Visual Neuroscience (2011), **28**, 371–379.
Copyright © Cambridge University Press, 2011 0952-5238/11 $25.00
doi:10.1017/S0952523811000162

Artistic representations: Clues to efficient coding in human vision

DANIEL J. GRAHAM[1,2] AND MING MENG[2]

[1]Department of Psychological Basic Research, University of Vienna, Vienna, Austria
[2]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire

(RECEIVED December 22, 2010; ACCEPTED April 10, 2011; FIRST PUBLISHED ONLINE May 16, 2011)

Abstract

In what ways is mammalian vision—and in particular, human vision—efficiently adapted to its ecology? We suggest that human visual artwork, which is made for the human eye, holds clues that could help answer this question. Paintings are readily perceived as representations of natural objects and scenes, yet statistical relationships between natural images and paintings are nontrivial. Although spatial frequency content is generally similar for art and natural images, paintings cannot reproduce the dynamic range of luminance in scenes. Through a variety of image manipulations designed to alter image intensity distributions and spatial contrast, we here investigate the notion that artists' representational strategies can efficiently capture salient features of natural images, and in particular, of faces. We report that humans perform near flawless discrimination of faces and nonfaces in both paintings and natural images, even for stimulus presentation durations of 12 ms. In addition, contrast negation and up-down inversion have minimal to no effect on performance for both image types, whereas 1/f noise addition significantly affects discrimination performance for art more than for natural images. Together, these results suggest artists create representations that are highly efficient for transmitting perceptual information to the human brain.

Keywords: Coding efficiency, Art perception, Face detection, Natural scene statistics, Cortex

Introduction

Humans appear to be a singular species in their ability to create and to understand representational pictures (Fagot et al., 2000; de Waal, 2001). What does this psychological innovation tell us about human vision? More specifically, could it be the case that artistic representations tap into fundamental efficient visual encoding strategies unique to the human brain? Here, we address these questions by investigating whether artwork is efficient at representing faces, a key class of visual stimulus to which the human brain devotes considerable neural resources (e.g., Haxby et al., 2000).

Investigations of human vision that employ art stimuli have recently emerged as an important research area (see Graham & Redies, 2010, for a review), much as researchers studying auditory processing, language, emotion, and other brain functions have long embraced investigations of music (see e.g., Patel, 2008 for an overview). This work follows on two decades of research concerning efficient visual processing of natural scenes (see e.g., Olshausen & Field, 2004) and proceeds in a similar spirit. Art is, moreover, especially relevant to the study of visual processing due to the fact that art is created to be seen by the human eye, much as classical visual stimuli (e.g., sine-wave gratings, depth illusions, etc.) are designed for their perceptual effect. As such, artworks can serve to bridge the existing gap between natural stimuli, which, though offering important insights into vision coding, are subject to sampling biases and are difficult to parameterize (see Felsen & Dan, 2005; Rust & Movshon, 2005; Pinto et al., 2008), and artificial stimuli, which, though easily parameterized, do not capture important nonlinear structure relevant to visual processing (Field, 1987, 1994). In addition, artworks often depict faces, which are common in the human visual environment but have yet to be considered in relation to possible efficient coding strategies in human visual cortex.

Here, we investigate whether artists' representational strategies constitute an efficient representation of perceptual data, with "efficiency" here taken to mean fluency or ease of processing (see e.g., Winkielman et al., 2003). If this is the case, we suggest such efficient representations could themselves relate to neural coding strategies employed in human visual cortex. That is, efficient artistic representations of faces, objects, and scenes could take advantage of neural coding schemes that employ similar representations. However, we restrict our present investigations to the question of whether artistic representations are efficient at eliciting rapid detection of faces, and we examine a number of image transforms that help to narrow down the array of possible representational strategies that could allow this efficiency.

At present, far less is known about the common structure of visual art and its possible relation to basic mechanisms of perception, compared to what is known about natural scenes (see Geisler, 2008, for a recent review of the natural scene literature). In past work, art images have been shown to possess statistical regularities that may be connected with visual processing (Graham & Field, 2007, 2008*c*; Redies et al., 2007*a*,*b*; Graham et al., 2010*a*,*b*; Graham & Meng, 2011). For example, it has been found that, like natural scenes, both

Address correspondence and reprint requests to: Daniel J. Graham, Department of Psychological Basic Research University of Vienna, Liebiggasse 5, Vienna 1010, Austria. E-mail: artstats@gmail.com

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

representational and nonrepresentational art images share many statistical regularities including (1) scale-invariant spatial frequency amplitude spectra[1] (Graham & Field, 2007, 2008c; Redies et al., 2007*b*); (2) disproportionate amounts of spatial frequency amplitude at cardinal orientations (Koch et al., 2010; Graham & Meng, 2011); and (3) similar higher order spatial statistics (Hughes et al., 2011). These findings support the notion that efficient neural coding strategies that take advantage of the same regularities in natural scenes could shape the basic structure of human artwork (Graham & Field, 2008*c*). However, paintings cannot capture the large dynamic range of luminances typical in natural scenes, necessitating nonlinear scaling (Graham & Field, 2008*a*; Graham et al., 2009). Moreover, artists have a great deal of freedom to modify scene structure. In this regard, representational art is of particular interest because it is designed to capture natural scenes and objects—and especially, faces in the form of portraits—for human viewing.

We present a suite of tests aimed at measuring human performance at rapid face/nonface discriminations in art and natural images. A host of experiments stretching back decades has demonstrated that humans can make extremely rapid judgments of scene content (*O* ~50 ms; see, e.g., Potter & Levy, 1969; VanRullen & Thorpe, 2001), and distinct neural processing signatures for targets (e.g., animals and faces) *versus* distractors can be detected using Electroencephalography within 150 ms (see e.g., Thorpe et al., 1996). Much work in this area has been devoted to uncovering the "diagnostic" features that are necessary for fast recognition of the "gist" of an image (Torralba & Oliva, 2003; Oliva, 2005; Li et al., 2007; Meng & Potter, 2008). The current paper advances this line of inquiry by extending these studies to painted representations in order to assess efficiency, and we employ a variety of image manipulations that disrupt selected features in the images.

