Nonaccidental Properties Underlie Human Categorization of Complex Natural Scenes
Dirk B. Walther and Dandan Shen
Psychological Science 2014 25: 851 originally published online 28 January 2014
DOI: 10.1177/0956797613512662

The online version of this article can be found at:
http://pss.sagepub.com/content/25/4/851

Published by:
http://www.sagepublications.com

On behalf of:
Association for Psychological Science

Additional services and information for Psychological Science can be found at:
Email Alerts: http://pss.sagepub.com/cgi/alerts
Subscriptions: http://pss.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

Version of Record - Apr 9, 2014
OnlineFirst Version of Record - Jan 28, 2014
What is This?
Nonaccidental Properties Underlie Human Categorization of Complex Natural Scenes

Dirk B. Walther and Dandan Shen
The Ohio State University

Abstract
Humans can categorize complex natural scenes quickly and accurately. Which scene properties enable people to do this with such apparent ease? We extracted structural properties of contours (orientation, length, curvature) and contour junctions (types and angles) from line drawings of natural scenes. All of these properties contain information about scene categories that can be exploited computationally. However, when we compared error patterns from computational scene categorization with those from a six-alternative forced-choice scene-categorization experiment, we found that only junctions and curvature made significant contributions to human behavior. To further test the critical role of these properties, we perturbed junctions in line drawings by randomly shifting contours and found a significant decrease in human categorization accuracy. We conclude that scene categorization by humans relies on curvature as well as the same nonaccidental junction properties used for object recognition. These properties correspond to the visual features represented in area V2.

Keywords
natural scenes, scene categorization, nonaccidental properties, line drawings, structural description, visual perception, vision

Received 4/19/13; Revision accepted 10/21/13

People can recognize their natural environment as, say, a forest or an office (Tversky & Hemenway, 1983), and they succeed at categorizing scenes even with extremely short presentation times (Fei-Fei, Iyer, Koch, & Perona, 2007; Greene & Oliva, 2009; Loschky, Hansen, Sethi, & Pydimarri, 2010; Potter & Levy, 1969) and possibly in the absence of attention (Li, VanRullen, Koch, & Perona, 2002; but see; Cohen, Alvarez, & Nakayama, 2011; Evans & Treisman, 2005). What are the properties of natural scenes that make this impressive feat possible?

Over the past few decades, several featural representations of scenes have been proposed. Marr (1982) thought of scenes as collections of surfaces in three-dimensional space; that is, people can infer scene content from the orientations and spatial relations in a 2.5-dimensional sketch of a scene. Other researchers found statistics of color (Goffaux et al., 2005; Oliva & Schyns, 2000) or oriented image energy as represented in the spatial frequency spectrum (Oliva & Torralba, 2001; Torralba & Oliva, 2003) to be indicative of scene category. Here, we argue that contour junctions, which are invariant to slight changes of viewing position (nonaccidental properties; Biederman, 1987), and curvature play an essential part in human scene categorization.

Computational analysis of natural scenes has been largely focused on the discovery of statistical regularities of the luminance patterns of photographs (in pixel space or frequency space) that allow for computational scene categorization and, in some cases, relate to human categorization performance (Fei-Fei & Perona, 2005; Oliva & Torralba, 2001; Renninger & Malik, 2004; Torralba & Oliva, 2003). The image patches preferred by the resulting filters tend to look like contrast edges or straight contour lines and, occasionally, like corners or curved contours (Fei-Fei & Perona, 2005; Renninger & Malik, 2004). The same types of filters were derived as an efficient encoding of natural images when sparsity (Olshausen & Field, 1996) or
statistical independence (Bell & Sejnowski, 1997) of the ensuing code was enforced. However, establishing explicit descriptions of the structure of scenes from these luminance patterns has proven to be challenging.

