ARTICLE IN PRESS

Knowledge-Based Systems xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Assessment model for perceived visual complexity of painting images

Xiaoying Guo[a,b], Yuhua Qian[*,b], Liang Li[c], Akira Asano[d]

[a] School of Software Engineering, Shanxi University, Taiyuan, China
[b] Institute of Big Data Science and Industry, Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, China
[c] College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Japan
[d] Faculty of Informatics, Kansai University, Osaka, Japan

ARTICLE INFO

Keywords:
Visual perception
Visual complexity
Image emotion knowledge
Complexity and aesthetic

MSC:
00-01
99-00

ABSTRACT

In the construction of the knowledge system, visual perception is the primary means of acquiring knowledge. Thus, it is very essential to solve the problems related to visual perception. Visual complexity, as a basic aspect of visual perception, is extremely important for human being to understand and perceive the visual stimuli. This leads to an interesting question: what factors affect visual complexity of images and how to evaluate the visual complexity objectively. In order to address this issue, we take digital painting images as the visual stimuli. We firstly conduct an experiment to collect the subjective complexity labels of painting images and then identify the factors that affect visual complexity perception. Three main factors that affect human visual complexity perception are identified, namely, distribution of compositions, colors, and contents. Secondly, we study theoretical and empirical concepts from psychology and art theory to design 29 global, local, and salient region features which represent the above three factors. Moreover, we provide two ways to estimate the visual complexity of painting images. One is to evaluate the visual complexity level of painting images by classifying the complexity level into three levels (low, middle and high complexity). Another one is to predict a complexity value for painting images by a regression model. The experimental results indicate that the proposed classification method (by Random Forest classifiers) can predict the visual complexity perception of paintings with an accuracy of 86.78%. By the comparisons, the proposed method outperforms other measurements of image complexity with a higher correlation coefficient between subjective complexity and objective measures of complexity. Furthermore, we apply the regression model of visual complexity to predict the other features of painting images. The results show that the regression model has a good ability of measuring aesthetic quality, beauty, and liking of color of the painting images involved in JenAesthetics dataset.

1. Introduction

For humans, 80% of the information and knowledge are acquired from human vision system. This makes the importance of visual perception is for more than other perceptions in the construction of the knowledge system. Moreover, with the development of Artificial Intelligence, it is increasingly necessary to endow a computer the ability of visual perception like a person. Visual complexity, as a basic aspect of visual perception, is extremely important for human being to understand and perceive the visual stimuli. Therefore, how to evaluate visual complexity objectively becomes a timely topic in the fields of psychology and computer science.

Visual complexity is regarded as a primary cue on judgments of visual appeal [1]. Nowadays, people can easily enjoy the paintings on the Internet without going to the museums [2]. If they select images only by visual feeling (e.g., visual appeal or pleasure) instead of specific keywords (e.g., rose), then visual complexity plays several central roles in composing the said feeling [3]. Hence, presenting an objective measure of complexity which is similar to human perception is exceedingly useful. In practice, objectively measuring visual complexity has a wide range of applications. From a psychological sense, measures of visual complexity are helpful for human viewers to analyze the effects of visual complexity on aesthetic judgments, and thus are useful for neuroscientists and psychologists who are interested in the mechanism of object perception and the process of learning and memory. From an applied sense, measures of visual complexity can be utilized by computer engineers to construct information systems and tools for the analysis, estimation, visualization and recognition of images, and could allow designers to anticipate consumers' and users' aesthetic and affective responses to the complexity of the products from wallpapers to

* Corresponding author.
E-mail addresses: guoxiaoying@sxu.edu.cn (X. Guo), jinchengqyh@sxu.edu.cn (Y. Qian).

https://doi.org/10.1016/j.knosys.2018.06.006
Received 15 November 2017; Received in revised form 8 June 2018; Accepted 9 June 2018
0950-7051/

Please cite this article as: Guo, X., Knowledge-Based Systems (2018), https://doi.org/10.1016/j.knosys.2018.06.006

webpages.

Previous research [3–15] has proposed a number of measures of image complexity such as information theory, pattern measure, fractal dimension, quad tree method and region of interest method. These methods can provide a computable and objective means to measure the image complexity. However, they merely consider the distribution of spatial frequencies of visual stimuli and disregard the mechanism of human visual perception.

Considering the nature of visual perception (i.e., assessing visual complexity is a constructive process of perception), we assume that visual complexity perception is significantly affected by visual features in the images, such as the features of colors, distribution of objects, and contents. Therefore, in this paper, we hope to achieve a model to estimate the visual complexity of painting images based on image features. In order to achieve this purpose, three steps are conducted: (1) Experiment of subjective complexity: labeling subjective complexity of paintings and identifying the factors that affect human visual complexity through a questionnaire survey;(2) Feature extraction: extracting a group of global, local, and salient features depending on the results of the questionnaire in step (1); and (3) Mapping stage: employing classification and regression methods to build the relationship between the visual complexity perceived by humans and the features extracted from the paintings.

In conclusion, our research identifies the main factors that affect visual complexity perception of visual stimuli, and provides computational methods to estimate visual complexity of painting images. Our research contributes to the "computational visual complexity" by:

- Identifying three main factors (distribution of compositions, colors, and contents) that affect visual complexity perception of painting images.
- Providing quantified methods to compute visual features that represent three factors and distinguishing influential visual features of perceived complexity.
- Validating the effectiveness of the proposed complexity assessment methods (visual complexity level classification and visual complexity score prediction) and applying it to predict other visual emotions (aesthetic, beauty, etc.).

The rest of the paper is organized as follows. Section 2 reviews the measures of visual complexity and its related works. Section 3 introduces the subjective complexity assessment experiment and identifies the main factors affecting visual complexity perception. Section 4 quantifies the visual features that extracted from global, local and salient regions. Section 5 builds the objective measures of complexity: complexity level classification and complexity score regression. Section 6 discusses the influential visual attributes that affect visual complexity of painting images, followed by conclusions in Section 7.

2. Measures of complexity

A variety of methods to measure complexity have been proposed in the fields of psychology and computer science.

In the field of psychology, several researchers mainly investigated the factors that affected human visual complexity perception. According to Oliva et al. [16], visual complexity was defined by the degree of difficulty in providing a verbal description of an image. In their study, 34 participants used the method of hierarchical grouping to classify indoor scenes. The results showed that visual complexity is represented by several dimensions, such as the number of objects, clutter, openness, symmetry, organization, and variety of colors. Pieters et al. [17] investigated the visual complexity of advertising. They distinguished two types of visual complexity (feature complexity and design complexity) in advertising and proposed an objective measure for each. Saleem et al. [18] studied the visual complexity of 3D shapes and introduced an approach based on view similarity to determine the perceived shape complexity. Purchase et al. [19] explored the visual complexity of images. They attempted to investigate whether visual complexity could be quantified to match a human's perception of complexity. Through an empirical study, they concluded that the subjective notion of complexity was consistent both in an individual and in a group but did not easily relate to the most obvious computational metrics.

