PHOG-Derived Aesthetic Measures Applied to Color Photographs of Artworks, Natural Scenes and Objects

Christoph Redies[2], Seyed Ali Amirshahi[1,2],
Michael Koch[1,2], and Joachim Denzler[1]

[1] Computer Vision Group, Friedrich Schiller University Jena, Germany
{seyed-ali.amirshahi,koch.michael,joachim.denzler}@uni-jena.de
http://www.inf-cv.uni-jena.de
[2] Institute of Anatomy I, Friedrich Schiller University,
Jena University Hospital, Germany
redies@mti.uni-jena.de
http://www.anatomie1.uniklinikum-jena.de

Abstract. Previous research in computational aesthetics has led to the identification of multiple image features that, in combination, can be related to the aesthetic quality of images, such as photographs. Moreover, it has been shown that aesthetic artworks possess specific higher-order statistical properties, such as a scale-invariant Fourier spectrum, that can be linked to coding mechanisms in the human visual system. In the present work, we derive novel measures based on a PHOG representation of images for image properties that have been studied in the context of the aesthetic assessment of images previously. We demonstrate that a large dataset of colored aesthetic paintings of Western provenance is characterized by a specific combination of the PHOG-derived aesthetic measures (high self-similarity, moderate complexity and low anisotropy). In this combination, the artworks differ significantly from seven other datasets of photographs that depict various types of natural and man-made scenes, patterns and objects. To the best of our knowledge, this is the first time that these features have been derived and evaluated on a large dataset of different image categories.

Keywords: Aesthetic, art, self-similarity, complexity, anisotropy, Birkhoff-like measure, Pyramid of Histograms of Orientation Gradients (PHOG).

1 Introduction

In recent years, there has been a growing interest in studying what image features characterize aesthetic images and distinguish them from non-aesthetic ones. In computational aesthetics, computer vision techniques are used in combination with mathematical approaches to assess the aesthetic quality of images and paintings [1]. Research in this field has led to the identification of several features

that can be related to the aesthetic quality of images, mainly for the assessment of the quality of photographs [2–4]. For example, lower-order features such as contrast, colorfulness, hue, saturation, rule of thirds, symmetry, saliency etc. are analyzed [2–5]. The extracted features are used in combination to distinguish between aesthetic and non-aesthetic images. Most of these features are derived from common knowledge about factors that affect the quality of photographs and paintings. For example, Datta et al. [3] extracted 56 features from photographs to evaluate and predict their aesthetic quality. Li et al. [2] used 40 features mainly related to the characteristics of color, brightness and composition for assessing the aesthetic quality of landscape paintings. Xue et al. [5] employed aesthetic measures (color histograms, spatial edge distribution, repetition identification etc.) to differentiate between aesthetic and non-aesthetic photographs taken by amateur and professional photographers.

Obviously, cultural factors play an important role in determining aesthetic preferences in humans. However, it has also been argued that aesthetic images display universal characteristics, which can potentially elicit aesthetic perception in all humans or may reflect basic functional properties of the human visual system [6]. In the ensuing line of research, visual scientists have investigated higher-order properties of images, for example, the spatial frequency spectrum and measures of order and complexity. The aim of this research is to identify one or a few properties that characterize aesthetic images and artworks [7–12] and can be related to specific aspects of information processing in the human visual system, as previously studied in natural scene research [10, 11, 13]. Some of the measured properties can be related to low-level visual coding, for example, to the efficient (sparse) coding of sensory input [9, 10]. In this paper, we combine the statistical approach with a modern computational method (Pyramid of Histograms of Orientation Gradients - PHOG) to calculate values that have been previously linked to aesthetic perception by different psychologists, e.g. Arnheim [14], such as self-similarity [8, 11, 15], complexity [7, 12] and anisotropy [16]. We compared these features for photographs of aesthetic artworks and seven different categories of natural scenes, patterns and man-made objects. Results reveal that the aesthetic images studied by us, which represent a large and diverse subset of paintings, are characterized by defined higher-order properties. To the best of our knowledge, this is the first time that such values linked to aesthetic perception have been automatically extracted from images and statistically evaluated on a dataset of different colored image categories consisting of 2763 images.

