How Self-Similar are Artworks at Different Levels of Spatial Resolution?

Seyed Ali Amirshahi[1,2*] Christoph Redies[2†] Joachim Denzler[1‡]
[1]Computer Vision Group, Friedrich Schiller University Jena, Germany
[2]Experimental Aesthetics Group, Institute of Anatomy I, Jena University Hospital, Germany

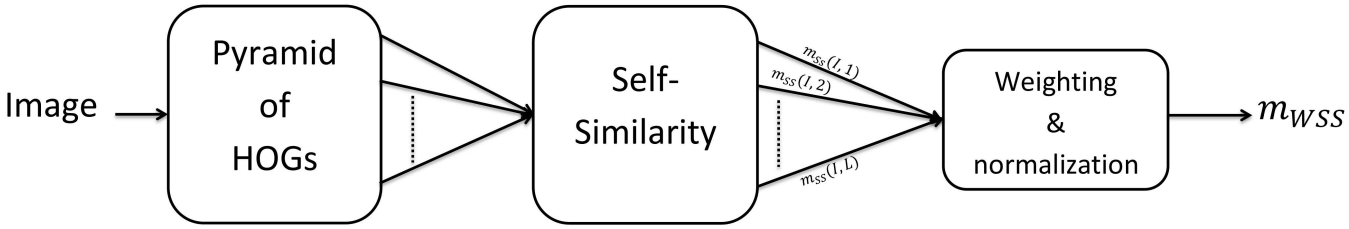


Figure 1: *Block diagram depicting the proposed Weighted Self-Similarity measure.*

Abstract

Recent research has shown that a large variety of aesthetic paintings are highly self-similar. The degree of self-similarity seen in artworks is close to that observed for complex natural scenes, to which low-level visual coding in the human visual system is adapted. In this paper, we introduce a new measure of self-similarity, which we will refer to as the Weighted Self-Similarity (WSS). Using PHOG, which is a state-of-the-art technique from computer vision, WSS is derived from a measure that has been previously linked to aesthetic paintings and represents self-similarity on a single level of spatial resolution. In contrast, WSS takes into account the similarity values at multiple levels of spatial resolution. The values are linked to each other by using a weighting factor so that the overall self-similarity of an image reflects how self-similarity changes at different spatial levels. Compared to the previously proposed metric, WSS has the advantage that it also takes into account differences between self-similarity at different levels of spatial resolution with respect to one another.

An analysis of a large image dataset of aesthetic artworks (the JenAesthetics dataset) and other categories of images reveals that artworks, on average, show a relatively high WSS. Similarly, high values for WSS were obtained for images of natural patterns that can be described as being fractal (for example, images of clouds, branches or lichen growth patterns). The analysis of the JenAesthetics dataset, which consists of paintings of Western provenance, yielded similar values of WSS for different art styles. In conclusion, self-similarity is uniformly high across different levels of spatial resolution in the artworks analyzed in the present study.

CR Categories: J.5 [Computer Applications]: Arts and Humanities—[fine arts] I.5.3 [Pattern Recognition]: Clustering—[Similarity measures] G.1.6 [Numerical Analysis]: Optimization—[Gradient methods];

Keywords: aesthetic quality assessment, visual art, paintings, Weighted Self-Similarity, Pyramid of Histograms of Orientation Gradients (PHOG)

1 Introduction

In recent years, computer vision experts and mathematicians have joined artists, philosophers, psychologists and other research groups in trying to define the basis of aesthetic perception. Attempts have been made in the field of computational aesthetics [Hoenig 2005] to find out what image features make an artwork more aesthetically pleasing than another. To this aim, aesthetic image features were extracted that allow differentiating between aesthetic and non-aesthetic images. Although good progress has been made in the case of aesthetic quality assessment of photographs [Datta et al. 2006; Li and Chen 2009; Bhattacharya et al. 2010; Xue et al. 2012] there is still ample room for progress in aesthetic quality assessment of other aesthetic visual stimuli, such as paintings.

Over the years, two different approaches have been used in the field of computational aesthetics. In the first approach [Datta et al. 2006; Li and Chen 2009; Xue et al. 2012], a number of different features that are assumed to have an effect on the aestheticness (defined here as the state of being aesthetic, or its aesthetic value) of an image are extracted. A number of different features are selected based on common knowledge in art surveys, user interactions, and features described in a variety of textbooks on photography and paintings. On the one hand, Datta et al. [Datta et al. 2006] extracted 56 different features, which are believed to be important for the aestheticness of photographs, while Li and Chen [Li and Chen 2009] used 40 features to assess the aesthetic quality of landscape paintings. This high number of features covered a wide variety of global as well as local properties. Examples for global features are average hue and saturation value as well as brightness contrast across the entire image. Examples for local features are average hue and saturation value for the three largest segments seen in the image or the coordinates and center of mass for these segments. In summary, the features used in these studies were mainly related to the role of color, brightness, and composition in an artwork. On the

*e-mail: seyed-ali.amirshahi@uni-jena.de
†e-mail: redies@mti.uni-jena.de
‡e-mail: joachim.denzler@uni-jena.de

CAe 2013, July 19 – 21, 2013, Anaheim, California.

other hand, Xue et al. [Xue et al. 2012] used features such as color histograms, repetition identification, spatial edge distribution, etc. to distinguish between photographs taken by amateurs and professional photographers (the latter having higher aesthetic quality than the former).

In the second approach, researchers try to identify universal statistical properties that are found among artworks in general. While it is believed that different factors such as the cultural background, age, gender, etc. of the observer play an important role in his decision whether or not an artwork is perceived as aesthetic by an individual, researchers have recently tried to find universal properties shared among artworks that are ranked as aesthetic, independent of the period during which they have been created, of the style they follow, the technique they use, or the subject matter they cover [Redies et al. 2007; Koch et al. 2010; Graham and Redies 2010; Amirshahi et al. 2012; Redies et al. 2012]. This type of research is in line with the notion that aesthetic artworks share specific and universal properties, which reflect functions of the human visual system in particular and of the human brain in general [Zeki 1999; Reber et al. 2004; Redies 2007]. Over the years, different research groups proposed several properties that characterize aesthetic paintings [Birkhoff 1933; Arnheim 1954; Berlyne 1974; Graham and Field 2007; Redies et al. 2007; Rigau et al. 2008; Forsythe et al. 2011]. For example, Redies et al. [Redies et al. 2007] and Graham and Field [Graham and Field 2007] have shown that, on average, log-log plots of the radially averaged 1d power spectrum of grey-scale images tend to drop according to a power law, similar to results that have been described for natural scenes [Field et al. 1987; Burton and Moorhead 1987]. This finding indicates that images of artworks and natural scenes share a scale-invariant (fractal-like) power spectrum. Other types of (less aesthetic) images created by humans, for example, images of typed or written text, do not show this property [Melmer et al. 2013]. A scale-invariant Fourier spectrum implies that, when zooming in and out of an image, the spatial-frequency profile remains constant, i.e., it is self-similar at different levels of resolution.