From the view of computer science, various methods have been proposed to measure complexity. Andrienko et al. [5] developed a complexity measure based on mean information gain of spatial correlations of 2-D patterns. Rigau et al. [20] proposed a new framework to investigate the complexity of an image by considering the number of partitioned regions and the compositional complexity of partitioned images. The Jensen-Shannon divergence was employed to calculate the compositional complexity of partitioned image. Patel and Holt [21] compared the pattern measure proposed by Klinger and Salingaros [6] with respondents' perceptions of the complexity of background image scenes; the results showed that a high positive correlation exists between mathematical measures and the subjects' perceptions. Furthermore, Murguia et al. [7] proposed a novel fuzzy approach to determine the complexity of an image based on the analysis of edge level percentage. Cardaci et al. [8] presented an experiment to obtain the perceived time of paintings. The aim of this experiment was to build the relationship between the objective measure of complexity and the perceived time. The results indicated that there is a strong correlation between psychological and computational results (statistical properties of the paintings). In their another work [9], they proposed a fuzzy mathematical model of visual complexity based on fuzzy measures of entropy. Their proposed method fitted well with the perceived time of images, but neglected the image color and other perceived features. Fractal dimension (FD) has often been applied as a parameter of complexity, related to, for example, surface. The previous research [10] showed that FD accounts for more of the variance in judgments of perceived beauty in visual art than measures of visual complexity alone, particularly in abstract and natural images. Besides, Donderi [11] found a correlation between subjective estimations of visual complexity and the size of compressed digital image files. Rosenholtz et al. [12] proposed two classical methods for image visual complexity measurement: Subband Entroy (SE) and Feature Congestion (FC). Additionally, compression based methods should be the simplest method of measuring the complexity of an image. A larger file size indicates high complexity [22]. However, these methods are abstract and difficulty in explaining why some images look more complex than others. Redies et al. [23] proposed a measure of image complexity (Com) based on the maximum gradient magnitudes of each pixel in the Lab color space. The gradient represents the local changes of lightness in an image. Thus, the higher the mean absolute gradient, the more complex an image is. Sun, Yamasaki and Aizawa [24] designed 114-dimension features to evaluate the image complexity, and then they extended the proposed method to the applications of beauty predication and quality assessment.

3. Experiment: subjective assessment of complexity

Unlike research on photographs, it is very difficult to find a website of paintings with complexity ratings by a large community. Consequently, to implement the first step of the proposed approach, we conduct an experiment to acquire the subjective assessment of the complexity of painting images.

3.1. Stimuli

In the experiment, 500 painting images are utilized, including 50 painting images obtained from the dataset of PaintingDb[1] [25], 150

[1] PaintingDb is a virtual art gallery with thousands of painting images and a

painting images selected from Jenaesthetics subjective dataset [26–29],[2] and 300 painting images from Wikiart.[3] The two main styles of paintings are abstract and impressionism. The set of painting images used in our experiment includes paintings with different complexity levels created by different artists. Our dataset associated with the complexity scores are available for the public and other research group from the authors on request.

The experimental images are selected carefully to minimize some possible affected factors (e.g. a painting's fame and a respondent's preference). In the selection process, several rules are considered: 1) the painting images are not famous paintings and are unknown to the participants, 2) the portraits paintings are excluded, and 3) the painting images with clear human faces are excluded. To reduce the influence of image size, we resize all experimental images with the same height in the experiment. In the experiment, the painting images are randomly displayed on a 65-inch plasma display panel TV one by one.

3.2. Participants

In the experiment, 68 participants from Shanxi University, with ages ranging from 21 to 30 years (mean = 21.5, SD = 1.35 , 33 males and 35 females), join in the experiment. They are required to sit two meters away from the screen and asked to provide responses according to the experiment rules. Before conducting the experiment, informed consents are obtained from all participants.

3.3. Procedure

Two parts (complexity rating and questionnaire establishment) are implemented in our experiment. After a brief introduction of the experiment, Part I is implemented before Part II. In Part I, all painting images are displayed twice. In the first instance, the respondent is required to view all images one by one with no time constraint. In the second instance, the respondent is asked to assess complexity using a Likert scale of 1–7, with the degrees ranging from very simple to very complex. In Part II (the questionnaire is shown in Appendix), the respondent is asked to complete a questionnaire that contains a list of possible factors that affect complexity assessment. Then, the respondent is asked to select the factors that are important for them to assess the complexity of a painting.

3.4. Results

The subjective complexity assessments of the 400 paintings are obtained from Part I. Fig. 1 shows several painting images marked as Low Complexity (LC), Middle Complexity (MC), and High Complexity (HC) by the subjects. To construct a ground truth, the mean and standard deviation (SD) values of complexity score for each painting are calculated among all respondents. Some examples of mean and SD values of complexity scores are shown in Fig. 1. Furthermore, the answers to the questionnaire in the experiment are ranked according to the frequency of options mentioned by the subjects. The results are shown in Table 1. The top three frequently mentioned options in the experiment are "distribution of compositions," "color", and "contents." Apparently, these three factors significantly affect the respondents' assessment of complexity.

4. Feature extraction

From the results of the subjective experiment, three factors that greatly affect subjective complexity assessment (distributions, color, and contents) are identified. Synthesizing the results of the questionnaire and human visual processing mechanism, we extract a total of 29 features to represent the three factors.

These features are separated into three categories: global features, local features, and salient region features. Global features refer to the characteristics of the first impression human beings acquire when they see a painting. Local features reflect the regional information of a painting. Salient region features represent the characteristics of the most visually important region of a painting.

4.1. Global features

When viewing something, people first develop a holistic impression of the item and then go into its segments and details [30]. Therefore, global features may influence the first impression of human visual perception. Each global feature extracted in this research is explained below.

4.1.1. Global color features

When humans view the paintings, color is an immediate and distinguished characteristic in subjective visual perception. As is defined in [31], color is the characteristic of a visible object or light source by which an observer may distinguish differences between two structure-free fields of the same size and same shape. Furthermore, it is identified that color compositions influence the painting price in [32]. Therefore, color is the first and foremost characteristic of paintings among all global features.

Color complexity measure

f [4] $_1 - f_4$: Color complexity measures in four levels of image pyramid are extracted.

To measure the complexity of the colors in a painting, we employ the method of color complexity measure (CCM) [33].

CCM is defined as follows

$$\psi(i,j) = \sum_{x,y\in\Omega_{(i,j)}} G_\alpha(\|c(x,y) - \bar{c}\|), \tag{1}$$

where (x, y) is the pixel that belongs to the local window $\Omega_{(i,\ j)}$, G_α denotes the Gaussian weighting function and $\bar{c}$ is an average color value within a local window size $\Omega_{(i,\ j)}$ centered at (i, j). $\|\ \|$ denotes the color difference measure.

The average color value within a local window is calculated by

$$\bar{c} = \frac{1}{N}\sum_{x,y\in\Omega_{(i,j)}} c(x,y), \tag{2}$$

where (x, y) is the pixel that belongs to the local window $\Omega_{(i,\ j)}$ and N is the number of pixels in the local window $\Omega_{(i,\ j)}$.

The color difference measure is important to represent the human visual perception of colors. Generally, the color difference is evaluated using the Euclidean distance between two color points in a color space. In this part, we adopt the CIELab color space which describes all the colors visible to the human eyes. But in this color space, a small Euclidean distance between two color points is proportional to the difference that human visual system perceives. It is identified in [33] that a larger Euclidean distance has no meaning but only large difference in human visual system. Thus, we employ the color difference proposed in [33]. It is defined as

$$D(c(i,j), c(x,y)) = 1 - exp\left[-\frac{E(c(i,j), c(x,y))}{\gamma}\right], \tag{3}$$

where γ is the normalized factor and $E(c(i, j), c(x, y))$ is the Euclidean distance in CIELab color space.

