Efficient visual system processing of spatial and luminance statistics in representational and non-representational art

Daniel J. Graham[*a], Jay D. Friedenberg[b] and Daniel N. Rockmore[a,c]

[a] Department of Mathematics, Dartmouth College, Hanover NH 03755 USA
[b] Department of Psychology, Manhattan College, Riverdale NY 10471 USA
[c] Department of Computer Science, Dartmouth College, Hanover, NH 03755 USA

ABSTRACT

An emerging body of research suggests that artists consistently seek modes of representation that are efficiently processed by the human visual system, and that these shared properties could leave statistical signatures. In earlier work, we showed evidence that perceived similarity of representational art could be predicted using intensity statistics to which the early visual system is attuned, though semantic content was also found to be an important factor. Here we report two studies that examine the visual perception of similarity. We test a collection of non-representational art, which we argue possesses useful statistical and semantic properties, in terms of the relationship between image statistics and basic perceptual responses. We find two simple statistics—both expressed as single values—that predict nearly a third of the overall variance in similarity judgments of abstract art. An efficient visual system could make a quick and reasonable guess as to the relationship of a given image to others (i.e., its context) by extracting these basic statistics early in the visual stream, and this may hold for natural scenes as well as art. But a major component of many types of art is representational content. In a second study, we present findings related to efficient representation of natural scene luminances in landscapes by a well-known painter. We show empirically that elements of contemporary approaches to high-dynamic range tone-mapping—which are themselves deeply rooted in an understanding of early visual system coding—are present in the way Vincent Van Gogh transforms scene luminances into painting luminances. We argue that global tone mapping functions are a useful descriptor of an artist's perceptual goals with respect to global illumination and we present evidence that mapping the scene to a painting with different implied lighting properties produces a less efficient mapping. Together, these studies suggest that statistical regularities in art can shed light on visual processing.

Keywords: natural scenes, visual system, scene perception, context, similarity, efficient coding, artist's look-up table, retina, nonlinearities, luminance compression.

1. INTRODUCTION

Studies of categorization for natural images have demonstrated relationships between basic image statistics and classes of natural scenes,[1] and between basic statistics and perceptual judgments.[2] These same statistics, which generally measure higher-order or nonlinear structure, are relevant to efficient visual coding in primates.[3] However, it is not clear whether variations in these statistics are used by the visual system to determine scene class, context (i.e., how one scene relates to others in its class), or other perceptual information. Art offers a promising arena for attacking this question. Since art is created expressly for human viewing, it may be especially germane to the understanding of efficient human visual processing.[4,5,6] Here we examine two examples of relationships between statistics and perception, which parallel those for efficient natural scene perception. In particular, we study the notion of similarity in representational and non-representational works. The first study maps the perceptual similarity space of a group of abstract paintings and relates the intrinsic dimensions of this space to statistics relevant to early visual coding. The second study attempts to statistically characterize a well-known artist's approach to luminance scaling in landscape painting. Together this work demonstrates that artworks as a group may possess consistent statistical properties, which match or efficiently transform

* daniel.j.graham@dartmouth.edu; phone: +1 603 646 9020; fax: +1 603 646 1312; http://math.dartmouth.edu/~dgraham/dg/

those of natural scenes. Moreover, variations in these statistics could be tied to basic judgments of similarity, both for the observer, in the case of non-representational art and for the artist himself, in the case of representational art.

1.1 Background

Converging evidence demonstrates that paintings show many statistical regularities. Some of these regularities match those of natural scenes,[4,5,7] and this match suggests an influence of early visual coding strategies in the production of art. That is, coding strategies in the retina and cortex that are efficiently matched to the regular statistics of natural scenes could bias the production of art towards those images that can be efficiently processed, i.e., towards art that contains natural scene-like regularities.[6]

Is art efficient, not just at conveying scene-like statistics, but also for conveying perceptual and cognitive information related to semantics, as well perhaps as expressive and aesthetic qualities? If art is an efficient way to communicate such qualities, it would seem a natural tool for the study of many levels of visual processing.[§] Preliminary evidence shows that basic statistics may influence higher level perceptual and cognitive responses: statistical properties appear to be good predictors of human judgments of the similarity of pairs of artworks,[8] and other groups have found statistics that appear relevant to higher-level cognitive properties including aesthetic[9,10,17] and affective responses.[11] In addition, other high-level aspects of visual art appear to show statistical regularities. Paintings of known authorship and/or provenance show consistent spatial and intensity statistics. For example, a growing body of work suggests that an artist's style may leave statistical fingerprints.[6,12,13] Consistent with earlier findings,[14,15] this work suggests that basic statistical measures may be linked to a variety of perceptual qualities, and they may also be useful compliments to tag-based approaches to image retrieval.

However, some statistics differ in fundamental ways from those of natural scenes due to the optical properties of paintings.[4,6] This fact is interesting on its own but we can use it to study the interplay of perception and nonlinear statistics. In particular, intensity statistics are markedly different for paintings and scenes. Transforming highly skewed, high-dynamic range scene luminances into more Gaussian-distributed, low dynamic range painting luminances requires nonlinear scaling, much as the retina must perform nonlinear scaling.[18] This finding suggests artists must transform natural scene statistics in order to achieve a satisfactory representation, and it may therefore constitute an efficient solution to the problem of scene representation. It should be noted, though, that artists may have varying degrees of interest in matching scene luminances efficiently.