Human artists can emphasize lines that describe object boundaries over those created by shadows or irrelevant textures (Sayim & Cavanagh, 2011). We can thus sidestep the issue of detecting contours from photographic images by extracting structural features from line drawings of natural scenes that were created by trained artists. As we have shown recently (Walther, Chai, Caddigan, Beck, & Fei-Fei, 2011), line drawings of natural scenes and color photographs elicit the same category-specific activation patterns in the scene-selective parahippocampal place area (PPA) and retrosplenial cortex. Neural activity in the PPA is tightly linked to scene categorization in humans (Walther, Caddigan, Fei-Fei, & Beck, 2009), which suggests that the scene structure preserved in line drawings plays an important role in scene categorization.

Line drawings provide ready access to several aspects of scene structure: orientation, length, and curvature of contours, as well as type and angle of contour junctions. To determine the differential contributions of these structural properties to the representation of scene categories, we tested their efficacy for both computational scene categorization and scene categorization by human subjects. Computational scene categorization based on line orientations yielded the highest accuracy among all the properties. However, human subjects’ patterns of errors in a scene-categorization task were best matched by a computational model based on summary statistics of curvature and junctions, but not orientation or length. The same pattern of results was obtained for photographs of natural scenes, which validated the idea that people rely on the same structural properties when categorizing scenes depicted in photographs. To determine whether contour junctions have a causal role in scene categorization, we repeated the behavioral experiments using line drawings in which the distribution of junctions was altered but all other properties were unchanged. If junctions are causally involved in scene categorization, then accuracy should be lower for such manipulated drawings than for intact line drawings. We indeed found such a decrease in categorization accuracy, confirming a causal relationship between junction properties and scene categorization by humans.

**Method**

**Participants**

Twenty-two subjects (12 male, 10 female; age range = 18–21 years) participated in the first experiment, and a separate group of 27 subjects (8 male, 19 female; age range = 18–23 years) participated in the second experiment, in both cases for partial course credit. All subjects had normal or corrected-to-normal vision and gave written informed consent. The experiments were approved by the institutional review board of The Ohio State University. For each experiment, the data of 4 subjects were excluded from the analysis because they did not complete the entire experiment or did not follow the instructions.

**Stimuli**

We used 432 color photographs from six real-world scene categories (three natural: beaches, forests, and mountains; three man-made: city streets, highways, and offices). These images were chosen from a set of 4,025 images downloaded from the Internet as the best exemplars of their categories according to ratings by an average of 137 observers per image (Torralbo et al., 2013). Images were resized to 800 × 600 pixels.

Line drawings were produced by trained artists at the Lotus Hill Research Institute (Wuhan, Hubei Province, People’s Republic of China), who traced outlines in the color photographs using a custom graphical user interface. The artists were given the following instructions:

For every image, please annotate all important and salient lines, including closed loops (e.g., boundary of a monitor) and open lines (e.g., boundaries of a road). Our requirement is that, by looking only at the annotated line drawings, a human observer can recognize the scene and salient objects within the image.

Each contour was traced as a series of straight lines connecting individual anchor points and thus closely approximated the contour with as few anchor points as possible. Line drawings were rendered by connecting the anchor points with black straight lines on a 800- × 600-pixel white background.

Contour-shifted line drawings were generated automatically by translating each contour by a random distance while ensuring that the entire contour would fit within the image boundaries. Horizontal and vertical translation distances were drawn from independent uniform distributions of allowable translations. Contour-shifted line-drawing images were rendered the same way as the intact line drawings at a size of 800 × 600 pixels.

We generated color masks by synthesizing random, scenelike textures separately for each color channel (red, green, and blue) using the texture synthesis method of Portilla and Simoncelli (2000). We used the color photographs from our stimulus set as input (see Loschky et al., 2010, for a rationale for using such masks). To create

---

Walther, Shen

Downloaded from pss.sagepub.com at OhioLink on April 28, 2014
each mask, we averaged six random textures, each synthesized from a randomly chosen photograph from a different scene category, to avoid category bias. We precomputed 100 such masks and selected from them at random for each trial. We used the same set of masks for color photographs and line drawings. See Figure S1 in the Supplemental Material available online for examples of scene stimuli and masks.