In view of these results, Amirshahi et al. [Amirshahi et al. 2012] assessed self-similarity in monochrome artworks and proposed a new measure to calculate self-similarity that is based on the Pyramid of Histograms of Orientation Gradients (PHOG) [Bosch et al. 2007] approach. Redies et al. [Redies et al. 2012] extended this line of research to colored paintings. Both studies revealed a relatively high degree of self-similarity in images of visual artworks as well as in natural patterns and scenes. Moreover, Redies et al. used the Histogram of Orientation Gradients (HOG) approach to calculate two other aesthetic measures, complexity and anisotropy. In support of previously published ideas, their results showed that intermediate levels of complexity characterize visual artworks on average [Berlyne 1974; Forsythe et al. 2011]. The anisotropy of orientation gradients in visual artworks is about as low as in images of diverse natural patterns and scenes (see also [Koch et al. 2010]).

Although research has been carried out with regard to the role of self-similarity (fractality) in the drip paintings by the American artist Jackson Pollock [Taylor et al. 2007] and other works of abstract expressionism [Mureika and Taylor 2012], to the best of our knowledge, no work has evaluated self-similarity at different levels of spatial resolution for artworks from different art periods and styles. In this work, we propose a new self-similarity measure, which from here on, we will refer to as the Weighted Self-Similarity (WSS). Based on the PHOG approach, WSS evaluates self-similarity at all levels of spatial resolution. The calculated values are then linked to each other using a weighting function, which is based on the coverage of the sub-regions at different spatial levels in the PHOG pyramid. Furthermore, WSS takes into account the changes in the self-similarity values between different levels of spatial resolution. Consequently, higher WSS are obtained with images that are highly self-similar at multiple levels of spatial resolution. The proposed measure is tested on 12 different datasets with a total of 5451 images. Some of the datasets used are downloaded from two public databases, a dataset of aesthetic painting [JenAesthetics ; Amirshahi et al. 2013] and a database of images previously used by Redies et al. [Jena ; Redies et al. 2012] other datasets are collected by our research group (see Table 1). Results show that the calculation of WSS models the self-similarity of images. Our results confirm the previous finding that artworks tend to have a degree of self-similarity that is about as high as that of images of natural patterns and scenes.

In the following, we will give an overview of the previous works in Section 2. Section 3 introduces the proposed WSS measure. The datasets used in the experimental results are described in Section 4. Section 5 is dedicated to the experimental results and, finally, the conclusions from the present work as well as possible future research is discussed in Section 6.

2 Previous Works

PHOG descriptors are spatial shape descriptors that were originally proposed for classifying images. The PHOG descriptor is based on calculating Histograms of Oriented Gradients (HOGs) [Dalal and Triggs 2005] over sub-regions of an image at different levels of spatial resolution resulting in a pyramid representation of the image. The bin values in each HOG vector correspond to the spatial distribution of edges in the image sub-region. The bin values for each HOG vector are then normalized so that images with strong texture are not weighted more than other images. The steps taken for calculating PHOG features are as follows:

1. HOG features are calculated for the global image (level zero).
2. The image is divided into cells at different spatial levels resulting with 2^l cells at level l (all cells at the same level have equal size).
3. HOG features are calculated for each sub-region at different levels in the image.
4. The concatenation of all HOG vectors in an image results in a pyramid representation of the HOG vectors (PHOG).

Using n bins in each HOG vector, the PHOG descriptor calculated up to level L is a vector with a dimension equal to $n\sum_{l=0}^{L} 4^l$. Amirshahi et al. [Amirshahi et al. 2012] used $L = 3$ and $n = 8$, in their calculation.

Using the HIK (Histogram Intersection Kernel) [Barla et al. 2002] function

$$\mathrm{HIK}(\mathbf{h}, \mathbf{h}') = \sum_{i=1}^{n} \min(h_i, h'_i), \tag{1}$$

Amirshahi et al. evaluated the similarity between two HOG vectors. In Equation (1), $\mathbf{h}$ and $\mathbf{h}'$ represent two sets of HOG vectors for two sub-regions in an image and h_i denotes the i^{th} bin in $\mathbf{h}$. To measure the self-similarity in image I at level L, the HIK is calculated between the HOG vector of each sub-region $S \in \text{Sub-regions}(I, L)$, represented by $\mathbf{h}(S)$ and the HOG vector of its parent region $\mathbf{h}(\Pr(S))$, $\Pr(S)$ corresponds to the parent region of sub-region S. Figure 2b represents the 16 sub-regions at level three. The colored regions in the figure correspond to the four parent regions of the 16 sub-regions. To measure the self-similarity of the image,

$$m_{\mathrm{SS}}(I, L) = \mathrm{median}(\mathrm{HIK}(\mathbf{h}(S), \mathbf{h}(\Pr(S)))) \tag{2}$$

(a) (b)

Figure 2: *A painting by Rosa Bonheur, 1849 (a), with its corresponding sub-regions and parent regions at level three (b). The colored regions correspond to the parent region of the sub-regions they cover. The painting is downloaded from the Google Art Project.*

is calculated.

Redies et al. extended the work of Amirshahi et al. and applied it to color images. Because an important step in the calculation of the HOG features is calculating the gradient image, they calculated ∇I_L,∇I_a and ∇I_b for the L, a, and b color channels, respectively. They then introduced a new gradient image

$$\mathbf{G}_{\text{new}}(x,y) = \max(\|\nabla I_L(x,y)\|, \|\nabla I_a(x,y)\|, \|\nabla I_b(x,y)\|). \quad (3)$$

The HOG based on $\mathbf{G}_{\text{new}}$ is then calculated in each sub-region. Results of both works show a high degree of self-similarity for artworks close to that of natural scenes.

