Multimedia Tools and Applications (2019) 78:19305–19323
https://doi.org/10.1007/s11042-019-7271-7

Selective clustering for representative paintings selection

Yingying Deng[1,2] · Fan Tang[1,2] · Weiming Dong[1] · Fuzhang Wu[3] · Oliver Deussen[4] · Changsheng Xu[1]

Received: 3 May 2018 / Revised: 20 December 2018 / Accepted: 24 January 2019 /
Published online: 9 February 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Selective classification (or rejection based classification) has been proved useful in many applications. In this paper we describe a selective clustering framework with reject option to carry out large-scale digital arts analysis. With the help of deep learning techniques, we extract content-style features from a pre-trained convolutional network for the paintings. By proposing a rejection mechanism under Bayesian framework, we focus on selecting style-oriented representative paintings of an artist, which is an interesting and challenging cultural heritage application. Two kinds of samples are rejected during the rejection based robust continuous clustering process. Representative paintings are selected during the selective clustering phase. Visual qualitative analysis on small painting set and large scale quantitative experiments on a subset of Wikiart show that the proposed rejection based selective clustering approach outperforms the standard clustering methods.

Keywords Digital arts analysis · Pattern mining · Rejection mechanism · Deep feature representation

1 Introduction

As digital image acquisition of painting arts has made rapid progress, millions of digital visual arts can be easily found on the internet [29, 43]. This makes computer-aided painting

✉ Weiming Dong
weiming.dong@ia.ac.cn

Oliver Deussen
oliver.deussen@uni-konstanz.de

Changsheng Xu
changsheng.xu@ia.ac.cn

[1] NLPR-LIAMA, Institute of Automation, Chinese Academy of Sciences, Beijing, China

[2] University of Chinese Academy of Sciences, Beijing, China

[3] Institute of Software, Chinese Academy of Sciences, Beijing, China

[2] University of Konstanz, Konstanz, Germany

analysis more relevant to social needs and to ordinary people who wish to get access to artworks. Scientists have made efforts to understand the aesthetic perception and connotation of digital arts in their own languages [22]. Researchers have come to a point where it is possible to perform computer-aided large-scale visual arts analysis to make a bridge between artworks and ordinary audiences.

As a significant characteristic of a painting, including texture, color, brush stroke [3], defines a specific genre for visual arts. During the past few years, researchers have made a tremendous progress in the automatic analysis of artworks targeting a diverse range of tasks, such as identifying western paintings [36], inferring painting styles [1, 6, 35], distinguishing authentic drawing from imitations [17], and automatically generating artworks in specific style [10, 25]. Furthermore, by using digital image processing and deep learning techniques, various methods on automatic painting analysis and generation have been studied by directly exploring powerful computational deep visual features [4, 7, 15, 29], but none deals with the problem of style-oriented visual art analysis by focusing on computer-aided representative painting selection.

As a particular case of the automatic art painting analysis, style-oriented representative painting selection focuses on exploring all the paintings of an artist and aims to select style specific visual arts. Indeed, artists immersed into the art paintings tend to express objects or scenarios in their own styles. However, external forms and techniques of the paintings from an artist may vary throughout all his/her painting works. Imaging you are trying to know more about the paintings of an artist, Vincent van Gogh for instance, and use "van Gogh" as keyword to search on *Google Images*.[1] The results, as shown in Fig. 1a contains many famous works of van Gogh, but are highly reduplicate in style. It's hard for audiences to recognize Fig. 1b by van Gogh. Therefore, when analysing digital art paintings, it is of utmost importance to analyse the preference evolution throughout one artist's painting career and find multiple artworks to represent an artist's creation characteristics. We believe the focus is interesting and challenging for the research community, since the statistic properties of art images are different with the ones of natural images.

In this work, we propose a novel method to select representative paintings of an artist. The proposed work aims for a global understanding of the painting characteristics of an artist throughout his career, and finding style-specific representative images in all the paintings under an unsupervised learning framework. A good clustering can help a data analyst to explore and understand a data set, what constitutes a good clustering may depend on domain-specific and application specific criteria [38]. Different from traditional clustering problems, we don't try to assign each image a correct label. Instead, we use rejection mechanism to pick out the unrepresentative and confused samples, so as to generate convincing clusters. Clustering centres are used as the original representative images. The rejection and clustering options are carried out literately.

The focus of representative painting selection has many potential applications, such as style-oriented style transfer between paintings and photos. Style transfer is used as a means to render an image in the artistic style of another one. An ideal style transfer algorithm should be able to extract and represent the semantic image content from the source image and then render the content in the style of the example image. The decomposition of content and style in artistic images is bound to the coupling between the source content and the example style. Previous image style transfer works only focus on expressing the artistic style of a specific painting [10, 25]. By adapting our method we can generate a series of stylized images with the artist's painting styles, as shown in Fig. 5.

[1] https://images.google.com/

(a) (b)

Fig. 1 **a** shows the results from *Google Image* when using "van Gogh" as key word; **b** is another painting chosen from more than 1900 images by Vincent van Gogh (collected from Wikiart.org [43])

In summary, our main contributions are:

- We introduce representative paintings selection as a new research direction in the field of cultural heritage.
- We adopt the progress in deep learning techniques to represent art paintings by content-style features.
- We integrate the concept of "rejection" and propose a selective clustering framework for art painting analysis. Two kinds of paintings are detected and rejected, outliers and confusing examples. Furthermore, we conduct a serious experiment to prove the soundness of the proposed method.

2 Related works

2.1 Automatic painting analysis

Taylor et al. [40] used "box counting" to analysis the fractal dimension of Pollock's drip paintings, showing that Pollock's drip paintings are fractal and the fractal dimensions increased during Pollocks career. Zhang et al. [44] measured the sophistication of computer generated abstract paintings by their computational complexity. Sartori et al. [32] analyzed the color combinations in abstract paintings and used this information to infer whether a painting would evoke positive or negative emotions in an observer. Sartori et al. [33] combined both visual and metadata features to learn statistical patterns associated with positive and negative emotions on abstract paintings. Shen [36] first introduced the problem of western painting identification by using a set of features to describe the color texture and shape of a painting in global and local level. Shamir et al. [35] extracted numerous visual descriptors to classify painters and schools of artworks. Arora et al. [1] showed that semantic-level information was more suitable for the task of fine-art genre classification. Liu et al. [26] learnt to predict the artistic styles. They learnt a dictionary representation for each style by taking artist-specific information into the learning progress.