**Procedure**

In Experiment 1, images were presented on a CRT computer monitor (screen dimensions = 1024 × 768 pixels; refresh rate = 150 Hz) at an image size of 800 × 600 pixels (approximately 23° × 18°, depending on the exact viewing distance). Images were centered on a 50%-gray background. At the start of each trial, a fixation cross was presented for 500 ms; it was followed by a brief presentation of an image (initially 250 ms). The image was replaced by a perceptual mask for 500 ms; finally, a blank screen appeared for 2,000 ms. Subjects performed a six-alternative, forced-choice (6AFC) categorization of each image by pressing one of six buttons ("s," "d," "f," "j," "k," and "l") on a computer keyboard. The mapping of categories to buttons was randomized for each subject, and subjects practiced the mapping until they achieved 90% accuracy. After practice, a staircase procedure was used to adjust the stimulus onset asynchrony (SOA) for each subject to achieve a criterion of 65% accuracy. A randomly selected subset of 12 color photographs from each category was used for practice and for the staircase procedure, which left 60 images per category for testing. A tone alerted participants to incorrect responses during these two phases.

In the testing phase of Experiment 1, half of the 360 test images were shown as photographs and half as line drawings, randomly intermixed. Each test image was shown only once, either as a photograph or as a line drawing. Trials were grouped into 18 blocks of 20 images each. The SOA was fixed to the final SOA achieved in the staircase phase for each subject. No feedback was given during the testing phase. Responses for photographs and line drawings were recorded in separate 6 × 6 confusion matrices. Accuracy was computed as the mean over the diagonal elements of the confusion matrix.

Experiment 2 followed the same procedure, except that line drawings were used for practice and for the staircase procedure, and original and contour-shifted line drawings were used for testing.

**Computational image analysis**

We computed properties of individual contours (length, orientation, and curvature) as well as junctions of contours (type and angle) from the line drawings (Fig. 1a). These properties were computed directly from the geometrical information of the line drawings. They are closely related to the structure present in the corresponding scene photographs. Details on the computation of properties from line drawings are provided in the Supplemental Material. We then constructed histograms of these properties to quantify their distribution for a given line drawing. Each line drawing had one histogram for each of the five properties. Figure 1b shows the average of these five histograms over all images of a given category.

For each scene, we concatenated the histograms of all properties into a 36-dimensional vector and performed 10-fold cross-validation classification of scene categories with a linear support-vector-machine classifier. Predictions of scene categories by the classifier were recorded in a confusion matrix (Fig. 2a). Contour-shifted images were classified the same way, except that the classifier was trained on the original drawings and tested on contour-shifted line drawings to mimic the behavioral experiment.

Agreement between behavioral and computational results was assessed by computing the Pearson correlation coefficient between the off-diagonal elements of the two confusion matrices, thus quantifying the match of error patterns. Significance of correlations was determined nonparametrically by repeating the analysis with all 720 permutations of the six category labels. Correlations with a p value less than .05 (fewer than 36 of the 720 permutations resulted in correlations larger than that of the correct order) were accepted as statistically significant.

**Results**

**Image properties**

To determine whether the five properties of line drawings (length, orientation, curvature, intersection types, and intersection angles) contained sufficient information to infer scene category, we computed summary statistics of all five properties over the line drawings. Figure 1b shows the average histograms of the properties for each scene category. Unsurprisingly, man-made scene categories (city streets, highways, offices) are dominated by horizontal and vertical orientations, whereas the distribution of orientations among natural scene categories (beaches, forests, mountains) is more nuanced. Images of beaches are dominated by horizontal orientation, and images of forests are dominated by vertical orientation; images of mountains have a more balanced distribution of orientations (but slightly more horizontal and near-horizontal orientations than vertical and near-vertical orientations). These differences in orientation between scene categories are consistent with results obtained by Torralba and Oliva (2003). Furthermore, T-junctions are
more prevalent in man-made than in natural scenes, which is also reflected in the dominance of 90° intersection angles. Contour curvature and length do not show obvious distinctions among the categories, although both reflect total line length in their overall histogram magnitude.