While the above-mentioned methods allow differentiating between images of high self-similarity and low self-similarity, they measure self-similarity at one single level of the spatial resolution only. With respect to the description of (fractal-like) self-similarity at different levels of resolution, calculating self-similarity at a single level does not seem to be the best possible option. For example, we found in preliminary experiments that the self-similarity of large-vistas natural scene photographs have lower self-similarity values at low levels in the PHOG pyramid and higher values at high levels. The opposite tendency is seen, for example, for photographs taken of building facades that have higher self-similarity values at low levels of the PHOG pyramid and lower values at high levels. Actual data for these examples are provided in Section 5 and Figure 7a (see below).

3 Weighted Self-Similarity

In the present work, we advance and extend the method for evaluating self-similarity based on a PHOG approach. Keeping the positive properties of the previous measure and in view of its drawbacks (see above), we propose to base the novel metric for calculating self-similarity on the following properties:

Property (1). Boundedness. This property allows to evaluate how different an image is to an image with maximum or minimum possible self-similarity.

Property (2). Calculation of self-similarity at different levels of spatial resolution and incorporation of the results for all levels into the measure.

Property (3). Sensitivity towards changes of self-similarity across different levels. A highly self-similar image with equal or similar values at all calculated spatial levels will be given a self-similarity value that is higher than an image with different values.

Amirshahi et al. showed that HOG vectors up to level 4 of the PHOG pyramid could be used to differentiate between images of high and low self-similarity. Because the images used in their calculations were all resized to 1024×1024 pixels, sub-regions on level 4 will be of 64×64 pixels size. They emphasize that going to even higher levels will result in even smaller sub-regions that have a more and more uniform distribution of luminance and, consequently, fewer gradients. For this reason, we propose to calculate self-similarity for all levels of the PHOG pyramid as long as the size of the width and length of the smallest sub-region is larger than 64 pixels (**Property 2**). This will result in having a self-similarity vector,

$$\mathbf{m}_{\text{SS}}(I) = (m_{\text{SS}}(I,1), \cdots, m_{\text{SS}}(I,z), \cdots, m_{\text{SS}}(I,L)), \quad (4)$$

for image I. In Equation (4), $m_{\text{SS}}(I,z)$ represents the self-similarity value for image I at an arbitrary level z, and L corresponds to the number of the highest level in the PHOG pyramid which follows the mentioned criteria.In this approach, we will be able to use different levels for images with different sizes.

The proposed measure of WSS is calculated by

$$m_{\text{WSS}}(I) = \frac{1-\sigma(\mathbf{m}_{\text{SS}}(I))}{\sum_{l=1}^{L}\frac{1}{l}} \sum_{l=1}^{L} (\frac{1}{l} \cdot m_{\text{SS}}(I,l)). \quad (5)$$

Becasue $m_{\text{SS}}(I,l)$ is calculated based on normalized bins in the HOG vectors, self-similarity values at lower levels of the pyramid represents self-similarity at larger regions of the image and so need higher weights to be assigned to them. Accordingly, we use a weighting factor for the self-similarity values at the different levels. In Equation (5), we use a weight of $\frac{1}{l}$, with l representing different spatial levels. These weighted self-similarity values are then added up and divided by $\sum_{l=1}^{L}\frac{1}{l}$ so that the WSS is normalized to one (**Property 1**). In Equation (5), $1-\sigma(\mathbf{m}_{\text{SS}}(I))$ is used as a measure of changes among the values in $\mathbf{m}_{\text{SS}}(I)$, $\sigma(\mathbf{m}_{\text{SS}}(I))$ representing the standard deviation among the values in the $\mathbf{m}_{\text{SS}}(I)$ vector (**Property 3**). This factor will result in a higher WSS for an image with small changes among the levels. To visually compare the $m_{\text{SS}}(I,3)$ and $m_{\text{WSS}}(I)$ values, Figures 3, 4, 5, and 6 represents rounded values (to two digits) of both measures. It should be emphasized that the proposed WSS measure is only a measure of (fractal-like) self-similarity at different spatial levels and not a measure of repetition (for example seen in the case of facades with regards to the placement of windows).

4 Image Database

To evaluate our results, two publicly available databases that cover a wide variety of images were used [JenAesthetics ; Amirshahi et al. 2013; Jena ; Redies et al. 2012] and our results were compared to previous published results [Redies et al. 2012]. In addition, other datasets of images are also used in our experiments. A summary of the datasets used in our work is provided in Table 1.

4.1 JenAesthetics Dataset

Previously, studies on statistical image properties of paintings have mostly used images of artworks scanned from high-quality art books [Amirshahi et al. 2012; Redies et al. 2012]. To analyze aesthetic paintings in the present work, we gathered the JenAesthetics dataset [JenAesthetics ; Amirshahi et al. 2013], which is a dataset of high-quality images from colored oil paintings that different museums made freely available for download from the Google

Table 1: *List of the datasets used*

Category	Dataset	Number of images	Description	Sample images
JenAesthetics	Aesthetic paintings	1621	section 4.1 [JenAesthetics ; Amirshahi et al. 2013]	Figure 3
Man-made structures and objects	Facades	175	section 4.2	Figure 4a-c
	Buildings	528	section 4.2	Figure 4d-f
	Urban scenes	225	section 4.2	Figure 4g-i
	Simple objects	207	[Redies et al. 2012]	Figure 4j-l
Highly self-similar natural patterns	Turbulent water	425	section 4.3	Figure 5a-c
	Lichen	280	[Redies et al. 2012]	Figure 5d-f
	Branches	302	[Redies et al. 2012]	Figure 5g-i
	Clouds	248	[Redies et al. 2012]	Figure 5j-l
Natural scenes	Plant patterns	331	[Redies et al. 2012]	Figure 6a-c
	Vegetation	525	[Redies et al. 2012]	Figure 6d-f
	Large vistas	584	[Redies et al. 2012]	Figure 6g-i

(a) $m_{\mathrm{WSS}}(I) = .61$, $m_{\mathrm{SS}}(I,3) = .83$. **(b)** $m_{\mathrm{WSS}}(I) = .85$, $m_{\mathrm{SS}}(I,3) = .88$. **(c)** $m_{\mathrm{WSS}}(I) = .83$, $m_{\mathrm{SS}}(I,3) = .84$. **(d)** $m_{\mathrm{WSS}}(I) = .73$, $m_{\mathrm{SS}}(I,3) = .87$. **(e)** $m_{\mathrm{WSS}}(I) = .86$, $m_{\mathrm{SS}}(I,3) = .90$. **(f)** $m_{\mathrm{WSS}}(I) = .70$, $m_{\mathrm{SS}}(I,3) = .84$.