(*footnote continued*)
streamlined user interface.

[2] A public dataset with subjective aesthetic scores, http://www.inf-cv.uni-jena.de/en/jenaesthetics

[3] http://www.wikiart.org

[4] **f denotes the feature.**

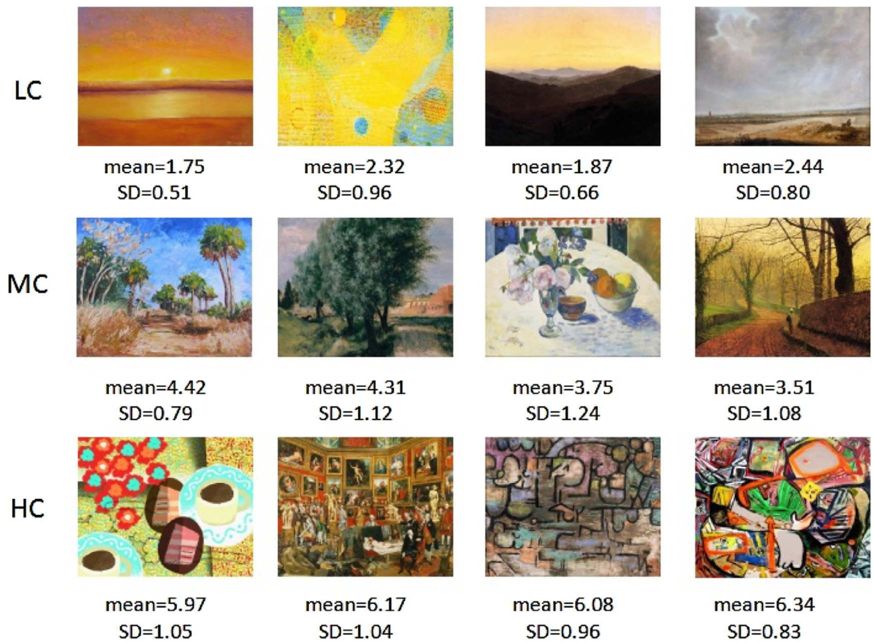

Fig. 1. Some painting images that are labeled as LC, MC, and HC.

Table 1
Top three frequently mentioned options of the questionnaire in the experiment.

Options	Frequency
Distribution of compositions (D)	58
Colors (A)	53
Contents (C)	50

$$E(c(i,j), c(x,y)) = \sqrt{(L_{ij} - L_{xy})^2 + (a_{ij} - a_{xy})^2 + (b_{ij} - b_{xy})^2}, \tag{4}$$

We employ Eq. (1) to calculate the color complexity of the local neighbor region centered at pixel of (i, j). For the entire image, we calculate the mean CCM of all pixels.

Several features hidden in this resolution are extracted in another resolution. In this study, we apply the Gaussian pyramid to the painting image and calculate the CCM ($f_1 - f_4$) based on the four pyramid levels.

$$f_i = CCM|G_j(x,y),\ i = 1, 2, 3, 4;\ j = 1, 2, 3, 4. \tag{5}$$

where i is the feature in a different level (j) of the pyramid.

Basic color features

$\mathbf{f_5} - \mathbf{f_7}$: The average hue ($f_5$), saturation ($f_6$) and brightness ($f_7$) of a painting are exacted based on HSL color space.

1. Hue
   Hue is the most obvious characteristic of a color [34]. In artistic sense, the average of hue and saturation more or less reflect the colorful keynote of a painting relative to the viewer's perception [2]. The average of hue is calculated by

$$f_5 = \frac{1}{MN}\sum_n \sum_m I_H(m,n), \tag{6}$$

   where M and N are the number of rows and columns of the image, and $I_H(m, n)$ is the hue value at the pixel (m, n).
2. Saturation
   The saturation of a painting somehow affects a human's feeling evoked by the painting. The average of saturation is calculated by

$$f_6 = \frac{1}{MN}\sum_n \sum_m I_s(m,n), \tag{7}$$

   where M and N are the number of rows and columns of the image, and $I_H(m, n)$ is the saturation value at the pixel (m, n).
3. Brightness
   Brightness is a measure of how light or dark a color is in the entire painting, and it reflects the tone of a painting. The average of brightness of a painting can be calculated as:

$$f_7 = \frac{1}{MN}\sum_n \sum_m L(m,n), \tag{8}$$

   where M and N are the number of rows and columns of the image, and $L(m,n) = (I_R(m,n) + I_G(m,n) + I_B(m,n))/3$. Here, I_R, I_G, and I_B respectively stand for red, green, and blue channels of the image.

4.1.2. Content features

Contents include points of interest and edges.

Points of interest

f_8: Points of interest are a reflection of the contents of an image. When many points of interest exist in an image, the complexity of the image is perceived to be high [9]. The feature of points of interest is extracted in this part.

Symmetry plays a relevant role in perception problems [35]. It is an interesting property in detecting the points of interest. Discrete symmetry transform (DST) is an algorithm to measure the local symmetry. It searches the area of interest in active vision. In this study, we employ DST method to extract the points of interest in the painting images. It is defined as

$$DST_{i,j} = (1 - std_k(T_{i,j}^k))E_{i,j}, \tag{9}$$

where std stands for the standard deviation for $k = 0, \ldots, n-1$ of the first-order moment relative to an axis with orientation $\alpha_k = k\pi/n$. Here, we assign $n = 4$.

$$T_{i,j}^k = \sum_{(x,y)\in C_r} \left| (x-i)sin\alpha_k - (y-j)cos\alpha_k \right| g_{x,y}, \tag{10}$$

where C_r is the disk center at (i, j) of radius r, α is the axis orientation, and $g_{x,y}$ is the value of pixel (x, y) located in C_r. $E_{i,j}$ measures the local smoothness of the image.

$$E_{i,j} = \sum_{(x,y)\in\delta C_r,(s,t)\in\delta C_{r+1}} \left| g_{x,y} - g_{s,t} \right|, |x-s| + |y-t| = 1, \tag{11}$$

where δC_r stands for the circular edge of C_r. If the image is locally flat, $E_{i,j} = 0$.

The points of interest (P) are extracted by

$$P_{ij} = \begin{cases} DST_{ij} & DSTd_{ij} \geq \mu + 2\sigma \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where μ and σ are respectively the mean and standard deviation of DST values calculated from the entire image.

Through the above DST extraction method, we extract the points of interest in the painting images.

$$f_8 = P \tag{13}$$

Edge density

f_9: The edge density of a painting can be determined by the ratio between the pixel number of the extracted edges and that of the entire image.

An image with lots of edges contains many objects, resulting in high perceived complexity [9]. Edge density can be determined as follows

$$f_9 = N_{edges}/N_{img}, \tag{14}$$

where f_9 is the edge density of an image, N_{edges} is the pixel number of the extracted edges, and N_{img} is the pixel number of the entire image. In our research, the Canny algorithm is used to extract the edges in this work. The high and low thresholds are automatically calculated based on the gradient histogram of the image.

