Different spatial frequency bands selectively signal for natural image statistics in the early visual system

Bruce C. Hansen,1 Aaron P. Johnson,2 and Dave Ellemberg3

1Department of Psychology & Neuroscience Program, Colgate University, Hamilton, New York; 2Department of Psychology, Concordia University, Montréal, Québec, Canada; and 3Centre de Recherche en Neuropsychologie et Cognition (CERNEC), Université de Montréal, Québec, Canada

Submitted 5 April 2012; accepted in final form 15 July 2012

Hansen BC, Johnson AP, Ellemberg D. Different spatial frequency bands selectively signal for natural image statistics in the early visual system. J Neurophysiol 108: 2160–2172, 2012. First published July 25, 2012; doi:10.1152/jn.00288.2012.—Early visual evoked potentials (VEPs) measured in humans have recently been observed to be modulated by the image statistics of natural scene imagery. Specifically, the early VEP is dominated by a strong positivity when participants view minimally complex natural scene imagery, with the magnitude of that component being modulated by luminance contrast. However, since natural scene imagery is broad band in terms of spatial frequency, it is not known whether the above-mentioned modulation results from a complex interaction within or between the early neural processes tuned to different bands of spatial frequency. Here, we sought to address this question by measuring early VEPs (specifically, the C1, P1, and N1 components) while human participants viewed natural scene imagery that was filtered to contain specific bands of spatial frequency information. The results show that the C1 component is largely unmodulated by the luminance statistics of natural scene imagery (being only measurable when such stimuli were made to contain high spatial frequencies). The P1 and N1, on the other hand, were observed to exhibit strong spatial frequency-dependent modulation to the luminance statistics of natural scene imagery. The results therefore suggest that the dependency of early VEPs on natural image statistics results from an interaction between the early neural processes tuned to different bands of spatial frequency.

natural scenes; image complexity; scene statistics; amplitude spectrum; visual evoked potentials

THE DIFFERENT EARLY visual evoked potentials (VEPs) elicited by sinusoidal luminance gratings that vary in spatial frequency (SF) and luminance contrast have been argued to reflect the underlying neural processes associated with the parvocellular or magnocellular visual pathways (Ellemberg et al. 2001; Hansen et al. 2011; Jemel et al. 2010; Vassilev et al. 1994, 2002). Specifically, the C1 VEP component (a negative deflection from baseline, typically peaking between ~60- and ~100-ms poststimulus onset time, PSOT): 1) appears at medium-to-high contrasts with high-SF stimuli, i.e., >4 cycles per degree (cpd); 2) increases slowly in amplitude with increasing contrast; and 3) exhibits a nonsaturating response across higher contrasts (Ellemberg et al. 2001). Conversely, the P1 (a positive deflection from baseline, typically peaking between ~100- and ~140-ms PSOT): 1) appears at lower contrasts at SFs as low as 0.8 cpd; 2) increases rapidly in amplitude with increasing contrast; and 3) saturates at medium contrasts (Ellemberg et al. 2001). Such response characteristics respectively follow the known contrast response functions of the parvocellular and magnocellular pathways to sinusoidal luminance gratings (Benardete et al. 1992; Benardete and Kaplan 1997a,b; Derrington and Lennie 1984; Merigan and Eskin 1986; Merigan and Maunsell 1993; Tootell et al. 1988). Additionally, both C1 and P1 components have been shown to be pharmacologically separable in the cat visual cortex (e.g., Arakawa et al. 1993; Zemon et al. 1980) and are differentially modulated by cathodal or anodal stimulation in transcranial direct current stimulation (tDCS) paradigms with human participants (Accornero et al. 2007; Antal et al. 2004). The pharmacological separability of the C1 and P1 VEP components in cats and differential modulation of the same components in humans using tDCS would both be expected if each component were generated by different neural populations. It is worth noting, however, that by 100-ms PSOT, several early extrastriate cortical areas are known to be active (e.g., Di Russo et al. 2001, 2005, 2012; Martínez et al. 2001; Vanni et al. 2004); thus the neural populations underlying the C1 and P1 components likely reflect cortical populations receiving signals (either directly or indirectly) relayed from either the parvocellular or magnocellular pathways.

Recently, we examined the extent to which the early VEP components mentioned above are modulated by well-established physical characteristics of natural scene imagery (Hansen et al. 2011). Concerning the well-established physical regularities of natural scenes, such imagery has been shown to possess a physical property where the contrast at different SFs, f, falls with increasing SF, following a 1/fα relationship (e.g., Billock 2000; Field 1987; Ruderman and Bialek 1994; Tolhurst et al. 1992; van der Schaaf and van Hateren 1996), with α typically observed to be near 1.0 on logarithmic axes. The exact α measured for a given image is typically referred to as the slope of the orientation averaged amplitude spectrum of