**Computational analysis**

Can one use these regularities to discriminate among natural-scene categories? To answer this question, we trained a support-vector–machine classifier to predict the category of a scene on the basis of the distribution of its

---

*Fig. 1.* Illustration of the computational image analysis. An example photograph of a city street scene (original by Flickr user aleph.pukk) is shown in (a) with corresponding color-coded line drawings indicating its image properties. The histograms in (b) show the average distributions of the five properties of line drawings for each of the six scene categories.
Fig. 2. Classifier predictions and results for Experiment 1. Each cell of the confusion matrix in (a) shows the proportion of images of one category (indicated by the row label) that were predicted by the classifier to be of another or the same category (specified by the column label). Diagonal elements indicate correct classifications; off-diagonal elements are confusions. The confusion matrix in (b) shows results for Experiment 1. Row labels indicate the image categories presented to the observers, and column labels indicate the observers’ responses. The graph in (c) shows the classifier’s accuracy (mean of the diagonal of the confusion matrix) in categorizing the line drawings as a function of image property and participants’ (n = 18) behavioral accuracy in classifying photographs and line drawings. The dashed line denotes chance level (1/6). Correlation of computational with behavioral errors for line drawings and photographs is depicted in (d); correlations were calculated separately for each of the five image properties. In (c), asterisks indicate a significant difference between behavioral accuracy for line drawings and photographs, and in (d), asterisks indicate significant correlations between computational and behavioral errors (**p < .01, ***p < .001). Error bars in (c) represent standard errors of the mean.
structural properties. Classifier errors were recorded in the off-diagonal elements of confusion matrices (Fig. 2a).

When all five properties were included in the analysis, predictions for scenes previously unseen by the classifier were correct for 83.6% of the test images (chance level: 16.7%; Fig. 2c). This suggests that the properties that were derived from the line drawings contained sufficient information to determine their category membership in most cases.

To determine the contribution of each of the five properties, we repeated the computational analysis for each property separately. Line orientation seems to be the most informative property; it allowed for correct classification of 71.6% of all line drawings. The other properties allowed for correct classification of between 47.4% and 56.9% of all images (Fig. 2c).

**Experiment 1**

To investigate how these computational results relate to human behavior, we performed a 6AFC categorization experiment. The staircase procedure resulted in SOAs between 16.7 and 86.7 ms ($M = 26.5$ ms). Subjects’ responses for line drawings and photographs were recorded in separate confusion matrices (Fig. 2b). Categorization accuracy (mean of the diagonal of the confusion matrix; chance: 16.7%) was 66.2% ($SEM = 3.1\%$ for line drawings and 77.3% ($SEM = 3.0\%$) for photographs, averaged over 18 subjects (Fig. 2c). The difference in accuracy between the two image types was significant, $p = 4.9 \times 10^{-9}$, $t(17) = 6.6$ (paired $t$ test). There was no significant difference in response times between color photographs and line drawings.

To assess the similarity in error patterns, we correlated the off-diagonal elements (i.e., the confusions) of the behavioral confusion matrices with those obtained from the computational analysis. For line drawings, we observed the highest correlation between human performance and the computational results for junction angles ($r = .730$, $p < .0014$), junction types ($r = .513$, $p = .0083$), and contour curvature ($r = .583$, $p = .0069$); computational categorization and human performance were not significantly correlated for line orientation ($r = .091$, $p = .33$) or contour length ($r = .169$, $p = .20$; Fig. 2d).

It is conceivable that this finding was an artifact of human participants viewing line drawings of natural scenes, as line intersections may be more readily accessible in line drawings than in photographs under natural lighting conditions. To address this possible confound, we correlated the patterns of errors made by human subjects when categorizing photographs and the classification errors in the computational analysis of the corresponding line drawings. We again found significant correlation of error patterns for junction angles ($r = .658$, $p = .0014$), junction types ($r = .576$, $p = .0097$), and curvature ($r = .511$, $p = .0097$) but not for line orientation ($r = .078$, $p = .34$) or contour length ($r = .023$, $p = .41$; Fig. 2d). Because we observed the same pattern of correlations for photographs and for line drawings, we conclude that the importance of junctions and curvature for human behavior transcends the specifics of the low-level representation of scenes.