The next sections of the article are organized as follows: Section 2 provides an overview on previous work. Section 3 gives an introduction to the computational approach used by us, in particular to the PHOG algorithm. Section 4 gives details of the databases analyzed. Experimental results are described in Section 5. Section 6 provides a short conclusion and proposals for future work.

2   Previous Work

Using Fourier analysis, Graham and Field [11] and Redies et al. [10, 13] showed that images of graphic artworks of Eastern and Western provenance share

statistical properties with images of complex natural scenes. Both types of images possess a scale-invariant Fourier spectrum, that is, the spatial frequency profile remains relatively constant when zooming in and out of the images. This result implies that the two image categories have spatial frequency profiles that are self-similar (fractal-like) at different levels of spatial resolution. Other image categories, such as photographs of faces, plants, and simple objects, do not possess this property [13]. Koch et al. [16] demonstrated that, in addition, artists seem to imitate natural scenes by preferring a similar degree of anisotropy of orientations in their creations. Cardinal (horizontal and vertical) orientations are more prominent in natural scenes as well as in aesthetic artworks than oblique orientations [16]. In conclusion, artworks and natural scenes share statistical image properties. It has therefore been proposed that the perception of the two types of images may be mediated by similar coding mechanisms in the human visual system, such as an efficient (sparse) sensory code [9, 13, 17].

The above findings and other results [8, 15] suggest that aesthetic images are self-similar at different levels of spatial resolution. To follow this hypothesis with a modern computational method, Amirshahi et al. [18] assessed self-similarity in large datasets of grey-scale artworks of Western provenance and other categories of images using PHOG features [19]. The PHOG descriptor represents an image by its local shape and spatial layout. PHOG was originally developed for object recognition and image categorization [19]. Amirshahi et al. calculated the self-similarity of Histograms of Oriented Gradients (HOGs) between different levels of the pyramid [18]. Their results confirmed that the datasets of aesthetic artworks resemble natural scenes in that they are more highly self-similar at different levels of spatial resolution than, for example, photographs of simple objects and faces.

3   Aesthetic Measures Based on a PHOG Representation of Images

In the present work, we will use PHOG to derive several novel measures for image properties that have been related previously to the aesthetic quality of images, in addition to self-similarity [18]. We will then validate these measures on datasets of color photographs of aesthetic artworks and other visual scenes and patterns of lesser or no aesthetic value.

To calculate PHOG [19], a pyramid approach is taken. In this method, the HOG [20] feature for the global image (level 0) is calculated in the first step. The image is then divided into 4 equally sized sub-images and the HOG feature is calculated for each sub-image (level 1). Each sub-image is then divided into 4 other equally sized sub-images resulting in 16 sub-images in total. HOG is again calculated for these sub-images as well (level 2). The division and calculation procedure can be continued as long as desired and the image permits. 8 bins are used for binning the orientation in the HOG feature. The normalized values of the bin represent the orientation strength in each direction.

For calculating PHOG for color images, each image is converted to the Lab color space. HOG values are then calculated based on the maximum gradient

magnitudes in the L, a and b color channels. For this, we will first generate a new gradient image $\mathbf{G}_{\max}$,

$$\mathbf{G}_{\max}(x, y) = \max(\|\nabla I_L(x, y)\|, \|\nabla I_a(x, y)\|, \|\nabla I_b(x, y)\|) \tag{1}$$

$\nabla I_L(x, y), \nabla I_a(x, y)$, and $\nabla I_b(x, y)$ are the gradients at pixel (x, y) for the L, a, and b color channels respectively. We will then calculate the HOG feature for the new gradient image, $\mathbf{G}_{\max}$ by calculating the HOG features in each sub-image. Let us denote with $h(S)$ the HOG histogram of a sub-image S of $\mathbf{G}_{\max}$.