Note that we refer to the first negative deflection in the VEP as C1, which has also been referred to as N1 in earlier work (including our recent report in Hansen et al. 2011) measuring VEPs in response to sinusoidal gratings. It is well-known that the C1 can be either positive or negative depending on whether the stimuli are presented above or below fixation. However, when large-field stimuli are presented at fixation (as in the current study), the C1 is consistently observed to be negative (e.g., Ellemberg et al. 2001; Hansen et al. 2011; Murray et al. 1987; Previc 1988; Vassilev et al. 1994, 2002).
that image. What this amounts to is that natural scene imagery generally contains much more contrast at lower SFs relative to higher SFs, with slope of the amplitude spectrum reflecting the degree of luminance contrast at lower relative to higher SFs. Another physical dimension along which natural scene stimuli have been shown to vary is the amount of edges/lines any given image contains, a property referred to as structural complexity (or structural sparseness, e.g., Hansen and Hess 2007). Essentially, structural complexity can be assessed (in a biologically inspired manner) by filtering a given stimulus in the spatial (i.e., image) domain with log-Gabor filters and thresholding the modulus of quadrature-pair filter responses at a number of different SFs and orientations. Next, the ratio of the total summed thresholded filter responses (within each scale and orientation) to the total image area is taken and then averaged across filter scale and orientation. The resulting averaged ratio therefore serves as an index of local filter activity (which is dependent on the number of edges/lines at different scales and orientations) and therefore corresponds to the relative density of edges and lines making up the structure of a given image. Such physical characteristics of scenes are important because both the slope of the amplitude spectra and structural complexity of natural scene stimuli have been shown to modulate behavioral performance on a number of different tasks involving the detection, discrimination, or identification of different visual stimuli (Bex et al. 2009; Hansen and Essock 2005; Hansen and Hess 2007; Párraga et al. 2000; Tadmor and Tolhurst 1994; Thomson and Foster 1997; Webster and Miyahara 1997). Correspondingly, Hansen et al. (2011) found an interaction between early VEP component magnitude that was dependent on both the structural complexity and amplitude spectrum slope. Specifically, the P1 component dominated the early VEP for scenes low in structural complexity and was virtually absent from the waveform for scenes high in structural complexity with the opposite being true for an apparent C1 component. Additionally, the overall magnitude of the P1 exhibited a reliable dependency on amplitude spectrum slope, whereas the apparent C1 component was not found to vary significantly with amplitude spectrum slope. Such a finding offers important insight into our understanding of how the human visual system processes natural scene imagery because up until now, the method by which the early visual system processes and encodes natural image content has almost entirely been driven by neural evidence from nonhuman vertebrates (e.g., Dan et al. 1996; Felsen et al. 2005; Maldonado and Babul 2007; Mante et al. 2005; Tolhurst et al. 2009; Vinje and Gallant 2000; Weliky et al. 2003).

Given the history of reports supporting the notion that the P1 VEP component is linked with signaling for low-SF contrast, with the C1 linked with signaling for high-SF contrast, the results of Hansen et al. (2011) suggest the possibility of a SF dependency for signaling different natural scene statistics in the human visual system. First, when viewing scenes that are low in structural complexity, the P1 seems to be the dominant signal for luminance contrast (i.e., carrying a signal that corresponds to amplitude spectrum slope) with the majority of that signal determined by contrast at low SFs. Second, when viewing scenes rich in edge/line content, an apparent C1 becomes the dominant signal modulator, suggesting that high-SF information governs the visual signal by possibly emphasizing structural content as opposed to general luminance contrast. However, since the natural scene imagery employed in that study were all broad band (i.e., they contained significant luminance contrast across all SFs), the early visual processes coding for specific SFs should, in theory, be simultaneously active. Since it is known that interactions occur within and between populations of neurons tuned to different SFs (Bau-man and Bonds 1991; DeAngelis et al. 1992; Morrone et al. 1982), it is not clear whether the C1 modulation caused by structural complexity is dependent on neural populations solely tuned to high SFs or whether that signal resulted from interactions between the neural processes associated with signaling for different SFs. Similarly, it is not clear whether the modulation of the P1 component by SF luminance contrast (controlled by the slope of the amplitude spectra of the stimuli) was solely due to low-SF-tuned neural populations or again resulted from an interaction between populations tuned to different SFs.

The current study therefore sought to address the above-mentioned issue by measuring C1 and P1 component modulation while participants viewed natural scene stimuli (varying in amplitude spectrum slope and structural complexity) for which the amount of SF contrast was systematically controlled with either band-pass or low-pass SF filters. That is, with band-pass filtering, scene stimuli can be made to contain only low- or only high-SF information, thereby allowing for a direct assessment of whether scene complexity was more or less influential in one or the other (or both) SF bands. The low-pass conditions are designed to vary the cutoff frequency of the filter, thereby allowing for a systematic assessment of component modulation as stimuli move from containing only low SFs to scenes containing increasingly higher SFs.

Following from the above, the current study consists of three experiments. The first experiment sought to gather benchmark C1 and P1 data from the two polar extremes of stimulus space, namely vertical sinusoidal luminance gratings (narrow band in SF and orientation content) and broad-band natural scenes (broad band in terms of SF and orientation content). The second experiment used a subset of natural scenes from the first experiment and subjected them to band-pass SF filtering before viewing. The third and final experiment employed a subset of stimuli from the second experiment and subjected those to variable low-pass SF filtering before viewing.

Altogether, the results of the current study suggest that when viewing broad-band natural scene imagery, the early VEP modulation resulting from structural complexity is limited to high SFs and, contrary to our previous findings, is most likely carried by an N1 component (peaking between ~150- and ~180-ms PSOT). The implication for such a result is that the human visual system relies on “lower-level” C1-P1 related processes for natural scenes sparse in structural content, and as that content becomes more dense and complex, a later N1 negativity becomes dominant and is dependent on the presence of higher SF content. Thus it seems that the visual system signals for the presence of different natural scene statistics in a SF-specific manner, with scene structural complexity regulating the dominance of signals arising from cortical populations tuned to different SFs. That is, the neural signals reflecting processes tuned to different SFs seem to signal selectively for different natural scene image properties. Given the supposed extrastriate contribution to the P1 and N1 components, the early neural signaling of natural scene statistics may be largely...
dominated by extrastriate visual cortical areas when that activity is measured via VEPs.