**Experiment 2**

If humans rely so heavily on intersection-related properties for categorizing scenes, then perturbing these properties should have a detrimental effect on categorization performance. To test this prediction, we randomly translated the contours of the line drawings within the image bounds, thereby preserving all properties that pertain to the contours themselves (length, orientation, and curvature) but altering junction angles and counts of junction types (Fig. 3a).

By definition, computational classification of line drawings using length, orientation, or curvature remained unaffected, because these properties of individual contours do not change when the contours are shifted. Classification using counts of junction types, on the contrary, dropped from 47.4% to 35.2%, and classification using junction angles dropped from 53.7% to 39.7%, which led to a decrease from 83.6% to 73.4% in classification error when the combination of all five properties was used (Fig. 3b). To test the effect of contour shifting on human performance, we repeated the 6AFC experiment with intact line drawings for practice and for the staircase procedure. We then randomly intermixed intact and contour-shifted line drawings for testing. Stimuli were presented for 13.3 to 133.3 ms ($M = 33.9$ ms). Performance was significantly lower for the contour-shifted line drawings (31.9%) than for the original line drawings (53.8%), $p = 4.8 \times 10^{-9}$, $t(22) = 9.26$ (paired $t$ test), which confirms both our ad hoc prediction and the computational analysis using intersection-related properties (Fig. 3b). The relatively large variation in SOA due to individual differences in performance during the staircase procedure did not affect the outcome of the experiment adversely. Participants with longer SOAs showed higher overall accuracy during testing, but accuracy was lower for contour-shifted drawings than for intact line drawings for all participants.

With junction properties no longer available for categorizing scenes, which properties did participants fall back on to make their decisions? To answer this question, we correlated participants’ errors (off-diagonal entries in the confusion matrix) when viewing contour-shifted line drawings with the computational classification errors using properties of intact line drawings (Fig. 3c). As
expected, there was no longer any correlation for junction properties (types: $r = .160, p = .20$; angles: $r = .141, p = .18$), but there were highly significant correlations for contour length ($r = .725, p < .0014$) and orientation ($r = .412, p = .0069$), and a marginal correlation for curvature ($r = .310, p = .057$). For categorization based on all properties, the correlation was significant ($r = .352, p = .017$).

**Discussion**

We have demonstrated that computational scene categorization is more similar to human scene categorization when based on structural properties describing junctions and curvature in images than when based on properties relating to orientation and length of contours. Moreover,
this similarity is not limited to humans’ categorization of line drawings; the effect is just as pronounced for color photographs, which suggests that people readily extract these properties from photographs when perceiving scenes. We have shown that, beyond mere correlation, contour junctions are causally involved in natural scene categorization by demonstrating that human categorization performance decreased significantly when the distribution of junction properties was altered. When deprived of junction properties, observers resorted to contour length and line orientation to categorize scenes.

These junction properties coincide with some of the nonaccidental properties identified as critical for object recognition by Biederman (1987): Y- and arrow junctions indicate the orientation of corners in three-dimensional space, and T-junctions occur along the contour of one surface occluding another. Note that our set of properties is bound to be incomplete. For practical reasons, we did not take into account L-junctions, which indicate corners, or curved Y-junctions, which indicate curved, cylinder-like objects. We did not distinguish between internal object contours and contours at the boundary between objects as would be necessary for an explicit representation of shape (Marr & Nishihara, 1978). Even more important, although junction properties are localized, our summary statistics ignored the junctions’ locations and spatial relations, which are believed to be critical for object and scene recognition (Biederman, 1987; Kim & Biederman, 2011) as well as the segregation of surfaces (Rubin, 2001). In light of these gross simplifications, it is even more remarkable that simple summary statistics of junctions and curvature had such a clear relation to natural-scene categorization by humans. One possible explanation is that summary statistics of the image properties may suffice for fast gist recognition, whereas spatial locations and relations may be more important for more detailed inspection of a scene. This interpretation is consistent with findings that scene categorization does not require focused attention (Li et al., 2002; but see Evans & Treisman, 2005; Cohen et al., 2011).