Though painting has a greatly restricted range of luminances with which to reproduce scene luminances, painting also allows greater variation than is available in photographs, since each area can be "dodged and burned" (i.e., locally under- or over-exposed) to effectively increase the dynamic range.[6,19,20,21] Paintings not only must work with smaller dynamic ranges, they also typically work with less sparseness in pixel and wavelet coefficients.[4] But reduced flexibility (at least in a statistical sense) may not be a disadvantage when the goal is to communicate information. Greatly reduced forms of representation such as drawings, and especially caricatures, in fact appear better at communicating object identities compared to photographs.[22,23] Taken together, this could mean that art in general is especially efficient at communicating information, and therefore the elements of efficient representation are well isolated in the corpus of art images. In other words, we can use the efficient coding hypothesis in reverse: the efficient coding hypothesis[3] states that regular statistics in the natural world are processed efficiently by the visual system since evolution will demand a maximally efficient solution, and since retinal processing is so well conserved across vertebrates,[24] and, especially, across mammals. To the extent that statistics of the typical visual environment are regular, one can seek to test the extent to which general processing strategies in the visual system are well matched to those regularities. This can be done with respect to information theoretic limits, a method which has recently been reviewed.[25] Here we argue the opposite: since art shows evidence of being efficient, not only at matching and transforming scene statistics, but also at communicating information and affect, perhaps it is characteristic of the *human* visual environment. That is, in the same way that one can analyze visual processing in the crab with respect to the crab's visual environment,[26] so we can analyze humans with respect to art.

In the present paper, we first examine abstract art from the perspective of similarity perception. Though humans can discriminate scenes quite rapidly,[27] it is unclear what factors influence an observer's judgment of similarity. Though semantic information relevant to cognitive processes is an important contributor, scene statistics may also play a role.[1,28,29] Abstract art is a useful class of stimuli with which to study this question since it is generally devoid of

---

[§] Analogous notions of the efficiency of human language for communicating information have also been proposed.[16]

semantic content but it is also designed for human viewing and may possess regular statistical properties which allow efficient visual processing. We investigate the role of global image statistics (both spatial statistics and luminance statistics) in judgments of similarity for abstract art, and we compare these results with our previous findings involving representational art.[30] This first experiment demonstrates that lacking other cues, humans discriminate among complex natural images in a way that correlates strongly with variations in nonlinear and higher-order image statistics. We argue that this finding is important to the study of scene perception in humans because abstract art, like all art, is designed for human viewing and may therefore be efficiently matched to human visual processing strategies.

But one of the main purposes of art is to represent objects and scenes from the world. From the perspective of efficient coding, landscape painting is an especially interesting class of artwork since the goal of such works is at some level to produce a novel but comprehensible scene from the natural world. In the second half of this paper, we investigate the role of global lighting in landscape painting, an area that has recently received popular attention.[31] In the 1970s, artist Joseph Beuys proposed that the dramatic effect of lighting depicted in the landscapes of Jacob Ruisdael and other Dutch Golden Age painters was largely a result of light reflections from the inland Zuider Zee waterway. Further, he argued that 1950s land reclamation, which sharply reduced the area of the waterway, destroyed this effect. Contemporary art historians describe this scenario as unlikely but the fact that global illumination would generate passionate debate is indicative of its importance for the visual properties of representational images, and specifically of landscapes. In our second study, therefore, we consider luminance statistics of landscape painting. Landscapes are useful from the point of view of visual perception for many reasons, but here we focus on the question of how artists are able to efficiently represent the large dynamic range of luminances present in the world. In particular, we examine Van Gogh's strategy for representing summer landscape in Arles, France.

2. STUDY 1: SIMILARITY IN NON-REPRESENTATIONAL ART

It has been proposed that context is an important determinant of object identification strategies in human vision,[32] and one area within this field of research has focused on the information observers use to determine context, i.e., to identify scenes. Discrimination of complex natural images by human observers is known to be reliable even for very brief presentations[27,33] but there is debate as to the contribution of top down and bottom up factors in scene perception.[34] Both scene statistics (e.g., saliency maps[35] and global statistics[28]) and higher-level cognitive mechanisms (e.g., learning, task-specific search strategies, naïve physics, and experience in similar environments[36]) may influence such discriminations. Though many statistics (e.g., local contrast, spatial frequency content) of fixated regions have been shown to differ from those of unfixated regions,[37] models based on these statistics are not generally successful at predicting fixations in active viewing tasks.[38] Much evidence supports the notion that basic statistical regularities that are "diagnostic" of scene categories could in principle be extracted by the early visual system, though the question of whether the brain uses such statistics remains open.

Here we show that basic statistics are likewise related to variations in the perceived similarity of artwork, and that this finding has bearing on similarity perception more generally. Art has special import for the visual system, since it is a class of stimuli designed for human viewing. Art also has special import for the study of scene perception because the act of evaluating a novel artwork is in large part about the establishment of context. Paintings, including abstract works, can be thought of as summaries or distillations of many natural scenes, compressed into a low dynamic range image. This idea is somewhat similar to Zeki's notion that art generally represents perceptual constancies.[39] That is, art captures not only perceptual constancies associated with object identification, but also statistical regularities that are shared by many perceptually related scenes.

Perceptual aspects of art such as judgment of similarity have been studied extensively by cognition researchers as a means of understanding how the brain represents relatedness.[40,41,42,43] In previous work, we have employed multidimensional scaling analysis of observers' similarity ratings for paired paintings, and we compared resulting dimensions to a host of statistical measures, modeled neural responses, and semantic variables derived from image metadata.[30] Three image sets were classified in a prior forced choice-test as "landscapes," "portraits/still-life," and "abstract." For both the landscapes and the portraits, one of the first two similarity dimensions was highly correlated with a measure of intensity distribution sparseness, and for landscapes this correlation explained a greater portion of data variance than did semantic variables. Spatial statistics such as spatial frequency amplitude spectrum slope were nearly as good predictors as semantic variables. Given these results, we have turned our attention here to abstract art.

It could be argued that the first human art was abstract art: the 75,000 year old Blombos fragment shows a carved pattern of regularly spaced ×s bordered on the top and bottom by horizontal lines.[‡] Let us suppose that abstract art possesses a similar level of "intrinsic interestingness"[45] compared to representational art. This simply means that it is created, like representational art, to be seen by other humans, and may therefore be expected to obey certain statistical regularities.