**EXPERIMENTAL PROCEDURES**

**Apparatus.** All stimuli were presented on a 19-in. ViewSonic CRT monitor (G90FB Graphics Series) driven by an Intel Core 2 Duo processor (2 GHz) equipped with 2 GB of RAM and a 256-MB NVIDIA GeForce 9400M graphics card with 8-bit grayscale resolution. Stimuli were linear with the output values of the digital camera from which they were obtained. Mean luminance output of the display monitor was 58 cd/m², the frame rate was set to 85 Hz, and the resolution was set to 1,024 × 768 pixels and was viewed from 1 m.

**Participants.** Nine observers (2 female) participated in all experiments of the current study. All participants (all right-handed) were naïve to the purpose of the study. The age of the participants ranged between 20 and 25 yr (median = 21 yr). All had normal (or corrected to normal) vision and were compensated for their participation. Research Ethics Board-approved informed written consent was obtained.

**Stimuli.** All stimuli consisted of grayscale digital images (512 × 512 pixels) windowed in the spatial domain with a circular edge-blurred aperture (ramped to mean luminance) to eliminate the sharp edges at the boundaries of the stimulus images when presented on the display screen. The diameter of the windowed region subtended 9.3° of visual angle viewed from 1 m. All stimuli had a root-mean-square contrast (rms) of 0.26 and a grayscale mean of 127 (which corresponded to 58 cd/m²) on the display monitor.

The conditions within experiment 1 served as baseline measures to set a within-study benchmark for interpreting the C1 and P1 VEP component modulations observed in experiments 2 and 3 of the current study. The baseline measure stimuli of experiment 1 are split into 2 separate sets of stimuli representing the 2 polar extremes of stimulus space. The stimuli employed in part I consisted of vertical sinusoidal luminance gratings (i.e., highly narrowband in stimulus space) set to an SF of either 0.8 or 8.0 cpd. The stimuli employed in part II of experiment 1 consisted of 50 broad-band natural scene images used in Hansen et al. (2011) and naturally varied in terms of both structural complexity and amplitude spectrum slope (i.e., highly broad band in stimulus space). Specifically, 25 images were quantified as having relatively few edges/lines (i.e., low structural complexity) and the other 25 as being dense with edge/line content (i.e., high structural complexity). Within each structural complexity set, there were 5 different images grouped according to their average amplitude spectrum slope. The average amplitude slope within the 5 different sets within each structural complexity set was as follows: mean = −0.76, SD = 0.020; mean = −0.87, SD = 0.018; mean = −1.04, SD = 0.025; mean = −1.24, SD = 0.035; and mean = −1.44, SD = 0.029 (the exact method for quantifying amplitude spectrum slope and structural complexity can be found in Hansen and Hess 2006 and 2007, respectively). Finally, it is worth noting that all natural scene stimuli (as well as the sinusoidal grating stimuli) contained approximately equivalent amounts of structural content across the spatial extent of each image.

Stimulus examples are shown in Fig. 1A.

For experiment 2, the stimuli consisted of a subset of natural scene images (n = 30) used in part II of experiment 1. Specifically, all images in the −0.76, −1.04, and −1.44 amplitude spectrum slope groups were used here with half of those in the low structural complexity set and the other half in the high structural complexity set. All 30 scene stimuli were then filtered in the Fourier domain to possess a fairly narrow band (i.e., 1 octave measured with respect to full-filter-width at half-filter-height) of SF content centered on either 1 or 8 cpd of visual angle, resulting in a total of 60 stimuli (30 filtered at 1 cpd and 30 filtered at 8 cpd). Details regarding the filter itself can be found in Hansen and Hess (2007). An example stimulus from each of the slope groups within each of the structural sparseness sets filtered at 1 cpd is shown in Fig. 1B, with Fig. 1C showing the same examples filtered at 8 cpd. It is important to note that differences in amplitude spectrum slope for the different band-pass-filtered conditions essentially translate to differences in rms contrast for band-pass-filtered stimuli (i.e., steeper amplitude spectrum slopes yield higher contrast images in the 1-cpd condition, whereas shallower slopes yield higher contrast images in the 8-cpd condition). That is, we did not attempt to hold rms contrast at a fixed level but allowed it to vary naturally with filter output (that is, rms contrast of the filtered stimuli was completely determined by the slope of the amplitude spectra of the original imagery).

For experiment 3, the stimuli consisted of a subset of the broadband stimuli (n = 10) employed in experiment 2 of the current study. Specifically, all stimuli from the −1.04 average amplitude spectrum slope set were selected, with half (n = 5) from the low structural complexity set and half (n = 5) from the high structural complexity set. Each of the 10 stimuli were low-pass filtered in the Fourier domain (at each of 5 different cutoff frequencies, 1, 2, 4, 8, and 16 cpd), resulting in a total 50 stimuli altogether. The filter cutoff was not hard-limited but instead consisted of a gradual Weibull function taper. An example stimulus from each low-pass filter cutoff for the low-complexity set is shown in Fig. 1D, top, with bottom showing the same examples from the high-complexity set.