Recently, deep learning models have shown potential for supporting approaches in the analysis of visual arts. Jangtjik et al. [19] built a data-set containing about one thousand paintings from 13 artists. They trained multi-scale deep networks by decomposing an image into multi-scale pyramid and aggregated the predictions from each layer adaptively. Chu et al. [4] used Gram matrix on deep features map as style vector to achieve image style classification and tested their feature using SVM on a data set containing about twenty thousand

oil paintings belonging to 17 styles. Mao et al. [29] learnt an efficacious deep representation for visual arts. Both content information and texture information were encoded in the final embedded feature vector. The deep architecture was trained by triplet-based ranking loss on *Art500k*, a large-scale visual arts dataset containing over 500,000 artworks, which were annotated with detailed labels of artist, art movement, genre, etc. Hicsonmez et al. [15] constructed a dataset of illustrations containing 6,468 distinct illustrations from 24 different illustrators. They compared serval different deep architectures on different classification tasks. Ma et al. [27] established a multi-task multi-range framework to identify authorship of art paintings. They used multiple correlated tasks such as style, genre, and date as related/auxiliary tasks to build a loss function jointly. They extracted features in multi-ranges: local features, region deep features, and global deep features. Furthermore, the authors took advantage of Random Forests to assess the importance of these three different kinds of features.

2.2 Representatives discovering

Representatives discovering, as a variant of pattern mining techniques that have been studied primarily amongst the data mining community, has a growing number of applications in the computer vision and graphics community. Wang et al. [42] proposed a sparse coding method which applied locality constrain to select similar basis of local image descriptors and learned a linear combination weight of these basis to reconstruct each descriptor. Doersch et al. [5] found meaningful mid-level elements that can represent a certain spatial area, such as Pairs. They built a sensitive detector to discover specific image patches of architectural style and considered an individual patch can solely serve as a representative of a location. Jain et al. [18] proposed to represent videos based on mid-level discriminative spatial-temporal patches. He et al. [16] focused on the representatives over a set of 3D shapes. They translated high-level style information into localized regions that characterize each style. Li et al. [24] combined CNN activations and a pattern mining technique to find clusters of image patches that were representative of, and which discriminated between, the contents of the relevant images. Noord et al. [41] trained a convolutional neural network on a large painting collection to perform the task of automatic artist attribution. Furthermore, they used the network to discover artist-specific visual features. Figure 2 shows the visualization of two artists' characteristics. The yellow regions are representative for Jan Luyken, whereas the blue regions are representative for Caspar Luyken. This kind of visualization can demonstrate that the learned network has the ability to distinguish dual authorship artworks, however, is not of much use to help the audience to know more about the artists' characteristic or preferences. Hicsonmez et al. [15] used two methods to find representative and discriminative elements out of all illustrations. They found meaningful patches using approach in [5] and representative image centroid following methods in [12].

In contrast to these previous works, our approach only focuses on image-level representative items selection. We aim at finding multiple artworks to represent an artist's creation characteristics. Compared with patch-level or pixel-level information, image-level representative items are more intuitive to generate an overall understanding of the artists' painting style.

2.3 Clustering

Clustering is an unsupervised method used for classification. MacQueen et al. [28] proposed K-means algorithm to divide a data set into clusters, while making the inter-class spacing

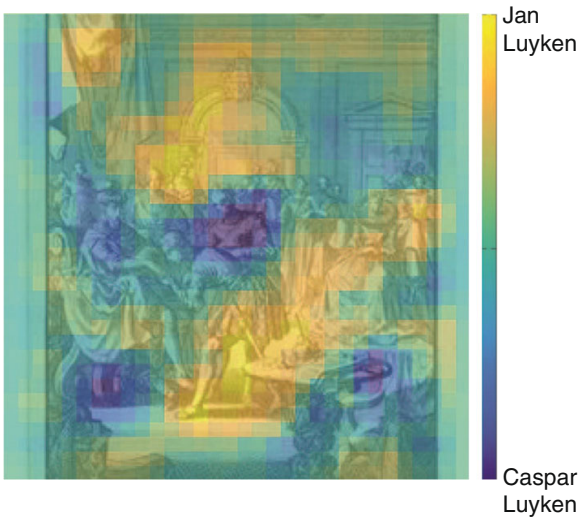


Fig. 2 Visualization in [41] of how characteristic each image region is for the artists Jan and Caspar Luyken

small and the extra-class spacing large. So far, many clustering tasks have chosen this classical algorithm. Karypis et al. [21] proposed a hierarchical clustering algorithm based on the idea of dynamic modeling (the chameleon algorithm), and also guided readers how to use hierarchical clustering on text data. The idea of spectral clustering algorithm comes from the theory of spectral division [8]. The optimal partition criterion based on graph theory is to make the internal similarity of the two subgraphs maximum and the similarity between the two subgraphs minimum. Soltanolkotabi et al. [37] developed a tractable algorithm that can provably cluster data points in a fairly challenging regime in which subspaces can overlap along many dimensions and in which the number of points per subspace is rather limited. Rodriguez et al. [31] propose an approach based on the idea that cluster centers are characterized by a higher density than their neighbors and by a relatively large distance from points with higher densities. Shah et al. [34] presented a clustering algorithm which used redescending robust estimators that allow even heavily mixed clusters to be untangled by optimizing a single continuous objective to handle large-scale data in high-dimensionality feature space. Compared to K-means and spectral clustering, robust continuous clustering does not need to know the number of clusters. Considering that hierarchical clustering algorithm has high computational complexity, adopting robust continuous clustering makes calculation more fast and easy to use. Therefore, in this paper, we employ robust continuous clustering as our clustering method.

2.4 Rejection mechanisms

Rejection-based classification (or selective classification) has been proposed by defining different optimal reject rules [2, 9, 14, 30, 39]. Recently, selective classification techniques have been considered in the context of deep neural networks. Geifman et al. [11] constructed a selective classifier by guaranteed risk control during testing phase of deep convolutional networks opening up possibilities to operate DNNs in mission-critical applications. Karmakar et al. [20] generated additional data points from the complement region and labeled them to a new class "Do not know" by estimating the domain of the training data. Then they trained the neural network with the training data and together with the generated data to equip the network with the ability to not make any judgment when it should not. Srivastava et al. [38] used rejection option when doing clustering analysis. They focused on the rejection on cluster level by a particularly simple feedback mechanism, in which an analyst can choose to reject individual clusters and request new ones. It is desirable to allow interactions

with users, but totally relying on users' intentions seems more or less subjective [45]. In this paper, we propose rejection based clustering analysis to achieve representative painting selection. We show that when selecting representative items, the results will be highly influenced by outliers or hard examples. We reject these two kinds examples by Bayes rules which has been widely used in traditional selective classification tasks [13, 23, 39].

