

Adaptive Sparse Coding for Painting Style Analysis

Zhi Gao^{1*}, Mo Shan^{2*}, Loong-Fah Cheong³, and Qingquan Li^{4,5}

¹ Interactive and Digital Media Institute, National University of Singapore, Singapore

² Temasek Laboratories, National University of Singapore

³ Electrical and Computer Engineering Department, National University of Singapore

⁴ The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China

⁵ Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen, China

{gaozhinus, sholmes9091}@gmail.com, eleclf@nus.edu.sg, qqli@whu.edu.cn

Abstract. Inspired by the outstanding performance of sparse coding in applications of image denoising, restoration, classification, etc, we propose an adaptive sparse coding method for painting style analysis that is traditionally carried out by art connoisseurs and experts. Significantly improved over previous sparse coding methods, which heavily rely on the comparison of query paintings, our method is able to determine the authenticity of a single query painting based on estimated decision boundary. Firstly, discriminative patches containing the most representative characteristics of the given authentic samples are extracted via exploiting the statistical information of their representation on the DCT basis. Subsequently, the strategy of adaptive sparsity constraint which assigns higher sparsity weight to the patch with higher discriminative level is enforced to make the dictionary trained on such patches more exclusively adaptive to the authentic samples than via previous sparse coding algorithms. Relying on the learnt dictionary, the query painting can be authenticated if both better denoising performance and higher sparse representation are obtained, otherwise it should be denied. Extensive experiments on impressionist style paintings demonstrate efficiency and effectiveness of our method.

1 Introduction

Painting analysis has been carried out by art connoisseurs and experts traditionally, and the procedure could be fairly costly and subjective. Even worse, the conclusions made in this way could be changed over time due to the emergence of new historical evidence. Therefore, assisting the authentication by less biased automatic method has been attracting increased attention from the communities of art, mathematics and engineering, etc. Although appealing, and much cross-disciplinary interaction of image analysis researcher and art historians has been

* denotes joint first author.

reported, a survey of the representative literature [1, 9, 11, 14, 20, 19, 21] reveals that the research of painting authentication using computer vision techniques is still in its early stage.

Currently, most published research of computerized painting analysis focuses on two tasks: to distinguish authentic painting from its forgery, given a set of authentic samples; to classify the query authentic one for dating challenge or stylometry comparison, given multiple sets of authentic samples. Intuitively, such identification and classification like tasks can be easily solved by exploiting the latest achievements in pattern classification. However, due on one hand to the requirement of sophisticated high-level art knowledge which should be conveyed and applied mathematically, on the other hand to the lack of sufficient well-prepared positive and negative samples¹, a classifier with satisfactory performance still demands extensive efforts.

Inspired by the outstanding performance of sparse coding which has achieved the state-of-the-art performance in a variety of applications including image denoising and restoration, action classification and recognition, face recognition and abnormal (or irregularity) detection, etc, it has been applied for painting analysis as well in recent years [9, 16, 15]. In the widely celebrated paper published on Nature [18], sparse coding is demonstrated with the capability of capturing well the localized orientation and spatial frequency information existed in the training images. Nevertheless, when applied to artistic analysis, sparse coding exploits very little artistic knowledge, perhaps due to the difficulty of incorporating the stylistic information into the standard training procedure. Moreover, such methods make the final decision based on the comparison of pertinent statistic, either among multiple query samples or among multiple sets of given authentic samples, for the authentication and dating challenge tasks respectively. However, it is not always the case to distinguish the authentic one from its imitations, and the more realistic problems is to determine authenticity for one query sample given a set of authentic samples. Consequently, several questions of practical importance could be raised: firstly, is it possible to import any artistic knowledge appropriately into the sparse coding model to capture more characteristics of the painting? Secondly, how to estimate a decision boundary to determine how much a query painting can deviate from the training data before it could be categorized as being too different to be authentic or consistent?

In this paper, we propose an adaptive sparse representation algorithm in which the DCT baseline is leveraged to overcome the aforementioned drawbacks of most available methods in some extent. Firstly, instead of brushstrokes or features, discriminative patches containing the most representative characteristics of the given authentic samples are extracted via exploiting the statistical information of their representation on the DCT basis. Moreover, such patches are sorted according to their distinctiveness. Subsequently, the strategy of adaptive

¹ Due on one hand to the copyright issue, the high quality reproductions of the paintings in the museums are rarely publicly available even for research purpose, on the other hand to the fact that museums usually have no interests to acquire and keep paintings that are known as forgeries.

sparsity constraint which assigns higher sparsity regularization weight to the patch with higher discriminative level is enforced to make the dictionary trained on such patches more exclusively adaptive to the authentic samples than dictionary trained via previous sparse coding algorithms. Relying on the learnt dictionary, the query painting can be authenticated if both better denoising performance and more sparse representation are obtained than the results obtained on the baseline DCT basis, otherwise it should be denied. Herein the DCT basis is chosen to represent the general style to some extent, which is used to set a baseline for comparison purpose. In this way, a decision boundary can be built. Besides the sparsity measure, here we further exploit a well proved conclusion in sparse coding community that the dictionaries learned from the data itself significantly outperform predefined dictionaries, such as DCT, curvelets, wedgelets and various sorts of wavelets in the generative tasks including denoising, restoration, etc, since the learnt dictionary is more suitably adapted to the given data [8, 5]. We apply such conclusion from another direction, namely, if the testing image can be better denoised by using the learnt dictionary than DCT, it can be determined as more consistent with the dictionary (basis) learned from the authentic samples. Extensive experiments including classification of van Gogh’s paintings from different periods, dating challenge and stylometry measure demonstrate the efficiency and effectiveness of our method. In particular, to our knowledge, the experiment of painting classification based on decision boundary is conducted in a computerized artistic analysis fashion for the first time.