**Procedure.** The experimental design for all experiments consisted of a standard block paradigm, with all stimuli viewed binocularly (the task was passive viewing). Each trial began by presenting a given participant with a mean luminance blank screen (500 ms) followed by a stimulus interval (500 ms), resulting in a total trial time of 1,000 ms. That is, all experiments investigated pattern-onset VEPs (with abrupt onset and offset of stimuli). All stimulus conditions within each experiment were randomly interleaved across EEG recording sessions. The total number of trials for each condition within each experiment well exceeds the minimum standard established by the International Society for Clinical Electrophysiology of Vision (ISCEV; Odom et al. 2010).

For experiment 1, part I, the total number of trials was 200 (100 trials for each grating SF) with 100 trials per recording session. For experiment 1, part II, there were 10 broad-band scene conditions (i.e., 2 levels of structural sparseness × 5 different amplitude slopes). The 10 conditions produced 1,000 total trials, with each of the 5 stimuli in each condition presented 20 times (100 trials per condition) with 10 recording sessions per participant (100 trials per recording session).

In experiment 2, there were 12 band-pass scene conditions (i.e., 2 central SFs × 2 levels of structural complexity × 3 different amplitude slopes). The 12 conditions resulted in 1,200 total trials with each of the 5 stimuli in each condition presented 20 times (100 trials per condition) with 12 recording sessions per participant (100 trials per recording session).

In experiment 3, there were 10 variable low-pass scene conditions (i.e., 2 levels of structural complexity × 5 different low-pass cutoff frequencies). The 10 conditions produced 1,000 total trials with each of the 5 stimuli in each low-pass cutoff condition presented 20 times (100 trials per condition) with 10 recording sessions per participant (100 trials per recording session).

**EEG recording and analysis.** Continuous EEGs were recorded in a Faraday and anechoic (sound- and echo-proof) chamber using the EGI acquisition system. All EEGs were obtained by means of a Geodesic Sensor Net consisting of a dense array of 128 channels (electrolytic sponges). The online reference was at the vertex (Cz), and the impedances were maintained below 40 kΩ (Tucker 1993). All EEG signals were amplified and analog band-pass-filtered from 0.1 to 100 Hz. The signal was digitized at 250 Hz and digitally filtered with a 0.1- to 50-Hz band-pass filter to reduce 60-Hz noise. All EEGs were subjected to algorithmic artifact rejection of voltage exceeding ±70 µV or transients greater than ±70 µV. Continuous EEGs were sampled from electrode Oz (as defined by the international 10–20 system), consistent with previous work investigating the C1 and P1 VEP components (e.g., Ellemberg et al. 2001; Kubová et al. 1995; Murray et al. 1987; Previc 1988). It is worth noting that optimized
Electrode vectors were assembled for each participant (see Pernet et al. 2011 and Rousselet and Pernet 2011 for further details) with all analyses conducted on electrodes accounting for the most variance in accordance with each experimental design and resulted in virtually identical results to those obtained from Oz. Thus, for the sake of simplicity, we only report the results of analyses conducted on Oz.

All continuous EEGs were divided into 1,000-ms epochs. Each epoch was rereferenced offline to the corresponding averaged mastoid epoch and baseline-corrected to the last 100 ms of the luminance blank-trial interval. Grand-average VEPs were then assembled by averaging all rereferenced and baseline-corrected epochs associated with each stimulus condition across participants. Additionally, topo-

---

**Fig. 1.** A: example natural scene stimuli from experiment 1 (see EXPERIMENTAL PROCEDURES). Each row indicates stimuli coming from the 2 different levels of structural complexity. The columns show imagery from the different amplitude slope conditions (marked at the top of the figure). B: example stimuli from the low-spatial frequency (SF; band-pass-filtered with a peak SF at 1 cycle per degree (cpd)) condition of experiment 2. The organization of the stimulus examples is identical to A. C: example stimuli from the high-SF (band-pass-filtered with a peak SF at 8 cpd) condition of experiment 2. The organization of the stimulus examples is identical to A. D: example stimuli from experiment 3. The top row is an example image from the low structural complexity set, and the bottom row is an example from the high structural complexity set. Each column shows the same stimulus filtered with a different low-pass SF cutoff. Note that larger cutoff SFs pass increasingly higher SFs into the image.
graphic plots were generated for all experimental conditions using EEGLAB (Delorme and Makeig 2004) version 8.0.3.5 in MATLAB (version R2011b). Last, the dependent variable entered into all statistical analyses consisted of the mean component magnitude assessed by averaging across all PSOT points spanning the width (from component onset to offset) of a given grand-average VEP component for each participant separately. That is, for each analysis, an analysis window was defined by the average width of a set of grand-average VEP components. Once established, VEP data was sampled from each participant within that analysis window and then averaged (again, on a participant-by-participant basis).

RESULTS

Experiment 1: benchmark measures. Experiment 1 established a within-study benchmark for interpreting the observed results in experiments 2 and 3. Grand-average VEPs from experiment 1, parts I and II, are shown in Fig. 2. In Fig. 2A, the first positive peak (i.e., P1) is seen in both SF conditions with the majority of the component width between ~100- and ~160-ms PSOT. The first negative peak (i.e., C1) is only seen in the high-SF condition with the majority of the component width observed between ~90- and ~110-ms PSOT. The appearance of the C1 with high-contrast 8.0-cpd grating stimuli and the P1 with both high-contrast 0.8- and 8.0-cpd grating stimuli is consistent with previous reports (Ellemberg et al. 2001; Hansen et al. 2011; Vassilev et al. 1994, 2002). Topographic plots for part I are shown in Fig. 3A. Figure 2B shows a large P1 component between ~100- and ~148-ms PSOT for broadband natural scenes low in structural complexity with the magnitude of that component modulated by amplitude spectrum slope. The modulation was verified with a one-way repeated-measures ANOVA on the averaged component magnitudes within the 100- to 148-ms PSOT window and revealed a significant main effect of amplitude spectrum slope, $F_{(4,32)} = 9.48, P < 0.001$, partial $\eta^2 = 0.54$, and a significant linear trend, $F_{(1,8)} = 13.61, P < 0.01$. That is, images with steeper amplitude spectra slopes (i.e., possess more contrast at lower
SFs compared with scenes with shallower amplitude spectra slopes) produce larger P1 magnitudes. Note that here the C1 component is virtually nonexistent in all VEPs.