Although line orientation and length are outweighed by junctions and curvature in their importance for human classification performance, they nevertheless allowed for fairly accurate categorization of natural scenes by computational algorithms (Fig. 2c). These statistical regularities can be exploited successfully by computational scene analysis in, for instance, the spatial-frequency domain (Oliva & Torralba, 2001). Why, then, do people rely predominantly on nonaccidental junction properties? There may be two, somewhat related reasons. First, nonaccidental properties afford stable representations under moderate three-dimensional transformations. In fact, the neural representation of a scene in the retrosplenial cortex generalizes over different views of the same scene (Park & Chun, 2009). Such limited viewpoint invariance may form the basis for interpreting scenes as arrangements of surfaces in three dimensions (Marr, 1982). Second, it is likely to be more efficient for people to make use of the same visual features (i.e., junctions and curvature) for scene perception as for a somewhat simpler visual task, object recognition (Biederman, 1987). Such convergence dispenses with the need for a separate scene-processing pathway and instead suggests a feature representation that is shared between scene and object perception.

Both contour junctions and curvature are representatives of third-order image properties in the nomenclature of Koenderink and van Doorn (1987), the zeroth order being luminance patterns, the first order being luminance gradients, and the second order being straight lines: Junctions describe the interaction between lines, and curvature describes the change of line orientation along contours. With contours and their junctions clearly delineated, line drawings may serve as a useful stand-in for the analysis of third-order scene structure. However, line drawings do not capture all of the information present in color photographs of natural scenes, as is illustrated by the fact that observers categorized line drawings significantly less accurately than color photographs in Experiment 1 (Fig. 2c). Aside from color, photographs contain additional texture information. Regularities in the textures of natural scenes have been modeled successfully at the level of visual area V1 (Bell & Sejnowski, 1997; Karklin & Lewicki, 2009; Olshausen & Field, 1996; Portilla & Simoncelli, 2000) as well as areas V2 and V4 (Freeman & Simoncelli, 2011). In the latter study, V2-type features were found to be critical for human scene perception: Observers were unable to distinguish between intact images and images with peripheral noise that mimicked the summary statistics of visual features at area V2. Nevertheless, when human observers were asked to apply scene category labels to the textures of scenes produced by Portilla and Simoncelli (2000), they performed only slightly above chance level (Loschky et al., 2010). Thus, more work is needed to determine how such texture information may be used by human observers for rapidly categorizing scenes.

An important question remains: How could the human visual system determine the statistics of structural scene properties during such brief image presentations? First, sensitivity to the structural features we found to be most important for human scene categorization (i.e., junctions and curvature) has been identified in visual areas V2 (e.g., Heeger & Van Essen, 2000; Peterhans & von der Heydt, 1989) and V4 (e.g., Pasupathy & Connor, 2002). Thus, these properties are available relatively early in the visual processing stream. Second, various visual features can be efficiently summarized by the visual system, which enables
Nonaccidental Properties Underlie Scene Categorization

D. B. Walther and D. Shen designed the computational analysis and experiments, collected and analyzed the data, and wrote the manuscript.

Acknowledgments

D. Shen is now at Google Inc., Mountain View, California. We thank Bart Larsen, Monica Craigmyle, and Ryan McClincy for help with running the experiments and Shaun Vecera, Wil Cunningham, and members of the Bernhard-Walther Lab, as well as two anonymous reviewers, for insightful comments on the manuscript.

Declarations of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was funded in part by the National Eye Institute (Collaborative Research in Computational Neuroscience Grant 1R01-EY019429).

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

References


Supplemental Material

Properties of Line Drawings

We defined properties of contours and contour junctions in line drawings as follows.

**Curvature** measures the extent to which a contour deviates from a straight line. In our work, each contour is approximated by a series of consecutive straight lines. We define the curvature of one segment as the change of orientation angle between the current line segment and the next (or previous, if there is no next segment) divided by its length. For example, for the contour in figure S2 A, which was drawn by connecting the six anchor points from $P_1$ to $P_6$ with five line segments, the curvature of segment $P_2P_3$ was computed as $\theta_2/\|P_2P_3\|$ and the curvature of segment $P_5P_6$ as $\theta_4/\|P_5P_6\|$.