In Experiment 1, we evaluate human performance at detecting faces in natural images compared to detecting faces in paintings using the original images (OI). If artistic representations efficiently capture face features necessary for fast recognition, despite artists' efforts to scale scene luminances and their freedom to modify scene structure, we would expect face detection performance in art to at least match that for natural images. Next, in Experiment 2, we test performance following a variety of image manipulations. These manipulations have been studied in detail by many investigators and were chosen in order to selectively alter different image properties. In particular, we manipulate local and global intensity and contrast using contrast negation (CN), up-down inversion (UD), 1/f noise addition (NO). The typical effects of the image manipulations on spatial frequency amplitude spectra, intensity distributions, and local contrast in images are displayed in Table 1. We propose that artists' representational strategies permit fast recognition, and therefore, we expect that global manipulations such as CN and UD will produce little decrease in detection performance for art images. Also, if local contrast scaling is important for achieving efficient representations, 1/f noise addition should lead to a relative decrease in performance for art images since amplitude spectra of art images are typically somewhat steeper than those of natural images. Finally, in Experiment 3, we test the effect of combinations of these manipulations on performance: noise plus contrast negation (NOCN) and noise plus up-down inversion (NOUD).

[1]The amplitude spectrum is the square root of the power spectrum, and its rotationally averaged slope on log-log axes is half that of the power spectrum.

General methods

Participants

Tests of detection performance were conducted using the same source images, with a different set of participants viewing each test condition. No participants were tested on more than one condition. All participants were undergraduate students at Dartmouth College with normal or corrected to normal vision and were naive as to the purpose of the experiment. Each participant received course credit for participation.

Apparatus

Images were displayed on a luminance-calibrated 17-inch cathode ray tube monitor (1152 × 870, 85 Hz). Routines from the Psychophysics Toolbox (Brainard, 1997) were used to present images with desired timing. The laboratory room was darkened during testing.

Stimuli

Thirty art images (15 portraits and 15 landscapes) were selected from the Herbert F. Johnson Museum of Art, Cornell University and from the Web Gallery of Art (www.wga.hu). Portraits were required to show frontal to ¾-view faces and were cropped and scaled to isolate the face. Landscapes were required to be devoid of human forms. For natural images, we selected 15 scenes from the McGill Database of Natural Images (Olmos & Kingdom, 2004) and 15 uncompressed face images from flickr.com that showed frontal to ¾-view faces in natural lighting, which were cropped and scaled to isolate the face.

Images of pixel dimension 400 × 512 and 200 × 256 were tested separately, though results were pooled across the two image sizes because it was found that this variability produced no significant differences in performance. Red-green-blue images were converted to gray scale intensities (using the YIQ transform: Intensity = 0.299 * R + 0.587 * G + 0.114 * B) and were then linearly normalized such that their mean pixel intensity was 127. The set of art images showed no significant difference from the natural images in the standard deviation of their intensities following normalization (mean s.d. of intensities: artwork = 56.3; natural images = 49.1). In order to preserve statistical regularities in image amplitude spectra that have been associated with face/nonface discriminations (Honey et al., 2008), we did not normalize the amplitude spectra between images containing faces and those not containing faces.

In Experiments 2 and 3, image manipulations were performed as follows: contrast negated images were created by linearly inverting intensity values (subtracting each pixel value from 255); noise addition consisted of adding a different 1/f noise pattern to each image (for each noise pattern, the spatial frequency amplitude spectrum falls as 1/f, and intensities are normally distributed around 127); and for images with both noise addition and CN, noise was added before CN. All manipulations preserved a mean pixel intensity of 127.

Procedure

The goal of all of the experiments was to compare performance for face/nonface discriminations in artwork and in natural images as a function of stimulus duration. We compared performance following the image manipulations described above. Each participant was randomly assigned to one of six experimental conditions,

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

Table 1. *The typical effect of image manipulations (compared to OI) on spatial frequency amplitude spectrum slope, intensity histogram shape, and local contrast*

	Experiment 2			Experiment 3	
	CN	UD	NO	NOUD	NOCN
Amplitude spectrum slope	No change	No change	Flatter	Flatter	Flatter
Intensity histogram	Inverted	No change	More Gaussian	More Gaussian	Inverted and more Gaussian
Local contrast	Equivalent up to 180 deg phase shift	No change	Lower	Lower	Lower

Here, local contrast denotes, for example, the response of a set of oriented Gabor functions at a given scale.

each testing a single condition. See Fig. 1 for examples of the stimuli.

During testing, images were centered on the screen and surrounded by a midgray background. Images subtended approximately 10 deg of visual angle. A fixation cross (duration = 1 s) was presented before each trial, and a mask consisting of 1/f noise was presented for 100 ms immediately following each trial. Within a given condition, each image was shown once to each participant in random order with a randomly assigned stimulus duration (12, 24, 36, 48 ms) such that over all participants within a given condition, combinations