(g) $m_{\mathrm{WSS}}(I) = .90$, $m_{\mathrm{SS}}(I,3) = .95$. **(h)** $m_{\mathrm{WSS}}(I) = .86$, $m_{\mathrm{SS}}(I,3) = .88$. **(i)** $m_{\mathrm{WSS}}(I) = .70$, $m_{\mathrm{SS}}(I,3) = .81$. **(j)** $m_{\mathrm{WSS}}(I) = .81$, $m_{\mathrm{SS}}(I,3) = .82$. **(k)** $m_{\mathrm{WSS}}(I) = .79$, $m_{\mathrm{SS}}(I,3) = .88$.

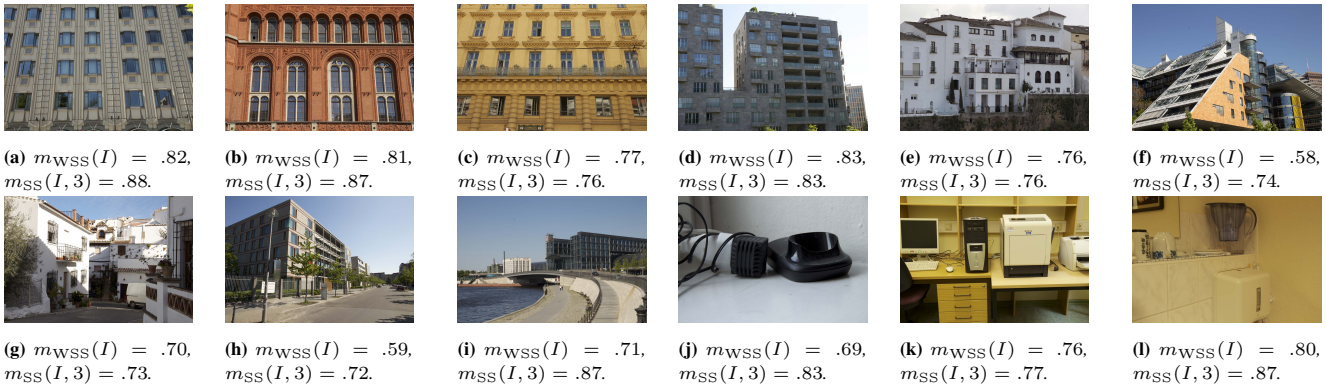
Figure 3: *Sample images from different art periods in the JenAesthetics dataset. Renaissance (a, Vincenzo Catena, 1510), Mannerism (b, El Greco, 1600), Baroque (c, Pieter de Grebber, 1623), Rococo (d, Antonio Canaletto, 1739), Classicism (e, Johann Zoffany, 1777), Romanticism (f, Hovhannes Aivazovsky , 1850), Realism (g, John Frederick Herring, 1847), Impressionism (h, Mary Cassatt, 1907), Symbolism (i, Henri Fantin-Latour, 1904), Post-Impressionism (j, Paul Cézanne, 1877), and Expressionism (k, Leo Gestel, 1914). The images were downloaded from the Google Art Project. Corresponding $m_{\mathrm{WSS}}(I)$ and $m_{\mathrm{SS}}(I,3)$ values are shown under each image (values are rounded to two digits).*

Art Project[1] through the Wikimedia Commons website[2]. We carefully selected 1621 high-quality images (most of a size greater than 3Mbytes) from 407 painters. The paintings cover 11 different art periods, such as Renaissance, Baroque, Classicism, Romanticism, Realism, Impressionism, etc. and a wide variety of subject matters. More information about the JenAesthetics dataset can be found in [JenAesthetics ; Amirshahi et al. 2013]. Figure 3 presents sample images from each art period in the dataset.

[1] www.googleartproject.com
[2] www.commons.wikimedia.org

4.2 Man-Made Structures and Simple Objects

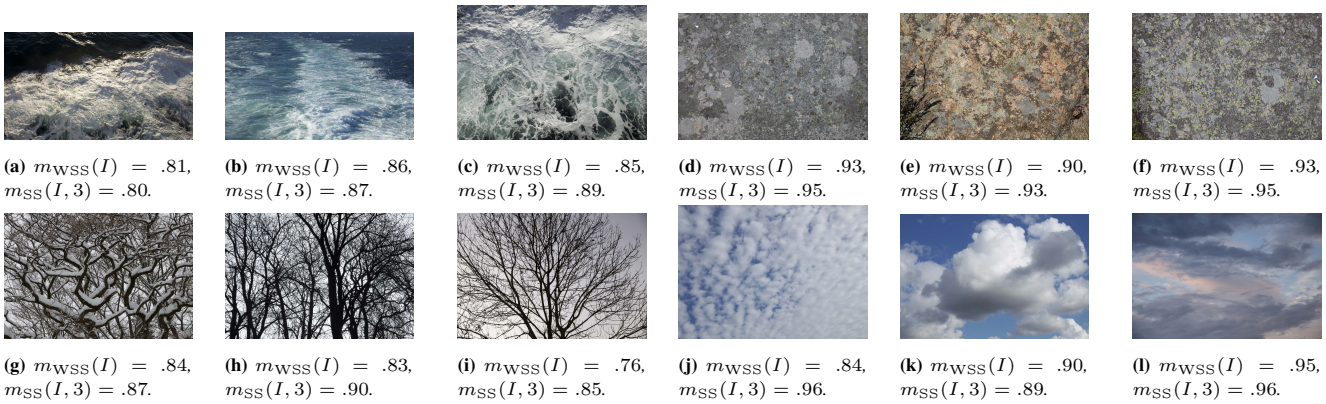
This set of images was analyzed to compare statistical regularities in aesthetic artworks with other man-made structures and objects. Similar to the database of natural scenes and objects introduced in [Redies et al. 2012], urban scenes and architecture were photographed as follows: (1) 175 photographs of building facades about 3-4 floors in height (Figure 4a-4c). (2) 528 photographs of entire buildings, mostly without the ground floors to avoid the inclusion of cars and people (Figure 4d-4f). (3) 225 photographs of urban scenes, including street views (Figure 4g-4i).