3 Methodology

In this section, we describe our rejection-based clustering framework for art painting analysis. We begin with an overview of the proposed framework. Paintings are first represented by deep content-style features (see Section 3.1). Then we cluster the images using robust continuous clustering [34] (see Section 3.2). However, different from traditional unsupervised problems where most of the data have unambiguous structure, painting styles and preferences of an artist are usually hard to define and recognize. Except for the outliers, sometimes there may exist some images that are close to multiple clusters. This will affect the accuracy of representative painting extraction. Therefore, we integrate the concept of "rejection" to find the images which cannot be clearly classified in each class under a Bayesian prior elicitation framework to generate more reliable and representative clusters (see Section 3.3). The cluster-rejection process can be repeated as many times as desired.

3.1 Content-style feature extraction

In order to categorize a series of paintings of an artist into different clusters and find the representative painting of each cluster, we should represent a painting image in an efficient way. When analyzing the paintings of an artist, we considered numerous factors, such as period, residence, interest and so on, which make the paintings exhibit various appearances. Apparently, the appearances are influenced by not only paintings' styles but also the contents on account of artist's preference. Therefore, we construct a feature vector f_{cs} to describe both the style (texture, color, brushwork, etc.) and content (object, scene, etc.) features of a painting image.

Content representation When expressing scenes of different contents, the painting styles of an artist often vary when using different strokes and colourizations. Our algorithm starts with precomputing feature maps by a VGG-19 network which is trained on the ImageNet database for object recognition. Let $F_j(x)$ be the activations of the jth layer (convolutional layer) when processing the image x, which is a feature map of shape $C_j \times H_j \times W_j$. In the jth layer, C_j represents the number of convolution kernel. H_j represents the height of the image, W_j represents the width of the image. We transform the feature map into a 1D vector f_c containing $C_j H_j W_j$ elements and use it as object-based representation of the image. f_c contains the image *content* information.

Style representation Define the Gram matrix $G_j(x)$ to be the $C_j \times C_j$ matrix whose elements are given by:

$$G_j(x)_{c,c'} = \frac{1}{C_j H_j W_j} \sum_{h=1}^{H_j} \sum_{w=1}^{W_j} F_j(x)_{h,w,c} F_j(x)_{h,w,c'}, \tag{1}$$

where $G_j(x)$ is the inner product between the vectorized feature maps. In our algorithm, the feature space consists of the correlations between the different filter responses, where the expectation is taken over the spatial extent of the feature maps. Gram matrix calculates the correlations between these features [10]. The elements on the diagonal of the Gram matrix reflect the numbers of times that the features appear. Thus, the Gram matrix can represent the painting style of an image. When measuring the style difference between two images, we can just compare the differences between their Gram matrices. We transform the Gram matrix into a 1D vector f_s containing $C_j C_j$ elements and use it as the feature vector of the image *style*.

Finally, we formulate the image content-style feature vector as $f_{cs} = [f_c, f_s]$, $f_{cs} \in \mathbb{R}^D$.

Implementation details In convolutional neural network, the low level features pay more attentions to image details, the high level features pay more attentions to image structures. So we use the first feature set of *con1*, *con2*, *con3*, *con4*,*con5* from VGG-19 network as style layers. We use the feature set of fully connection layer *fc7* as content layer.

3.2 Robust continuous clustering

We use standard robust continuous clustering described in [34], namely RCC, to generate initial clusters. Based on robust estimation, RCC uses redescending robust estimators that allow even heavily mixed clusters to be untangled by optimizing a single continuous objective to handle large-scale data in high-dimensionality feature space. This method is fast, easy to use and doesn't need to know the number of clusters. Consider the input of RCC is n images represented by a set of content-style features, denoted as $\mathcal{F}_{cs} = \{f_{cs}^1, f_{cs}^2, \ldots, f_{cs}^n\}$, the goal of RCC is to find a set of representatives for $\mathcal{F}_{cs}$ which embeds all data points into a partial connection diagram ε and coalesces to reveal the cluster structure latent in the data by solving the object function in (2).

$$\begin{aligned} \mathbf{U}, \mathbb{L} = arg \min_{U,\varepsilon} \sum_{1}^{n} \| f_{cs}^i - u^i \|^2 \\ + \lambda \sum_{(p,q)\in\varepsilon} w_{p,q}(l_{p,q} \| u^p - u^q \|_2 + \Psi(l_{p,q})), \end{aligned} \tag{2}$$

where $\mathbf{U} = \{u^1, u^2, \ldots, u^n\}$ is the desired representative set embedded as nodes in graph ε. $l_{p,q}$ is the connection $(p, q) \in \varepsilon$, and $\mathbb{L} = \{l_{p,q}\}$. Here $\Psi(l_{p,q})$ is a penalty on ignoring the connection (p, q) to force the graph ε to observe the same latent cluster to collapse into a single point. Two data points in ε are considered connected if their cosine similarity is larger than clustering threshold δ(set to 0.9 by default according to paper [34]). $\Psi(l_{p,q})$ tends to zero when the connection is active. λ and $w_{p,q}$ are balanced items. Among them,

$$\Psi(l_{p,q}) = \mu(\sqrt{l_{p,q}} - 1)^2, \tag{3}$$

where μ is a scale parameter.

Optimization This algorithm uses iterative solution of linear least-squares systems to enable efficient and scalable optimization. Equation (2) can be solved by setting $\mathbf{U}$ and $l_{p,q}$ fixed and optimized respectively. Over the iterations, $w_{p,q}$ is precomputed, μ is automatically decreased, gradually introducing nonconvexity into the objective, and the value of λ is updated automatically according to this formula after every update of μ. Then the diseased

clusters can be picked out from the data graph ε as well as the learned representations of the data points $\mathbf{U}$.

3.3 Rejection operation

In this section we describe the rejection operation used in our framework. Centroid-based clustering methods are particularly vulnerable to misclustered samples. This will affect the accuracy of representative painting extraction. Based on such observation, we adopt rejection operation to find the images which cannot be clearly classified in each cluster to generate more reliable and representative clusters. We formalize this process into a Bayesian framework, in which we view class-conditional-probability density of each cluster generated in Section 3.2 as a mechanism for prior elicitation probability distribution to find the rejected images. Denote the representative feature set of all the paintings as $U = \{u^i | 1 \leq i \leq n\}$, where n is the size of paint set and $u^i \in \mathbb{R}^D$. Assume the elements in cluster U_M are mutually independent and obey the Gaussian distribution, then for each cluster M, we formulate the class-conditional-probability density of an image feature vector u^i as:

$$P(u^i|M) = \frac{1}{\sqrt{(2\pi)^D |\sum_M|}} e^{-\frac{1}{2}(u^i-\mu_M)^T \sum_M^{-1}(u^i-\mu_M)}, \tag{4}$$

where μ_M and $\sum_M$ are computed by maximum similarity estimation. Thus, we can get the original probability $P(M|u^i)$ by