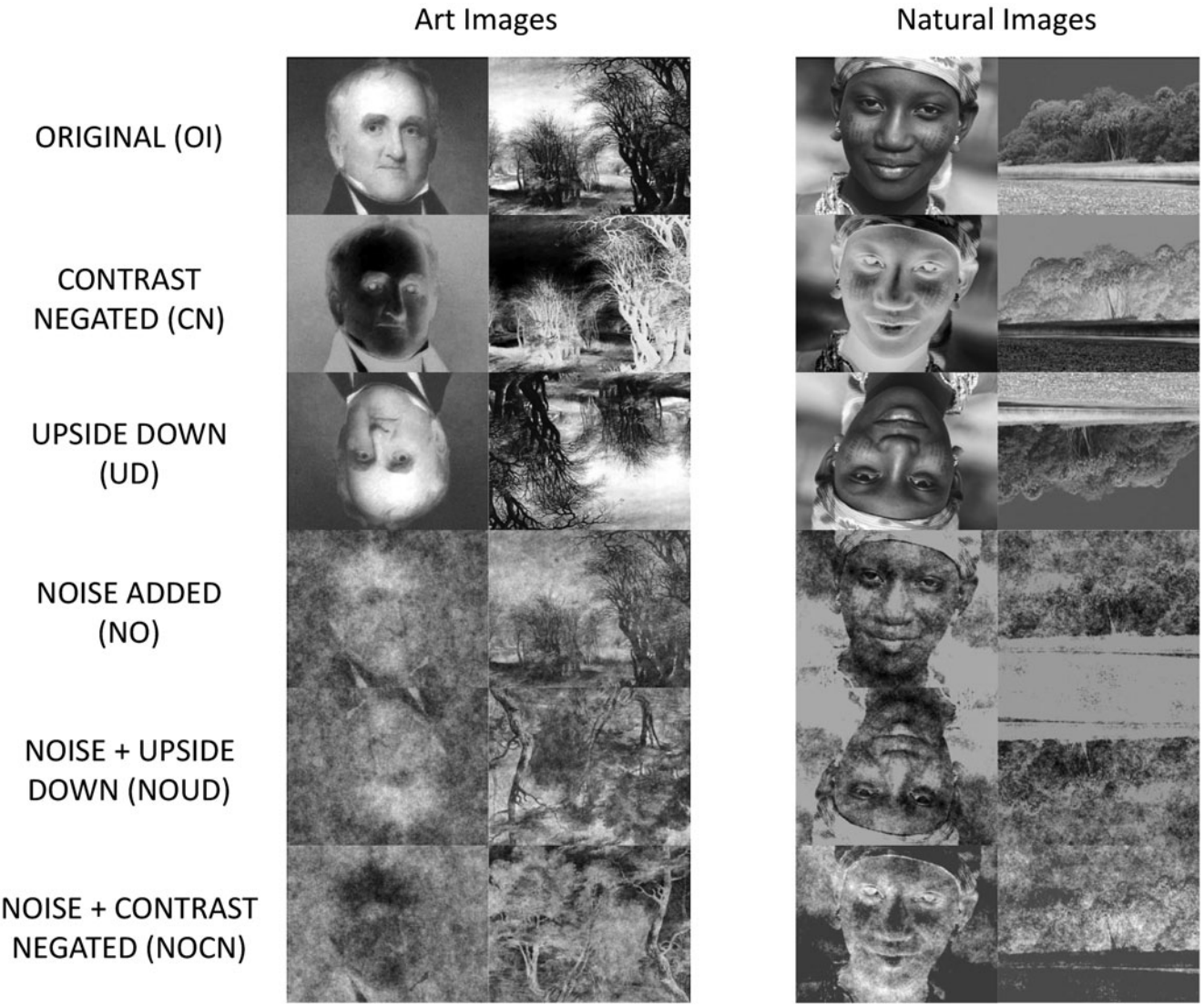


Fig. 1. Examples of art images and natural images used in this study, along with the effect of image manipulations. Image credits (left to right): Henry Inman, *Portrait of Benjamin Mumford* (1830), Gift of Mr. and Mrs. Quinto Maganini Accession, Herbert F. Johnson Museum of Art, Cornell University; Denis van Alsloot, *Winter Landscape*, (1610), Musée du Louvre, Paris; flickr.com; and McGill Calibrated Colour Image Database (http://tabby.vision.mcgill.ca/).

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

of images and durations were adequately sampled. Because these durations required single-frame reliability, stimulus timing accuracy was confirmed *via* software checks and through tests of the display measuring voltage changes across a photodiode as a function of time.

Experiment 1

Methods

Images without manipulation (OI) were tested on 16 participants using procedures described above.

Results and discussion

Participants performed at ceiling for the OI condition across all stimulus durations (art percent correct = 98.3% s.e. = 0.0059; natural image percent correct = 97.5% s.e. = 0.0068), with all d' values in this condition 3.0 or greater. There were no significant differences in performance at the durations tested (*t*-test: $P > 0.10$). See Figs. 2A and 3A.

These results show that faces can be detected in our diverse but biased corpora of artistic and natural images with high accuracy at extremely brief stimulus durations, even those as short as 12 ms. This is a rather striking finding in at least two respects. First, and most significantly, faces in art images can be detected with the same accuracy, despite the fundamental differences in the manner of creation of artistic and natural images. We note that although there are significant differences in spatial frequency content between art and natural images (see Table 2), performance is essentially the same for the two groups. Second, these results confirm the fact that faces in artwork and in natural images can be reliably detected during extremely brief presentations in the absence of color information.

Experiment 2

Methods

Image sets were manipulated as described above in order to achieve CN, UD, and 1/f noise addition (NO) and were tested on separate sets of participants (CN: 14 participants; UD: 19 participants; NO: 18 participants).

Results and discussion

CN condition

Performance (percent correct) was at ceiling for the CN condition (art percent correct = 96.6% s.e. = 0.0079; natural image percent correct = 95.7% s.e. = 0.0012), with all d' values in this condition 3.0 or greater. There were no significant differences (*t*-test, $P > 0.10$) in performance between art and natural images at any stimulus durations tested in the CN condition. See Figs. 2B and 3B.

Because CN drastically changes intensity distributions, but not amplitude spectra or phase relationships, the observed near flawless performance for the CN condition suggests that modifications of the global intensity distribution may not be crucial for rapid face detection. While measuring and characterizing the precise nature of