(a) $m_{\mathrm{WSS}}(I) = .82$, $m_{\mathrm{SS}}(I,3) = .88$. **(b)** $m_{\mathrm{WSS}}(I) = .81$, $m_{\mathrm{SS}}(I,3) = .87$. **(c)** $m_{\mathrm{WSS}}(I) = .77$, $m_{\mathrm{SS}}(I,3) = .76$. **(d)** $m_{\mathrm{WSS}}(I) = .83$, $m_{\mathrm{SS}}(I,3) = .83$. **(e)** $m_{\mathrm{WSS}}(I) = .76$, $m_{\mathrm{SS}}(I,3) = .76$. **(f)** $m_{\mathrm{WSS}}(I) = .58$, $m_{\mathrm{SS}}(I,3) = .74$.

(g) $m_{\mathrm{WSS}}(I) = .70$, $m_{\mathrm{SS}}(I,3) = .73$. **(h)** $m_{\mathrm{WSS}}(I) = .59$, $m_{\mathrm{SS}}(I,3) = .72$. **(i)** $m_{\mathrm{WSS}}(I) = .71$, $m_{\mathrm{SS}}(I,3) = .87$. **(j)** $m_{\mathrm{WSS}}(I) = .69$, $m_{\mathrm{SS}}(I,3) = .83$. **(k)** $m_{\mathrm{WSS}}(I) = .76$, $m_{\mathrm{SS}}(I,3) = .77$. **(l)** $m_{\mathrm{WSS}}(I) = .80$, $m_{\mathrm{SS}}(I,3) = .87$.

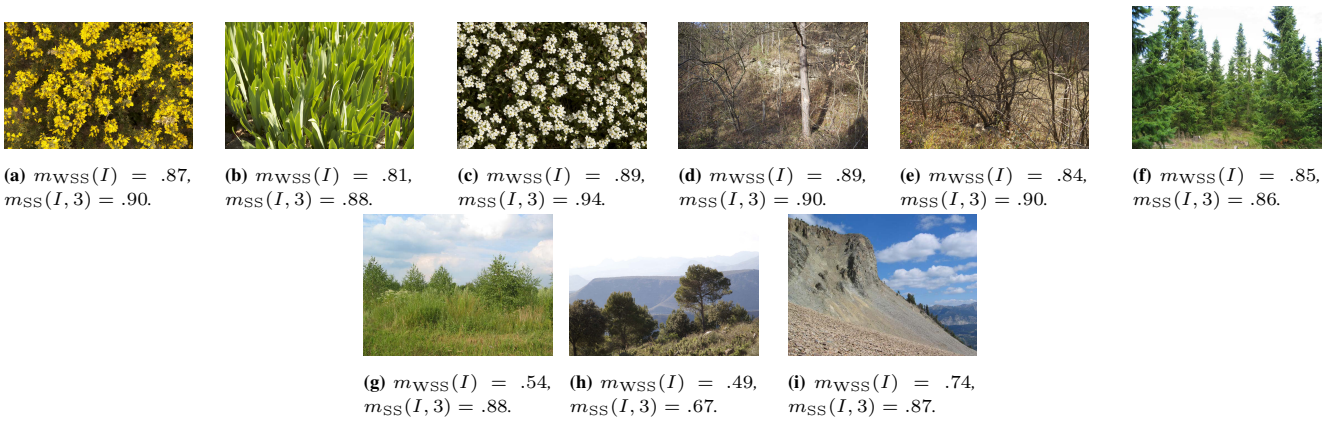
Figure 4: *Sample images from the man-made structures and objects category. Facades (a-c), buildings (d-f), urban scenes (g-i), and simple objects (j-l). Corresponding $m_{\mathrm{WSS}}(I)$ and $m_{\mathrm{SS}}(I,3)$ values are shown under each image (values are rounded to two digits).*



(a) $m_{\mathrm{WSS}}(I) = .81$, $m_{\mathrm{SS}}(I,3) = .80$. **(b)** $m_{\mathrm{WSS}}(I) = .86$, $m_{\mathrm{SS}}(I,3) = .87$. **(c)** $m_{\mathrm{WSS}}(I) = .85$, $m_{\mathrm{SS}}(I,3) = .89$. **(d)** $m_{\mathrm{WSS}}(I) = .93$, $m_{\mathrm{SS}}(I,3) = .95$. **(e)** $m_{\mathrm{WSS}}(I) = .90$, $m_{\mathrm{SS}}(I,3) = .93$. **(f)** $m_{\mathrm{WSS}}(I) = .93$, $m_{\mathrm{SS}}(I,3) = .95$.

(g) $m_{\mathrm{WSS}}(I) = .84$, $m_{\mathrm{SS}}(I,3) = .87$. **(h)** $m_{\mathrm{WSS}}(I) = .83$, $m_{\mathrm{SS}}(I,3) = .90$. **(i)** $m_{\mathrm{WSS}}(I) = .76$, $m_{\mathrm{SS}}(I,3) = .85$. **(j)** $m_{\mathrm{WSS}}(I) = .84$, $m_{\mathrm{SS}}(I,3) = .96$. **(k)** $m_{\mathrm{WSS}}(I) = .90$, $m_{\mathrm{SS}}(I,3) = .89$. **(l)** $m_{\mathrm{WSS}}(I) = .95$, $m_{\mathrm{SS}}(I,3) = .96$.

Figure 5: *Sample images from the category of highly self-similar natural patterns. Turbulent water (a-c), lichen (d-f), branches (g-i), and clouds (j-l). Corresponding $m_{\mathrm{WSS}}(I)$ and $m_{\mathrm{SS}}(I,3)$ values are shown under each image (values are rounded to two digits).*



(a) $m_{\mathrm{WSS}}(I) = .87$, $m_{\mathrm{SS}}(I,3) = .90$. **(b)** $m_{\mathrm{WSS}}(I) = .81$, $m_{\mathrm{SS}}(I,3) = .88$. **(c)** $m_{\mathrm{WSS}}(I) = .89$, $m_{\mathrm{SS}}(I,3) = .94$. **(d)** $m_{\mathrm{WSS}}(I) = .89$, $m_{\mathrm{SS}}(I,3) = .90$. **(e)** $m_{\mathrm{WSS}}(I) = .84$, $m_{\mathrm{SS}}(I,3) = .90$. **(f)** $m_{\mathrm{WSS}}(I) = .85$, $m_{\mathrm{SS}}(I,3) = .86$.