However, our previous findings also show that abstract art has a significantly shallower spatial frequency power spectrum compared to representational art, and both are more shallow than natural scene spectra.[4,5,6,7] In other words, abstract art has relatively less low spatial frequency amplitude compared to representational art and natural scenes. We have suggested that this may be due to the lack of shading typical of abstract works, which tends to decrease high frequency amplitude and boost low frequencies. Moreover, abstract art does not appear to possess luminance statistics that will efficiently transform natural scenes. As reported in an earlier paper,[6] a set of luminance-calibrated natural scenes was scaled such that each scene was histogram-matched to the ensemble intensity histogram for a diverse set of art categories, including Western works, Eastern works, American landscapes, abstract works, and other categories. Of the six categories tested, the set of scenes transformed to match the abstract art histogram was rated with the lowest preference by observers (we will return to the idea of nonlinear luminance scaling in Study 2).

It is as yet unclear how these special spectral and luminance properties may interact but it seems safe to say that abstract art is compelling[46] despite the fact that it differs in basic statistical ways compared to representational art and to natural scenes. Indeed, even significant deviance from the manner in which representational art transforms luminance and matches spectral power can still achieve a work of art satisfactory enough to hang in a museum. But recent work[11] suggests that some images that deviate strongly from natural statistics are not favored by observers. Through a suite of studies, they found that abstract art that is statistically abnormal in terms of spatial frequency power made observers uncomfortable. They suggest that such aversion may be related to stimuli that can induce seizure and migraine in those susceptible to such conditions. This finding is consistent with the notion that perception (e.g., discrimination) of stimuli such as white noise is difficult because the visual system is not designed to operate in such an environment. The finding of a link with aversion shows that certain unnatural statistics may be more problematic for the visual system than others.

In order to classify the space of abstract art in a basic perceptual sense (as with, e.g., textures[47]), we tested observer ratings of simple similarity for a set of digitized abstract art images. Basic, nonlinear, and higher-order statistics were then compared to the principal axes of the similarity data (classical multidimensional scaling, which is equivalent to principal components analysis). Color is a surely an important factor in scene perception but as with our previous paper, we focus here on the role of spatial and luminance statistics.

2.1 Methods

Six observers divided a set of 140 digitized paintings from a major university collection (Herbert F. Johnson Museum of Art, Cornell University, Ithaca, NY USA; see earlier paper for description of database[4]) into content categories using a three-alternative forced choice paradigm (choices were "landscape," "portrait/still-life," or "abstract"). Images classified by the majority of observers as abstract (18 images) were then included in a larger test of similarity perception. Participants in the similarity perception testing included 27 undergraduate students from Manhattan College. A survey of the number of art-related courses taken showed a variety of levels of exposure, though the vast majority had taken one or zero art-related courses. Image pairs were presented at roughly .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 centered vertically and their horizontal centers were equidistant from the center 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. Randomizing the order of the pairs for each participant controlled anchoring and ordering effects.

2.2 Results

We analyzed the similarity data using multidimensional scaling,[48,49] as in our previous work.[30] We again found that classical and non-classical MDS produced similar solutions and correlations with statistical measures, though we only report values for the classical MDS solution (see Table 1). To an even greater degree than was found with

---

[‡] D'Errico et al. call this artifact "the most ancient irrefutable evidence for symbolic behavior."[44]

representational art (i.e., landscapes and portraits),[30] higher-order statistics were strongly correlated with the first two MDS scales, which together explain 45% of the overall data variance. This is a greater portion of variance explained for these two scales compared to representational art, which suggests that the leading factors in the determination of similarity are relatively more important for non-representational works. With little or no representational content, it follows that there are fewer perceptual dimensions along which images can vary. This conclusion is supported by another finding: the distribution of similarity ratings fell off steeply but linearly, and was well fit with a line of slope 81.3 (R^2=.995), which stands in contrast to our tests of landscapes and portraits. Landscapes fell off proportional to exp(-x), where x is the similarity rating, while ratings of portraits fell off linearly but at a faster rate (-140.6). Representational works also had lower mean similarity ratings than non-representational works. In other words, pairs of non-representational works are more similar to one another than are representational works.

We found that a sum and a difference of the activity fraction measure of sparseness[◊,50] and the log-log slope of the spatial frequency amplitude spectrum explain roughly a third of the overall variance in the similarity data. The summing and subtracting operations serve to rotate the image data so as to better align with empirical dimensions (i.e., the first two MDS scales).[Δ] Figure 1 shows images plotted with respect to the sum (horizontal axis) and difference (vertical axis) of the activity fraction and the amplitude spectrum slope. MDS scale 1 and the summed statistics have a correlation of .85 ($p < 0.0001$); MDS scale 2 and the subtracted statistics have a correlation of . 81 ($p < 0.0001$). A linear fit of the two relationships explained 72% and 66% of the data variance, respectively.

Table 1.Correlations strengths are listed with regard to the first 2 MDS scales and various statistics. All are Pearson correlation, except for the symbolic content variable, which is Spearman correlation.

	Linear Images		**Log-transformed Images**	
Statistic/Variable	MDS Scale 1	MDS Scale 2	MDS Scale 1	MDS Scale 2
Mean	*0.5364*	**0.6655**	*0.5466*	0.7287
Variance	*-0.5291*	*-0.5092*	*-0.5878*	*-0.5114*
Skew	*-0.4724*	**-0.6288**	x	**-0.6073**
Kurtosis	**0.5962**	x	*0.4776*	x
Activity Fraction	*0.5814*	**0.7028**	*0.5109*	*0.5819*
Amplitude Spectrum Slope	0.8054	x	x	*-0.5457*
DoG filtering: Variance	**-0.6512**	x	-0.7839	x
DoG filtering: Skewness	x	-0.7292	x	*-0.6976*
DoG filtering: Kurtosis	x	**-0.6652**	0.5707	x
Symbolic Content	x	x	N/A	N/A

NOTE: *p < .05*, **p < .01**, p < .0001, x = no significant correlation

◊ *S*, the activity fraction,

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

is measured over n pixels each with intensity r_i and S 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 pixels have high intensities and the rest show very low intensity.