4.2. Local features

Local features represent the detailed information of paintings. To extract the local features in the painting, we first segment the image into several parts and then analyze the characteristics in segments.

4.2.1. Image segmentation

In this study, an initial segment is required to partition the image into small regions for merging. We select the method of mean shift for the initial segment because it creates less over segmentation when compared with the method of watershed. We use a free software, EDISON System [36], to obtain the initial segmentation map. After the initial segmentation, an image is subdivided into many small regions.

In human visual perception, several regions with similar color or spatial adjacency should be merged into one region. These regions need to be represented by several feature descriptors and a rule for region merging must be defined. We employ the method of color histogram similarity [37] to calculate the similarity between two adjacent regions. Each color histogram is quantized into 16 levels and then into a total of 4096 bins in each region. The color histogram similarity ρ is calculated between two adjacent regions (e.g. regions P and Q) using Bhattacharyya coefficient. Bhattacharyya coefficient is an approximate measurement of the amount of overlap between two statistical samples. The coefficient can be utilized to determine the relative closeness of the two samples. It is defined as:

$$\rho(P, Q) = \sum_{u=1}^{4096} \sqrt{Hist_P^u * Hist_Q^u}, \tag{15}$$

where $Hist_P$ and $Hist_Q$ are the normalized histograms of adjacent regions P and Q, and u denotes their *uth* bin of them. A high Bhattacharyya coefficient between P and Q implies a close similarity between the two adjacent regions. We use the region adjacency graph to store the similarity of the pair of regions.

After calculating the similarity between two neighbor regions, we use the Region Adjacency Graph (RAG) [38] to store the similarity of the pair of regions. In this work, the merging rule is defined as follows:

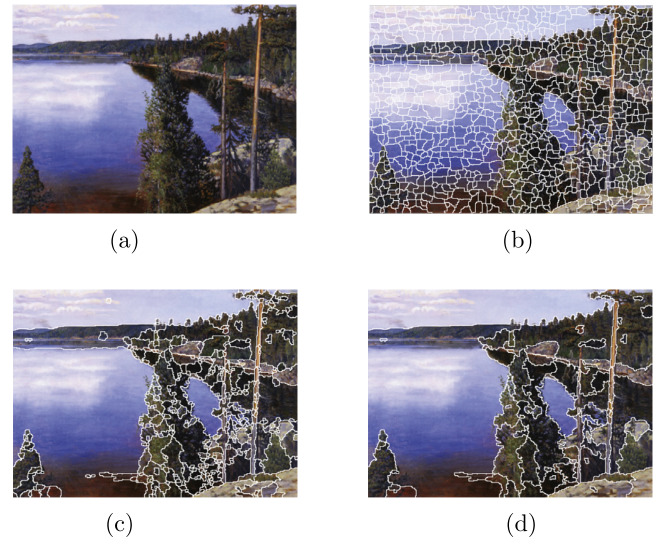

Fig. 2. (a) Example of a painting image in our dataset. (b)Initial watershed segmentation. (c) Initial mean shift segmentation. (d) Finial segmentation after region merging.

if the similarity (Bhattacharyya coefficient) between adjacent regions P and Q is the highest among all the similarities, then regions P and Q are merged. Following this merging rule, we obtain the final segmentation map. In this experiment, we iterate twice for merging the adjacent regions. The initial mean shift segmentation and final segmentation images are shown in Fig. 2.

Human vision is sensitive to the large segments in images. Thus, in this experiment, we mainly extract the features in the first largest segment (FLS) and the second largest segment (SLS).

4.2.2. Local color features

Color features are important not only for the global impression but also for the local details. In this part, the local color features are represented through the average of hue, saturation and lightness for the two largest segments and the color contrast between the largest segments and their neighbor segments.

Hue, saturation and lightness of the two largest segments

$f_{10} - f_{15}$: Hue (f_{10}), saturation (f_{12}), and lightness (f_{14}) of the FLS. Hue (f_{11}), saturation (f_{13}), and lightness (f_{15}) of the SLS.

A total of 6 features are calculated based on HSL color space in this part. Hue, saturation and lightness are calculated as follows:

$$f_{9+i} = \frac{1}{Area_{R_i}} \sum_{(m,n)\in R_i} I_H(m, n), i = 1, 2 \tag{16}$$

$$f_{11+i} = \frac{1}{Area_{R_i}} \sum_{(m,n)\in R_i} I_S(m, n), i = 1, 2 \tag{17}$$

$$f_{13+i} = \frac{1}{Area_{R_i}} \sum_{(m,n)\in R_i} I_L(m, n), i = 1, 2 \tag{18}$$

where $Area_{R_i}$ is the area of the segment R_i. $I_H(m, n)$, $I_S(m, n)$, and $I_L(m, n)$ are the hue, saturation, and lightness value of the pixel (m, n), respectively.

Color contrast between segments

$f_{16} - f_{18}$: Contrast of hue (f_{16}), saturation (f_{17}), and lightness (f_{18}) between the FLS and its neighbor segments.

In this part, we analyze the contrast of hue, saturation and lightness between the largest segment and its neighbor segments. The calculations are listed as follows.

The hue, saturation and lightness contrasts of the largest segment are calculated as follows:

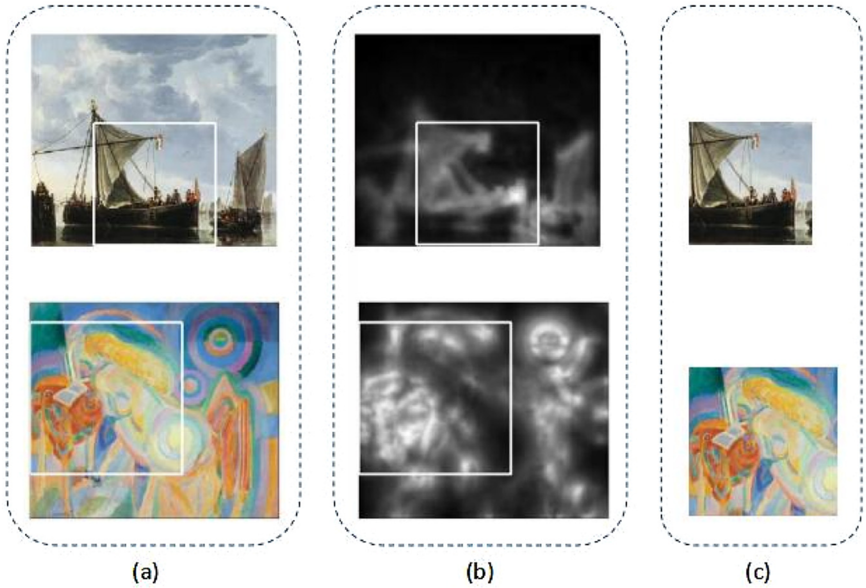

Fig. 3. The salient region extracted from the painting images. (a) The input painting images, (b) The saliency maps of the input images, (c) The cropped salient regions.

$$f_{16} = max|H_{largest} - H_i|,\ i \in \Omega_{nei} \tag{19}$$

$$f_{17} = max|S_{largest} - S_i|,\ i \in \Omega_{nei} \tag{20}$$

$$f_{18} = max|L_{largest} - L_i|,\ i \in \Omega_{nei} \tag{21}$$

where Ω_{nei} is the set of neighbor segments around the largest segment. $H_{largest}$ is the hue value of the largest segment, and H_i is the hue value of the *ith* neighbor segment. $S_{largest}$ is the saturation value of the largest segment, and S_i is the saturation value of the *ith* neighbor segment. $L_{largest}$ is the lightness value of the largest segment, and L_i is the lightness value of the *ith* neighbor segment.