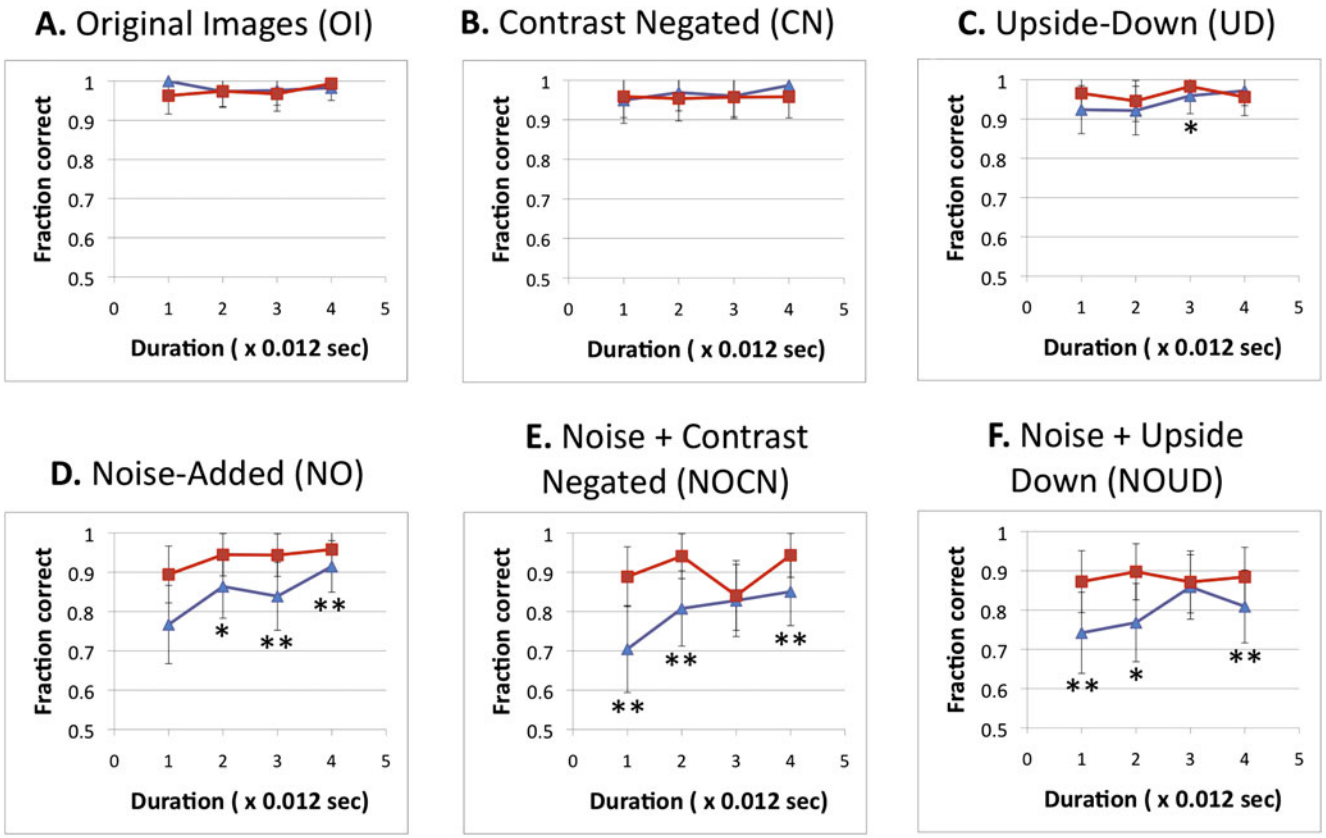


Fig. 2. Performance across experimental conditions (fraction correct). Squares (red) represent natural images and triangles (blue) represent art images. * Indicates significant difference in performance (using *t*-test) for given duration at $P < 0.10$; ** indicates significance at $P < 0.05$.

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

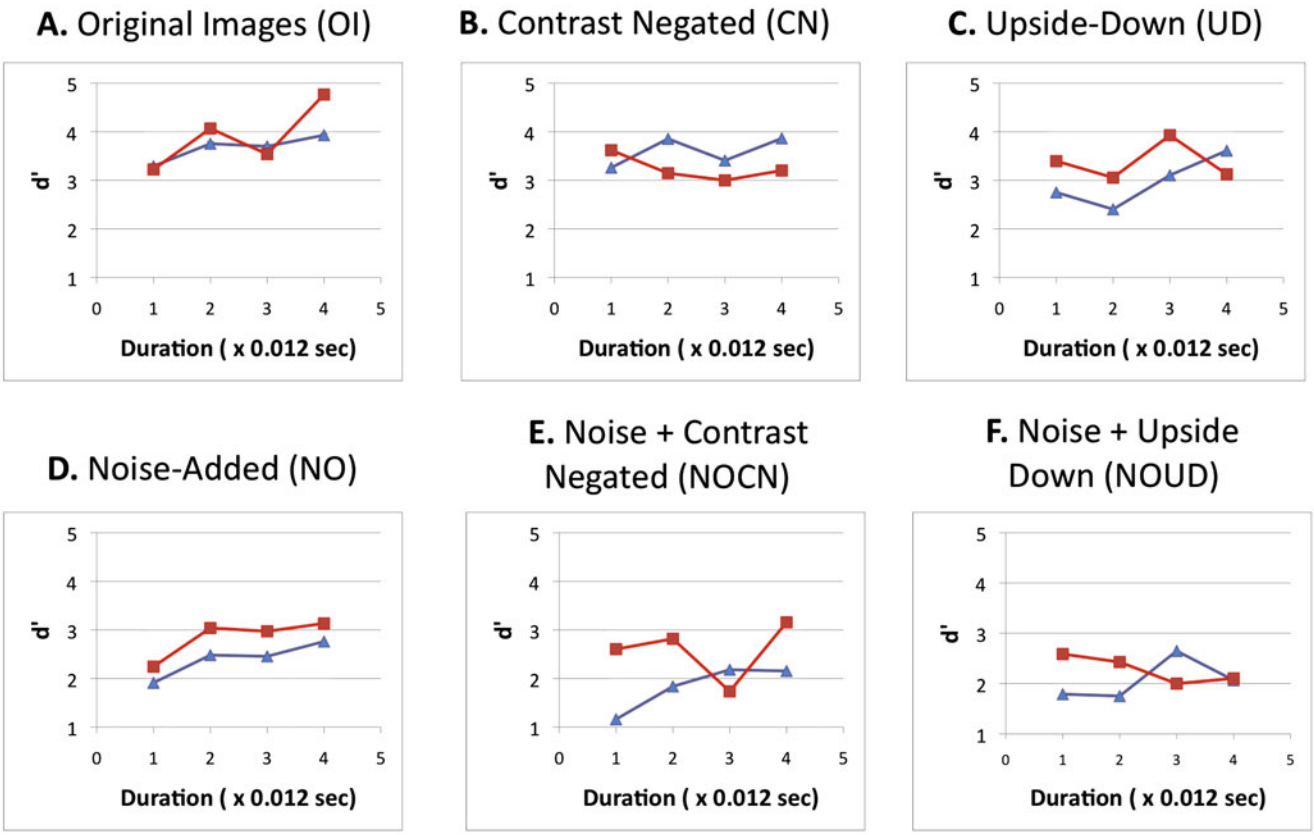


Fig. 3. Values of d' across experimental conditions. Squares (red) represent natural images and triangles (blue) represent art images.