(g) $m_{\mathrm{WSS}}(I) = .54$, $m_{\mathrm{SS}}(I,3) = .88$. **(h)** $m_{\mathrm{WSS}}(I) = .49$, $m_{\mathrm{SS}}(I,3) = .67$. **(i)** $m_{\mathrm{WSS}}(I) = .74$, $m_{\mathrm{SS}}(I,3) = .87$.

Figure 6: *Sample images from the natural scenes category. Plant patterns (a-c), vegetation (d-f), and large vistas (g-i). Corresponding $m_{\mathrm{WSS}}(I)$ and $m_{\mathrm{SS}}(I,3)$ values are shown under each image (values are rounded to two digits).*

4.3 Highly Self-Similar Natural Patterns

In this category of images, apart from the images introduced by Redies et al., we introduce the turbulent water dataset. This dataset consists of 425 images of water turbulences and wave patterns taken from a ship during an ocean crossing (Figure 5a-5c). Due to the nature of turbulent water the images are highly self-similar with fractal-like characteristics.

5 Experimental Results

As a first step in our analysis, we calculate the $\mathbf{m}_{\mathrm{SS}}(I)$ vector for each image. Figure 7a represents median self-similarity values for different levels of spatial resolution for the 12 different image datasets. As can be seen in the figure, the self-similarity values for artworks are similar at all levels. Note that for some images, the $\mathbf{m}_{\mathrm{SS}}(I)$ vector has a length of seven because the size of the images

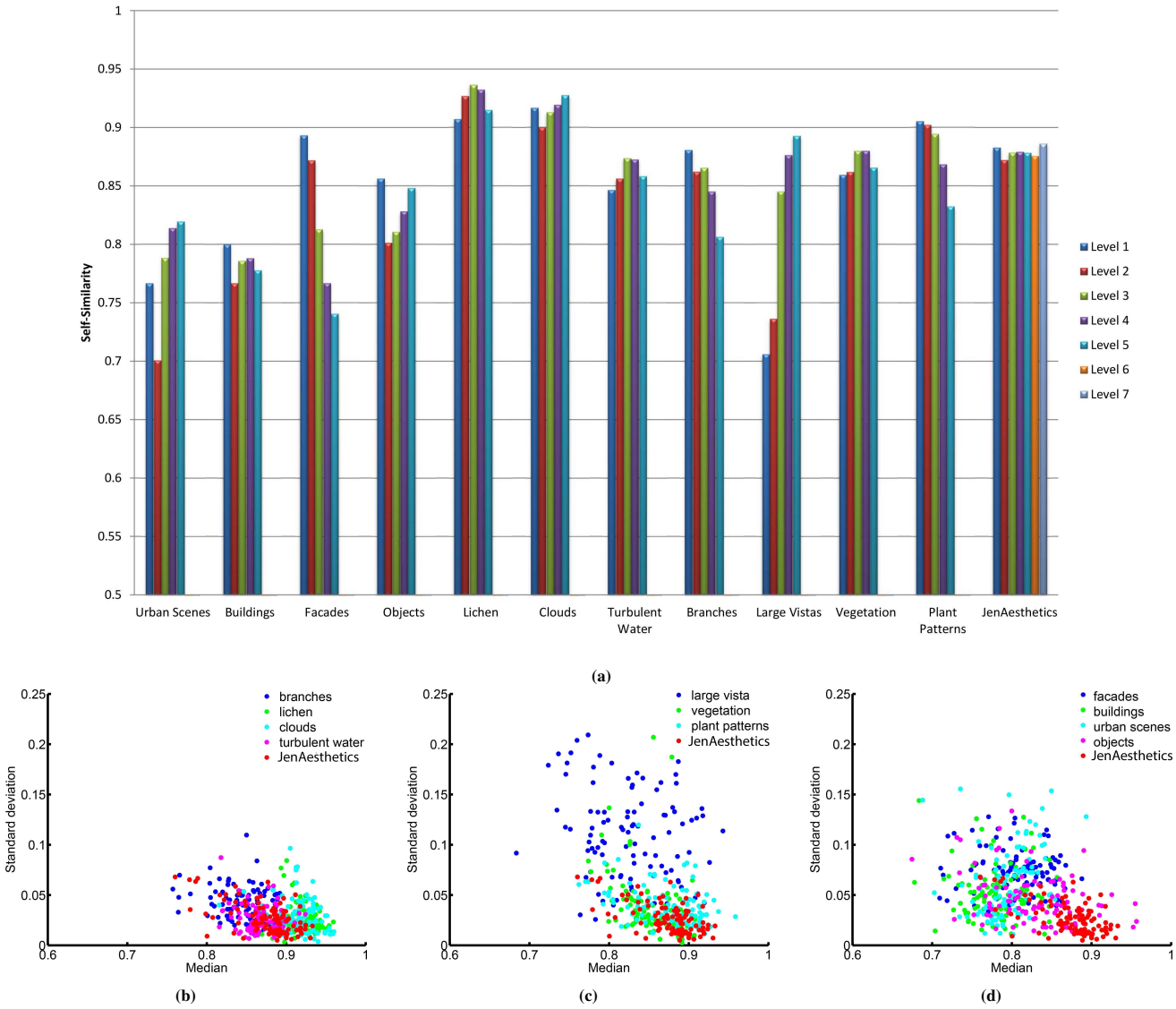


Figure 7: *a) median self-similarity values for different levels of spatial resolution for the 12 different image dataset. b-d) Median and standard deviation of $\boldsymbol{m}_{\mathrm{SS}}(I)$. Each dot represents one of 100 images randomly selected from each of the 12 image datasets. Results for the JenAesthetics dataset (artworks) (red dots) are compared with photographs of highly self-similar natural patterns (b), natural scenes and plants (c), and man-made structures and simple objects (d).*

in the JenAesthetics dataset is larger than in the other datasets, allowing to analyze self-similarity at higher levels of resolution. Using the same weight for all levels of the pyramid (averaging the values) will result in self-similarity values that will not be accurate.

To compare the results for the different image categories, Figure 7b, c, and d show scatter diagrams of the standard deviation of the self-similarity values across all levels of resolution, plotted as a function of the median value of the $\mathbf{m}_{\mathrm{SS}}(I)$ vector. Values for 100 randomly selected images are shown for each of the different datasets with each dot representing the result for one image.