4.2.3. Distribution features

In the experiment of subjective assessment, the distribution of the components in the painting image is found to be an important aspect that affects visual complexity perception. This aspect refers to the distribution of the segments, the number of all segments and the shapes of the major segments.

Number of all segments

f_{19}: Number of all segments. In general, a larger number of segments in a painting image implies high complexity of the image.

In the image segmentation procedure, the painting image is merged according to color similarity, and finally segmented into small regions. The number of these small regions is the number of all segments.

Areas of top two largest segments

$f_{20} - f_{21}$: Areas of the FLS (f_{20}) and the SLS (f_{21}). A larger segment denotes a highly homogenous region. This condition creates gentle visual perceptions.

Shape complexity of the first two largest segments

$f_{22} - f_{23}$: Shape complexities of the FLS (f_{22}) and the SLS (f_{23}).

The contours of the regions with different shapes (rectangle, circle or fractal) elicit different visual perception [39], which is used to measure the shape complexity of the top two largest segments.

The shape complexity of each segment is calculated as follows:

$$f_{21+i} = \frac{P_{R_i}^2}{4\pi A_{R_i}},\ i = 1, 2 \tag{22}$$

where P_{R_i} and A_{R_i} are the perimeters and areas of both the first and second largest regions.

4.3. Salient region features

Salient region is the most visually important region in an image [40]. When the observers see a painting, they are easily attracted to the visual salient region at the first glance. In this part, we introduce the features extracted from the salient region.

4.3.1. Salient region extraction

Our purpose is to obtain a local salient region, in which the human viewers' attentions are highly attracted. In order to achieve this purpose, we first generate the saliency map and then crop the saliency map into a local region.

In the step of saliency map generation, we adopt the method of context-aware saliency proposed in [40] which detects the visually important parts of the scene. This saliency is based on four principles observed in the psychological literature: local low-level considerations, global considerations, visual organizational rules, and high-level factors. By this method, we can obtain the saliency map of an input painting, which is shown in Fig. 4(b).

Based on the saliency map, the most visually important region is automatically cropped using the method mentioned in [41]. This method crops the region with an optimum rectangle which preserves at least 70% of total energy of the saliency map. The aspect ratio of the rectangle can be set as different ratios, such as, 16/16, 9/16, and 16/9. In our research, the aspect ratio of rectangle is set as 16/16 (a square). The samples of the input images and the cropped regions are shown in Fig. 3(a) and (c).

Based on the cropped salient region, we extract four color features ($f_{24} - f_{27}$) and two content features ($f_{28} - f_{29}$). The definition of each feature is defined as follows.

4.3.2. Color features

The color complexity measure and the basic colors are extracted.

Color complexity measure

f_{24}: The CCM of the salient region. It is calculated according to Eq. (1).

Basic colors

$f_{25} - f_{27}$: Hue (f_{25}), saturation (f_{26}), and lightness (f_{27}) of the salient region. The calculation formulas are similar to the Eq. (16)-(18). Here, $Area_{R_i}$ refers to the area of the salient region.

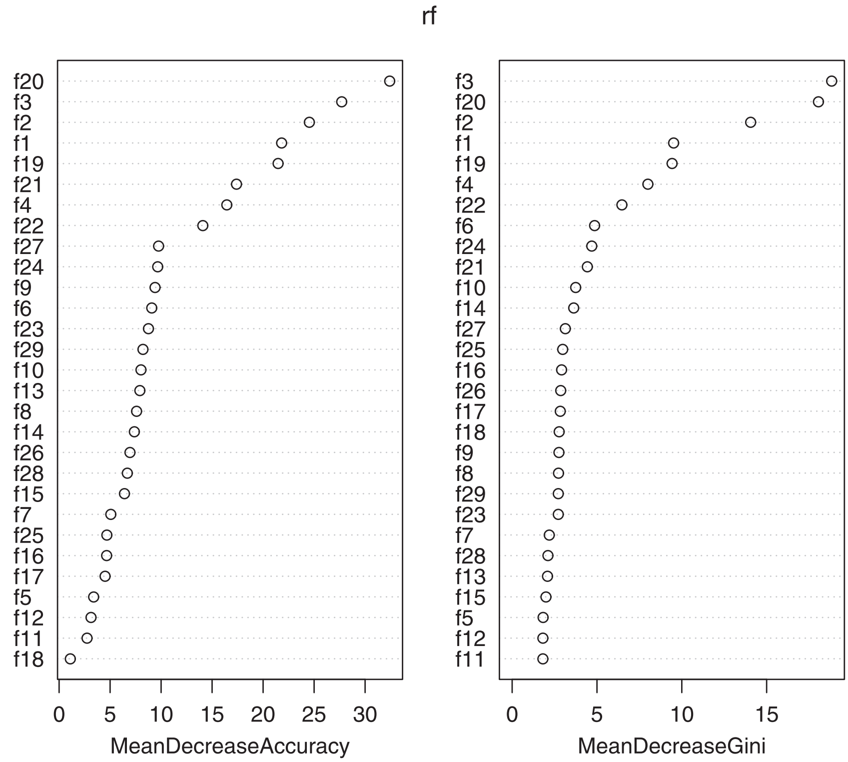

Fig. 4. Ranking variable importance associated with visual complexity of painting images by RF.

4.3.3. Content features

For the content features, the edge density and the area of the salient region are calculated.

f_{28}: The edge density of the salient region. The computation is similar as f_9.

f_{29}: The ratio between the area of the salient region and the area of the whole image.

A total of 29 global, local, and salient region features are extracted to represent the factors identified in the subjective experiment.

5. Classification and regression models of perceived complexity

The purpose of this work is to propose a computational measure of estimating visual complexity of paintings. To achieve this purpose, we propose two ways to objectively measure the complexity of painting images: 1) to regard the complexity evaluation as a three-class pattern classification problem, and 2) to predict a complexity score for a painting image using regression.

5.1. Complexity classification by random forest and support vector machine

In the classification process, we classify the paintings into three classes: LC, MC, and HC. From the subjective complexity experiment, we calculate the histogram of the scores for all images. The peaks of labels are concentrated around “2”, “4”, and “5.8”. Accordingly, the range of “1-2.5”, “3.5-4.5”, and “5.5-7” are respectively set as the classes of LC, MC, and HC. In this case, a large gap is assigned between different levels. In total, 450 painting images are labeled as LC, MC, and HC. s There are many classifiers used for image classifications [42–45]. In this paper, we employ Random Forest (RF) and Support Vector Machine (SVM) for complexity classification. RF is a powerful machine learning classifier that is first introduced in [42] and then developed in [43]. The RF classifiers are operated by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees. SVMs [46] separate classes by maximizing the margin between a hyperplane and the nearest training examples, called support vectors. We implement RF algorithm with randomForest [47] toolbox and SVM with e1071 toolbox in R (version 3.3.0.). The parameters for RF are set as 1000 trees and a minimum node size of 10. We employ 70% of data as training data and the rest 30% of data as testing data. For each time of training and testing, the data are randomly selected. The average accuracy (86.78% ± 1.01%) is calculated among 5 times of classifications. In addition, we make experiments to examine the performance when 30% of the data is used as the training data. The average accuracy (83.86% ± 1.44%) is obtained among five times of classifications. Compared with the situation that 70% of the data is used as the training data, the classification accuracy becomes less than before. But the difference is slight. For SVM classifiers, a Radial Basis Function is set to assign the complexity labels. To avoid the overfitting, we carry out 5-fold cross validation and obtain the average accuracy (78.88% ± 0.02%). Compared with SVM, RF classifiers perform higher accuracy in predicting the visual complexity of painting images.