artistic luminance transforms is beyond the scope of the present article, and while not all images in our corpora are luminance calibrated (only the Cornell and McGill sets are known to be linear with respect to luminance), it is clear that a wholesale reordering of the intensity distribution (subsequent to possible film gammas applied during image acquisition for the uncalibrated images) is not sufficient to decrease performance for art images relative to natural images. But local intensity relations are preserved in CN, up to a phase shift. That is, the scaling achieved by CN preserves local contrast gradients and only changes their sign. Local intensity relationships, presumably preserved by artists using local luminance scaling strategies, could thus be highly informative about salient or "diagnostic" contours in faces.

UD condition

For the UD condition, mean performance (percent correct) for artwork was 94.4% s.e. = 0.013 and mean performance for natural images was 96.2% s.e. = 0.0079. There was a small difference in performance at stimulus durations of 36 ms that was significant at the $P < 0.10$ level (art performance = 95.9% s.e. = 0.046, natural image performance = 98.3%, s.e. 0.030). All other durations showed no significant differences in performance ($P > 0.10$). However, d' values for this condition suggest that performance for artwork is lower than that for natural images. In particular, art images show a d' of less than 3.0 for stimulus durations of 12 and 24 ms, while d' values at these durations for natural images are above 3.0. Together with the raw performance results (percent correct), this finding suggests that UD disrupts face detection in art images modestly more than it does in natural images. See Figs. 2C and 3C.

In the case of UD, performance may be degraded for many of the same reasons as those that degrade performance for face photographs, namely the altering of global configural information. The finding that art is less robust to this manipulation (as indicated by d' values below 3.0 for art at shorter durations and at least 3.0 for natural images at all durations) can be viewed in at least two ways. First, the lower d' values could be due to delays or feedback from face-selective cortical areas that receive conflicting information derived from higher areas and from bottom-up cues. Alternatively, artists' local scaling strategies could be less robust to UD compared to natural images because of directional biases in these strategies—that is, because paintings are not generally subject to UD during the painting process, "encoding" informative contours could be orientation dependent. In any case, the differences in performance are relatively small, suggesting that global orientation is less critical to fast recognition compared to local contrast relations (in agreement with Gilad et al., 2009).

Table 2. *Spatial frequency amplitude spectrum slopes for art and natural images containing faces and not containing faces*

Amplitude spectrum slope	Faces	Nonfaces
Art images	−1.75*	−1.42*
Natural images	−1.63	−1.31

Slopes are calculated *via* rotational averaging of the two-dimensional amplitude spectra and linear fitting on log-log axes.
*Indicate significant differences (*t*-test, $P < 0.05$) in slopes between art and natural images.

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

NO condition

The NO condition resulted in mean performance (percent correct) of 84.6% (s.e. = 0.031) for artworks and 93.5% (s.e. = 0.014) for natural images. Performance was significantly lower (*t*-test) for art images at the $P < 0.10$ level for stimulus duration of 24 ms and at the $P < 0.05$ level for durations of 36 and 48 ms. Values of d' were also lower by at least 0.33 across all durations for art images and d' increased monotonically as a function of stimulus duration for both art and natural images. These results indicate that 1/f noise addition causes performance to approximate a typical psychometric function for both art and natural images, with significant decrements in performance for artwork at most durations. See Figs. 2D and 3D.

For images with added 1/f noise (NO), we are disrupting global intensity and spatial statistics as well as local intensity and spatial statistics. Adding noise of this sort (which has a Gaussian intensity distribution) tends to make intensity distributions more Gaussian, and it generally makes the amplitude spectrum of natural images more flat (since natural scenes typically scale as $1/f^{1.2}$). Because 1/f noise is scale invariant in terms of contrast energy, it affects contrast information across spatial scale, though it may have a greater effect at high spatial frequencies, where the signal-to-noise ratio is lower at the outset (see Atick & Redlich, 1992). Since art images in our sample have proportionally less high frequency content at the outset (i.e., steeper spectra, see Table 2), the effect of noise would be greater compared to natural images. That is, the loss of a proportionally greater amount of spatial frequency information at high frequencies in art after noise addition may lead to a consistent degradation of discrimination across a range of stimulus presentation times.

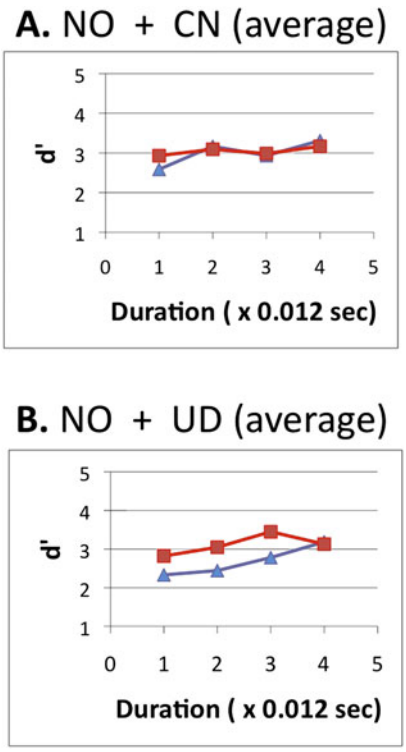


Fig. 4. Values of d' averaged between the NO and UD conditions (**A**) and NO and CN conditions (**B**). Squares (red) represent natural images and triangles (blue) represent art images.