Based on the PHOG analysis of the gradient image, $\mathbf{G}_{\max}$, we will calculate the following measures, which are all assumed or even accepted to correlate with aesthetic appeal:

1. *Self-similarity* (M_{SeSf}). As a measure of self-similarity, we compare the HOG features of each sub-image at level 3 with the HOG features of the entire image at the ground level. Level 3 was chosen because the different image categories showed more diverse results than lower levels; above level 3, self-similarity values are not robust and reliable [18]. To measure the similarity between the HOG features of sub-images at different levels, we use the Histogram Intersection Kernel [21],

$$\text{HIK}(\mathbf{h}, \mathbf{h}') = \sum_{i=1}^{m} \min(h(i), h'(i)) \tag{2}$$

In Eq. (2), $\mathbf{h}$ and $\mathbf{h}'$ are two different normalized histograms and m is the number of bins present in the HOG features. To calculate the self-similarity of an image, the median value of the HIK values at each level is calculated,

$$M_{\text{SeSf}}(I, L) = \text{median}(\text{HIK}(h(S), (h(\text{Pr}(S)))) \mid \text{Pr}(S) \in \text{Sections}(I, L)) \tag{3}$$

to be robust with respect to outliers (see also [18]). In Eq. (3), M_{SeSf} is the self-similarity value, I corresponds to the image, L represents the level, at which we are assessing the HOG features (in our work we use $L = 3$), $h(S)$ is the HOG value for a sub-image in the Sections(I, L) which corresponds to the sections in the image I in level L and $\text{Pr}(S)$ corresponds to the parent of sub-image S.

2. *Complexity* (M_{Co}). The aesthetic appeal of images may depend on their degree of complexity [7, 12, 15, 22]. For example, the "savanna hypothesis" [23] states that images of moderate complexity, similar to those of the natural habitat of our ancestors, have a higher aesthetic appeal than highly complex images. To calculate the complexity of an image, we will calculate the mean norm of the gradient across all orientations over $\mathbf{G}_{\max}(x, y)$ as shown in Eq. (1).

$$M_{\text{Co}}(\mathbf{G}_{\max}) = \frac{1}{N \cdot M} \sum_{(x,y)} \mathbf{G}_{\max}(x, y) \tag{4}$$

In this equation, M_{Co} corresponds to complexity, and N and M are the height and width of the new gradient image, $\mathbf{G}_{\max}$. Since image gradients

(a)   (b)   (c)   (d)   (e)   (f)   (g)   (h)
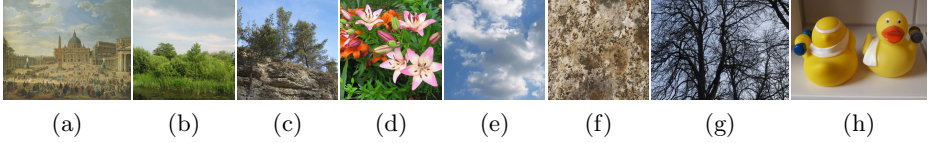
Fig. 1. Examples of images from the 8 databases used in this study. The image in (a) is reproduced with permission by the Staatliche Museen zu Berlin, Gemäldegalerie, 2012.

represent the changes of lightness in an image, we assume that calculating the mean gradient over the L channel will give us a good prediction on image complexity. The higher the mean absolute gradient, the more complex an image is.

