VISUAL COGNITION, 2010, 18 (4), 559–573

Psychology Press
Taylor & Francis Group

# Mapping the similarity space of paintings: Image statistics and visual perception

Daniel J. Graham

*Department of Mathematics, Dartmouth College, Hanover, NH, USA*

Jay D. Friedenberg

*Department of Psychology, Manhattan College, Riverdale, NY, USA*

Daniel N. Rockmore

*Departments of Mathematics and of Computer Science, Dartmouth College, Hanover, NH, and Santa Fe Institute, Santa Fe, NM, USA*

David J. Field

*Department of Psychology, Cornell University, Ithaca, NY, USA*



What makes two images look similar? Here we test the hypothesis that perceived similarity of artwork is related to basic image statistics to which the early visual system is attuned. In two experiments, we employ multidimensional scaling (MDS) analysis of paired-image similarity ratings from observers for paintings. Two sets of images, classified as "landscapes" and "portraits/still-life", are tested separately. For the landscapes, we find that one of the first two MDS scales of similarity is strongly correlated with a basic greyscale image statistic, whereas the other dimension can be accounted for by a semantic variable (representation of people). For portrait/still-life, the first two MDS scales of similarity are most highly correlated with semantic variables. Linear combinations of statistical and non-statistical features achieve improved predictive values for the first two MDS scales for both sets. The statistics that play the largest role in shaping similarity judgements in our tests are the activity fraction measure of sparseness and the log-log slope of the spatial frequency amplitude spectrum. We discuss these results in the context of scene perception and in terms of efficient coding of statistical regularities in scenes.

***Keywords:*** Perception; Vision; Art; Natural Scenes; Statistics.

---

Please address all correspondence to Daniel J. Graham, Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA. E-mail: daniel.j.graham@dartmouth.edu

This work was supported by a National Science Foundation Small Grant for Exploratory Research (DMS-0746667).


http://www.psypress.com/viscog DOI: 10.1080/13506280902934454

How are natural scenes discriminated by human observers? It has long been known that humans can make simple discriminations of scenes very quickly (Potter & Levy, 1969; see Henderson & Hollingworth, 1999, for a review) and recent studies have suggested that an efficient visual system could collect basic statistics early in processing in order to "diagnose" context or scene type (Oliva & Schyns, 2000; Oliva & Torralba, 2001; Torralba & Oliva, 2003). If there are statistical features that can be used to predict perceptual judgements, is it the case that these quantities are themselves involved in the neural coding of such perceptions, or are they correlates of some other quantity of interest to the visual system? With this paper, we intend to spur quantitative investigation into these questions using visual art. We offer preliminary findings that do not fully resolve these questions, but do shed light on the relationship between visual perception and basic image statistics. In particular, we test two collections of paintings to determine whether statistics relevant to efficient visual system processing are related to perceptual judgements of those images.

Visual art comprises a rich laboratory in which to evaluate links between efficient visual processing and basic image statistics (Graham & Field, 2007). Paintings are a special class of images, both because of their expressive and aesthetic qualities, but also because of their relationship to natural scenes. As a group, paintings have been found to exhibit predictable statistical properties, including power law fall-off in spatial frequency amplitude spectra (Graham & Field, 2007, 2008a; Redies, Hasenstein, & Denzler, 2007) and nonlinear tone reproduction curves (Graham & Field, 2007, 2008a, 2008b). The regularity of statistics in art and the relationships between regularities in art and those in natural scenes suggests an influence of early visual coding strategies in the production of art. In particular, it has been suggested that the notion that visual systems are efficiently matched to the regular statistics of natural scenes biases production of art towards those images that can be efficiently processed, i.e., towards images that contain natural scene-like regularities (Graham & Field, 2008b).

Other statistics are also shared by art and by natural scenes. For example, both classes of images (including abstract art) produce high statistical sparseness in a population of modelled neurons (Graham & Field, 2007).[1] The visual system, as well as many other brain areas, has been shown to respond in a statistically sparse manner in response to the natural environment (see Graham & Field, 2006; Olshausen & Field, 2004). With regard to art, the median statistical sparseness of modelled ganglion cell

[1] Sparseness is a statistic measured over responses to a given stimulus set. A code in which each unit fires at a low rate for most or all stimuli would have a low statistical sparseness, whereas a code where a handful of units fire at high rates for each stimulus would have high statistical sparseness.

responses to natural scenes is found to be similar to that for a large, diverse collection of artwork (Graham & Field, 2007).