However, for natural scene imagery dense in edge/line content (Fig. 2C), the dominant P1 gives way to a significant negativity between ~120- and ~160-ms PSOT that is also significantly modulated by amplitude spectra slope, $F_{(4,32)} = 6.98, P < 0.001$, partial $\eta^2 = 0.47$, and yielded a significant linear trend, $F_{(1.8)} = 9.46, P < 0.05$. Thus stimuli with shallower amplitude spectra slopes (i.e., possess more contrast at higher SFs compared with scenes with steeper amplitude spectra slopes) yielded a larger negativity ~150-ms PSOT. Averaged component magnitudes within the 100- to 148-ms PSOT window for each amplitude spectrum slope in the low-complexity condition are shown in Fig. 2D and within the 120- to 160-ms PSOT window for the high-complexity condition shown in Fig. 2E. The data in Fig. 2D largely replicate those reported by Hansen et al. (2011). However, Hansen and colleagues (2011) did not observe a statistically significant modulation of the negativity near ~150 ms with amplitude spectrum slope, possibly due to their participant sample size (which is almost twice as large in the current study). Topographic plots for part II are shown in Fig. 3B.

Experiment 2: band-pass-filtered scenes. Grand-average VEPs from experiment 2 are shown in Fig. 4. Topographic plots for several conditions can be found in Fig. 5. Starting with the 1.0-cpd (low SF) conditions (Fig. 4, A and B), regardless of stimulus complexity, the first positive component (i.e., P1) occurs between ~100- and ~160-ms PSOT: A 2 (structural complexity) $\times$ 3 (amplitude spectrum slope) two-way repeated-measures ANOVA on the averaged component magnitudes within the 100- to 160-ms PSOT window did not support a main effect of structural complexity or amplitude spectrum slope nor was there a significant interaction ($P > 0.05$). Averaged component magnitudes within the 100- to 160-ms PSOT window for each amplitude spectrum slope and level of complexity are shown in Fig. 3C. Again, note that due to the band-pass nature of the stimuli, differences in amplitude spectrum slope translate to subtle changes in rms contrast. We next compare the results reported in Fig. 4, A and B, with the 0.8-cpd vertical grating condition reported in experiment 1.

Since amplitude spectrum slope (i.e., rms contrast differences) did not yield any significant main effects, we collapsed VEPs across amplitude spectrum slope within the low-complexity and high-complexity conditions with all VEPs shown in Fig. 4D. The slight differences in P1 magnitude between the 0.80-cpd grating conditions and the two 1-cpd band-pass-complexity conditions were not found to be statistically significant. Thus it seems that the addition of content at all orientations, regardless of structural complexity, does not significantly modulate the P1. The lack of modulation by rms differences (corresponding to amplitude spectrum slope differences) suggests that the P1 modulation reported in Fig. 2B was not the result of interactions within neural populations tuned to low SFs. We will return to this issue in experiment 3.

For the 8.0-cpd (high SF) conditions (Fig. 4, E and F), regardless of stimulus complexity, a subtle negative component (i.e., an apparent C1) occurs between ~100- and ~130-ms PSOT. Again, the same analysis mentioned above was conducted on the averaged component magnitudes within the 100- to 130-ms PSOT window did not support a main effect of structural complexity or amplitude spectrum slope nor was there a significant interaction ($P > 0.05$). From the data shown in Fig. 4, A and B, it appears that neither the C1 nor the P1 is significantly modulated by stimulus structural complexity or amplitude spectrum slope. However, structural complexity (but not amplitude spectrum slope) was found to modulate significantly, $F_{(1.8)} = 6.28, P < 0.05$, partial $\eta^2 = 0.44$, the waveform between 148- and 180-ms PSOT, which corresponds to a second negativity (i.e., an N1 component). The same N1 component was also apparent in the low-SF conditions (as a negative going positivity between ~150 and ~180 ms) but is not significantly modulated by structural complexity (or amplitude spectrum slope). Thus the N1 component appears to be linked with stimulus structural complexity but only when that structure is limited to higher SFs. A clearer picture can be gleaned by examining the data shown in Fig. 4, E and F, against the 8.0-cpd vertical grating condition VEP reported in experiment 1 (shown in Fig. 4I). Again, since there was no main effect of
amplitude spectrum slope, the VEPs have been collapsed across amplitude spectrum slope within each structural complexity condition. From Fig. 4, it is clear that the magnitude of the N1 becomes increasingly large as structural complexity moves from a single orientation high-SF grating to spatially dense multiple high-SF orientation natural image patterns. We will return to that issue in Discussion.