5.1.1. Importance of each feature

By RF classification, we obtain the importance of each feature for predicting visual complexity of painting images. As shown in Fig. 4, MeanDecreaseAccuracy defines how much the model fit decreases when a variable is dropped. The greater the drop the more significant the variable is. According to the MeanDecreaseAccuracy value of each feature, we list the top ten features and their corresponding meanings in Table 2. From this table, it is clearly shown that f_{20} and f_3 play very important roles in classifying the visual complexity of painting images. The importance of the ten features can also be identified from the figure of MeanDecreaseGini in Fig. 4. MeanDecreaseGini is a measure of variable importance based on the Gini impurity index which is used for the calculation of splits during training [46]. Both Meandecreaseaccuracy and MeanDecreaseGini reflect the feature importance.

In order to check the performance in case the features are decreased, we replicate the experiment by 29 times while reducing one feature for each time. A total of 29 classification accuracies are shown in Fig. 5. In

Table 2
Top ten important features for visual complexity perception of painting images.

Features	Definitions	MeanDecreaseAccuracy	Group
f_{20}	Area of the first largest segment	32.42	Local
f_3	Color complexity of the third-level Gaussian pyramid	27.69	Global
f_2	Color complexity of the second-level Gaussian pyramid	24.54	Global
f_1	Color complexity of the first-level Gaussian pyramid	21.83	Global
f_{19}	Number of all segments	21.46	Local
f_{21}	Areas of the second largest segment	17.38	Local
f_4	Color complexity of the forth-level Gaussian pyramid	16.44	Global
f_{22}	Shape complexity of the first largest segment	14.094	Local
f_{27}	Lightness of the salient region	9.75	Salient
f_{24}	Color complexity of the salient region	9.66	Salient

this figure, the “yellow line” shows the accuracy that when we use all features for classification. Each subsequent mark of horizontal axis means the deletion of the current feature in each classification. It is clearly shown that when the feature number is decreased, the classification accuracy is decreased.

Moreover, we replicate the experiment by using different features combinations: global features, local features and salient features. The average accuracies are calculated among 5 times of classifications, which are shown in Table 3. In this table, all features used for classification yield the highest classification accuracy.

5.1.2. Comparisons with other measures of complexity

We compare the proposed method with the conventional measures of complexity. Redies et al. [23] proposed a measure of image complexity based on the Maximum Gradient Magnitudes (MGM) of each pixel in the Lab color space. Sun, Yamasaki and Aizawa [24] predicted image complexity based on 114-dimension image features. We realize the above computational methods of complexity and calculate Pearson’s Correlation Coefficient (PCC) between subjective complexity

Table 3
Feature combinations and performances.

Features	performance(Accuracy)
All features	86.78% ± 1.01%
Global features	77.66% ± 1.41%
Local features	81.86% ± 2.92%
Salient features	72.25% ± 2.72%

Table 4
Comparison of correlations between subjective complexity and objective measures of complexity for different methods.

	PCC
Redies-MGM [23]	0.521 *
Sun-114Features [24]	0.716 *
Proposed	0.777*

$*=p < 0.001$

assessments and objective measures of complexity. The comparisons are shown in Table 4. In this table, it is indicated that the proposed method outperforms other measures with a higher correlation (r = 0.777) between subjective complexity and objective measure of complexity.

5.2. Complexity regression by random forest and support vector regression

As complexity is a continuous value, a regression is suitable to predict a complexity score for a painting image. For each painting image, the corresponding feature vector is properly normalized to the range of [0,1]. We employ support vector regression (SVR) and random forest (RF) for regression. For both regression methods, the performance is measured by mean square error (MSE) and PCC. The parameters of SVR are learned by performing 5-fold cross validation on the training set. Table 5 shows the comparison results between SVR and RF. From this table, RF regression model is better than SVR.

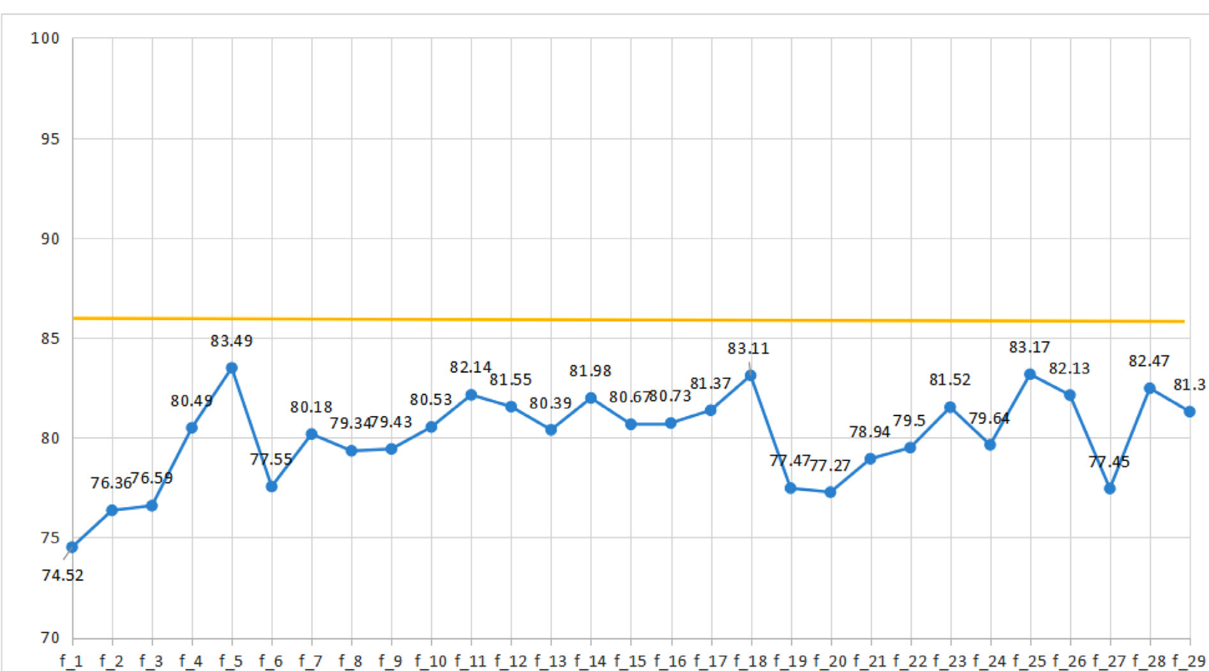

Fig. 5. Feature combination with one feature deleted and the corresponding accuracy.

Table 5
Complexity regression by SVR and RF.