Given these relationships between neural coding and certain image statistics, could such statistics play a role in basic perceptual judgements? For natural scenes, the answer could be yes. Torralba and Oliva (2003) have demonstrated that two-dimensionl spatial frequency amplitude spectra contribute to predictions of scale and content categorization for natural scenes, and they have proposed that rapid processing of scenes depend at least in part on the analysis of such statistics. Following on earlier findings regarding colour statistics (e.g., Swain & Ballard, 1991), a study of image search algorithms based on behavioural data (Neumann & Gegenfurtner, 2006) showed that a simple method based on similarity of pixel colour statistics alone performed well above chance in predicting judgements of perceptual similarity for a large collection of photographs. Moreover, combinations of colour statistics, luminance statistics, and amplitude spectrum statistics had even higher performance. Rogowitz, Frese, Smith, Bouman, and Kalin (1998) found that perceptual similarity data and a model of similarity based on colour statistics resulted in qualitatively similar mappings of similarity for a large collection of photographs. Also, the perception of glossiness has been shown to be highly correlated with pixel intensity skewness (Motoyoshi, Nishidam, Sharan, & Adelson, 2007). Together, these studies suggest that for basic perceptual judgements of scenes, simple statistics may capture much of the information necessary for making those judgements.

However, it remains unclear if the visual system specifically makes use of these statistics, or if these statistics are merely correlated with some other measure employed by the visual system to make such judgements. Though we suspect that both kinds of processes may be at work, this paper has a narrower focus. Here we consider if paintings, a subclass of natural images with special import for human viewers, can elicit perceptual judgements that are predictable using basic image statistics. Though there are surely a great many factors that influence such judgements, this may be a less daunting challenge than one might imagine. Despite the wide range of art styles, paintings require a number of statistical restrictions not found with natural scenes. For example, the maximum dynamic range of luminances for paintings is only a fraction of what is possible for natural scenes. It has been demonstrated that artists must employ some form of a nonlinear luminance compression scheme in order to depict the world, which can be modelled to a first approximation as a log-like nonlinearity[2] (Graham & Field, 2008a).

[2] This nonlinearity has also been proposed as a model of photoreceptor luminance response.

Moreover, image statistics have been shown to provide useful data for distinguishing high-level cognitive properties of art, including style. Lyu, Rockmore, and Farid (2004) used wavelet-transform statistics to separate authentic drawings of Pieter Bruegel the Elder from those of Bruegel's imitators. Statistics derived from box-counting data (related to the power spectrum) have been used to authenticate paintings thought to be by Jackson Pollock (Taylor et al., 2007), though some of these conclusions have been put in doubt for art historical (Landau & Cernuschi, 2007) and technical (Graham & Field, 2007; Jones-Smith & Mathur, 2006) reasons. Stylometric analysis on sets of works of different provenance (i.e., place of origin) show consistent pixel distribution statistics for Eastern and Western Hemisphere works (Graham & Field, 2008a). Together, this work suggests that low-level statistics may be useful for understanding high-level properties of art. However, we are not aware of any attempts to connect perceived similarity of art with image statistics.

In the current study, we map the similarity space of two sets of paintings (20 images each), which were selected from a group of images previously judged by observers to fall into classes labelled "landscape" and "portrait/still-life". We apply multidimensional scaling (MDS) analysis and attempt to ascribe units—either semantic or statistical—to the perceptual dimensions that account for the largest portion of the data variance. Because the two classes of images tested were intended to represent different perceptual features (large vs. small implied depth of field for landscapes and portrait/still-life, respectively) and different semantic features (portrait/still-life often includes people, landscape often does not), statistical measures and semantic classifiers would be expected to contribute in corresponding proportions for the two image sets.