3. *Birkhoff-like measure* (M_{BLM}). Birkhoff [12] proposed that the aesthetic appeal of objects relates to the ratio of order and complexity in images. Different attempts have been made to model this measure [7, 24]. In the present work, we propose that self-similarity may be a suitable indicator of order in an image. We calculate a Birkhoff-like measure (M_{BLM}) according to Eq. (5) where M_{SeSf} represents the self-similarity in the image calculated in Eq. (3) and M_{Co} represents the complexity introduced in Eq. (4).

$$M_{\mathrm{BLM}} = \frac{M_{\mathrm{SeSf}}}{M_{\mathrm{Co}}} \tag{5}$$

4. *Anisotropy* (M_{AnI}). Koch et al. [16] found that the Fourier spectrum is more uniform across orientations (that is, less anisotropic) in gray-scale artworks than in natural scenes [16]. This prompted us to calculate the variance of gradient strength in the HOGs across its bin entries,

$$M_{\mathrm{AnI}}(L) = \sigma(H(L)) \tag{6}$$

as a measure of anisotropy. In this equation, M_{AnI} represents the anisotropy in the image at level L, $H(L)$ corresponds to a vector which is consisted of all the HOG value at level L, and σ is the variance. Consistent with the calculation of self-similarity, anisotropy is calculated at level 3.

4   Image Databases

The previous study by Amirshahi et al. [18] was restricted to gray-scale images of graphic art. However, in many categories of visual artworks, especially in oil paintings and watercolors, color plays a crucial factor for their aesthetic appeal. In the present work, we therefore chose to apply the PHOG analysis to colored artworks and compared them with color photographs of natural scenery and patterns formed by plants. Because artworks have been shown to be self-similar (see above), we compared them also with images of highly self-similar natural

scenes and patterns (photographs of clouds, water turbulences, growth patterns of lichen, and branch patterns of trees). Furthermore, we asked in how far these categories of images differ from images of simple man-made objects.

Except for the database of aesthetic color paintings, all the other 7 databases comprise photographs that were taken in RAW format with a 15.1 megapixel digital camera (EOS 500D with EF-S15-85mm f/3.5-5.6 IS USM lens; Canon, Tokyo, Japan) by one of the authors. A small number of the photographs of large vistas, vegetation and plant patterns were taken also by a 4-megapixel Digital Ixus 400 digital camera (Canon). A total of 2,763 color photographs of artworks, natural patterns and man-made objects were analyzed. Sample images from each database are shown in Fig. 1. The images were scaled such that one side of the image was 1024 pixels. Because the aspect ratio of an image plays an important role in the aesthetic evaluation [3], we used isotropic scaling. The databases will be released for public use in due course.

4.1 Aesthetic Paintings

This database consisted of 854 different colored paintings of Western provenance that represented a large variety of styles (Renaissance, Baroque, Romanticism, Realism, Impressionisms, Modern Art, etc.), about 200 different painters, and examples from the 15th to 21st centuries (Fig. 1a). We assume that the paintings were aesthetic because they were from famous artists and prestigious museums. All paintings were scanned from high-quality art books using a digital scanner (Perfection 3200 Photo, Epson).

4.2 Natural Scenes, Vegetation and Plants

This category comprised three different databases of photographs (Fig. 1b-d): (1) 289 photographs of large-vista natural scenes, including horizon. (2) 289 photographs of vegetation (bushes, trees, etc.) taken from a distance of about 5-50m. (3) 316 close-up photographs of one type of plant. These photographs were taken at a distance of about 0.5-5 m; each photograph displayed a uniform pattern of leaves or blossoms (plant patterns).

4.3 Highly Self-similar, Natural Patterns, and Man-Made Objects

This category of photographs was analyzed to study whether aesthetic paintings differ from highly self-similar patterns found in nature or man-made objects. It comprised four databases (Fig. 1e-h): (1) 268 photographs of clouds. (2) 245 photographs of various species of crustose lichen growing on (thomb)stones. (3) 301 photographs of branching patterns of trees and bushes, taken during wintertime. (4) 201 photographs of simple household and laboratory objects.