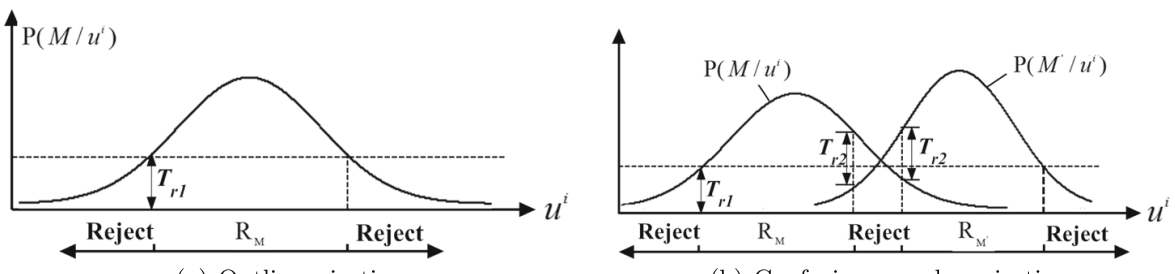
$$P(M|u^i) = \frac{P(u^i|M)P(M)}{P(u^i)}, \tag{5}$$

which indicates the probability of u^i belonging to class M. However, the original probability may be significantly affected by the paintings that is hard for style distinguishment. To boost the effectiveness of representative painting selection, we adopt *rejection* mechanism. As illustrated in Fig. 3, image which is represented as u^i classified to class M will be rejected:

- if $P(M|u^i)$ is not higher than a threshold T_{r1}, which means image u^i is more likely to be an outlier of M, as shown in Fig. 3a;
- if the difference between the probability of u^i belonging to class M and class M' is not higher than a threshold T_{r2}, which means u^i is easy to be confused, as shown in Fig. 3b.

In our experiments, T_{r1} and T_{r2} are respectively set as 40% and 20% of the peak value of $P(M|u^i)$. These values are obtained through numerous experiments. And we show an example experiment in Section 4.2.

Once this rejection is completed, we cluster the data again, modifying the probability distribution of each cluster. The cluster-rejection process are repeated until the stop criterion



(a) Outlier rejection (b) Confusing samples rejection

Fig. 3 Two kinds of rejection mechanisms

meets. Then, we get the representative image with the highest $P(M|u^i)$ in each cluster. The pseudocode for the proposed method is shown in Algorithm 1.

Algorithm 1 Selective clustering for representative paintings selection.

Require: art paintings represented by content-style features $\mathcal{F}_{cs} \in \mathbb{R}^D$

Initial
  $u^i \leftarrow f^i_{cs}$, build graph ε according to **U**
repeat
  Update μ_M and $\sum_M$ for cluster M by maximum similarity estimation
  for painting p $\in$ cluster M **do**
    calculate $P(M|p)$ and $P(M'|p)$ by (5)
    if $P(M|p) < T_{r1}$ **then**
      reject p as outlier
    else if $P(M|p) - P(M'|p) < T_{r2}$ **then**
      reject p as confusing example
    else
      cluster p to cluster M
    end if
  end for
until stopping criterion is met
Select cluster centers as representative paintings

4 Results and evaluation

In this section, we demonstrate the proposed algorithm and compare it to existing techniques. We start from showing representative painting selection results and qualitative analysis. Furthermore, we conduct quantitative evaluations of our technique, as well as comparisons with standard centroid-based clustering methods without rejection option proposed in our work (Figs. 4, 5 and 6).

In Figs. 4 and 6, we show representative paintings selection result by using the paintings of van Gogh and monet. Figure 5 shows four style transfer results using the selected

Fig. 4 Our representative painting selection result for Vincent van Gogh. Layout by CollageIt (http://www.collageitfree.com/), the same below

Fig. 5 Style transfer results using different representative paintings for Vincent van Gogh

representative paintings using the proposed method. Our program automatically selected 20 representative paintings from all the 1928 digital artworks by Vincent van Gogh. Compared with the retrieval result in Fig. 1a, our result shows more different styles, including some well-known artworks. Moreover, our selection covers more kinds of contents, such as still life, self-portrait and landscape. We show more representative painting selecting results in Figs. 7 and 8 together with style transfer results using different representative paintings.

4.1 Qualitative analysis

In Fig. 9, we show qualitative analysis visually on two small painting sets. We randomly selected 10 images from each of the three styles: "Rayonism", "Gothic" and "Gongbi". We compare our clustering results with the ones of standard K-means method in Fig. 9 on 30 images. Standard centroid-based K-means is usually used to generate meaningful clusters and the clustering centroid is the most representative samples for each cluster. Note the number of clusters are manually set to 3 for K-means but need not be known in advance for our method. Compared with standard K-means, our method can automatically detect fail cases in style "Rayonism" and "Gothic" that are misclustered to "Gongbi" and reject these confusing samples.

Verify the validity of rejection mechanism To reveal the performance of the proposed rejection mechanism intuitively, we compare the results of K-means clustering with and without rejection mechanism. Figure 10a shows the representative images before rejection, Fig. 10b shows the representative images after rejection. We use paintings of "Francis Picabia" to demonstrate the validity of rejection mechanism. In Fig. 10a, we can see that the result is highly influenced by hard examples (see C3). By integrating rejection mechanism, all the noisy samples are picked out, and the representative images can accurately present the creation characteristics of the artist.

Fig. 6 Our representative painting selection result for Claude Monet

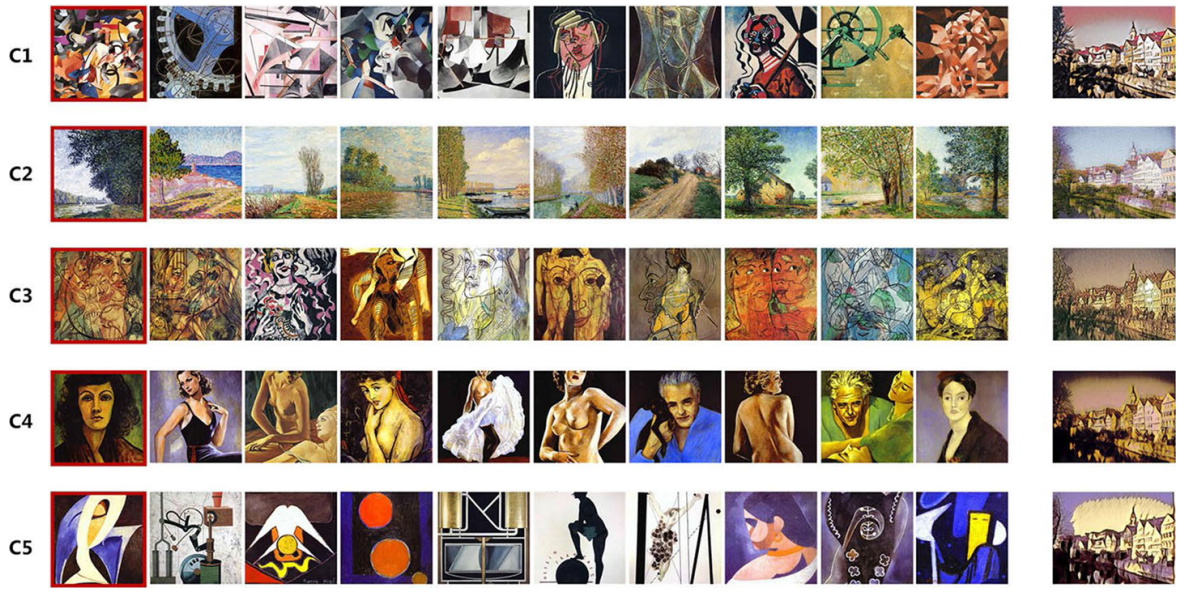


Fig. 7 Clustering and style transfer results using Francis Picabia's paintings. The images with red frames are representative images of each cluster. The last column is style transfer results