## METHODS

The idea of measuring the similarity space for a set of stimuli has a long tradition in psychology (Kruskal, 1964; Shepard, 1962). Here we used a paired-image rating task for all pairs of images to measure perceptual similarity. Multidimensional scaling analysis can display the matrix of similarities in a Euclidean space while preserving as best as possible the absolute distances (measured as dissimilarities) among data points. We calculated a classical MDS solution, which is equivalent to principal components analysis, as well as nonclassical MDS. Our nonclassical MDS solution was calculated over Kruskal sum-squared stress of inter-point distances (Kruskal & Wish, 1978).

## Stimuli

Observers ($N = 6$) divided a large set (140 images) of digitized paintings from a major university collection (Herbert F. Johnson Museum of Art, Cornell University, Ithaca, NY, USA; see Graham & Field, 2007, for description of database) into general content categories using a three-alternative forced choice paradigm (choices were "landscape," "portrait/still-life", or "abstract"). From the groups of images whose set membership was agreed upon by the majority of the observers, 20 paintings each from the landscape and portrait/still-life sets were selected. The selected paintings were chosen such that at least 20% of the images in a set would contain one or more of the following: Bodies of water (landscapes), humans (landscapes), women (portraits), still life without humans (portraits). In both sets, at least 20% of the images had pre-twentieth-century production dates, and at least 20% were of Eastern hemisphere provenance. Binary semantic variables were constructed for each of these semantic classes based on metadata provided by the museum (see Table 1).

## Statistical measures

Images were converted from RGB to YIQ colour space and we performed statistical analysis on the Y (luminance) channel data. We tested a host of statistical measures of the images against the MDS scales derived from mean similarity ratings. These included the first four statistical moments of image pixel intensity histograms (mean, variance, skewness, kurtosis); the slope of the spatial frequency amplitude spectra averaged over orientation and plotted on log-log axes; statistical moments of modelled retinal (difference-of-Gaussians, DoG; see, e.g., Graham, Chandler, & Field, 2006) and cortical (Gabor wavelet; see e.g., Jones & Palmer, 1987) responses to the images; a pixel sparseness measure $S$, the "activity fraction",

$$S = \frac{\left(\frac{1}{n}\sum_{i}^{n} r_i\right)^2}{\frac{1}{n}\sum_{i}^{n} r_i^2}$$

(see also Rolls & Tovee, 1995); as well as the same measures of log-transformed images. The activity fraction $S$ over $n$ pixels each with intensity $r_i$ has a range of 0 to 1. This measure is typically used to gauge the sparseness of neural population responses. Small values of $S$ (near zero) correspond to a highly sparse, heavy-tailed distribution of intensities, where only a few

TABLE 1
Metadata for paintings used in experiment

| *Image Number* | *Artist* | *Title* | *Date (CE)* | *Water* | *Humans* | *East/West (E = 1)* |
|---|---|---|---|---|---|---|
| LANDSCAPES | | | | | | |
| 1 | Zhao | Thousands of Mountains Invite Hermit Scholars | 1604 | 0 | 0 | 1 |
| 2 | Mistuyoshi | Eight Views of Nara | Unknown | 1 | 0 | 1 |
| 3 | Hopper | Monhegan Landscape | 1916–1919 | 1 | 0 | 0 |
| 4 | Smith | Long Island Sound | 19th C. | 1 | 0 | 0 |
| 5 | Mahoney | Sunday Afternoon | 1932 | 0 | 1 | 0 |
| 6 | Rosa | Landscape with Philosophers | 16th C. | 0 | 0 | 0 |
| 7 | Daubigny | Les Champs au Mois de Juin | 1874 | 0 | 0 | 0 |
| 8 | Post | Brazilian Landscape | 1665 | 1 | 0 | 0 |
| 9 | Bierstadt | Swiss Mountain Scene | 1859 | 1 | 0 | 0 |
| 10 | O'Keefe | Red Hills, Blue Sky | 1937 | 0 | 0 | 0 |
| 11 | Benton | The Artist's Show, Washington Square, New York | 1946 | 0 | 1 | 0 |
| 12 | Childe | Rocks and Sea, Isle of Shoals | 1912 | 1 | 0 | 0 |
| 13 | Constable | Netley Abbey | 18th–19th C. | 0 | 0 | 0 |
| 14 | Kensett | The Rocks at Newport, Rhode Island | 1862 | 1 | 0 | 0 |
| 15 | Guler School | The Devi Attacks a Demon Army | C. 1800–1820 | 0 | 1 | 1 |
| 16 | Guler School | King Suratha and Samadhi | 18th C. | 0 | 1 | 1 |
| 17 | Pahari Style | Krishna and the Milk Maids | late 18th C. | 1 | 1 | 1 |
| 18 | Unident. (India) | Lady Implores Krishna to go to Radha | C. 1700–1715 | 1 | 1 | 1 |
| 19 | Wood | Untitled Moutainscape | 1928 | 0 | 0 | 0 |
| 20 | Dove | Landscape at Cagnes | 1909 | 0 | 0 | 0 |