Experiment 3

Methods

We tested detection performance for images with two combinations of the image manipulations tested above: NOCN (17 participants) and NOUD (18 participants).

Results and discussion

Mean performance across stimulus duration (percent correct) in the NOCN condition was 79.8% (s.e. = 0.032) for art images and 90.4% (s.e. = 0.024) for natural images (Fig. 2E). In the NOUD condition, mean performance across stimulus duration was 79.4% (s.e. = 0.025) for art images and 88.1% (s.e. = 0.006) for natural images (Fig. 2F). Performance was significantly lower for art images ($P < 0.05$) at 12, 24, and 48 ms in the NOCN condition and at 12 and 48 ms in the NOUD condition. Performance at 24 ms in the NOUD condition was significantly lower for art images at the $P < 0.10$ level. Values of d' were lower for art images at most stimulus durations (Fig. 3E and 3F).

Trends in d' values indicate that the effects of these manipulations add nonlinearly. Taking the average of d' for the image manipulations separately (averaging d' for NO and CN conditions and for NO and UD conditions, see Fig. 4), one may expect that overall performance would be essentially the same for art and natural images in the NOCN condition but lower for art images in the NOUD condition. Our results do not fully agree with these expectations. Contrary to what a linear average model would predict, noise combined with CN decreases d' performance for art more than for natural images. One possibility is that the NOCN manipulation could interfere with artists' local representational strategies primarily through disrupting their representation of "diagnostic" phase relations in salient contours (since the NOCN manipulation first alters phase relations across the image and then gives them a uniform 180 deg phase shift). Alternatively, natural images are less affected by this manipulation perhaps because they simply possess a greater number of salient contours, of which some fraction presumably is robust to the transform. It remains unclear why the NOUD manipulation produces lower d' performance for both art and natural images compared to what would be expected if the NO and UD manipulations added linearly.

Viewed alongside results from Experiment 2, it appears clear that noise addition (with or without other manipulations) hinders performance on this task to a greater extent in art images compared to natural images.

General discussion

A main finding of this study is that artistic and natural images of faces can be discriminated from nonface images with the same high accuracy even for stimulus durations of 12 ms. This finding is consistent with a number of other studies showing that basic visual discriminations can be performed for extremely brief stimulus durations. Importantly, it indicates that artists are highly effective at capturing the informative features of faces, even though artists must perform a large degree of scaling (which is accomplished in a local piecewise manner using individual brushstrokes) in order to create an image. We found in Experiment 2 that the effect of CN on face detection cannot be differentiated for natural images and art. CN drastically alters global intensity distributions but preserves local Fourier phase relations (and amplitude spectra). Therefore,

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

putative regularities of intensity distributions in artistic representations appear not to be critical for rapid face detection[2].

We note that while only the Cornell and McGill stimuli are known to be linear with respect to luminance, it has been shown that film gammas (i.e., the luminance scaling function applied to a captured image) have little to no effect on spatial statistical regularities such as the amplitude spectrum slope (Redies et al., 2007a). However, we cannot rule out the possibility that image manipulations performed during image acquisition (e.g., gamma correction, manual color adjustments) may have affected performance to some degree for the uncalibrated images. Therefore, we cannot draw unequivocal conclusions regarding the causes of efficiency in artists' representational strategies at present. Nevertheless, the consistency of our findings—and the fact that our image manipulations in Experiments 2 and 3 (such as CN) engendered far greater modifications of image statistics than those that would be caused by, for example, subtle adjustments to color—suggests that the artistic transforms play a leading role in the efficiency of the representation. The fact that 1/f noise addition caused a differential drop in performance for artwork further suggests that representations of local features in artwork are chief candidates for achieving this efficiency. Current work is aimed at measuring the spatial and statistical transforms that artists use in practice, through large-scale photometric measurements of paintings and their associated natural scene tableaux, and tests of the efficiency of such strategies are ongoing.

Results related to the present study have been obtained by Velisavljevic and Elder (2009), who found that shape cues (contours) in natural scenes were stronger cues for rapid animal/nonanimal discrimination compared to luminance, color, and texture cues[3]. In their experiments, both machine-selected and human observer-selected contours produced similar performance. Human artists appear to have likewise taken advantage of what one could call the "primacy of form" in order to produce representational artwork. For example, research has shown that human artists have adopted strategies for representing certain object contours that are highly consistent over thousands of years (Biederman & Kim, 2008). Our present findings suggest that artists' strategies for selecting and representing salient scene contours could tap into the same neural coding mechanisms that generate rapid detection of animals, as in the work of Velisavljevic and Elder (2009).

One may wonder how an artistic image that could be less "photorealistic" compared to a digital photograph could yet be efficient for capturing the diagnostic features of that scene. Consider that digital images are generally not subject to local contrast scaling. Indeed, though digital image processing tools (and their emulsion film forebears, such as burning and dodging; see e.g., McCann, 2005) can be used to modify contrast in local regions, such methods are imprecise: they are difficult to apply to objects with fractal edges (such as are common in natural images; see Field et al., 1993), multiple sources of lighting, and/or complex backgrounds. Paintings, on the other hand, are created without such restrictions: scaling can be performed selectively across the image and on face and object contours. Moreover, scene composition can be modified freely. Artists' diverse methods for achieving transforms of this sort could thus be highly efficient in a representational sense (see Graham & Field, 2008*b*) since salient features can be emphasized in ways difficult to achieve with photographic representations. In other words, a painting can perhaps be seen as a natural scene that has been in a sense "optimized" for the *human* visual system but which yet retains statistical regularities to which *mammalian* visual coding is efficiently adapted. Of course not all paintings are thus—but to the extent that many paintings use similarly effective representational strategies, these shared strategies deserve concerted attention from visual neuroscience researchers.

