in the Thorpe et al. (2001) study, accuracy at 70° eccentricity was in a similar range. Our results are also consistent with Larson and Loschky's (2009) conclusion that scene gist recognition can be accomplished by peripheral vision and extend their results beyond the 27° eccentricity in their study.

Although scene categorization was still above chance at large eccentricities (50°, 70°), the highest performance occurred for scenes presented at smaller eccentricities. This result suggests that, though low-resolution is sufficient for scene categorization, scene categorization is aided by higher spatial frequencies, colors and other features that are available nearer to central vision where grouping processes may still operate.

The fact that scene categorization was still above chance at an eccentricity where visual resolution is poor is in agreement with scene-centered models claiming that scene recognition proceeds from coarse-to-fine spatial scales (i.e., from holistic descriptors of scene structure (Greene & Oliva, 2009a) and not only on local objects). However, not all of the putatively "global" properties tested in Experiment 1 could be classified in the far periphery. Indeed, compared to other global properties, accuracy was at chance in the indoor/outdoor categorization at 70° eccentricity. This indoor/outdoor categorization likely required finer analysis than discrimination based on natural/urban scenes due to the necessity to identify objects (to determine whether such objects are found indoor or outdoor). Peripheral vision could impair the processing of local information via the increased sensitivity to low spatial frequency information (Hilz & Cavonius, 1974) and increased crowding (Levi, 2008, 2011; Pelli, 2008; Strasburger, Rentschler, & Jüttner, 2011).

Numerous studies have shown that the perceptual availability of basic-level category information is privileged relative to superordinate- or subordinate-level information (e.g., Rosch et al., 1976; Tanaka, 2001) but exceptions to the primacy of basic-level category have been reported in studies using scenes as stimuli (Macé et al., 2009; Murphy & Wisniewski, 1989), possibly because entry-level categories for scenes have not been systematically established. For instance, as images in the same category tend to share global properties (Greene & Oliva, 2009a; Oliva & Torralba, 2001), we cannot preclude the possibility that basic-level categorization performance could be based on the perception of diagnostic global properties. In categorization tasks on photographs of natural scenes, several studies have reported a processing speed advantage for broadly defined categories (i.e., natural vs urban scenes) over more basic-level scene categories (i.e., mountains, forests, beach, Joubert et al., 2007; Rousselet, Joubert, & Fabre-Thorpe, 2005). In the current experiments, we have demonstrated that both basic-level and global property classification can be performed accurately in the far periphery. However, further studies will be needed to determine whether basic-level categorization is made possible through the perception of diagnostic global properties, or by other mechanisms.

Neuroimaging studies have demonstrated that object categories that engage the analysis of fine details such as faces and words are associated with central high magnification factor representations whilst objects whose recognition involves integration of visual information over large retinal distances are mapped onto more peripheral low magnification factor representations. For instance, Levy et al. (2001) examined whether there was a relationship between the activation in response to faces and buildings at specific eccentricity distances. Both faces and buildings were presented at the same location in the visual field (central vs 16° eccentricity). The results showed that faces were consistently associated with central visual field bias whilst buildings were associated with peripheral visual field bias. Based on these findings, Malach, Levy, et al. (2002; see also Grill-Spector & Malach, 2004) proposed an organizing principle of human high order object areas that is based on an orderly layout of visual field eccentricity. Two eccentricity maps were proposed: a posterior dorsal one (LO) located in lateral occipital cortex and a ventral anterior one (VOT) in the ventral occipito-temporal cortex. Objects whose recognition depends on analysis of fine details are associated with center-biased representations whereas objects whose recognition entails large-scale integration will be more peripherally biased (though Jebara et al., 2009 showed that the center-face versus peripheral-building bias is modulated by task demands).

In line with these studies on humans, animal studies indicate that the periphery can also be involved to some degree in shape perception (Rosa et al., 2009). Moreover, Palmer and Rosa (2006) found that the representation of the far periphery in area MT receives specific connections from parahippocampal and retrosplenial areas, structures that support scene perception (Epstein, 2008, 2011).