TABLE 1 (*Continued*)

| *Image Number* | *Artist* | *Title* | *Date (CE)* | *Female sitter* | *Humans* | *East/West (E = 1)* |
|---|---|---|---|---|---|---|
| PORTRAITS | | | | | | |
| 1 | Lucioni | Tocata in Green | 1973 | 0 | 0 | 1 |
| 2 | Baishi | Portrait of Lin Daiyu | early 20th C. | 1 | 1 | 0 |
| 3 | Asantey | Black Cowboys | 1972–1974 | 0 | 1 | 1 |
| 4 | Droungas | Nike | 1982 | 0 | 1 | 1 |
| 5 | Lewis | Portrait of Lord Carlow (Portrait of a Smiling Gentleman) | 1939 | 0 | 1 | 1 |
| 6 | Bouguereau | Madonna and Child with Saint John | 1882 | 1 | 1 | 1 |
| 7 | Dix | Liegende auf Leopardenfell | 1927 | 1 | 1 | 1 |
| 8 | Unident. (Japan) | Amida Raigo | 1185–1392 | 0 | 1 | 0 |
| 9 | Bouguereau | The Goose Girl | 1891 | 1 | 1 | 1 |
| 10 | Henri | Portrait of Carl Sprinchom | 1910 | 0 | 1 | 1 |
| 11 | Kuniyoshi | Charade | 1948 | 1 | 1 | 0 |
| 12 | Unident. (Europe) | An Ecclesiastic | 16th C. | 0 | 1 | 1 |
| 13 | Bailly | Vanitas | C. 1650 | 0 | 0 | 1 |
| 14 | Hogarth | Portrait of Dainel Lock F.S.A. | 1762 | 0 | 1 | 1 |
| 15 | Spaendonck | Still life with Flowers | 1793 | 0 | 0 | 1 |
| 16 | Unident. (India) | Portrait of a Nobleman | late 18th C | 0 | 1 | 0 |
| 17 | Unident. (India) | Portrait of (or by) Sultanu | 1813 | 0 | 1 | 0 |
| 18 | Unident. (India) | The Horse Jagjeth | early 19th C. | 0 | 1 | 0 |
| 19 | Cassatt | Small Profile (Head of a Girl) | c. 1874 | 1 | 1 | 1 |
| 20 | Kulicke | Untitled (Pear) | 1964 | 0 | 0 | 1 |

1 = presence of semantic features, 0 = absence.


**Figure 1.** Extreme values of the activity fraction for images tested in this paper. Note that a low value of the activity fraction indicates that most pixels are not active (i.e., black) and only a few are very active pixels (light tones), as in the image on the left. A high value indicates that most pixels are active at a mid-level intensity, as in the image on the right. Left: *Nike* by Achilles Droungas, Gift of Mr. and Mrs. Stelios St. Joannou. Right: *Eight Views of Nara* by Tosa Mitsuyoshi, George and Mary Rockwell Collection. Images courtesy of the Herbert F. Johnson Museum of Art, Cornell University. To view this figure in colour, please see the online issue of the Journal.



pixels have high intensities and the rest show very low intensity. Extreme values of $S$ for our images are shown in Figure 1.

## Perceptual testing

Participants in the perceptual testing were undergraduate students from Manhattan College. There were a total of 24 participants for the landscape test and 19 for the portrait/still-life test. Participants were naïve to the purpose of the experiment. A pretest survey of art experience (in terms of art classes taken) showed a variety of levels of exposure, though most had taken one or zero art classes.