Experiment 3: variable low-pass-cutoff-filtered scenes. Grand-average VEPs from experiment 3 are shown in Fig. 6. Topographic plots for several conditions are shown in Fig. 7. Starting
with the low structural complexity conditions (Fig. 6A), regardless of the low-pass SF cutoff, the first positive component (i.e., P1) occurs between ~100- and ~148-ms PSOT. A one-way repeated-measures ANOVA on the averaged component magnitudes within the 100- to 148-ms PSOT window revealed a significant main effect of SF cutoff, $F_{(4,32)} = 7.97, P < 0.001$, partial $\eta^2 = 0.50$, as well as a significant linear trend, $F_{(1,8)} = 9.23, P < 0.05$. Thus, as increasingly high SFs are included in the stimuli, the magnitude of the P1 decreases. Averaged component magnitudes within the 100- to 148-ms PSOT window for each SF cutoff are shown in Fig. 6C. The C1, typically observed around 60- to 100-ms PSOT, is not apparent in any of the VEPs for this condition. Interestingly, the downward deflection observed between ~150 and 180 ms in experiment 2 (i.e., the N1) was again observed in this condition and was significantly modulated in a similar way by SF cutoff, $F_{(4,32)} = 10.71, P < 0.001$, partial $\eta^2 = 0.90$. However, since there is no clear differential modulation between that observed with the P1 and

Fig. 5. Topographic plots for grand-average VEPs from experiment 2; we are only showing plots from the amplitude spectrum slope ~1.04 condition for simplicity. Columns are topographic plots at different temporal intervals (10-ms steps) from 80- to 200-ms PSOT. The top 2 rows are from the low-SF condition with the bottom 2 rows corresponding to the high-SF condition. Electrode Oz is located centrally near the bottom of the skirt region outside the head outline. The scale of the color bar is in microvolts.

Fig. 6. Grand-average VEP data from experiment 3. A: VEP data from the low structural complexity image conditions. Each trace is the averaged VEP obtained at different low-pass SF cutoff frequencies. B: same as A but for imagery possessing a high degree of structural complexity. C: plots the average component magnitude (as described in Fig. 2) shown in A. Error bars are ±1 SE. D: same as C but assessed for data shown in B. The values on the ordinate of all VEP plots are in microvolts.
corresponding modulation observed with the later N1, we suspect that the N1 signal (while likely reflecting different neural mechanisms than P1) is covarying with the P1 signal (cf. Fig. 4I for clear differential modulation) for that condition. To verify this observation statistically, a one-way repeated-measures analysis of covariance (ANCOVA) was conducted on the averaged component magnitudes within the 152- to 180-ms PSOT window (i.e., the N1 component) with the averaged P1 component magnitudes (obtained by averaging across the 100- to 148-ms time window on a participant-by-participant basis) for each slope entered as covariates. The result was a nonsignificant main effect of amplitude spectrum slope \( (P > 0.05) \). Thus controlling for the P1 modulation by amplitude spectrum slope eliminated the amplitude spectrum slope modulation on the N1.

For the high structural complexity conditions (Fig. 6B), a one-way repeated-measures ANOVA on the averaged component magnitudes within the 100- to 160-ms PSOT window revealed a significant main effect of SF cutoff, \( F_{(4,32)} = 12.25, P < 0.001 \), partial \( \eta^2 = 0.61 \), and yielded a significant linear trend, \( F_{(1,8)} = 15.33, P < 0.01 \). The modulation shifts from a P1 for SF cutoffs at and below 4 cpd to the N1 observed in experiment 2 at and above 8 cpd (see DISCUSSION for further details). Averaged component magnitudes within the 100- to 160-ms PSOT window for each SF cutoff are shown in Fig. 6D.

Comparing the data shown in Fig. 6A with Fig. 2B (i.e., the low structural complexity conditions from experiment 1) reveals an almost identical pattern of results. Specifically, Fig. 2B shows that as the slope of the amplitude spectra increases from \(-0.76\) to \(-1.44\), the magnitude of the P1 VEP component increases. Since images with steeper amplitude spectra slopes have more physical contrast at lower SFs compared with images with shallower amplitude spectra slopes, the magnitude of that component appears dependent on luminance contrast at low SFs. However, given that similar contrast changes at low SFs did not seem to modulate reliably the P1 in experiment 2 (i.e., Fig. 4, A–D), the likely modulator may be linked to the presence of higher SFs. The current experiment (which used images with an amplitude spectrum slope in the middle of that range, i.e., \(-1.04\)) lends support for such a notion. That is, as higher SFs are included in a given image, the magnitude of the P1 component decreases almost proportionately, suggesting possible inhibitory influences from neural populations tuned to higher SFs on those tuned to lower SFs. Interestingly, a similar trend is apparent in Fig. 6B (i.e., the high structural complexity condition in the current experiment). However, the reduction in P1 magnitude is much more severe in Fig. 6B as higher SFs are included in the stimuli, so much so that as SFs at and above 8 cpd are included in the stimuli, the component inverts from positive to negative (and more closely resembles the VEPs shown in Fig. 2C). Given that the only difference between the two conditions in the current experiment is level of structural complexity, it appears as though the negativity observed between \(-120\) and \(-160\) ms is linked with both the overall complexity of the image in the spatial domain (i.e., edge/line density) and the amount of higher SFs present in the imagery. Furthermore, considering Fig. 6B along with Figs. 4I and 2C, it seems clear that the high-SF, high structural complexity–triggered negative VEP component between \(-120\)- and \(-160\)-ms PSOT (i.e., the N1) occurs independent of the amount of low-SF information present in the imagery.