6. Conclusion

Previous studies suggested that scene gist recognition can be accomplished with the low resolution of peripheral vision but photographs were presented within the macular region. Using a panoramic screen covering the whole visual field we demonstrated that categorization of both global scene properties as well as basic-level categories can be accomplished with high accuracy (70% correct) at 70° eccentricity.

Acknowledgments

The authors are grateful to Claire Lemaitre and Claire Barbet for testing the participants and to Aude Oliva for her suggestions. The study was funded by a grant from the French National Research Agency to the first author. The study was approved by the ethical committee (CPP Nord-Ouest IV).

References

Azzopardi, P., & Cowey, A. (1993). Preferential representation of the fovea in the primary visual cortex.. *Nature, 361*(6414), 719–721.

Bayle, D. J., Schoendorff, B., Hénaff, M. A., & Krolak-Salmon, P. (2011). Emotional facial expression detection in the peripheral visual field. *PLoS One, 6*(6), e21584.

Boucart, M., Naili, F., Despretz, P., Defoort, S., & Fabre-Thorpe, M. (2010). Implicit and explicit object recognition at very large visual eccentricities: No improvement after loss of central vision. *Visual Cognition, 18*(6), 839–858.

Bouma, M. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177–178.

Cannon, M. W. (1985). Perceived contrast in the fovea and periphery. *Journal of the Optical Society of America, A2*, 1760–1768.

Chui, T. Y., Song, H., & Burns, S. A. (2008). Adaptive-optics imaging of human cone photoreceptor distribution. *Journal of the Optical Society of America, 25*(12), 3021–3029.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). *Natural-scene Perception Requires Attention, 22*(9), 1165–1172.

Curcio, C. A., Allen, K. A., et al. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *Journal of Comparative Neurology, 312*(4), 610–624.

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex. Correlates with acuity thresholds. *Neuron, 38*, 659–671.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Science, 12*(10), 388–396.

Epstein, R. A. (2011). Cognitive neuroscience: Scene layout from vision and touch. *Current Biology, 21*(11), R437–R438.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception & Performance., 31*(6), 1476–1492.

Greene, M. R., & Oliva, A. (2009a). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20*(4), 464–472.

Greene, M. R., & Oliva, A. (2009b). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology, 58*(2), 137–179.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*, 649–677.

Guo, K., Liu, C. H., & Roebuck, H. (2011). I know you are beautiful even without looking at you: Discrimination of facial beauty in peripheral vision. *Perception, 40*(2), 191–195.

Hess, R. F., & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature, 390*(6660), 602–604.

Hilz, R., & Cavonius, C. R. (1974). Functional organization of the peripheral retina: Sensitivity to periodic stimuli. *Vision Research, 14*(12), 1333–1337.

Horton, J. C., & Hoyt, W. F. (1991). The representation of the visual field in human striate cortex. A revision of the classic Holmes map. *Archives of Ophthalmology, 109*, 816–824.

Jebara, N., Pins, D., Despretz, P., & Boucart, M. (2009). Face or building superiority in peripheral vision reversed by task requirements. *Advances in Cognitive Psychology, 5*, 42–53.

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research, 47*, 3286–3297.

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin, 112*(1), 24–38.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*, 255–279.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision, 9*(10), 1–16. 6.

Leat, S. J., Li, W., & Epp, K. (1999). Crowding in central and eccentric vision: The effects of contour interaction and attention. *Investigative Ophthalmology & Vision Science, 40*(2), 504–512.

Levi, D. M. (2008). Crowding-an essential bottleneck for object recognition: A mini review. *Vision Research, 48*(5), 635–654.

Levi, D. M. (2011). Visual crowding. *Current Biology, 21*(18), R678–R679.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*, 533–539.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America, 99*(14), 9596–9601.

Macé, M. J. M., Joubert, O. R., Nespoulous, J. L., & Fabre-Thorpe, M. (2009). The time course of visual categorization: You spot the animal faster than the bird. *PLoS ONE, 4*(6), e5927. http://dx.doi.org/10.1371/journal.pone.0005927.