Δ We note, however, that MDS does not establish the orientation of the principal axes.

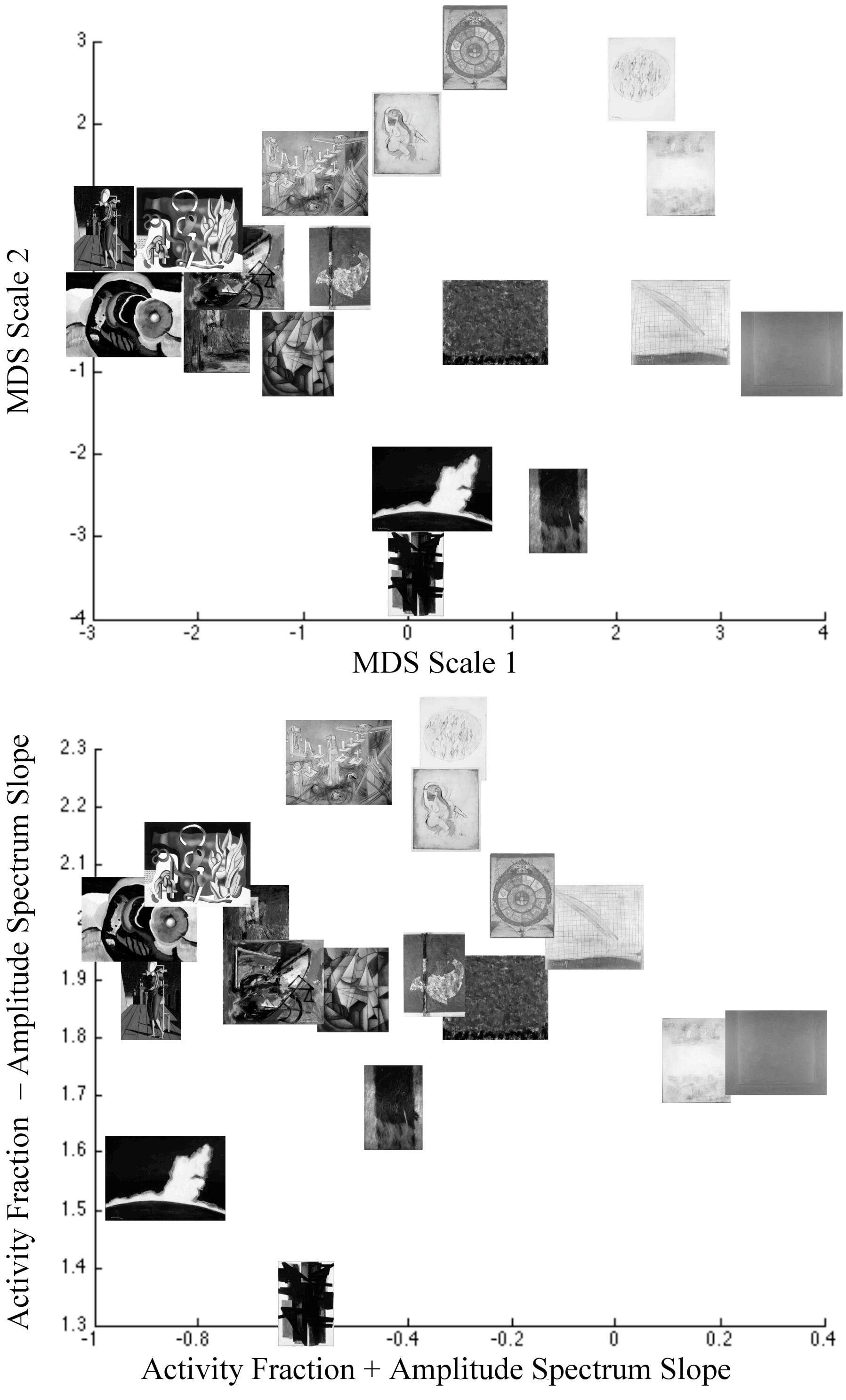

Figure 1. Top plot shows images plotted with respect to the first two classical MDS (PCA) scales. Bottom plot shows images plotted with respect to the sum (horizontal axis) and difference (vertical axis) of the activity fraction and the amplitude spectrum slope. MDS scale 1 and the summed statistics have a correlation of .85 ($p < 0.0001$); MDS scale 2 and the subtracted statistics) have a correlation of . 81 ($p < 0.0001$).

We also tested the correlations of these MDS scales with statistics calculated after simulated photoreceptor processing (log-like nonlinearity) and after simulated retinal ganglion cell filtering. Filters were constructed based on physiologically determined parameters of a 2D difference-of-Gaussians (DoG) model and convolved with the paintings (an earlier work describes the filtering procedure[4]). Following these transforms, we found that the mean intensity statistic showed a stronger correlation with MDS scale 2 compared to the activity fraction of the linear image (see Table 1). The same is true of the skew statistic after filtering the images with retinal ganglion cell-like center-surround filters. Both of these statistics are strongly correlated with the activity fraction and with each other (correlation of at least 0.75, $p<0.0001$). Including the transforms, all three of these statistics are measures of nonlinear intensity statistics in the paintings, which supports the notion that nonlinear statistics play an important role in visual coding and perception.[3,51] Moreover, it has been proposed that neural mechanisms in the early visual system could in principle be sensitive to skewness statistics. In particular the perception of glossiness has been found to be correlated with skew, in a study that we note employs art objects as stimuli.[2]

It should be noted that although some of the images contained representational content, the presence of this content (which was gauged using image metadata) was not significantly correlated with any of the first three MDS dimensions, suggesting that content had been abstracted sufficiently to make it a weak factor in similarity judgment.