The images of natural patterns in Figure 7b share high median self-similarity values and low standard deviations with artworks, as is also evident from Figure 7a. For images of lichen growth patterns, clouds and branches, this high and uniform degree of self-similarity is expected because the images represent examples of natural patterns with known fractal structure. The same applies to images of water turbulences (Figure 7b). Images of vegetation and plant patterns are also highly and uniformly self-similar across different levels of spatial resolution (Figure 7c). In contrast, there is much less overlap between results for artworks and large-vista natural scenes (Figure 7c). For large-vista natural scenes, the self-similarity value at the lower two levels is rather low while it increases as one goes up the spatial pyramid (Figure 7a). This non-uniformity across low levels is likely caused by the fact that images of large-vista scenes contain regions that differ strongly in their structure (for example, sky, rocks, trees and meadows). Finally, images of architecture and simple objects show distinct differences when compared to the artworks dataset (Figure 7d). On average, self-similarity values are lower and less uniform across levels.

To test the validity of the proposed WSS method, we calculated $m_{\mathrm{WSS}}(I)$ for each image in the 12 datasets. Figure 8a represents the calculated $m_{\mathrm{WSS}}(I)$ values for all datasets. The highest values

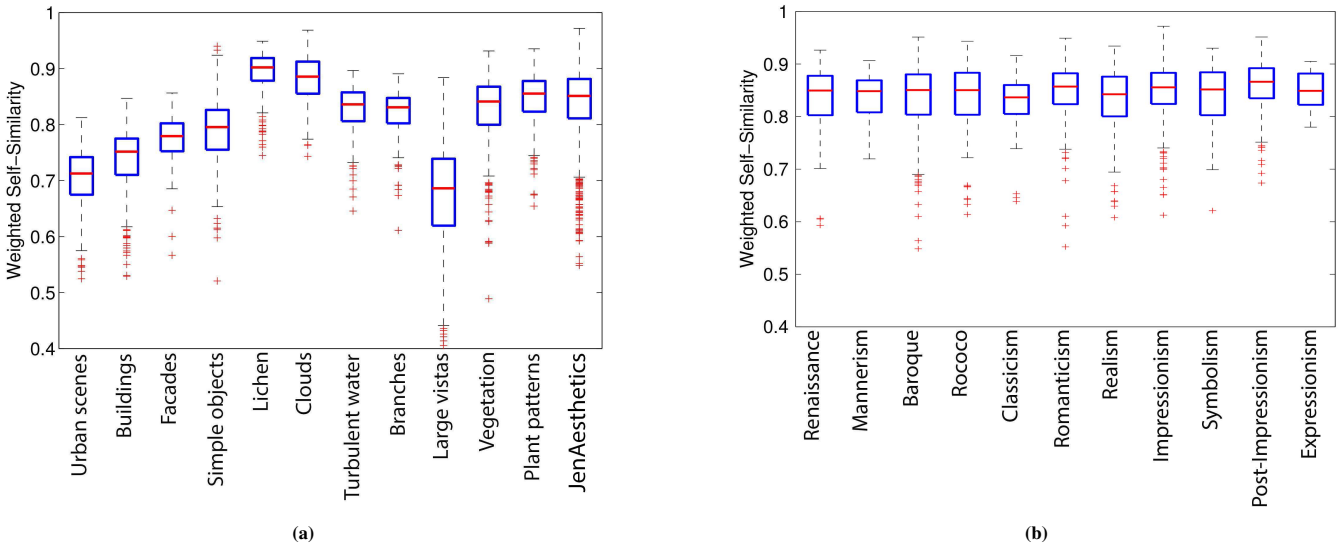


Figure 8: $m_{\mathrm{WSS}}(I)$ *values for different datasets (a), and for paintings of different art periods (b).*

are obtained for the image categories with well-established fractal characteristics (lichen growth patterns, clouds, branches and turbulent water datasets). This result provides evidence that the proposed metric indeed measures self-similarity across different levels of spatial resolution. In accordance with previous findings by Amirshahi et al. [Amirshahi et al. 2012] and Redies et al. [Redies et al. 2012], artworks tend to have a high degree of self-similarity close to those of some of the natural patterns (vegetation and plant patterns). In addition, we here show that artworks show fractal characteristics, i.e. they are highly self-similar across all levels of the spatial resolution. Together, these two characteristics result in high $m_{\mathrm{WSS}}(I)$ value.

Surprisingly, low $m_{\mathrm{WSS}}(I)$ values were obtained for the large vista dataset. With previous self-similarity measures [Amirshahi et al. 2012; Redies et al. 2012], this image category was ranked as highly self-similar. However, measurements were carried out at high levels of the PHOG pyramid (levels three or four). The present work yielded lower values in the first two levels, thereby increasing the standard deviation of the $m_{\mathrm{SS}}(I)$ values and decreasing the $m_{\mathrm{WSS}}(I)$ value. To compare the two values please refer to Figure 6g-i. Finally, low $m_{\mathrm{WSS}}(I)$ values can be seen for the images of architecture and simple objects. All differences in $m_{\mathrm{WSS}}(I)$ values between the different image categories were statistically significant (Kruskal-Wallis one-way ANOVA, with Dunns comparison post test, $p < .05$), except for the following comparisons: urban scenes/large vistas, facades/simple objects, lichen/clouds, turbulent water/branches, turbulent water/vegetation, branches/vegetation, and plant patterns/artworks.

Finally, we asked whether paintings from major art periods differ in their $m_{\mathrm{WSS}}(I)$ values. Figure 8b shows that values are very similar for the different art periods on average, suggesting that the fractal structure of paintings does not reflect a painting style of a particular art period. This finding supports a previous claim by [Wallraven et al. 2009] that low-level measures are not sufficient to differentiate between different art styles. Cross-cultural similarities were observed previously for the scale-invariant Fourier spectral properties of graphic artwork of Western provenance and East Asian paintings [Redies et al. 2007; Graham and Field 2007; Melmer et al. 2013].

6 Conclusion and Future Work

In conclusion, a novel self-similarity measure denoted as the Weighted Self-Similarity (WSS) is introduced in this paper. WSS takes advantage of a previously introduced self-similarity measure [Redies et al. 2012], which is based on a PHOG [Bosch et al. 2007] representation of an image. WSS reflects the level and uniformity of self-similarity at different levels of spatial resolution.