Methods	MSE	PCC
SVR	0.662 ± 0.060	0.692 ± 0.0.038 *
RF	0.645 ± 0.035	0.777 ± 0.015 *

$* = p < 0.001$

5.2.1. Performance on MART dataset

MART dataset [48] is an art collection of the MART museum in Rovereto, Italy (the digitalized images of the artworks are contained in the electronic archive of MART). In this work [48], the MART dataset is public with the subjective negative and positive scores from 1 to 7, where 1 meant a highly negative emotion and 7 meant a highly positive emotion. In this paper, we use MART dataset as the comparison dataset. We firstly collect the subjective complexity scores through a website and build a new MART database with complexity scores[5], and then perform our methods on this database. The results are shown in Table 6.

5.2.2. Predicting aesthetic quality of paintings in jenaesthetics dataset using proposed method

Visual complexity has been proved to be highly related to visual aesthetic of images. In this section, we try to apply the proposed regression method to predict the other features (aesthetic, beauty, Liking of Color et al.) of the painting images involved in JenAesthetics dataset [26–29].

We select 600 painting images from JenAesthetics dataset for the experiment. Visual complexity features are extracted by the measures illustrated in Section 4. In the prediction process, random forest regression is used for predicting the aesthetic quality of the painting images. To verify the performance of our method on evaluating other features, PCC between predicted values and subjective values are calculated and shown in Table 7. According to the coefficients in this table, it reflects that our proposed method has a good ability of predicting aesthetic quality, beauty, and liking of color of the painting images.

6. Discussions

We study theoretical and empirical concepts from psychology and art theory to design the features that represent the above three factors. A total of 29 global, local, and salient region features are extracted from the painting images. These features are finally applied in a RF classifier to predict the complexity of paintings into three levels, namely, HC, MC, and LC. By RF classification, we also investigate the importance of each feature for classification. In Table 2, ten important features and their corresponding meanings are listed. These results can help us further understand about which features are more powerful for the visual complexity assessment. The meanings of these features from the theories of human visual perception and psychology are explained as follows.

According to the results of subjective experiment, “distribution of components” is regarded as one of the important factors that affect visual complexity perception of painting images. f_{19}, f_{20}, f_{21} and f_{22}, the features that represent the factor of “distribution of components”, play important roles in the RF classification model. As shown in Fig. 4, when the variable f_{20} is removed from the feature set, the fit of the RF classification model decreases very quickly. The greater the drop the more significant the variable is. Thus, f_{20} plays the most important role in RF classification model. Human visual perception is sensitive to the largest segment of the image. The importance of f_{20} in the proposed model also proves that the area of the largest segment significantly affects subjective visual complexity of paintings. Obviously, a painting image with a large number of segments seems to have high visual complexity. Furthermore, f_{19} and f_{21} also prove to be of vital importance for predicting visual complexity of paintings. Besides, for the largest segment, its shape complexity also affects subjective visual complexity perception. It is shown in Fig. 4 that f_{22} is the sixth important feature associated with visual complexity of painting images by RF.

Moreover, f_1, f_2, f_3 and f_4 are the average color complexity measures of a painting in four levels of image pyramid. They are measured based on the color variation in a local region of a painting. A high color variation in a local region indicates that this local region has high complexity. For a whole painting, its color complexity highly affects visual complexity of paintings. This has been confirmed in our subjective experiment that color is an important factor affecting human perception of visual complexity in paintings. In addition, different levels of image pyramid contain different features. Some features hidden in this resolution are obvious in another resolution. In Fig. 4, it shows that f_3 plays a more important role than f_2, f_1, and f_4 in classifying the visual complexity of paintings. The color feature (f_{24}) in the salient region is found to be significant in visual complexity prediction. Furthermore, f_{27} shows lightness of the salient region. It is important for visual complexity classification.

f_9 is the edgy density of a painting image. It represents the content characteristic of the painting images. Generally, the more contents exist in a painting, the more complex the painting is. It is revealed in Fig. 4 that f_9 is significant for classification.

Table 6
Performance on MART database.

Methods	MSE	PCC
SVR	0.423 ± 0.038	0.551 ± 0.036*
RF	0.389 ± 0.015	0.580 ± 0.022*

$* = p < 0.001$

Table 7
Comparison of correlations between predicted values and subjective values.

	PCC
Aesthetic Quality	0.768*
Beauty	0.786*
Liking of Color	0.734*
Liking of Composition	0.174
Liking of Content	0.260

$* = p < 0.001$

7. Conclusions and future work

Visual complexity is extremely important for human being to understand and perceive the visual stimuli. It has a wide range of applications both in the fields of psychology and computer science. Therefore, how to evaluate visual complexity from human visual system needs to be solved timely and adequately. In this paper, we firstly identify the factors that affect visual complexity of painting images. Based on these factors, we study theoretical and empirical concepts from psychology and art theory to design 29 global, local, and salient region features which are highly associated with the human visual perception system. Furthermore, we propose two ways to assess visual complexity of paining images. One is to evaluate the visual complexity level of painting images by classifying the complexity level into three levels (low, middle and high complexity). Another one is to predict a complexity value for painting images by regression model. By the comparisons, the proposed RF classification method performs better than other measurements of image complexity in predicting the visual complexity of painting images. Moreover, we apply the RF regression

[5] For future information about the collected complexity scores, please contact guoxiaoying@sxu.edu.cn or 19850102eagle@163.com

model to predict the other features of painting images. The results show that the proposed method has a good ability of evaluating aesthetic quality, beauty, and liking of color of the painting images involved in JenAesthetics dataset. This work might help to understand the visual complexity of visual stimuli and contribute to measure complexity objectively from human visual perception system.

In the future, we will focus on enlarging our image database to increase the robustness of our proposed method and extending our method to other applications (e.g. painting style classification based on image complexity).

Acknowledgments

The authors also thank the subjects who participated in the experiment. This work is supported by the National Natural Science Foundation of China [Grant Nos. 61603228, 61672332, 61322211, 61432011].

Appendix A. The questionnaire in the experiment of subjective assessment of paintings.

The questionnaire used in the experiment of subjective perception of complexity is as follows:

Please answer the following two questions.

1) which factors affect your judgment of the visual complexity of paintings(multiple choices allowed)?

A. Colors
B. Strength of the color changes
C. Content (elements, objects, people)
D. Distribution of composition (regular, or not)
E. Understandability-Abstract
F. Symmetry
G. Contrast
H. Familiarity
I. · (if you have any other reasons)