In order to capture the wide range of possible curvature values, we constructed an eight-bin histogram on a logarithmic scale ($\log_{10}(\text{deg/pixel}+1)$) with bin centers at 0.14, 0.42, 0.71, 0.99, 1.27, 1.55, 1.83, and 2.12. When summing over the occurrences of curvatures in a particular bin, curvature values were weighted by the length of their corresponding line segments.

**Contour length** was computed as the sum of the lengths of the individual line segments of a contour. It was recorded in an eight-bin histogram on a logarithmic scale ($\log_{10}(\text{pixels}+1)$) with bin centers at 0.51, 0.92, 1.33, 1.74, 2.15, 2.56, 2.97, and 3.38. Bin counts were weighted by contour length so that each pixel in a line drawing contributed equally to the histogram.

**Orientation** of a line segment is given by its counter-clockwise angle from the horizontal. Orientation angles between 0 and 180 degrees were captured in an eight-bin cyclic histogram with bin centers at 0°, 22.5°, 45°, 67.5°, 90°, 112.5°, 135°, and 157.5°. Orientation values for individual line segments, weighted by the length of the line segments, were added up in the corresponding bins.

**Junctions** were detected wherever at least two line segments intersected, other than two consecutively drawn line segments within the same contour. Junctions that occurred within a distance of 3 pixels were combined into a single junction. X, Y, T and arrow junctions were defined based on the number of intersecting lines and based on the angles between them (fig. S2 B). Assuming that the intersecting contours divide a disk with its center at the intersection point into several sectors, we defined the four junction types as follows. If the disk was divided into four sectors, and the angles of all sectors were between 0° and 150°, then we labeled this junction as an **X junction**. If the disk was divided into three sectors, then we distinguished junction types based on the largest angle covered by any of the three sectors. If this angle was smaller than 160°, then the junction was assumed to be a **Y junction**; if it was between 160° and 200°, it was a **T junction**; and if the angle was larger than 200°, we assumed an **arrow junction**. If the junction did not fall into any of these categories, then we did not label it and did not consider it in our further analysis (25% of all junctions across all line drawings). Note that the specific angles used in these definitions are based on heuristics. They were selected so that the detected junctions agreed with the intuitive definitions of these junction types (e.g., Biederman, 1987). Note also that we did not detect L junctions. Since contours were defined as successions of straight lines, L junctions would have been too numerous to be useful, and their number would have been confounded with the number of line segments and therefore with contour length and curvature. After detecting the X, Y, T and arrow junctions within a line drawing, we recorded the number of each junction type in a four-bin histogram.

DOI:10.1177/0956797613512662
We defined the **junction angles** of X, Y, T, and arrow junctions as the smallest angle covered by any of the disk sectors. Since at least three lines are intersecting, intersection angles range from 0 to 120 degrees, which we divided into eight histogram bins with centers at 7.5°, 22.5°, 37.5°, 52.5°, 67.5°, 82.5°, 97.5°, 112.5°. Bins contained the total number of intersections with the corresponding intersection angles in a line drawing.

**References**

Supplemental Figures

Figure S1. Example images for photographs, line drawings, and contour-shifted line drawings (from left to right) of beaches, forests, mountains, city streets, highways, and offices (from top to bottom). Last column: six of the 100 texture masks used in the experiment. Original images by flickr users David K, Nicholas A. Tonelli, francois, Norris Wong, David Herrera, and Mo Riza.

DOI:10.1177/0956797613512662
**Figure S2.** Illustration of the computation of line drawing properties. A. Example for the computation of curvature. This contour was drawn in five segments from point P1 to point P6. Curvature of each segment is computed as the ratio of the change of orientation along the segment over the length of the segment. B. Intersecting lines partition a disk into several sectors. Types of junctions are defined based on the number of intersecting lines and on the angles of the disk sectors.