We speculate that artistic transforms constitute a strategy that efficiently matches human neural encoding of complex local face and object features. Such coding strategies are at present poorly understood. Further, regularities in artistic treatment of local features could map onto the representational building blocks (basis functions) of midlevel vision. One indication that this may be a useful approach is that midlevel visual processing is substantially concerned with "normalized" contrast relations (due to contrast gain control: see e.g., Bex et al., 2007 or due to neural response nonlinearities: see Olshausen & Field, 2005) rather than with absolute intensity relations. As such, the notion that artists perform local scaling can be taken as an indication that they have discovered a family of efficient ways to represent important contours, ones that could take advantage of the brain's own coding strategies. We encourage future investigations of midlevel visual representations that employ art stimuli. Such work could parallel other novel approaches to uncovering response properties of neurons tuned to complex features, for example, Yamane et al. (2008).

We note that one group performed experiments using rapid categorization that explored the processing of abstraction in painting (Wallraven et al., 2007). These experiments showed that the "indeterminate" artwork of Robert Pepperell showed significant differences in terms of reaction time for categorization compared to more representational artworks. Despite lacking key features such as eyes, Pepperell's indeterminate work gives an impression of "representationality" by capturing contours suggestive of human forms and other objects. This result is consistent with the idea that efficient representations of key facial features—and not necessarily representations of global form or global statistics—are the most critical features for rapid face detection in art. We are currently exploring the role of artistic representations of different face features in rapid detection by systematically adding and removing features.

In addition, the role of color is not studied in the present experiments. Previous studies (Velisavljevic & Elder, 2009) have suggested that color may not be a principal feature used for extremely rapid accurate detection of objects (see also Bindemann & Burton, 2009). With regard to art, it should be noted that representation of natural colors is limited not only by optical constraints similar to those that limit the dynamic range of luminance in art but also by the availability of diverse pigments and by lighting incident on the painting. For most of human history, available pigments could capture only a relatively small variety of natural colors—and they could do so only with relatively low reflectance—until the rise of oil-based paint in the Renaissance (see e.g., Clarke, 2001). Moreover, color diversity due to different lighting has been shown to produce important effects on perception (Pinto et al., 2006).

To summarize, there may be a long list of what one can learn about perception by studying visual representations in art. In

[2]It should be recognized also that linear contrast inversion of the kind employed here and in most studies to date does not account for the nonlinear (log-like) luminance response of cone photoreceptors. This is a potentially large problem in the case of natural images, given that their luminance distributions are typically non-Gaussian, multimodal, and highly skewed. Liu et al. (1999) partially addressed this question with computer-generated faces, but to date, we are not aware of such studies involving calibrated natural images.

[3]Though this study draws conclusions regarding luminance regularities, it should be noted that the images employed were not luminance calibrated.

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

demonstrating that art images show a similar degree of efficiency in terms of rapid detection of faces (even after undergoing CN and UD), the present study adds support to the view that artwork can help bridge the existing gap between studies using natural and artificial stimuli. Like natural scenes, artworks contain a wealth of statistical structure at many orders, and they generally depict scenes from the human visual environment; like artificial stimuli, they are designed for human viewing. While art images lack the parameterization of artificial stimuli (e.g., sine-wave gratings), they are in principle more constrained compared to natural scenes due to luminance compression. Moreover, art often contains representations of human faces, which are common in the natural visual environment of humans but rare in standard natural scene collections (e.g., no faces are included in the McGill database). A related argument has been made for the relevance of film to the question of how the visual system processes temporal dynamics of the natural visual world (Hasson et al., 2004, 2009; Cutting et al., 2010). Studies of art and film stimuli may thus help resolve questions regarding visual system coding that studies of traditional natural and artificial stimuli could not as clearly address.

Conclusions

We have found that humans perform near flawless discriminations of faces and nonfaces in both paintings and natural images, even for stimulus presentation durations of 12 ms. In addition, CN and UD have minimal to no effect on performance for both image types, though 1/f noise addition significantly affects discrimination performance for art more than for natural images. These results suggest that artists create representations that are efficient for transmitting "diagnostic" information in faces. While the full mechanisms that allow faces in artwork to be rapidly and accurately detected require further exploration, our results support the view that artists' strategies for local contrast scaling of face features contribute to this ability. As such, these strategies could aid our understanding of neural coding in human visual perception: the strategies that artists employ to control local contrast could succeed at transmitting perceptual information because they map onto human visual processing strategies aimed at extracting environmental cues necessary for object recognition.

Acknowledgment

This work was supported by a Martens Family Research Award.

References

Atick, J.J. & Redlich, A.N. (1992). What does the retina know about natural scenes? *Neural Computation* **4**, 196–210.

Bex, P.J., Mareschal, I. & Dakin, S.C. (2007). Contrast gain control in natural scenes. *Journal of Vision* **7**, 1–12.

Biederman, I. & Kim, J.G. (2008). 17,000 years of depicting the junction of two smooth shapes. *Perception* **37**, 161–164.

Bindemann, M. & Burton, A.M. (2009). The role of color in human face detection. *Cognitive Science* **33**, 1144–1156.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision* **10**, 433–436.

Clarke, M. (2001). The Art of All Colours: Mediaeval Recipe Books for Painters and Illuminators. London: Archetype.

Cutting, J.E., DeLong, J.E. & Nothelfer, C.E. (2010). Attention and the evolution of Hollywood film. *Psychological Science* **21**, 440–447.

de Waal, F. (2001). The Ape and the Sushi Master: Cultural Reflections by a Primatologist. New York: Basic Books.

Fagot, J., Martin-Malivel, J. & Depy, D. (2000). What is the evidence for an equivalence between objects and pictures in birds and non-human primates? In *Picture Perception in Animals*, ed. Fagot, J., Hove, UK: Psychology Press.

Felsen, G. & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience* **8**, 1643–1646.

Field, D.J. (1987). Relations between the statistics of natural images and the response profiles of cortical cells. *Journal of the Optical Society of America. A, Optics and Image Science* **4**, 2379–2394.

Field, D. (1994). What is the goal of sensory coding? *Neural Computation* **6**, 559–601.

Field, D.J., Hayes, A. & Hess, R. (1993). Contour integration by the human visual system: Evidence for a local 'association field'. *Vision Research* **33**, 173–193

Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology* **59**, 167–192.