2) Which two factors that are the most important?·

References

[1] D. Berlyne, Studies in the New Experimental Aesthetics, Washington DC: Hemisphere Pub. Corp., 1974.
[2] C. Li, T. Chen, Aesthetic visual quality assessment of paintings, IEEE J. Sel. Top Signal Process. 3 (2) (2009) 236–253.
[3] X. Guo, T. Kurita, C.M. Asano, A. Asano, Visual complexity assessment of painting images, image processing (ICIP), 2013 20th IEEE International Conference on Sept. 15–18, (2013), pp. 388–392.
[4] J. Perkio, A. Hyvarinen, Modelling image complexity by independent component analysis, with application to content-based image retrieval, 19th International Conference on Artificial Neural Networks: Part II, Limassol, Cyprus, (2009), pp. 704–714.
[5] Y.A. Andrienko, N.V. Brilliantov, J. Kurths, Complexity of two-dimensional patterns, Eur. Phys. J. B 15 (3) (2000) 539–546.
[6] A. Klinger, N.A. Salingaros, Environment and planning b, Plann. Design 27 (4) (2011) 537–547.
[7] M.I.C. Murguia, A.D.C. Saenz, R.S. Rodriguez, A fuzzy approach on image complexity measure, Computacion y Sistemas 10 (3) (2007) 268–284.
[8] M. Cardaci, V.D. Gesù, M. Petrou, M.E. Tabacchi, Attentional vs computational complexity measures in observing paintings, Spat. Vis. 22 (3) (2009) 195–209.
[9] M. Cardaci, V.D. Gesù, M. Petrou, M.E. Tabacchi, A fuzzy approach to the evaluation of image compexity, Fuzzy Sets Syst. (2009) 1474–1484.
[10] A. Forsythe, M. Nadal, N. Sheehy, C.J. Cela-Conde, M. Sawey, Predicting beauty: fractal dimension and visual complexity in art, Brit. J. Psychol. (2011) 49–70. (London, England:1953)
[11] D.C. Donderi, An information theory analysis of visual complexity and dissimilarity, Perception 35 (6) (2006) 823–835.
[12] R. Rosenholtz, Y. Li, L. Nakano, Measuring visual clutter, J. Vision 7 (2) (2007) 1–22. 17
[13] J.E. Cutting, J.J. Garvin, Fractal curves and complexity, Percept. Psychophys. 42 (4) (1987) 365–370.
[14] M. Jamzad, F. Yaghmaee, Achieving higher stability in watermarking according to image complexity, Scientia Iranica J. 13 (4) (2006) 404–412.
[15] M.P.D. Silva, V. Courboulay, P. Estraillier, Image complexity measure based on visual attention, Proc. ICIP, (2011), pp. 3281–3284.
[16] A. Olive, M.L. Mack, M. Shrestha, A. Peeper, Identifying the perceptual dimensions of visual complexity of scenes, Proceeding of the 26th Annual Meeting of the Congnitive Society, (2004), pp. 1041–1046.
[17] R. Pieters, M. Wedel, R. Batra, The stopping power of advertising: measures and effects of visual complexity, J. Mark. 74 (5) (2010) 48–60.
[18] W. Saleem, A. Belyaev, D. Wang, H.P. Seidel, On visual complexity of 3d shapes, Comput. Graph. 35 (3) (2011) 580–585.
[19] H.C. Purchase, E. Freeman, J. Hamer, Predicting visual complexity, in: Predicting perceptions, The 3rd International Conference on Appearance, Edinburgh, Scotland, 2012, pp. 17–19.
[20] J. Rigau, M. Feixas, M. Sbert, An information-theoretic framework for image complexity, in proceedings of the euro-graphics workshop on computational aesthetics in graphics, Visual. Imaging (2005).
[21] L.N. Patel, P. Holt, Testing a computational model of visual complexity in background images', Advanced Concepts for Intelligent Systems, Baden, 2000, pp. 119–123.
[22] C. Stickel, M. Ebner, A. Holzinger, The XAOS metric - understanding visual complexity as measure of usability', HCI in Work and Learning, Life and Leisure: 6th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering, USAB 2010, Klagenfurt, Austria, Proceedings, (2010), pp. 278–290.
[23] C. Redies, S.A. Amirshahi, M. Koch, J. Denzler, PHOG-derived aesthetic measures applied to color photographs of artworks, Int. Conf. Comput. Vision (2012) 522–531.
[24] L. Sun, T. Yamasaki, K. Aizawa, Relationship between visual complexity and aesthetics: application to beauty prediction of photos, Eur. Conf. Comput. vision 2014 Workshops (2014) 20–34.
[25] Paintingdb, 2015, http://paintingdb.com/, assessed 8 September.
[26] Jenaaesthetics dataset, 2015, http://www.inf-cv.uni-jena.de/en/jenaesthetics, assessed 8 October.
[27] S.A. Amirshahi, G.U. Haynleichsenring, J. Denzler, C. Redies, Jenaesthetics subjective dataset: analyzing paintings by subjective scores, Eur. Conf. Comput. Vision (2014) 3–19.
[28] S.A. Amirshahi, C. Redies, J. Denzler, How self-similar are artworks at different levels of spatial resolution? Proceedings of the Symposium on Computational Aesthetics, ACM, 2013.
[29] S.A. Amirshahi, J. Denzler, C. Redies, JenAesthetics-a public dataset of paintings for aesthetic research. Tech. rep. Computer Vision Group, University of Jena Germany, 2013.
[30] R. Arnheim, Art and Visual Perception: a Psychology of the Creative Eye(expand and revised edition), University of Califormia Press, 1974.
[31] G. Wyszecki, W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, Wiley, New York, 1982.
[32] E. Stepanova, The impact of color palettes on the prices of paintings (July 9), 2015, Available at SSRN: https://doi.org/10.2139/ssrn.2807443.
[33] K.J. Yoon, I.S. Kweon, Color image segmentation considering of human sensitivity for color pattern variations, Proc. SPIE, (2001), pp. 269–278.
[34] M.c. system, 2015, http://en.wikipedia.org/wiki/Munsell_color_system, assessed 4 October.
[35] V.D. Gesù, C. Valenti, The Discrete Symmetry Transform in Computer Vision, Technical Report DMA 011 95, Palermo Univ., 1995.
[36] EDISON software, 2014, http://coewww.rutgers.edu/riul/research/code/EDISON/index.html, assessed 12 December.
[37] A. Tremeau, Regions adjacency graph applied to color image segmentation, IEEE Trans. Image Process. 9 (4) (2000) 735–744.
[38] A.B. Watson, Perimetric complexity of binary digital images, Math. J. (2011) 1–40.
[39] S. Goferman, L. Zelnikmanor, A. Tal, Context-aware saliency detection, IEEE Trans. Pattern Anal. Mach. Intell. 34 (10) (2012) 1915–1926.
[40] J. Chen, G. Bai, S. Liang, Z. Li, Automatic image cropping : a computational complexity study, 2016 Comput. Vision Pattern Recognit. (2016) 507–515.
[41] T.K. Ho, Random decision forests, proceedings of the 3rd international conference on document analysis and recognition, 1995, Montreal, QC, 14C16 August. 278–282.
[42] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[43] A. Liaw, M. Wiener, Classification and regression by randomforest, R News 2 (3) (2002) 18–22.
[44] F. Lu, J. Huang, K. Zhan, Boosting classifiers for scene category recognition, J. Inf. Hiding Multimed. Signal Process. 6 (4) (2015) 708–717.
[45] X. Xie, B. Li, X. Chai, Kernel-based nonparametric fisher classifier for hyperspectral remote sensing imagery, J. Inf. Hiding Multimed. Signal Process. 6 (3) (2015) 591–599.
[46] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 3 (2) (2011). 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm .
[47] L. Breiman, Random forests, J. Mach. Learn. 45 (2001) 5–32, http://dx.doi.org/10.1023/A:1010933404324.
[48] V. Yanulevskaya, J.R. Uijlings, E. Bruni, A. Sartori, E. Zamboni, F. Bacci, D. Melcher, N. Sebe, In the eye of the beholder: employing statistical analysis and eye tracking for analyzing abstract paintings, ACM Multimed. (2012) 349–358.