In the absence of a consistent scheme for classifying affective response to color, line, etc., (but see Van Dantzig[52]) we are left with only these global statistics to perform an objective analysis of similarity for non-representational painting. Similarity is admittedly a potentially ambiguous property, and one potentially dependent on viewing environment and other factors. Nevertheless, our results suggest that human visual processing could use global scene statistics in the efficient encoding of basic perceptual properties of complex scenes.

3. STUDY 2: NONLINEAR LUMINANCE SCALING IN VAN GOGH LANDSCAPE

We now turn our attention from abstract painting to landscape painting, a class of artwork that is interesting for stylistic and statistical reasons. Stylistically, landscapes call for a particularly strong representational intention on the part of the artist, meaning that every region of space is typically filled with detail in order to reduce ambiguity of the scene. Consider that minimalist, caricatured or cubist treatment of landscape is relatively rare (at least in Western art history), which may be due in part to the fact such images are likely to be more ambiguous than a comparably treated portrait/still-life of small numbers of objects or people. It bears mentioning that in Europe, the depiction of naturalistic landscape was mostly unknown until late medieval times, and early techniques for portraying such scenes were adapted from those used for depicting portrait sitters and other individuated, near depth objects. Clark, for instance, writes: “The art of painting, in its early stages, is concerned with things which one can touch, hold in the hand, or isolate in the mind from the rest of their surroundings. This is the feeling which underlies the advice of Cennino Cennini, the last spokesman of the medieval tradition of painting, when he says, ‘if you wish to acquire a good manner of depicting mountains and make them look natural, get some large stones, which should be rough, and not cleaned, and portray them from nature, applying the lights and darks according as reason permits you.’”[53] Needless to say, this strategy was soon modified. From the Renaissance onward, artists developed new techniques that were necessary for naturalistic depictions of rural and urban landscape, but not for earlier types of painting (e.g., atmospheric and linear perspective).

In terms of statistics, it is possible to find the tone mapping of a landscape painting to a natural scene representative of the depicted scene[6] using histogram specification (also called histogram matching[54]). The *artist’s look-up table* (ALUT), i.e., the resulting transform, can be used to describe an artist’s use of luminance scaling. Nonlinear scaling is required to compress the high-dynamic range of luminances present in natural scenes into the far smaller range available in paint. The ALUT summarizes the perceptual relationship between the tableau and the painting. If artworks are indeed efficient with respect to early visual system processing, it follows that artists will often employ highly efficient tone mapping strategies, which would be reflected in the ALUT. In Study 2, we recreate a scene depicted by Van Gogh and compare the luminance transforms of this scene for 1.) the painting of this scene and 2.) a painting of a different scene, one with very different implied lighting.

Vincent Van Gogh worked in Arles, Southern France, from February, 1888 until May, 1889, a period that was among his most prolific.[55] He was particularly interested in portraying the dramatic lighting of the area, especially having spent his formative years in Holland, nearly 10 degrees of latitude to the north, and a large portion of his output from this period was landscape works. Van Gogh is also interesting because he has a painterly approach to representation, rather than photo-realist style. We therefore attempted to reconstruct a tableau as similar as possible to those Van Gogh painted.

3.1 Methods

Of course, much has changed in Arles in more than a century. But it did prove possible, given the vagaries of the weather, to find representative scenes. One tableau (which can be viewed from the Cimetière des Neuf Collines northeast of Arles) in particular was mostly unchanged from its depiction in “Harvest Landscape,” F412 (1888). This image was painted in high summer, 1888[56]; our image of the scene was captured in late summer, 2008. Calibration of painting and scene luminances was performed as follows: uncompressed greyscale TIFF-format scans of ektachrome slides, which included Q-13 color strips (Eastman Kodak, Rochester, NY), were provided by the Van Gogh Museum in Amsterdam. We converted these from RGB to CIELAB colorspace using Adobe Photoshop (this conversion assumes a D50 illuminant, which is the one used to capture the ektachrome). We then fit the value of the lightness (L) channel for each color strip grey level to its corresponding reflectance. We applied this transform to the whole image to produce a reflectance map. For a painting illuminated by a single diffuse source, we can model reflectance as being proportional to luminance. This assumption will not hold for all images, nor for all paintings (especially those with thick, glossy impasto), but it is a reasonable simplification for Van Gogh’s paintings.

The scene tableaux were captured in RAW format using a Canon PowerShot S60 camera (Canon Inc., Tokyo, Japan). These images were calibrated with respect to luminances measured from grey levels on a ColorChecker chart (GretagMacbeth, New Windsor, NY) held 10 m from the camera, as well as from two spot measurements near the horizon, using a standard photometer (LS-100, Konica-Minolta, Tokyo, Japan). The ALUT was calculated by histogram matching the luminance-calibrated scene to the histogram of luminances for the corresponding painting. A previous paper gives further details about calibration procedures and calculation of the ALUT.[6]

3.2 Results

The recreated tableau of F412 is shown in on the left in Figure 2, along with its luminance histogram, and F412 and its histogram are shown on the right. The scene has a skewed, high dynamic range luminance distribution (dynamic range of 3,200:1) while the painting has a more Gaussian, low dynamic range distribution (dynamic range of 32:1). Transformation between these images requires a compressive nonlinearity like the ALUTs shown in Figure 3.