Verify the validity of content-style features Paintings of an artist are different both in style and content. In this work, we aim at selecting style-oriented representative paintings of an artist jointly with the analysis of painting content. If no content information is taken into consideration, we may get the wrong clustering results. We select 20 paintings by "Claude Monet" in Fig. 11a, of which half are water lilies and half are paintings of sea in impressionism style. Without the content feature, all these paintings are clustered into one class and the selected representative painting is shown in Fig. 11b. By adopting our approach, these paintings are gathered into two clusters and the representative paintings are shown in Fig. 11c. Therefore, both style and content feature are important to artist paintings clustering.

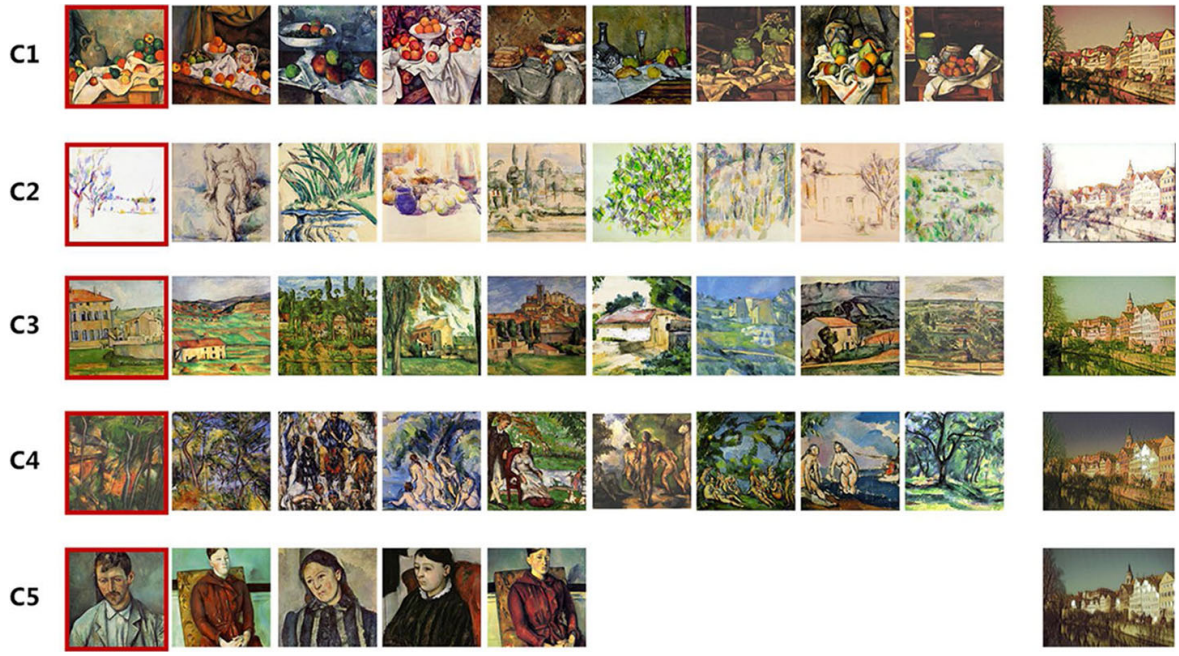
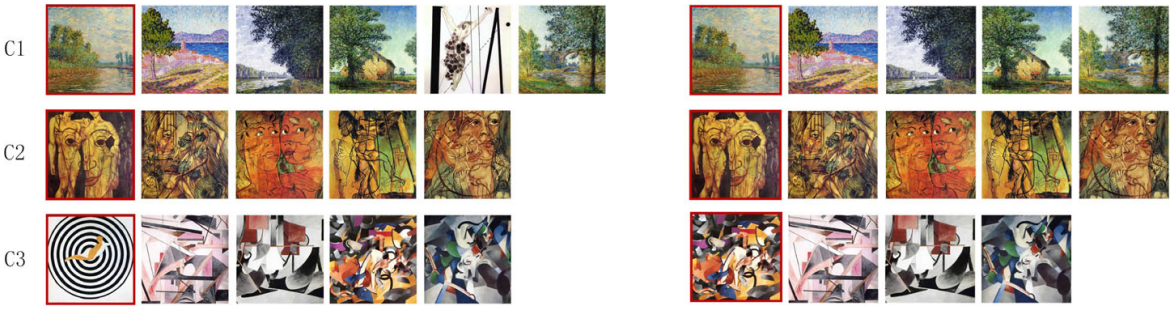


Fig. 8 Clustering and style transfer results using Paul Cezanne's paintings. The images with red frames are representative images of each cluster. The last column is style transfer results

(a) Result generated by standard K-means

(b) Results generated by our method with rejection

Fig. 9 Visual results for comparison with standard K-means. Clusters are indicated by different background colors (zoom in to see details)



(a) Representative images before rejection

(b) Representative images after rejection

Fig. 10 Rejection examples using K-means clustering. The images with red frames are representative images

(a) 20 paintings selected by "Claude Monet"