5 Experimental Results

In the present work, we calculate the introduced PHOG-derived measures for image properties introduced in Sec. 1. We ask whether these features assume

Table 1. Mean values for self-similarity, complexity, Birkhoff-like measure and anisotropy

Database	number of images	self-similarity ($\pm\sigma$)	complexity ($\pm\sigma$)	Birkhoff-like measure ($\times10^{-3} \pm \sigma$)	anisotropy ($\times10^{-6} \pm \sigma$)
artworks	854	0.77 ± 0.08	8.5 ± 3.8	106 ± 42	126 ± 40
large vistas	289	0.60 ± 0.13^{b}	9.9 ± 4.0^{b}	71 ± 37^{b}	182 ± 60^{b}
vegetation	289	0.79 ± 0.07^{ns}	19.7 ± 4.4^{b}	42 ± 9^{b}	94 ± 30^{b}
plant patterns	316	0.83 ± 0.07^{b}	17.2 ± 6.7^{b}	57 ± 26^{b}	87 ± 45^{b}
objects	201	0.55 ± 0.07^{b}	3.5 ± 1.4^{b}	179 ± 72^{b}	289 ± 76^{b}
clouds	268	0.68 ± 0.08^{b}	1.4 ± 0.4^{b}	527 ± 157^{b}	146 ± 27^{b}
lichen	245	0.90 ± 0.03^{b}	18.4 ± 4.3^{b}	52 ± 12^{b}	42 ± 11^{b}
branches	301	0.81 ± 0.05^{b}	23.7 ± 4.7^{b}	36 ± 8^{b}	93 ± 22^{b}

[a,b] significantly different from artworks ([a] $p < 0.01$, [b] $p < 0.001$; Kruskal-Wallis one-way ANOVA, with Dunn's post test).
[ns] not significantly different from artworks.

specific values in aesthetic artworks and how they compare to the other categories of photographs. Table 1 lists mean values for the calculated PHOG measures for each image category as well as the statistical results after running the Kruskal-Wallis one-way ANOVA, with Dunn's post test, between the datasets and the artwork dataset. Results show that most of the values in different datasets are significantly different to values calculated for the artworks dataset. Fig. 2 shows 2d plots of self-similarity versus image complexity (Fig. 2a, c) and of the Birkhoff-like measure versus anisotropy (Fig. 2b, d) for all image categories. Fig. 2 illustrates the good separation of artworks (red dots) from the other image categories (other colors), although partial overlap of the different image populations is observed for some of the PHOG-derived aesthetic measures.

5.1 Self-similarity

For validation, we calculated this measure for photographs of natural growth patterns, such as lichen, plants, vegetation and branches, all of which have a highly self-similar (fractal) structure. Consistent with our claim, these image categories show high self-similarity values. The only exception are images of clouds that may be less self-similar due to the presence of areas of homogeneous sky in some of the images. In contrast, photographs of simple objects show low self-similarity in general and, correspondingly, they result in low self-similarity values. The PHOG-derived self-similarity measure yields high values for colored artworks, confirming previous results from gray-scale artworks [18].