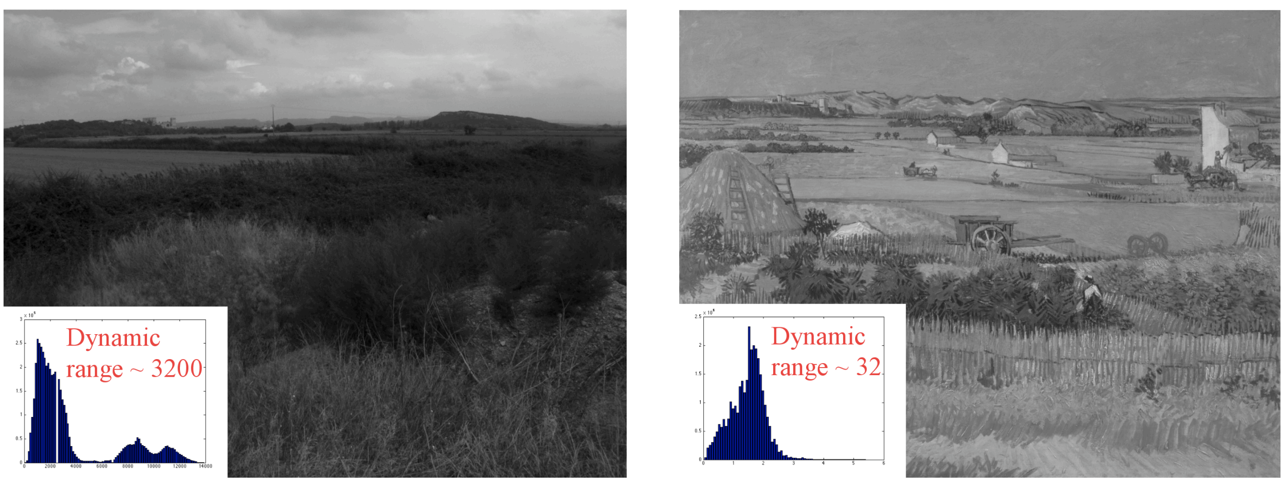

Figure 2. On left, “Harvest Landscape” scene, captured September 2008. On right, F412, “Harvest Landscape” by Van Gogh (June, 1888). Scene was calibrated with respect to a standard reflector to give luminance (candelas/meter^2). Painting was calibrated as described in text. Luminance histograms of the two images are inset. Note that the scene has a skewed, high dynamic range luminance distribution (dynamic range of 3,200:1) while the painting has a more Gaussian, low dynamic range distribution (dynamic range of 32:1). Transformation between these images requires a compressive nonlinearity. F412 image © Van Gogh Museum, Amsterdam.

As we have noted in earlier papers, the ALUT is not a full model of the painting process, or even of luminance scaling (it neglects local contrast adjustments, for example), but we point out that it does have explanatory power with regard to Van Gogh’s treatment of the luminance range. In Figure 3, we show our modern tableau of the F412 image matched to the F412 histogram and to the histogram of another Arles work, F486, “Les Alyscamps” (1888). Though the ALUTs differ only slightly, their effect is noticeably different. In particular, notice that the shadows under the near-ground bushes are detailed, but in the image mapped with the non-matching transform (F486), they are dark.

The consistency of the two ALUTs (Figure 3) also suggests that Van Gogh’s general strategy of using sigmoidal scaling is robust for two images with very different illumination: F412 depicts an open scene with sky, while F486 depicts a forested scene with no sky. That is, Van Gogh may have attempted to achieve consistent luminance statistics in his paintings despite changing luminance conditions in depicted scenes. We suggest that this maybe an important component of an artist’s style. We are investigating ways of using the ALUT to perform stylometric analysis, which could potentially be employed to assist in determination of authorship or historical ordering of works of unknown provenance.

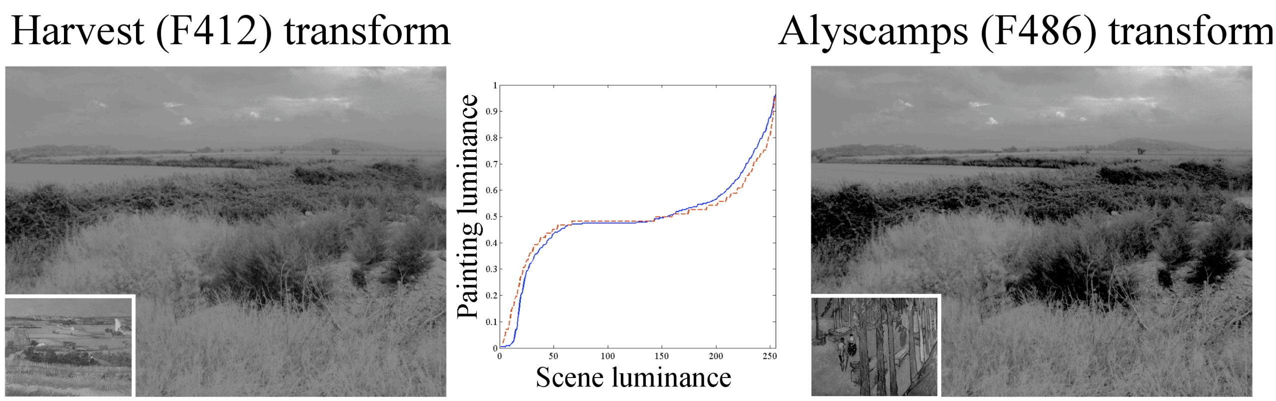

Figure 3. On left, scene transformed (tone mapped) to F412 “Harvest Landscape,” (inset). This transform is shown as the red (dotted) line is the center plot. Blue (solid) line is the ALUT transforming the same scene to F486 (“Les Alyscamps”), and the result of this transformation is on the right, with F486 inset. In center plot, horizontal axis represents scene luminances (linearly scaled to 0:255), vertical axis represents painting luminances (linearly scaled to 0:1). F412 and F486 images © Van Gogh Museum, Amsterdam.

Note that the ALUT for F412 can be inverted (so as to transform painting luminances to world luminances), which gives a familiar sigmoidal shape. Sigmoidal tone mapping functions are used often in imaging when detail in dark areas needs to be preserved.[57] The inverted ALUT can be thought of as the global nonlinear luminance stretching that would be applied to a painting in order to scale it to match corresponding scene luminances. This has strong parallels in the high dynamic range (HDR) graphics literature. For instance, one group[58] has used a scheme involving a sigmoidal-shaped look-up table (along with retina-like spatial filtering) in order to simulate “painting” images with a HDR using a low dynamic range display. Van Gogh—and other painters—would appear to employ a representational approach that is similar to a modern HDR tone mapping strategy.