(b) selected most representative painting only using style features

(c) selected representative paintings using content-style features

Fig. 11 Selecting representative paintings without and with the content features

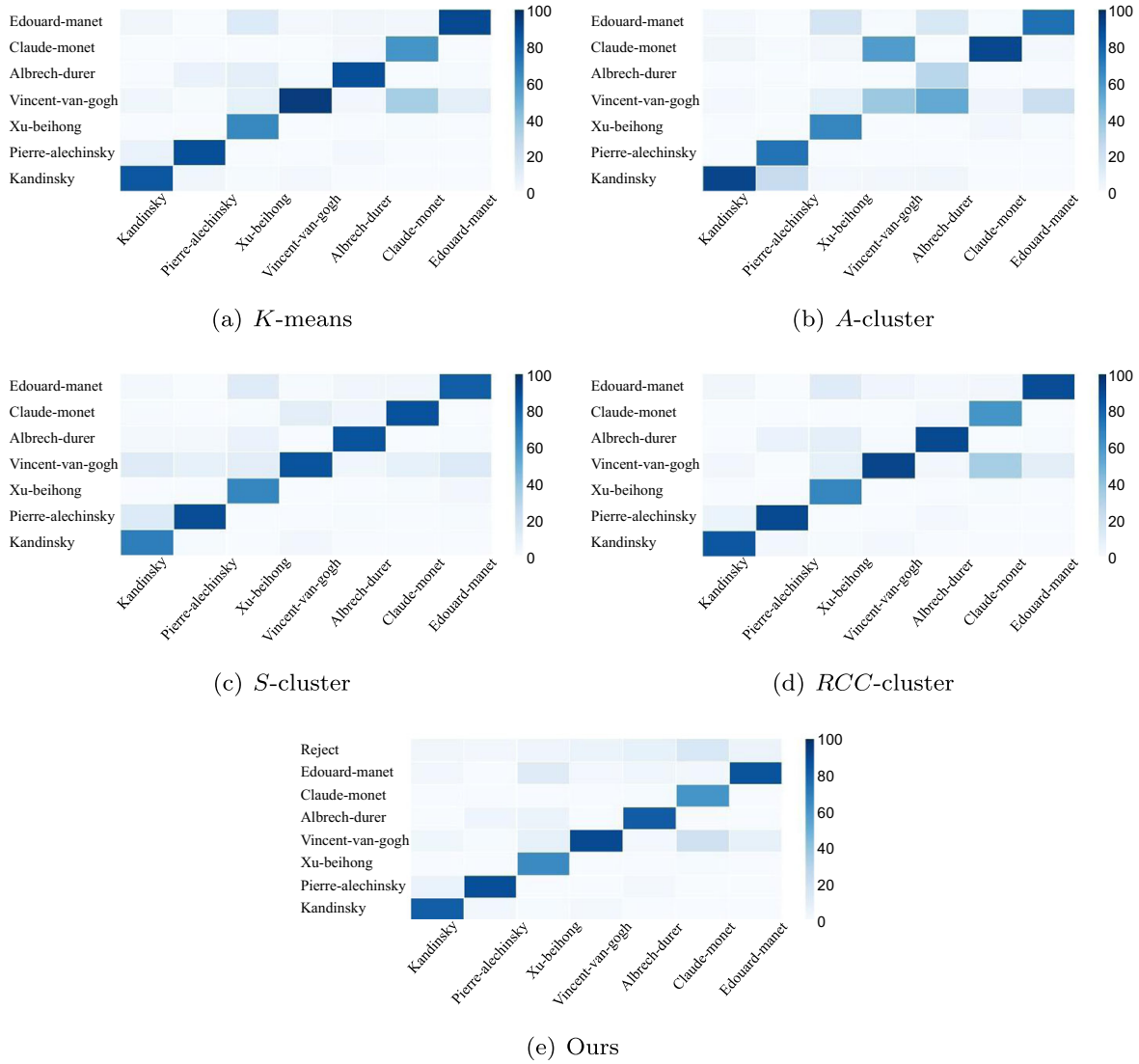


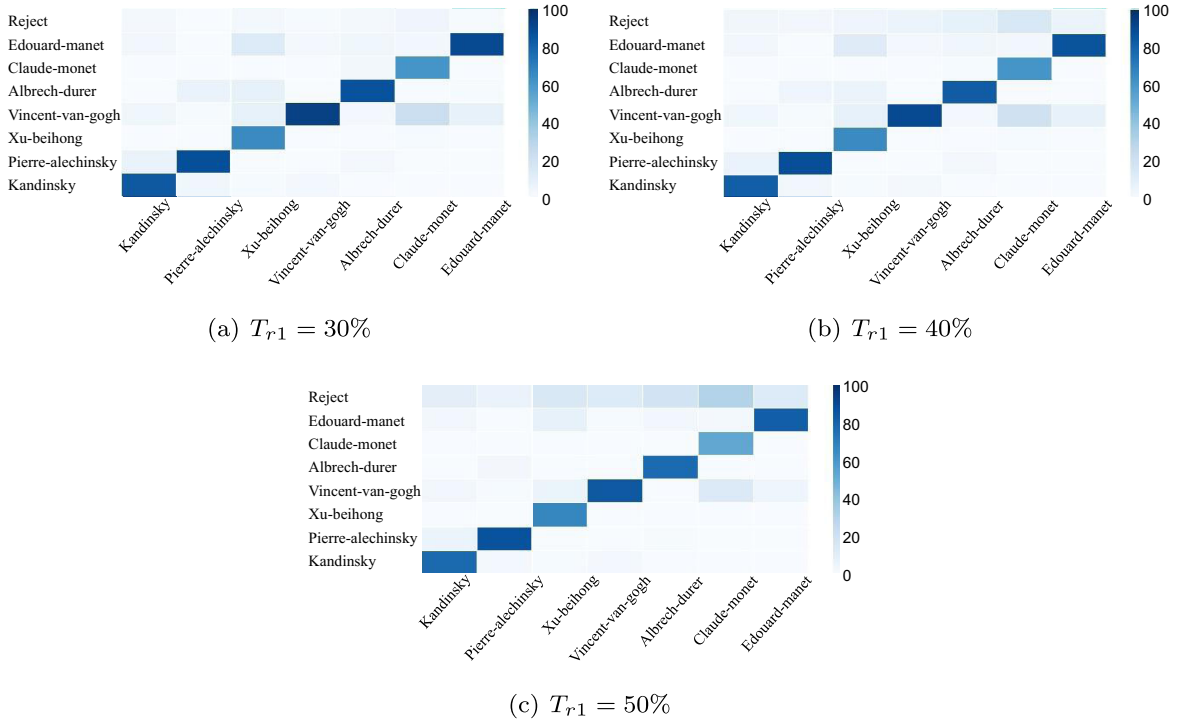
Fig. 12 Clustering results for testing on different clustering methods

4.2 Quantitative analysis

In this section, we test the proposed selective clustering method on large-scale painting collection. We select 1667 paintings on Wikiart.org [43] by seven artists: "Beihong Xu" (97 of 97), "Kindinsky" (218 of 218), "Pierre Alechinsky" (220 of 220), "Edourad Manet" (232 of 232), "Albrecht Durer" (300 of 838 randomly), "Monet" (300 of 1339 randomly), "Vincent van Gogh" (300 of 1928 randomly). We compared our method with three traditional clustering methods: standard K-means, agglomerative clustering (A-cluster) and spectral clustering (S-cluster). Paintings are all represented by the content-style feature

Table 1 Statics for different clustering methods

	K-means	A-cluster	S-cluster	Ours
Average Recall	82.15%	66.85%	81.30%	79.72%
Average Precision	85.54%	74.64%	83.58%	91.42%



(a) $T_{r1} = 30\%$ (b) $T_{r1} = 40\%$

(c) $T_{r1} = 50\%$

Fig. 13 Statistical results for testing on different rejection thresholds

and we adopt Euclidean distance to measure the similarity between features. Figure 12 shows the clustering results and Table 1 shows the statics. In Fig. 12 we can observe that our result contains the least wrong samples. This demonstrates that rejection mechanism can well recognize the misclustered samples. From Table 1 we can observe that the proposed method has higher average precision while lower average recall compared with other methods. It is evident that the proposed method implements selective clustering by rejecting the hard samples. By increasing clustering precision while slightly decreasing ratio of recall, our method can generate more reliable clusters for representative painting selection.