Image pairs were presented at roughly 0.5 m on 17-inch computer displays surrounded by a neutral grey background. Both images in each pair were scaled to roughly the same horizontal size and randomly placed on the left- or right-hand side. Images were displayed such that they were centred vertically and their horizontal centres were equidistant from the centre of the screen. Participants input integer ratings from 1 (= "not similar") to 9 (= "very similar") on a computer keyboard for each pair. The test was self-paced and participants were permitted to change their response if desired before proceeding to the next trial. Anchoring and ordering effects were controlled by randomizing the order of the pairs for each participant.

## RESULTS

We performed classical and nonclassical MDS on the mean similarity data across participants (after reversing the similarity index to generate dissimilarities). Generally speaking, classical and nonclassical MDS scalings produced essentially the same solutions (except for some axis inversions), and correlations of axes with statistics (described later) were similar for both solutions. The first three dimensions in the classical MDS (CMDS) solution for rating means averaged across participants accounted for approximately 20%, 14%, and 11% of the data variance, respectively, for both image sets. Scales 1 and 2 for the two image sets are plotted in Figure 2.

### Assigning units to MDS scales

As expected, both semantic and statistical measures were found to be strongly correlated with the first or second dimensions of the CMDS solutions. In particular, for landscapes the first dimension was strongly correlated with inclusion of humans in the image foreground (Spearman's $r = .79$, $p < .0001$), whereas the second dimension was strongly correlated with the activity fraction (Pearson's $r = .79$, $p < .0001$). It should be noted that the amplitude spectrum slope of the images was nearly as correlated with the second dimension of CMDS as was the activity fraction (Pearson's $r = .77$, $p < .0001$). The third scale had nearly the same correlation strength for both the absence of water in the image (Spearman's $r = .44$, $p < .05$) and for the DoG filtered images' intensity histogram skewness (Pearson's $r = .56$, $p < .05$).

For portraits, the first dimension of the CMDS solution was strongly correlated with the provenance of the work, i.e., whether it was produced in the eastern or western hemisphere (Spearman's $r = .79$, $p < .0001$), whereas the second dimension was strongly correlated with the presence of people in the image (Spearman's $r = .61$, $p < .005$). However, the activity fraction was also strongly correlated with the first CMDS scale (Pearson's $r = .67$, $p < .005$). No significant correlations could be found for the third CMDS scale for portraits.

Linear regressions of the first two dimensions of the CMDS solutions with the proposed scales are shown in Figure 3. The percentage of variability ($R^2$) explained by linear regression of CMDS scales and statistical measures or semantic classifiers is shown in the frame of the figures.

We found that combinations of statistical measures and/or semantic variables produced stronger correlations compared to measures considered separately (Figure 4). In particular, for landscapes, the sum of the activity fraction and the amplitude spectrum slope for each image had a correlation