4. DISCUSSION

Here we have presented two approaches that employ art to study efficient visual processing in humans. We have demonstrated that artists appear to represent similarity in an efficient way, such that basic perceptual information can be extracted early in visual processing. For representational art, this similarity is measured against the benchmark of the natural world, and one well-known artist appears to reproduce modern approaches to HDR tone mapping. For non-representational art, similarity is measured with respect to other artistic products, and we find that image statistics implicated in shaping efficient visual coding are associated with perceptual judgments.

We note that our analysis also implies that some nonlinear transformation of the images is desirable. Nonlinear operations are involved in calculation of the amplitude spectrum slope, the activity fraction, the mean intensity of the log-transformed image, and the skew of the DoG filtered images. In addition, the ALUT for Van Gogh (and for other artists) is clearly nonlinear. As more becomes known about the role of nonlinearities in efficient early visual system coding,[51] and in higher level perceptual and cognitive systems,[59] we may find candidate systems that could extract perceptual information rapidly and efficiently using statistical regularities.

5. CONCLUSION

The notion that basic nonlinear statistics are computed by the early visual system for use in perceptual judgments has been criticized as being unproven and perhaps motivated by mathematical convenience.[60] On the other hand, we have found robust relations between basic statistics and perceptual judgments, and these relations prove to be stronger when semantic content is removed, as in abstract art. Moreover, we have presented evidence that luminance compression in landscape painting may serve to produce an image that efficiently captures the perceptual effect of natural scene illumination. Together, this work shows that artworks are useful objects of study for vision scientists, and we suggest that if artworks as a group are characteristic of the human visual environment, their shared statistical properties could provide clues about efficient visual coding and perception.

Acknowledgements: This work was supported by a National Science Foundation Small Grant for Exploratory Research (DMS-0746667) and a grant from the William H. Neukom 1964 Institute for Computational Science. We wish to thank Eric Postma, Rick Johnson, Louis van Tilborgh, Ella Hendriks, Jim Coddington, David Field, Peter Nestor, Wilco Moerman, and the Van Gogh Museum for their generous assistance with this project.

6. REFERENCES

[1] A. Torralba. and A. Oliva, "Statistics of natural image categories," Netw.: Comput. Neural Syst. **14**, 391-412 (2003).

[2] I. Motoyoshi, S. Nishidam, L. Sharan. and E. H. Adelson, "Image statistics and the perception of surface qualities," Nature **447**, 206-209 (2007).

[3] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," Ann. Rev. Neurosci. **24**, 1193-1216.

[4] D. J. Graham and D. J. Field, "Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities," Spatial Vision **21**, 149-164 (2007).

[5] D. J. Graham and D. J. Field, "Variations in intensity statistics for representational and abstract art, and for art from the eastern and western hemispheres," Perception **37**, 1341-1352 (2008).

[6] D. J. Graham and D. J. Field, "Global nonlinear compression of natural luminances in painted art," Proc. SPIE **6810**, 68100K (2008).

[7] C. Redies, J. Hasenstein, and J. Denzler, "Fractal-like image statistics in visual art: similarity to natural scenes," Spatial Vision **21**, 137-148 (2007).

[8] D. J. Graham, J. D. Friedenberg, D. N. Rockmore and D. J. Field, "Mapping the similarity space of paintings: image statistics and visual perception," under revision for Visual Cognition.

[9] C. Redies, J. Hänisch, M. Blickhan and J. Denzler, "Artists portray faces with the Fourier statistics of complex natural scenes," Netw.: Comput. Neural Syst **18**, 235-248 (2007).

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

[11] D. Fernandez and A. J. Wilkins, "Uncomfortable images in art and nature," Perception **37**, 1098-1113 (2008).

[12] S. Lyu, D. N. Rockmore and H. Farid, "A digital technique for art authentication," Proc. Nat. Acad. Sci. USA, **101**, 17006-17010 (2004).

[13] J. M. Hughes, D. J. Graham and D. N. Rockmore, "Stylometric classification of Van Gogh paintings of undetermined provenance using sparse coding," In preparation for *Pattern Recognition Letters*.

[14] D. Neumann and K. R. Gegenfurtner, "Image retrieval and perceptual similarity," ACM Trans. Appl. Perception **3**, 31-47 (2006).

[15] M. J. Swain. and D. H. Ballard, "Color indexing," Intl. J. Computer Vision **7**, 11–32 (1991).