Test on rejection threshold During the rejection process, we used thresholds to detect two kinds of hard examples. We test the thresholds T_{r1} by setting it to 30%, 40%, 50%, respectively. Figure 13 shows the statistics. We can see that when the threshold T_{r1} is at a low level, paintings will be clustered into wrong class. As the increase of threshold T_{r1}, the rejection rate raises while the misclustering rate decreases. However, when the threshold T_{r1} is too high, some right samples will be rejected. Thus, we set T_{r1} as 40% of the peak value of $P(M|u^i)$ taking the median. The principle can apply to T_{r2} equally. If the threshold T_{r2} is too low, paintings will be clustered into wrong class. If the threshold T_{r2} is too high, some right samples will be rejected. Therefore, we set T_{r2} as 20% of the peak value of $P(M|u^i)$.

5 Conclusion and future work

In summary, this paper presents a new focus of art painting analysis. We propose a rejection-based selective clustering framework to select multiple representative painting images from the artworks of an artist. Two kinds of paintings are rejected during the clustering

process: outliers and confusing examples. Experimental results show that the proposed method can generate more meaningful samples for audience with little domain knowledge to comprehend more about an artist's painting styles.

However, in many cases, the representative artworks of an artist can not only be revealed by style-oriented analysis. It may be influenced by a wide range of factors. Therefore, we will try to take more issues into consideration in the future work. Besides, we will explore a more intelligent style transfer method that can combine the artistic styles of an artist, by using the representative paintings as guidance.

Acknowledgements This work was supported by National Natural Science Foundation of China under nos. 61832016, 61672520 and 61702488, as well as the independent research project of National Laboratory of Pattern Recognition.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Arora RS, Elgammal A (2012) Towards automated classification of fine-art painting style: A comparative study. In: Proceedings of the 21st international conference on pattern recognition (ICPR), pp 3541–3544
2. Bartlett PL, Wegkamp MH (2008) Classification with a reject option using a hinge loss. J Mach Learn Res 9:1823–1840
3. Chen CC, Amato G, Boujemaa N, Inria I, Pitas I, Alexander K, Kiernan K, Li CS, Wang ZJ (2002) Report of the delos-nsf working group on digital imagery for significant cultural and historical materials
4. Chu WT, Wu YL (2016) Deep correlation features for image style classification. In: Proceedings of the 2016 ACM on multimedia conference, MM '16. ACM, New York, pp 402–406
5. Doersch C, Singh S, Gupta A, Sivic J, Efros AA (2012) What makes paris look like paris? ACM Trans Graph 31(4):101:1–101:9
6. Dumoulin V, Shlens J, Kudlur M (2017) A learned representation for artistic style. In: International conference on learning representations (ICLR). arXiv:1610.07629
7. Dumoulin V, Shlens J, Kudlur M (2017) A learned representation for artistic style. In: ICLR, pp 9
8. Fiedler M (1973) Algebraic connectivity of graphs. Czechoslov Math J 23(23):298–305
9. Fumera G, Roli F, Giacinto G (2000) Reject option with multiple thresholds. Pattern Recogn 33(12):2099–2101
10. Gatys LA, Ecker AS, Bethge M (2016) Image style transfer using convolutional neural networks. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 2414–2423
11. Geifman Y, El-Yaniv R (2017) Selective classification for deep neural networks. In: Advances in neural information processing systems 30, pp 4878–4887. Curran Associates, Inc
12. Golge E, Duygulu P (2015) FAME: Face association through model evolution. In: IEEE conference on computer vision and pattern recognition workshops (CVPRW), pp 43–49
13. Grandvalet Y, Rakotomamonjy A, Keshet J, Canu S (2009) Support vector machines with a reject option. In: Advances in neural information processing systems 21, pp 537–544. Curran Associates, Inc
14. Herbei R, Wegkamp MH (2006) Classification with reject option. Can J Stat 34(4):709–721
15. Hicsonmez S, Samet N, Sener F, Duygulu P (2017) DRAW: Deep networks for recognizing styles of artists who illustrate children's books. In: Proceedings of the 2017 ACM on international conference on multimedia retrieval, ICMR'17. ACM, New York, pp 338–346
16. Hu R, Li W, Kaick OV, Huang H, Averkiou M, Cohen-Or D, Zhang H (2017) Co-locating style-defining elements on 3d shapes. ACM Trans Graph 36(3):33:1–33:15
17. Hughes JM, Graham DJ, Rockmore DN (2010) Quantification of artistic style through sparse coding analysis in the drawings of pieter bruegel the elder. Proc Natl Acad Sci 107(4):1279–1283
18. Jain A, Gupta A, Rodriguez M, Davis LS (2013) Representing videos using mid-level discriminative patches. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 2571–2578