Gilad, S., Meng, M. & Sinha, P. (2009). Role of ordinal contrast relationships in face encoding. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5353–5358.

Graham, D.J. & Field, D.J. (2007). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision* **21**, 149–164.

Graham, D.J. & Field, D.J. (2008*a*). Global nonlinear luminance compression in painted art. *Proceedings of SPIE* **6810**, 68100K.

Graham, D.J. & Field, D.J. (2008*b*). Natural images: Coding efficiency. In *Encyclopedia of Neuroscience*, **Vol. VI**, ed. Squire, L.R., pp. 19–27. Oxford: Academic Press.

Graham, D.J. & Field, D.J. (2008*c*). Variations in intensity statistics for representational and abstract art, and for art from the eastern and western hemispheres. *Perception* **37**, 1341–1352.

Graham, D.J., Friedenberg, J.D., McCandless, C.H. & Rockmore, D.N. (2010*a*). Preference for artwork: Similarity, statistics, and selling price. *Proceedings of SPIE* **7527**, 75271A.

Graham, D.J., Friedenberg, J.D. & Rockmore, D.N. (2009). Efficient visual system processing of spatial and luminance statistics in representational and non-representational art. *Proceedings of SPIE* **7240**, 72401N.

Graham, D.J., Friedenberg, J.D., Rockmore, D.N. & Field, D.J. (2010*b*). Mapping the similarity space of paintings: Image statistics and visual perception. *Visual Cognition* **18**, 559–573.

Graham, D.J. & Meng, M. (2011). Altered spatial frequency content in paintings by artists with schizophrenia. *i-Perception* **2**, 1–9.

Graham, D.J. & Redies, C. (2010). Statistical regularities in art: Relations with visual coding and perception. *Vision Research* **50**, 1503–1509.

Hasson, U., Malach, R. & Heeger, D. (2009). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences* **14**, 40–48.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640.

Haxby, J.V., Hoffman, E.A. & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences* **4**, 223–233.

Honey, C., Kirchner, H. & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of Vision* **8**, 1–13.

Hughes, J.M., Graham, D.J., Jacobsen, C.R. & Rockmore, D.N. (2011). Higher-order spatial statistics and perceptual judgments in the stylometric analysis of art. In *Proceedings of EUSIPCO 2011 (19th European Signal Processing Conference), Barcelona, Spain, 29 Aug 2011.*

Koch, M., Denzler, J. & Redies, C. (2010). 1/f2 Characteristics and Isotropy in the Fourier Power Spectra of Visual Art, Cartoons, Comics, Mangas, and Different Categories of Photographs. PLoS ONE 5(8): e12268. DOI: 10.1371/journal.pone.0012268.

Li, F.-F., Iyer, A., Koch, C. & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision* **7**, 1–29.

Liu, C.H., Collin, C.A., Burton, A.M. & Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research* **39**, 4003–4009.

McCann, J. (2005). Rendering high-dynamic range images: Algorithms that mimic human vision. *Proceedings of the AMOS Technical Conference* **2005**, 19–28.

Meng, M. & Potter, M.C. (2008). Detecting and remembering pictures with and without visual noise. *Journal of Vision* **8**:7, 1–10.

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press

Oliva, A. (2005). Gist of the scene. In *Encyclopedia of Neurobiology of Attention*, ed. Itti, L., Rees, G. & Tsotsos, J.K., pp. 251–256. San Diego, CA: Elsevier.

Olmos, A. & Kingdom, F.A.A. (2004). McGill Calibrated Colour Image Database. Available from http://tabby.vision.mcgill.ca.

Olshausen, B.A. & Field, D.J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology* **14**, 481–487.

Olshausen, B.A. & Field, D.J. (2005). How close are we to understanding V1? *Neural Computation* **17**, 1665–1699.

Patel, A.D. (2008). Music, Language, and the Brain. Oxford: Oxford University Press.

Pinto, N., Cox, D.D. & DiCarlo, J.J. (2008). Why is real-world visual object recognition hard? *PLoS Computational Biology* **41**, e27.

Pinto, P.D., Linhares, J.M.M., Carvalhal, J.A. & Nascimento, S.M.C. (2006). Psychophysical estimation of the best illumination for appreciation of Renaissance paintings. *Visual Neuroscience* **23**, 669–674.

Potter, M.C. & Levy, E.I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology* **81**, 10–15.

Redies, C., Hänisch, J., Blickhan, M. & Denzler, J. (2007*a*). Artists portray faces with the Fourier statistics of complex natural scenes. *Network: Computation in Neural Systems* **18**, 235–248.

Redies, C., Hasenstein, J. & Denzler, J. (2007*b*). Fractal-like image statistics in visual art: Similarity to natural scenes. *Spatial Vision* **21**, 137–148.

Rust, N.C. & Movshon, J.A. (2005). In praise of artifice. *Nature Neuroscience* **8**, 1647–1650.

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

Torralba, A. & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems* **14**, 391–412.

VanRullen, R. & Thorpe, S.J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception* **30**, 655–668.

Velisavljevic, L. & Elder, J.H. (2009). Cue dynamics underlying rapid detection of animals in natural scenes, *Journal of Vision* **9**, 1–20.

Wallraven, C., Kaulard, K., Kürner, C., Pepperell, R. & Bülthoff, H.H. (2007). Psychophysics for perception of indeterminate art. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization APGV'07, 115-122*. New York: ACM Press.

Winkielman, P., Schwarz, N., Reber, R. & Fandeiro, T.A. (2003). Cognitive and affective consequences of visual fluency: When seeing is easy on the mind. In *Persuasive Imagery: A Consumer Response Perspective*, ed. Scott, L.M. & Batra, R., p. 152. Mahwah, NJ: Lawrence Erlbaum Associates.

Yamane, Y., Carlson, E.T., Bowman, K.C., Wang, Z. & Connor, C.E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nature Neuroscience* **11**, 1352–1360.

https://doi.org/10.1017/S0952523811000162 Published online by Cambridge University Press