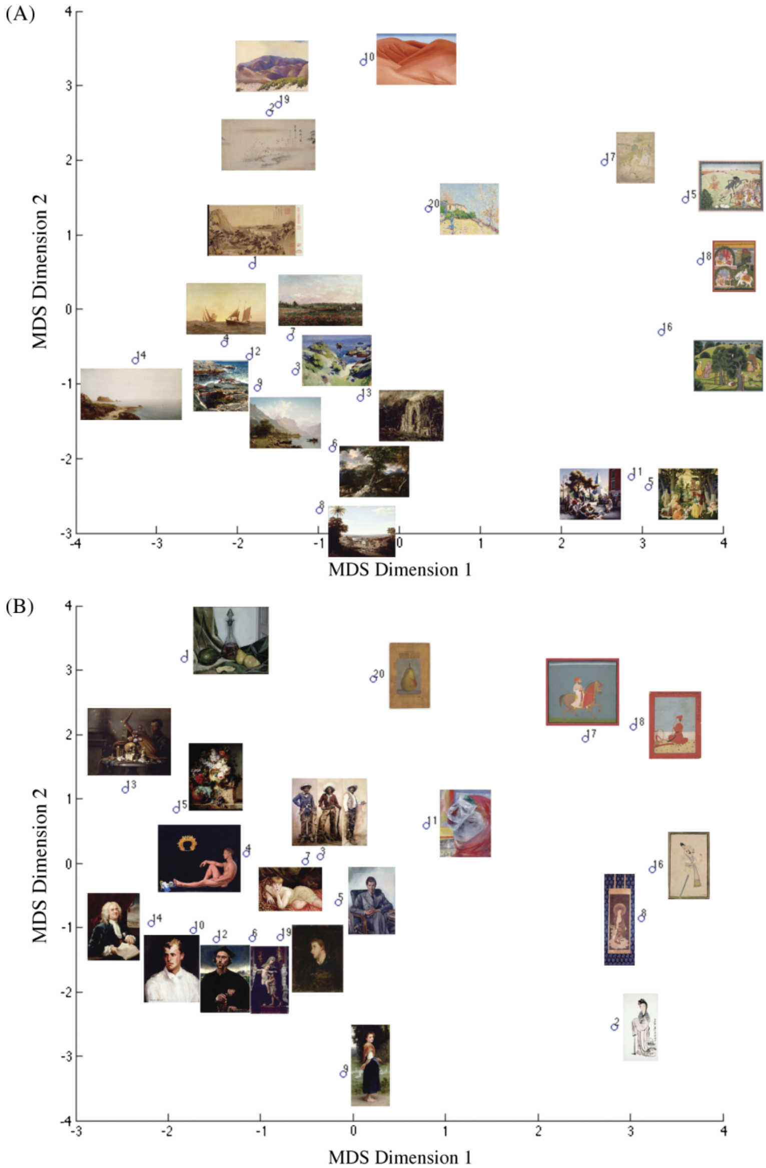


**Figure 2.** (A) Landscape paintings are plotted in a plane according to their value along the first two classical MDS dimensions. MDS was computed over mean dissimilarity (i.e., perceptual distance) ratings of 24 participants. (B) Portrait/still-life paintings are plotted in a plane according to their value along the first two classical MDS dimensions. MDS was computed over mean dissimilarity (i.e., perceptual distance) ratings of 19 participants. All images courtesy of the Herbert F. Johnson Museum of Art, Cornell University. To view this figure in colour, please see the online issue of the Journal.

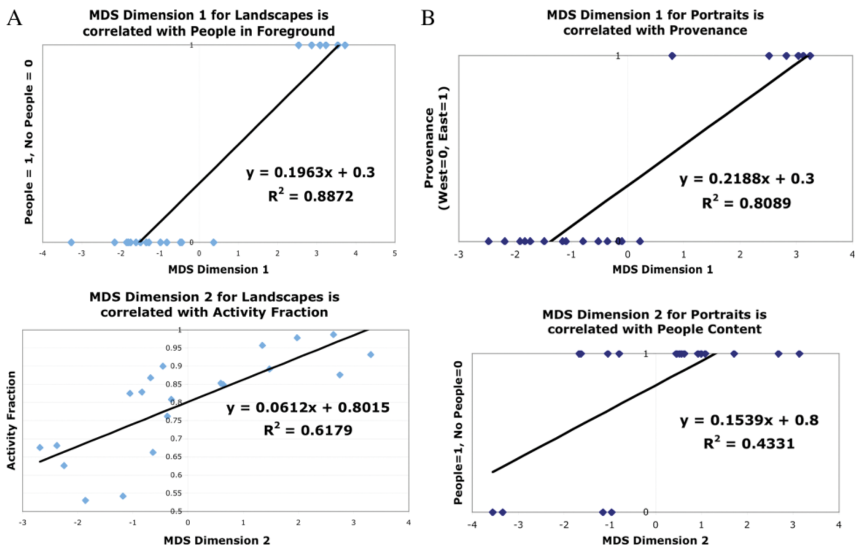


**Figure 3.** (A) Regressions for landscapes. (B) Regressions for portraits. To view this figure in colour, please see the online issue of the Journal.

of .84 ($p < .0001$) with the second CMDS scale. For portraits, the correlation of the first CMDS scale with the sum of the provenance variable and the activity fraction was .91 ($p < .0001$).