5.2 Complexity

The complexity measure used by us is valid because it yields low values for images of low complexity (clouds and simple objects) and high values for images
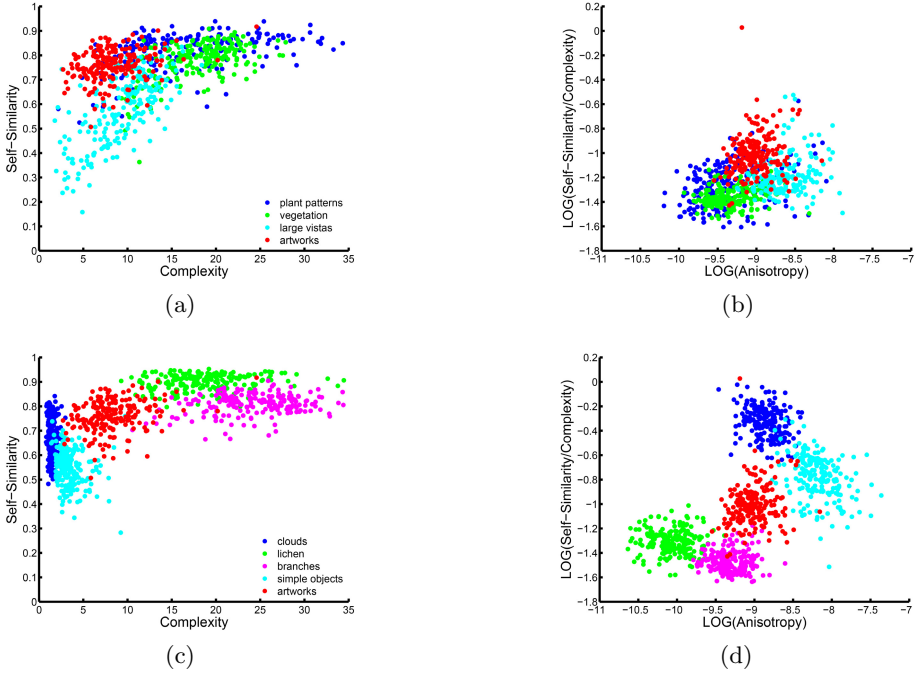
Fig. 2. Results from the PHOG analysis. Each dot represents one of 200 images randomly selected from each of the 8 image categories. Results for artworks (red dots) are compared with photographs of natural scenes and plants (a, b), and of highly self-similar natural patterns and man-made objects (c, d), as indicated by different colors. Values for self-similarity versus complexity are plotted in a, c, and for the Birkhoff-like measure versus anisotropy (log-log scale) in b, d.

of high complexity images (lichen, plant patterns and branches). On average, photographs of artworks assume intermediate values that are similar to those for photographs of large-vista natural scenes. We conclude that the images of colored artworks analyzed by us have a degree of complexity that resembles that of large-vista natural scenes.

5.3   Birkhoff-Like Measure

Intermediate values for artworks are also obtained for the Birkhoff-like measure. Values for objects and clouds are higher, whereas values for all other categories of natural and man-made objects are lower. Fig. 2b, d reveals a large degree of overlap between the image categories for this measure. Consequently, our Birkhoff-like measure does not seem to be well suited for distinguishing aesthetic artworks from most of the other image categories in a systematic way.

5.4   Anisotropy

Interestingly, artists create their works with a degree of anisotropy that is similar to that in natural patterns. Mean anisotropy in photographs of simple objects and large-vista natural scenes, in which the horizon introduces anisotropy, is larger than in artworks. As would be expected, lichen growth patterns on stone surfaces have the lowest anisotropy values because their growth patterns are fractal and do not have any direction.

In summary, the mean values for 3 PHOG-derived measures (Table 1) differ significantly between photographs of artworks and photographs of natural scenes and objects (large vistas, plants, clouds, lichen, and branches), with the exception of photographs of vegetation that have similar self-similarity values. Likewise, photographs of simple objects differ in all 3 measures from artworks.

6   Conclusion and Outlook

In conclusion, a combination of three PHOG-derived measures (self-similarity, complexity and anisotropy) characterizes aesthetic color artworks of Western provenance and distinguishes them from seven other categories of images that represent various types of natural scenes, patterns and objects. Our results support the notion that subsets of artworks possess well-defined statistical properties [9–11, 17]. Results show that the values in different datasets are statistically significant when compared to the artworks dataset. Whether the same measures allow distinguishing also aesthetic artworks from other cultures remains to be studied. Even within Western culture, extreme styles of art, for example monochrome artworks, will deviate substantially from the mean values calculated in the present study. It is not our claim that the aesthetic quality of images can be predicted by any of the measures in isolation. Rather, in combination, they define a specific subspace of image features, in which most of the aesthetic images analyzed by us are located. Whether all images in this subspace are aesthetic remains to be studied.

Moreover, it may be argued that the measures introduced by us relate to properties that reflect artistic technique rather than constraints of image composition that are followed during the creation of aesthetic artworks. However, high degrees of self-similarity are observed both in colored paintings (this study) and gray-scale graphic art [18] and these two categories of artworks comprise rather different techniques (e.g., oil paintings, woodcuts and pencil drawings). With respect to moderate complexity and natural pattern-like anisotropy, it is difficult to imagine how these features could relate to artistic technique. Alternatively, we propose that they represent principles of image composition that relate to aesthetic perception [6, 9, 15]. In future work, it will be of interest to study how these features relate to visual coding in the human brain.