The proposed measure was tested on the JenAesthetics dataset [JenAesthetics ; Amirshahi et al. 2013], a dataset of aesthetic artworks (colored oil paintings), and eleven other datasets that cover a wide variety of images and subject matters (self-similar natural patterns, natural scenes, and man-made structures and simple objects). Results show that aesthetic artworks exhibit high self-similarity values close to those of some highly self-similar natural patterns. Moreover, in contrast to large-vista natural scenes, self-similarity is uniformly high across different levels of spatial resolution for artworks. The finding that different Western art periods share a similar degree of WSS supports the notion [Redies et al. 2007; Graham and Redies 2010; Amirshahi et al. 2012; Redies et al. 2012] that high self-similarity may be an indicator of a universal feature among artworks from different art periods and cultures.

In future work, we will test the notion of universality by applying the WSS measure to other aesthetic images, e.g., to artworks of different cultural background or other artistic techniques. Also, weighting factors that model the functions of the human visual system in a more physiological way would have their merits. Finally, combining the WSS measure with other features that have been related to aesthetic images could eventually help us in finding a computational paradigm that can assist the evaluation of the aesthetic quality of visual artworks by human observers.

References

AMIRSHAHI, S. A., KOCH, M., DENZLER, J., AND REDIES, C. 2012. PHOG analysis of self-similarity in aesthetic images. In *IS&T/SPIE Electronic Imaging*, International Society for Optics and Photonics, 82911J–82911J.

AMIRSHAHI, S. A., DENZLER, J., AND REDIES, C. 2013. JenAesthetics—a public dataset of paintings for aesthetic research. Tech. rep., Computer Vision Group, University of Jena Germany.

ARNHEIM, R. 1954. *Art and visual perception: A psychology of the creative eye*. University of California Press.

BARLA, A., FRANCESCHI, E., ODONE, F., AND VERRI, A. 2002. Image kernels. *Pattern Recognition with Support Vector Machines*, 617–628.

BERLYNE, D. E. 1974. *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Hemisphere Publishing Corporation.

BHATTACHARYA, S., SUKTHANKAR, R., AND SHAH, M. 2010. A framework for photo-quality assessment and enhancement based on visual aesthetics. In *Proceedings of the International Conference on Multimedia*, ACM, 271–280.

BIRKHOFF, G. D. 1933. Aesthetic measure. *Cambridge, Mass*.

BOSCH, A., ZISSERMAN, A., AND MUNOZ, X. 2007. Representing shape with a spatial pyramid kernel. In *Proceedings of the 6th ACM International Conference on Image and Video Retrieval*, ACM, 401–408.

BURTON, G., AND MOORHEAD, I. R. 1987. Color and spatial structure in natural scenes. *Applied Optics 26*, 1, 157–170.

DALAL, N., AND TRIGGS, B. 2005. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, vol. 1, IEEE, 886–893.

DATTA, R., JOSHI, D., LI, J., AND WANG, J. 2006. Studying aesthetics in photographic images using a computational approach. *Computer Vision–ECCV 2006*, 288–301.

FIELD, D. J., ET AL. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America 4*, 12, 2379–2394.

FORSYTHE, A., NADAL, M., SHEEHY, N., CELA-CONDE, C. J., AND SAWEY, M. 2011. Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology 102*, 1, 49–70.

GRAHAM, D. J., AND FIELD, D. J. 2007. Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision 21*, 1-2, 149–164.

GRAHAM, D., AND REDIES, C. 2010. Statistical regularities in art: Relations with visual coding and perception. *Vision Research 50*, 16, 1503–1509.

HOENIG, F. 2005. Defining computational aesthetics. In *Computational Aesthetics in Graphics, Visualization and Imaging*, The Eurographics Association, 13–18.

JENA. Jena computational aesthetics group. http://www.inf-cv.uni-jena.de/en/aesthetics.

JENAESTHETICS. JenAesthetics dataset. http://www.inf-cv.uni-jena.de/en/jenaesthetics.

KOCH, M., DENZLER, J., AND REDIES, C. 2010. $1/f^2$ characteristics and isotropy in the Fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PloS One 5*, 8, e12268.

LI, C., AND CHEN, T. 2009. Aesthetic visual quality assessment of paintings. *Selected Topics in Signal Processing, IEEE Journal of 3*, 2, 236–252.

MELMER, T., AMIRSHAHI, S. A., KOCH, M., DENZLER, J., AND REDIES, C. 2013. From regular text to artistic writing and artworks: Fourier statistics of images with low and high aesthetic appeal. *Frontiers in Human Neuroscience 7*, 106.

MUREIKA, J., AND TAYLOR, R. 2012. The abstract expressionists and les automatistes: A shared multi-fractal depth? *Signal Processing 93*, 3, 573–578.

REBER, R., SCHWARZ, N., AND WINKIELMAN, P. 2004. Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality and Social Psychology Review 8*, 4, 364–382.

REDIES, C., HASENSTEIN, J., AND DENZLER, J. 2007. Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision 21*, 1-2, 137–148.

REDIES, C., AMIRSHAHI, S. A., KOCH, M., AND DENZLER, J. 2012. PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In *Computer Vision–ECCV 2012. Workshops and Demonstrations*, Springer, 522–531.

REDIES, C. 2007. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision 21*, 1-2, 97–117.

RIGAU, J., FEIXAS, M., AND SBERT, M. 2008. Informational aesthetics measures. *Computer Graphics and Applications, IEEE 28*, 2, 24–34.

TAYLOR, R., GUZMAN, R., MARTIN, T., HALL, G., MICOLICH, A., JONAS, D., SCANNELL, B., FAIRBANKS, M., AND MARLOW, C. 2007. Authenticating Pollock paintings using fractal geometry. *Pattern Recognition Letters 28*, 6, 695–702.

WALLRAVEN, C., FLEMING, R., CUNNINGHAM, D., RIGAU, J., FEIXAS, M., AND SBERT, M. 2009. Categorizing art: Comparing humans and computers. *Computers & Graphics 33*, 4, 484–495.

XUE, S., LIN, Q., TRETTER, D., LEE, S., PIZLO, Z., AND ALLEBACH, J. 2012. Investigation of the role of aesthetics in differentiating between photographs taken by amateur and professional photographers. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 8302, 7.

ZEKI, S. 1999. Art and the brain. *Journal of Consciousness Studies 6*, 6-7, 76–96.