Mäkelä, P., Nasanen, R., Rovamo, J., & Melmoth, D. (2001). Identification of facial images in peripheral vision. *Vision Research, 41*, 599–610.

Malach, R., Levy, I., et al. (2002). The topography of high-order human object areas. *Trends in Cognitive Science, 6*(4), 176–184.

May, K. A., & Hess, R. F. (2007). Ladder contours are undetectable in the periphery: A crowding effect? *Journal of Vision, 7*(13), 1–15.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory & Cognition, 15*(4), 572–586.

Nääsänen, R., & O'Leary, C. (1998). Recognition of band-pass filtered hand-written numerals in foveal and peripheral vision. *Vision Research, 38*, 3691–3701.

Naili, F., Despretz, P., & Boucart, M. (2006). Colour recognition at large visual eccentricities in normal observers and patients with low vision. *Neuroreport, 17*(15), 1571–1574.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353–383.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal in Computer Vision, 42*, 145–175.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research: Visual perception, 155*, 23–36.

Palmer, S. M., & Rosa, M. G. (2006). A distinct anatomical network of cortical areas for analysis of motion in far peripheral vision. *European Journal of Neuroscience, 24*(8), 2389–2405.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Review Neuroscience, 5*(4), 291–303.

Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology, 18*(4), 445–541.

Peyrin, C., Chauvin, A., Chokron, S., & Marendaz, C. (2003). Hemispheric specialization for spatial frequency processing in the analysis of natural scenes. *Brain & Cognition, 53*(2), 278–282.

Popovic, Z., & Sjostrand, J. (2001). Resolution, separation of numeral ganglion cells, and cortical magnification in humans. *Vision Research, 41*, 1313–1319.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology Human Learning, 2*(5), 509–522.

Rigoulot, S., D'Hondt, F., Defoort-Dhellemmes, S., Despretz, P., Honoré, J., & Sequeira, H. (2011). Fearful faces impact in peripheral vision: Behavioral and neural evidence. *Neuropsychologia, 49*(7), 2013–2021.

Rosa, M. G., Palmer, S. M., Gamberini, M., Burman, K. J., Yu, H. H., Reser, D. H., et al. (2009). Connections of the dorsomedial visual area: Pathways for early integration of dorsal and ventral streams in extrastriate cortex. *Journal of Neuroscience, 29*(14), 4548–4563.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get the "gist" of real-world natural scenes? *Visual Cognition, 12*(6), 852–877.

Saarinen, J., Rovamo, J., & Virsu, V. (1987). Texture discrimination at different eccentricities. *Investigative Ophtalmology and Vision Science, 30*, 293–296.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependant scene recognition. *Psychological Science, 5*, 195–200.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of Vision, 11*(5), 1–82. 13.

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast threshold for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49*, 495–508.

Swanson, W. H., Pan, F., & Lee, B. B. (2008). Chromatic temporal integration and retinal eccentricity: Psychophysics, neurometric analysis and cortical pooling. *Vision Research, 48*(26), 2657–2662.

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General, 130*(3), 534–543.

Thompson, B., Hansen, B. C., Hess, R. F., & Troje, N. F. (2007). Peripheral vision: Good for biological motion, bad for signal noise segregation? *Journal of Vision, 25*;7(10):12.1-7.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522.

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bulthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience, 14*, 869–876.

To, M. P., Gilchrist, I. D., Troscianko, T., & Tolhurst, D. J. (2011). Discrimination of natural scenes in central and peripheral vision. *Vision Research, 51*(14), 1686–1698.

Tran, T. H., Rambaud, C., Despretz, P., & Boucart, M. (2010). Scene perception in age-related macular degeneration. *Investigative Ophthalmology & Vision Science, 51*(12), 6868–6874.

Velisavljevic, L., & Elder, J. H. (2008). Visual short-term memory for natural scenes: Effects of eccentricity. *Journal of Vision, 8*(4), 28, 1–17.