References

1. Hoenig, F.: Defining computational aesthetics. EG Assoc., Goslar (2005)
2. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. IEEE J. Sel. Topics Signal Process. 3, 236–252 (2009)
3. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying Aesthetics in Photographic Images Using a Computational Approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part III. LNCS, vol. 3953, pp. 288–301. Springer, Heidelberg (2006)
4. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. In: Proceed. CVPR, pp. 419–426 (2006)
5. Xue, S.F., Lin, Q., Tretter, D., Lee, S., Pizlo, Z., Allebach, J.: Investigation of the role of aesthetics in differentiating between photographs taken by amateur and professional photographers. In: Proceed. SPIE, vol. 8302, p. 83020D (2012)
6. Zeki, S.: Art and the brain. J. Conscious Stud. 6-7, 76–96 (1999)
7. Rigau, J., Feixas, M., Sbert, M.: Informational aesthetics measures. IEEE Comput. Graph Appl. 28, 24–34 (2008)
8. Taylor, R.P.: Order in Pollack's chaos - computer analysis is helping to explain the appeal of Jackson Pollock's paintings. Sci. Am. 287, 116–121 (2002)
9. Redies, C.: A universal model of esthetic perception based on the sensory coding of natural stimuli. Spat Vis. 21, 97–117 (2007)
10. Graham, D., Redies, C.: Statistical regularities in art: Relations with visual coding and perception. Vision Res. 50, 1503–1509 (2010)
11. Graham, D.J., Field, D.J.: Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. Spat. Vis. 21, 149–164 (2007)
12. Birkhoff, G.: Aesthetic Measure. Harvard University Press, Cambridge (1933)
13. Redies, C., Hasenstein, J., Denzler, J.: Fractal-like image statistics in visual art: similarity to natural scenes. Spat. Vis. 21, 97–117 (2007)
14. Arnheim, R.: Art and Visual Perception: A Psychology of the Creative Eye. University of California Press (2004)
15. Taylor, R.P., Spehar, B., Van Donkelaar, P., Hagerhall, C.: Perceptual and physiological responses to Jackson Pollock's fractals. Front. Hum. Neurosci. 5, 60 (2011)
16. Koch, M., Denzler, J., Redies, C.: $1/f^2$ characteristics and isotropy in the Fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. PLoS one 5(8), e12268 (2010)
17. Redies, C., Hänisch, J., Blickhan, M., Denzler, J.: Artists portray human faces with the Fourier statistics of complex natural scenes. Network 18(3), 235–248 (2007)
18. Amirshahi, S.A., Koch, M., Denzler, J., Redies, C.: PHOG analysis of self-similarity in esthetic images. In: Proceed SPIE (HVEI XVII), vol. 8291, p. 82911J (2012)
19. Bosch, A., Tisserman, A., Munoz, X.: Representing shape with a spatial pyramid kernel. In: Proceed. CIVR (2007)
20. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceed. CVPR, pp. 886–893 (2005)
21. Barla, A., Franceschi, E., Odone, F., Verri, A.: Image Kernels. In: Lee, S.-W., Verri, A. (eds.) SVM 2002. LNCS, vol. 2388, pp. 83–96. Springer, Heidelberg (2002)
22. Jacobsen, T., Hofel, L.: Aesthetic judgments of novel graphic patterns: analyses of individual judgments. Percept. Mot. Skills 95, 755–766 (2002)
23. Orians, G.: An ecological and evolutionary approach to landscape aesthetics. Allen and Unwin, London (1986)
24. Boselie, F., Leeuwenberg, E.: Birkhoff revisited: beauty as a function of effect and means. Am. J. Psychol. 98(1), 1–39 (1985)