19. Jangtjik KA, Yeh MC, Hua KL (2016) Artist-based classification via deep learning with multi-scale weighted pooling. In: Proceedings of the 2016 ACM on Multimedia Conference, MM '16. ACM, New York, pp 635–639
20. Karmakar B, Pal NR (2018) How to make a neural network say "don't know". Inf Sci 430:444–466
21. Karypis G, Han EH, Kumar V (1999) Chameleon: Hierarchical clustering using dynamic modeling. IEEE computer society press
22. Kim D, Son SW, Jeong H (2014) Large-scale quantitative analysis of painting arts. Sci Report 4:7370:1–7370:7
23. Li Q, Vempaty A, Varshney LR, Varshney PK (2017) Multi-object classification via crowdsourcing with a reject option. IEEE Trans Signal Process 65(4):1068–1081
24. Li Y, Liu L, Shen C, Hengel AV (2017) Mining mid-level visual patterns with deep cnn activations. Int J Comput Vis 121(3):344–364
25. Liao J, Yao Y, Yuan L, Hua G, Kang SB (2017) Visual attribute transfer through deep image analogy. ACM Trans Graph 36(4):120:1–120:15
26. Liu G, Yan Y, Ricci E, Yang Y, Han Y, Winkler S, Sebe N (2015) Inferring painting style with multi-task dictionary learning. In: Proceedings of the 24th international conference on artificial intelligence, IJCAI'15, pp 2162–2168. AAAI Press
27. Ma D, Gao F, Bai Y, Lou Y, Wang S, Huang T, Duan L (2017) From part to whole: Who is behind the painting?. In: Proceedings of the 25rd ACM international conference on multimedia, pp 1174–1182. ACM
28. Macqueen J (1965) Some methods for classification and analysis of multivariate observations. In: Proceedings of berkeley symposium on mathematical statistics and probability, pp 281–297
29. Mao H, Cheung M, She J (2017) Deepart: Learning joint representation of visual arts. In: Proceedings of the 25rd ACM international conference on multimedia, pp 1183–1191. ACM
30. Pudil P, Novovicova J, Blaha S, Kittler J (1992) Multistage pattern recognition with reject option. In: 11th IAPR international conference on pattern recognition, 1992. Vol. II. conference B: Pattern recognition methodology and systems, Proceedings, pp 92–95. IEEE
31. Rodriguez A, Laio A (2014) Clustering by fast search and find of density peaks. Science 344(6191):1492–1496
32. Sartori A, Culibrk D, Yan Y, Sebe N (2015) Who's afraid of itten: Using the art theory of color combination to analyze emotions in abstract paintings. In: Proceedings of the 23rd ACM international conference on multimedia, MM'15. ACM, New York, pp 311–320
33. Sartori A, Yan Y, Özbal G, Salah AAA, Salah AA, Sebe N (2015) Looking at mondrian's victory boogie-woogie: What do i feel?. In: Proceedings of the 24th international conference on artificial intelligence, IJCAI'15, pp 2503–2509. AAAI Press
34. Shah SA, Koltun V (2017) Robust continuous clustering. Proc Natl Acad Sci 114(37):9814–9819
35. Shamir L, Macura T, Orlov N, Eckley DM, Goldberg IG (2010) Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. ACM Trans Appl Percept 7(2):8:1–8:17
36. Shen J (2009) Stochastic modeling western paintings for effective classification. Pattern Recogn 42(2):293–301
37. Soltanolkotabi M, Elhamifar E, Candès EJ (2013) Robust subspace clustering. Ann Statist 42(2):669–699
38. Srivastava A, Zou J, Sutton C (2016) Clustering with a reject option: Interactive clustering as bayesian prior elicitation. In: ICML workshop on human interpretability in machinelearning(WHI). New York, NY, USA
39. Tax DM, Duin RP (2008) Growing a multi-class classifier with a reject option. Pattern Recogn Lett 29(10):1565–1570
40. Taylor RP, Micolich AP, Jonas D (1999) Fractal analysis of pollock's drip paintings. Nature 399(6735):422–422
41. van Noord N, Hendriks E, Postma E (2015) Toward discovery of the artist's style: Learning to recognize artists by their artworks. IEEE Signal Process Mag 32(4):46–54
42. Wang J, Yang J, Yu K, Lv F, Huang T, Gong Y (2010) Locality-constrained linear coding for image classification. In: Computer vision and pattern recognition, pp 3360–3367
43. Wikiart.org - visual art encyclopedia. https://www.wikiart.org//. Accessed January 30, 2018
44. Zhang K, Harrell S, Ji X (2012) Computational aesthetics: on the complexity of computer-generated paintings. Leonardo 45(2):243–248
45. Zhang X, Hu BG (2012) Learning in the class imbalance problem when costs are unknown for errors and rejects. In: 2012 IEEE 12th international conference on data mining workshops (ICDMW), pp 194–201. IEEE

Yingying Deng received the BSc degree in Automation from University of Science and Technology Beijing in 2017. She is currently working toward the PhD degree in National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. Her research interests include multimedia analysis, computer vision and machine learning.

Fan Tang received the BSc degree in Computer Science from North China Electric Power University in 2013. He is currently working toward the PhD degree in National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. His research interests include computer graphics, computer vision and machine learning.

Weiming Dong is a Professor in the Sino-European Lab in Computer Science, Automation and Applied Mathematics (LIAMA) and National Laboratory of Pattern Recognition (NLPR) at Institute of Automation, Chinese Academy of Sciences. He received his BSc and MSc degrees in Computer Science in 2001 and 2004, both from Tsinghua University, China. He received his PhD in Computer Science from the University of Lorraine, France, in 2007. His research interests include image synthesis and image recognition. Weiming Dong is a member of the ACM and IEEE.

Fuzhang Wu received the PhD candidate in the Sino-European Lab in Computer Science, Automation and Applied Mathematics (LIAMA) and National Laboratory of Pattern Recognition (NLPR) at Institute of Automation, Chinese Academy of Sciences (CASIA). He received his BSc in computer science in 2010 from Northeast Normal University, P. R. China. His research interests include image synthesis and image analysis. He is currently a Post-Doctoral Scholar with Institute of Software, Chinese Academy of Sciences.

Oliver Deussen graduated at Karlsruhe Institute of Technology in 1996 and worked as a postdoctoral researcher at University of Magdeburg on Non-Photorealistic Rendering. In 2000 he was appointed as a professor for Computer Graphics and Media Design by Dresden University of Technology, since 2003 he professor for Computer Graphics and Media Informatics at University of Konstanz. He is interested in a number of areas in computer graphics and information visualization. He has published several books and over 100 refereed publications, he is member of ACM Siggraph, Eurographics and Gesellschaft fuer Informatik, he is associate editor of Computer Graphics Forum and Informatik Spektrum (German Journal for Computer Science). He organized several conferences, was papers co-chair of Eurovis 2004, Eurogrpahics Symposium of Rendering 2005, NPAR 2007, Computational Aesthetics 2009 and papers Co-chair of Eurographics 2011.

Changsheng Xu is a Professor in National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and Executive Director of China-Singapore Institute of Digital Media. His research interests include multimedia content analysis/indexing/retrieval, pattern recognition and computer vision. He has hold 30 granted/pending patents and published over 200 refereed research papers in these areas. Dr. Xu is an Associate Editor of IEEE Trans. on Multimedia, ACM Trans. on Multimedia Computing, Communications and Applications and ACM/Springer Multimedia Systems Journal. He received the Best Associate Editor Award of ACM Trans. on Multimedia Computing, Communications and Applications in 2012 and the Best Editorial Member Award of ACM/Springer Multimedia Systems Journal in 2008. He served as Program Chair of ACM Multimedia 2009. He has served as associate editor, guest editor, general chair, program chair, area/track chair, special session organizer, session chair and TPC member for over 20 IEEE and ACM prestigious multimedia journals, conferences and workshops. He is IEEE Fellow, IAPR Fellow and ACM Distinguished Scientist.

View publication stats