## DISCUSSION

We have shown that basic image statistics describing spatial structure and intensity distributions of artwork can play an important role in predicting the judgement of similarity, at least for images perceived as landscapes. Although semantic variables appear to be the dominant dimensions that observers use to determine similarity, one of two major axes of similarity ratings for landscapes is well predicted by intensity statistics. We also find that combinations of semantic classifiers and/or statistical measures can generate an even stronger correlation for both painting sets.

### Perception and regular statistics in paintings

Though viewing context is an important factor, paintings may be especially amenable to simple, objective similarity prediction algorithms because they are a distinct class of images. Indeed, they are distinct because they share many statistics with natural scenes, which separates them statistically from

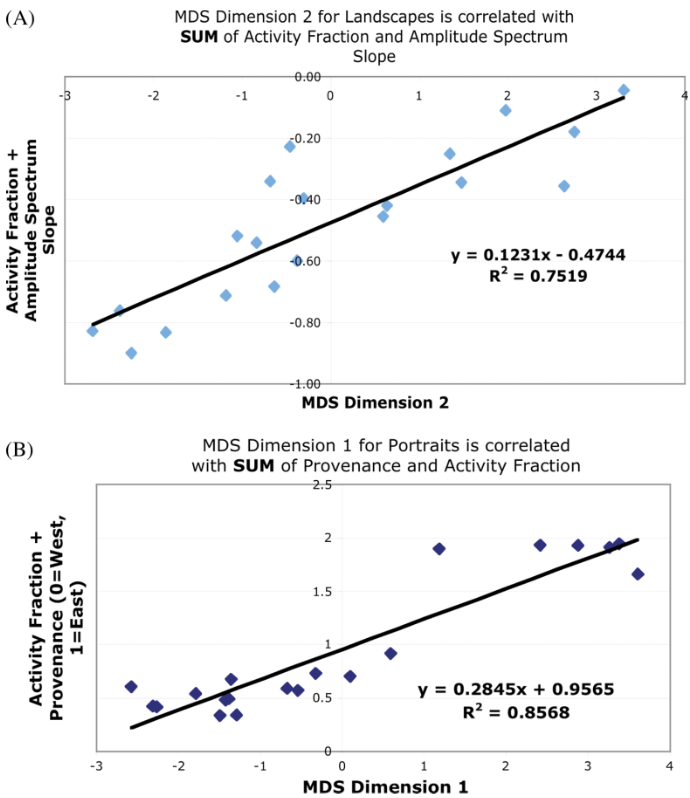


**Figure 4.** (A) Regression for combination of statistics (landscapes). (B) Regression for combination of statistical and semantic variables (portraits). To view this figure in colour, please see the online issue of the Journal.

random images, and they are also distinct because their production requires restrictions that are not imposed on natural scenes. Though artists are often free to create whatever image appeals to them, they seem generally inclined to create perceptible works. A perceptible work will possess many of the same statistical characteristics as are found in natural scenes, since the visual system is thought to be well adapted to the processing of natural scene statistics (see e.g., Field, 1987). As has been previously shown, this means artwork across the centuries generally possesses 1/*f*-shaped spatial frequency amplitude spectra (Graham & Field, 2008b), as well as other statistical regularities, since these regularities match or efficiently transform regularities in natural scenes.[3] Moreover, some of these same regularities are found

[3] Other authors suggest the regularity of the amplitude spectrum is due to an innate preference for fractal-like scaling (Redies, 2007).

to vary systematically with diagnostic content in natural scenes, as discussed earlier. Therefore, by exhibiting many of the regularities of natural scenes, art would appear to be similarly categorizable using statistics.

Statistical limitations to which paintings, but not scenes, are confined further restrict the statistical space occupied by paintings. Though samples of the mean amplitude spectrum slopes of art and natural scenes are found to be significantly different, a more crucial distinction between natural scenes and art involves luminant intensities, which are restricted to a far small dynamic range for art (Graham & Field, 2007, 2008a, 2008b). This perhaps explains why the activity fraction, a global measure of intensity characteristics, proved to be most useful in gauging perceived similarity for paintings.

No colour information was used in these experiments. As noted, other groups have shown that colour statistics can play an important role in similarity judgements of natural scenes. Our goal in the present study was to determine the contribution of spatial statistics and luminant intensity statistics only.