### DISCUSSION

Recent VEP work (Hansen et al. 2011) has suggested a possible means by which the early visual mechanisms may operate when processing broad-band scene imagery by examining the modulations of the C1 and P1 VEP components produced when humans view such imagery. Specifically, for broad-band scenes low in structural complexity (i.e., possess few edges/lines in the spatial domain), the P1 is the dominant component and appears to code for SF luminance contrast. On the other hand, for scenes rich in edge/line content, an apparent C1 VEP component dominates, suggesting that high-SF infor-

---

**Fig. 7.** Topographic plots for grand-average VEPs from experiment 3. Columns are topographic plots at different temporal intervals (10-ms steps) from 80- to 200-ms PSOT. The top 3 rows are from the low structural complexity condition (showing 3 different SF cutoff conditions) with the bottom 3 rows corresponding to the high structural complexity condition (showing 3 different SF cutoff conditions). Electrode Oz is located centrally near the bottom of the skirt region outside the head outline. The scale of the color bar is in microvolts.
mation governs the visual signal. The implication being that the early visual system may signal for the presence of different natural scene statistics in a SF-specific manner with scene structural complexity regulating the dominance of signals arising from cortical populations tuned to different SFs. However, as mentioned in the Introduction, the natural scene stimuli employed by Hansen and colleagues (2011) were all broad band, leaving an open question pertaining to whether said phenomena resulted from interactions within or between neural populations selective for lower vs. higher SFs.

The results of the current study show that the modulation resulting from highly complex stimuli (compared with stimuli low in structural complexity) is predominantly carried by neural populations tuned to high SFs. However, by carefully examining the VEPs plotted in Fig. 4, E and F, it seems that the driving negativity is not the C1 (as defined in the Introduction) but the downward deflection observed between ~150- and ~180-ms PSOT (i.e., N1). Although there are numerous caveats when examining VEP component peak latency from grand-average waveforms, it is interesting to note that there does appear to be a latency difference between the C1 observed with vertical sinusoidal luminance gratings and the peak of the N1 (see Fig. 4). Specifically, although the C1 and P1 are apparent in the high band-pass-filtered stimulus conditions, the significant modulation is taking place ~40 ms later in the waveform (i.e., N1). Furthermore, when considering all conditions reported here where lower SFs were present in the stimuli (either in isolation or together with higher SFs), the C1 component is virtually absent from the VEP waveform. When factoring in that latter observation, it appears that the presence of low SFs within a stimulus is enough to eliminate the C1 component from the broad-band VEP, and as stimulus complexity increases (within a broad-band image), the P1 component drops out and is replaced by an N1 component near ~150-ms PSOT. However, an alternative account would suggest that the modulation observed between ~120- and 160-ms PSOT (in experiments 2 and 3) is arising from a significantly lagged C1 component (given its well-known dependency on high SFs), serving to dampen the P1 positivity and partially sum with the N1 negativity when stimulus complexity is high, resulting in a rather largely negative at later PSOTs. Specifically, given that the parvocellular associated C1 would be a component reflecting visual acuity, as more content is added to a stimulus, the more the neural processes associated with the C1 would have to resolve, resulting in a lagged signal. Although such an account is plausible, there is little empirical evidence to support it. Specifically, the C1 component has consistently been reported to occur between ~80 and ~100 ms (give or take 10 ms) for a large range of stimuli varying in SF bandwidth (i.e., single sinusoidal gratings to various broad-band visual patterns). Furthermore, the broad-band stimuli that have been shown to elicit a C1 component in posterior VEPs have ranged from simple line configurations (alternating in luminance or color contrast) to spatially complex check or checkerboard patterns (also alternating in luminance or color contrast) to human face stimuli (e.g., Baseler and Sutter 1997; Butler et al. 2007; Foxe et al. 2008; Mitsudo et al. 2011; Scheckter et al. 2005). Granted, Vassilev and colleagues (2002) have shown the C1 to be lagged by as much as 50 ms from the typical PSOT range for high-SF gratings, but that was only true when those stimuli were very low in Michelson contrast (~5 to 15%). The Michelson contrast in the high-SF range for all of the stimuli employed in the current study (whether “nested” within broad-band stimuli or presented in isolation) was always well above 30%, thereby making an argument for a lagged C1 less appealing. Nevertheless, it is interesting that the C1 is reliably observed for the types of stimuli mentioned above but largely absent from the VEPs in many of the conditions reported here using natural scene imagery. We are currently exploring this issue in detail.