[16] R. Levy and T. F. Jaeger, "Speakers optimize information density through syntactic reduction," Proc. Neural Inform. Processing Syst. **19**, 1-6 (2006).
[17] C. M. Hagerhall, T. Purcell and R. Taylor, "Fractal dimension of landscape silhouette outlines as a predictor of landscape preference," J. Enivonmental Psych. **24**, 247-255 (2004).
[18] D. A. Baylor, B. J. Nunn. and J. L. Schnapf, "Spectral sensitivity of the cones of the monkey *Macaca fascicularis*," J. Physiol. **390**, 145-160 (1987).
[19] J. McCann, "Rendering high-dynamic range images: algorithms that mimic human vision," Proc. AMOS Technical Conf., 19-28 (2005).
[20] A. Adams, *The Negative*, New York Graphic Society, Boston (1948).
[21] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images," In SIGGRAPH **2002**, 267–276 (2002).
[22] T. A. Ryan and C. B. Schwartz, "Speed of perception as a function of mode of representation," Am. J. Psychol. **69**, 60-69 (1956).
[23] J. A. Ferwerda, "Three varieties of realism in computer graphics," Proc. SPIE **5007**, 290-297 (2003).
[24] B. L. Finlay, L. C. de Lima Silveira and A. Reichenbach, "Comparative aspects of visual system development," In: *The Structure, Function and Evolution of the Primate Visual System*, ed. Jan Kremers., John Wiley, New York (2004).
[25] D. J. Graham and D. J. Field "Efficient coding of natural images," In *Encyclopedia of Neuroscience* ed. Larry R. Squire, Academic Press, Oxford (2008).
[26] J. Zeil and J. M. Zanker, "A glimpse into crabworld," Vision Res. **37**, 3417-3426 (1997).
[27] M. C. Potter and E. I. Levy, "Recognition memory for a rapid sequence of pictures," J. Exp. Psychol. **81**, 10–15 (1969).
[28] A. Oliva and P. G. Schyns, "Diagnostic colors mediate scene recognition," Cognitive Psychol. **41**, 176-210 (2000).
[29] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," Intl. J. Computer Vision **42**, 145-175 (2001).
[30] D. J. Graham, J. D. Friedenberg, D. N. Rockmore and D. J. Field, "Mapping the similarity space of paintings: image statistics and visual perception," Visual Cognition, under revision. Available at http://tinyurl.com/9eldpe
[31] M. de Kroon and P-R. de Kroon, "Dutch light," (film) RNTV distribution, Amsterdam (2003).
[32] I. Biederman, R. J. Mezzanotte and J. C. Rabinowitz, Scene perception: detecting and judging objects undergoing relational violations, Cognitive Psychol. **14**, 143-177 (1982).
[33] L. Fei-Fei, A. Iyer, C. Koch and P. Perona, "What do we perceive in a glance of a real-world scene?" J. Vision *7*, 1-29 (2007).
[34] D. C. Richardson and M. J. Spivey, "Eye tracking: research areas and applications. In: *Encyclopedia of Biomaterials and Biomedical Engineering*, Eds., G. Wnek. and G. Bowlin, Marcel Dekker, New York (2008).
[35] L. Itti and C. Koch, "Computational modeling of visual attention," Nature Rev. Neurosci. **2**, 194-203 (2001).
[36] J. M. Henderson and A. Hollingworth, "High-level scene perception," Ann. Rev. Psychol **50**, 243-271 (1999).
[37] S. K. Mannan, K. H. Ruddock and D. S. Wooding, "Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images," Spatial Vision **11**, 157-178 (1995).
[38] J. M. Henderson, J. R. Brockmole, M. S. Castelhano and M. Mack, "Visual saliency does not account for eye movements during visual search in real world scenes," In: *Eye movements: a window on mind and brain*, Eds. R. P. G. van Gompel, M. H. Fischer, W. S. Murray and R. L. Hill, Elsevier, Oxford, 99-124 (2007).
[39] S. Zeki, "Art and the brain," Dædalus **127**, 71-103 (1998).
[40] F. G. Ashby and N. A. Perrin, "Toward a unified theory of similarity and recognition," *Psychol. Rev.* **95**, 124-150 (1988).
[41] F. G.Ashby and W. W. Lee, "Predicting similarity and categorization from identification," J. Experimental. Psychol.: General **120**, 150-172 (1991).
[42] S. Edelman. Representation of similarity in 3D object discrimination. Neural Comput. **7**, 407-422 (1995).
[43] Z. Solan and E. Rubin, "Similarity perception: a window to brain organization," J. Cognitive Neurosci. **13**, 18-30 (2001).
[44] F. d'Errico, C. Henshilwood, G. Lawson, M. Vanhaeren, M. Soressi, F. Bresson, A. M. Tillier, B. Maureille, A. Nowell, L. Backwell, J. A. Lakarra and M. Julien, "The search for the origin of symbolism, music and language: a multidisciplinary endeavour," J. World Prehistory **17**, 1-70 (2003).
[45] J. Hochberg, "Visual art and the structures of the mind," In: *The Arts, Cognition, and Basic Skills*, S. Madeja ed. CEMREL, St. Louis, MO., 151-172 (1978).
[46] J. Elkins, *Pictures and Tears*, Routledge, London, 1-19 (2001).

[47] A. R. Rao and G. L. Lohse, "Towards a texture naming system: identifying relevant dimensions in texture," IEEE Visualization **1993**, 220-227 (1993).
[48] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," Psychometrika **29**, 1-27 (1964).
[49] R. N. Shepard, "The analysis of proximities: multidimensional scaling with an unknown distance function. I.," Psychometrika **27**, 125-140 (1962).
[50] E. T. Rolls and M. J. Tovee, "Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex," J. Neurophysiol. **73**, 713-726 (1995).
[51] B. A. Olshausen and D. J. Field, "How close are we to understanding V1?" Neural Comput. **17**, 1665-1699 (2005).
[52] M. M. van Dantzig, *Pictology*, E. J. Brill, Leiden, (1973).
[53] K. Clark, *Landscape Into Art*. John Murray Publishers, London (1976).
[54] R. C. Gonzalez, R. E. Woods and S. L. Eddins, *Digital Image Processing Using Matlab*, Prentice-Hall, Upper Saddle River, N.J. (2004).
[55] R. Pickvance, *Van Gogh in Arles*, Harry Abrams, New York (1984).
[56] J. Hulsker, *New Complete Van Gogh: Paintings, Drawings, Sketches*, John Benjamins, Amsterdam (1996).
[57] P. Shirley, M. Ashikhmin, M. Gleicher, S. Marschner, E. Reinhard, K. Sung, W. Thompson and P. Willemsen, *Fundamentals of Computer Graphics*, A. K. Peters, Natick, MA (2005).
[58] M. Colbert, E. Reinhard and C. E. Hughes, "Painting in high dynamic range," J. Visual Comm. and Image Representation **18**, 387-396 (2007).
[59] J. D. Friedenberg, *Dynamical Psychology*, ISCE Publishing, Boston (2009).
[60] M. S. Landy, "A gloss on surface properties," Nature **447**, 158-159 (2007).