Note that while the amplitude spectrum slope and activity fraction were significantly correlated with each other in both sets of images (for landscapes, Pearson's $r = .60$, $p < .005$; for portraits, Pearson's $r = .47$, $p < .05$), the strength of the correlation was less than that between each statistic and the corresponding MDS scale. The same correlation strengths (and $p$-values) between spectrum slope and activity fraction held within the larger sets of images from which the 20 portrait/still-life and 20 landscape images were selected.

We note that our analysis implies that some nonlinear transformation of the images is desirable (calculation of both the amplitude spectrum slope and the activity fraction involves nonlinear operations). If similarity relationships are determined by images' relative distance along a nonlinear manifold within an orthonormal perceptual space, such structure would not be uncovered using Euclidean distances among untransformed images. In that case, techniques like local linear embedding (LLE) or other schemes could help identify the nonlinear manifold (e.g., Tenenbaum, de Silva, & Langford, 2000).

In summary, we find that basic image statistics predict a considerable portion of the variance in similarity judgements of representational art. It remains unclear whether the brain uses this information to encode perceptual qualities of scenes. In principle, such statistics could be collected by the early visual system, which would assist rapid discrimination based on context. The relevant image statistics are ones that have previously been implicated in shaping strategies for efficient coding of natural scenes. Because art images as a group comprise a statistically and perceptually circumscribed subset of the larger class of natural images, we suggest that they are especially useful for the study of relationships between perceptual judgements and basic statistics relevant to early vision.

## CONCLUSION

If art is an efficient way to communicate visual qualities, it would seem a natural tool for the study of many levels of visual processing. Our results suggest that basic statistics could influence perceptual and cognitive responses: Statistical properties appear to be good predictors of human judgements of the similarity of pairs of artworks, at least for landscape painting. An efficient visual system could make a quick and reasonable guess as to the relationship of a given image to others (i.e., the context of the image) by extracting these basic statistics early in the visual stream. This may hold for natural scenes as well as art.



## REFERENCES

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*, *4*, 2379–2394.

Graham, D. J., Chandler, D. M., & Field, D. J. (2006). Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields? *Vision Research 46*, 2901–2913.

Graham, D. J., & Field, D. J. (2006). Sparse coding in the neocortex. In J. H. Kaas & L. A. Krubitzer (Eds.), *Evolution of nervous systems* (Vol. 3, pp. 181–187). Oxford, UK: Academic Press.

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. (2008a). Global nonlinear luminance compression in painted art. *Proceedings SPIE: Computer Image Analysis in the Study of Art, 6810*, 189–200.

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

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*, 1233–1258.

Jones-Smith, K., & Mathur, H. (2006). Fractal analysis: Revisiting Pollock's drip paintings. *Nature*, *444*, E9–E10.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1–27.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Thousand Oaks, CA: Sage.

Landau, E., & Cernuschi, C. (2007). *Pollock matters*. Chicago: University of Chicago Press.

Lyu, S., Rockmore, D. N., & Farid, H. (2004). A digital technique for art authentication. *Proceedings of the National Academy of Sciences, USA*, *101*, 17006–17010.

Motoyoshi, I., Nishidam, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, *447*, 206–209.

Neumann, D., & Gegenfurtner, K. R. (2006). Image retrieval and perceptual similarity. *ACM Transactions in Applied Perception*, *3*, 31–47.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176–210.

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

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

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

Redies, C. (2007). A universal model of esthetic perception based on the sparse and efficient coding of natural stimuli. *Spatial Vision*, *21*, 97–117.

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

Rogowitz, B. E., Frese, T., Smith, J. R., Bouman, C. A., & Kalin, E. (1998). Perceptual image similarity experiments. *Proceedings SPIE: Human Vision and Electronic Imaging*, *3299*, 576–590.

Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, *73*, 713–726.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, *27*, 125–140.

Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, *7*(1), 11–32.

Taylor, R. P., Guzman, R., Martin, T. P., Hall, G. D. R., Micolich, A. P., Jonas, D., et al. (2007). Authenticating Pollock paintings using fractal geometry. *Pattern Recognition Letters*, *28*, 695–702.

Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, *290*, 2319–2323.

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

*Manuscript received August 2008*
*Manuscript accepted March 2009*
*First published online July 2009*