Before moving forward, it is worth addressing the passive nature of the task the participants engaged in during all experiments reported above. Since the task involved the passive viewing of stimuli, it may be possible that the component modulation described above was partly driven by the influences of attention. That is, since some of the natural scene stimuli in the high-complexity conditions could be considered “more interesting” than those presented in the low-complexity conditions, it may be possible that participants attended more to the “interesting” scene stimuli. In terms of the C1 component, there exists substantial evidence (e.g., Clark and Hillyard 1996; Di Russo et al. 2003, 2012; Heinez et al. 1994; Martínez et al. 1999; Noesselt et al. 2002) that the C1 is not modulated by attention. However, several recent reports have suggested that the C1 can be modulated by attention (reviewed in Rauss et al. 2011). However, the modulation seems mostly limited to stimuli presented in the upper visual field (our stimuli were presented at the fovea and largely extended into both upper and lower visual fields) and seems to be highly dependent on interactions of perceptual and attentional load (e.g., Fu et al. 2010a,b). Furthermore, when attention modulation of C1 was observed, it was typically observed at electrodes different from the one focused on in the current study (i.e., Oz). On the other hand, there is much agreement that the P1 and N1 components can be modulated by directed attention (e.g., Clark and Hillyard 1996). However, excepting Clark and Hillyard (1996), who report a 0.5-μV attention modulation effect under directed attention for the N1 at electrode Oz (IPz in their montage), most studies report the modulation of either P1 or N1 at electrodes largely displaced from Oz (i.e., typically the laterally displaced temporal electrodes with some reports from electrodes displaced anteriorly from Oz) (e.g., Clark and Hillyard 1996; Di Russo et al. 2003; Hillyard et al. 1998; Martínez et al. 1999). This is important because there is no reason to believe that a P1 (or N1) component observed at electrodes largely displaced from Oz reflect the same underlying neural processes as those producing P1 or N1 components measured at Oz. Furthermore, consider the topographical plots shown in Figs. 3, 5, and 7. For all conditions, the P1 and N1 at Oz (visible centrally near the bottom of the topographic “skirt”) behave very differently as a function of different image statistics and spatial filtering from the VEP components displaced temporally or anteriorly from Oz. Furthermore, the modulation of either the P1 or N1 component reported to be driven by directed attention usually involves just an increase in peak magnitude of either of those components. Here, the critical finding reported is a complete inversion of component magnitude (an observed P1 to an N1 with the transition largely apparent in Fig. 6B) as a function of stimulus complexity and spatial scale, a modulation that has not been reported to result from directed attention. Last, our participants viewed the stimuli of varying amplitude spectrum slope and structural com-
plexity in an interleaved manner, without any precueing (i.e., in the absence of experimenter-induced directed attention). That is, the majority of studies investigating the influence of attention on the P1 or N1 obtain the modulation after directly cueing participants to shift their attention. It therefore seems unlikely that participants would systematically (and reliably) attend to images with different amplitude spectra slopes (i.e., the P1 modulation) in a graded fashion. The same can be said for the N1. That is, whereas the current study did not measure VEPs to stimuli with intermediate complexity in the current study, VEPs to said stimuli were recorded by Hansen et al. (2011, Fig. 12), who reported intermediate N1 magnitudes (intermediate to low and high structural complexity). Again, this would imply that participants were systematically and reliably attending in a graded fashion to stimuli with varying degrees of structural complexity in an interleaved and noncued paradigm (which seems unlikely).

Regarding the underlying source of the three components observed in the current study (i.e., C1, P1, and N1), there is overwhelming agreement that the C1 is largely derived from neural activity within primary visual cortex (see Di Russo et al. 2001, 2005 for reviews). On the other hand, the underlying source for the P1 component is largely debated, being argued to arise from either neural activity in primary visual cortex or extrastriate cortical areas. However, more recent source localization techniques (mostly functional MRI-constrained) have shown the earlier portion of the P1 to arise from more dorsal extrastriate areas with the later portion of that component arising from more ventral extrastriate cortical areas (e.g., Di Russo et al. 2001, 2005, 2012; Martínez et al. 2001; Vanni et al. 2004). The N1 observed in the current study (i.e., measured over the occipital pole) has been linked to the same neural sources as the P1 (e.g., Di Russo et al. 2001, 2012). It therefore seems that the apparent SF selectivity of the P1 and N1 observed in the current study may reflect extrastriate cortical mechanisms receiving predominant input from either the magnocellular or parvocellular visual pathways, respectively (see also Schechter et al. 2005 for a similar observation). Furthermore, the N1 (but not the P1) becomes the dominant component as stimulus structural complexity increases along with increasing amounts of high-SF content, lending further support for the purported underlying extrastriate sources as receiving predominant input from either the parvocellular or magnocellular visual pathways.

To conclude, based on the results from the current study, together with the source localization literature, it seems that the early representation of natural scenes in the human visual system involves the expression of different SF-tuned processes that are dependent on the edge/line density within natural scene imagery. That is, the P1 component (which is well-known to signal for low SFs) seems to be the primary component for signaling changes in the distribution of contrast across SF (i.e., changes in amplitude spectrum slope). However, it only seems to be able to do this (or at least is only measurable) when the stimuli are relatively limited in edge/line density and when there are both low and high SFs present within the stimulus. As edge/line density increases, the N1 component (peaking between ~150- and 180-ms PSOT) becomes the dominant signal and seems largely tied to higher SFs (higher than ~4 cpd). The implication for such a result suggests that the human visual system relies on at least three different processes when viewing stimuli with variable complexity. The first two correspond to lower-level C1- and P1-like processes for simple narrowband stimuli and higher P1-type processes for broad-band natural scenes sparse in structural content. The third process emerges as scene content becomes more dense and complex and is reflected by a later negativity (N1) that becomes dominant and is dependent on the presence of higher SF content. The role of such a component may be to emphasize structural content over general luminance contrast (as it was observed to be only minimally modulated by amplitude spectrum slope). It therefore seems as though when viewing natural scene imagery, the amount of structural complexity (particularly at high SFs) may regulate dominance between early visual components (differentially reflecting either magnocellularly or parvocellularly driven mechanisms) and possibly serve as a constraint on popular spatial channel selection models for visual recognition (e.g., Bar et al. 2006; Hughes et al. 1996; Morrison and Schyns 2001) as well as selective vs. nonselective pathway usage in visual search (Wolfe et al. 2011).

GRANTS
This study was supported by a Colgate Research Council Grant to B. C. Hansen, a Natural Sciences and Engineering Research Council (NSERC) Grant to A. P. Johnson, and NSERC and Canada Foundation for Innovation Grants to D. Ellemberg.

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

REFERENCES
Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure.