The main contributions of our work may be summarized as follows. Firstly, a novel patch extraction method is developed based on the statistics on the DCT basis. Secondly, an improved sparse coding algorithm is proposed for stylistic assessment of the paintings, incorporating the prior statistics of extracted patches. Thirdly, a decision boundary is estimated based on a baseline measure leveraging the DCT basis, enabling the attribution of a single query painting. Last but not the least, insightful observations are drawn for the analysis of the paintings by van Gogh produced in different periods.

2 Related Work

Research efforts based on computational techniques to study art and cultural heritages have emerged in the recent years, interested readers can refer to such surveys [20, 19, 21] for a comprehensive introduction. Here, we focus on the most relevant works. Currently, according to the characteristics extracted from digital paintings and on which the following process performs, the main approaches of computerized painting analysis can be roughly divided into three categories: feature based, brushstroke based and sparse coding based.

In order to perform the high-level analysis on paintings, various methods have been proposed to capture their perceivable characteristics based on the features of color and texture [23], fractal dimension [22], etc. Without attempting to be exhaustive, other features include HOG 2×2 , local binary pattern (LBP), dense SIFT (scale-invariant feature transform) have also been tested [4]. However,

such feature based methods can hardly be generalized to deal with paintings of different artists, since they rely heavily on the prior knowledge of different artists to select proper features to represent their artworks, and the challenge of this regard has been pointed out in [11].

As suggested by art connoisseurs and experts that the pattern of brushwork is an important indicator of styles, methods have been developed tailored to the brushstroke analysis. The work [14] published on PAMI applied the brushwork traits to authenticate and date the paintings of van Gogh. However, brushstrokes are too difficult to be crisply segmented out using off-the-shelf detector. Several brushstroke extraction methods have been proposed [2, 14], but codes of such methods are not publicly available. Based on brushstroke detection, a variety of mathematic tools, such as neural network [17], hidden Markov Models (HMM) [10], multiresolution HMM [13], have been applied to represent and model the information of brushstroke width, length, curvature, shapes, etc. Subsequently, SVM classifier is trained on the given data to perform a specific task. In summary, researchers in this study identified several challenges deserving more efforts. Firstly, it is very difficult to automatically extract the strokes and art theory required to be exploited. Secondly, there is an urgent demand for more sophisticated mathematical models which are capable of capturing the subtle visual characteristics in paintings. Thirdly, it is difficult to train a satisfactory classifier without sufficient well-prepared positive and negative samples.

Motivated by its outstanding performance in various applications in computer vision and signal processing, sparse coding has been applied to the recognition of authentic paintings as well. In 2010, the work [9] applied the sparse coding model for the first time on quantification of artistic style to determine authentication of drawings of Pieter Bruegel the Elder and received a lot of attention in the media. Therein, the dictionary is firstly learned based on the given authentic paintings according to the sparse coding model, and the authentication is performed on the sparsity measure of representations on the learnt dictionary. The one whose representation coefficients are more sparse will be determined as authentic. In [15], sparse coding is applied to perform stylometry measure on van Gogh’s paintings of different periods. Instead of measuring the sparsity of the representation coefficient, the similarity measure is performed on the dictionary itself. The strength of the sparse coding algorithm is that it can capture the visual features in the painting through the overlapping patches extracted. However, the standard dictionary training adopted previously does not integrate the discriminative features that are indicative of styles, such as brushstrokes, possibly because the extraction of brushstrokes and the modeling involves extensive domain knowledge.

3 Our Adaptive Sparse Coding Algorithm

3.1 Notation

To facilitate the description, we first introduce the notation adopted in this paper. Matrices and vectors are in bold capital, bold lower-cased fonts respectively.

We define for $q \geq 1$, the ℓ_q -norm of a vector \mathbf{x} in \mathbb{R}^m as $\|\mathbf{x}\|_q \triangleq (\sum_{i=1}^m |\mathbf{x}[i]|^q)^{1/q}$, where $\mathbf{x}[i]$ denotes the i -th entry of \mathbf{x} . The ℓ_0 -pseudo-norm is defined as the number of nonzero elements in a vector. We consider the Frobenius norm of a matrix \mathbf{X} in $\mathbb{R}^{m \times n}$: $\|\mathbf{X}\|_F \triangleq (\sum_{i=1}^m \sum_{j=1}^n |\mathbf{X}[i, j]|^2)^{1/2}$. \mathbf{x}_i or $\mathbf{X}(:, i)$ represents the i -th column of \mathbf{X} , $\mathbf{X}(j, :)$ is the j -th row of \mathbf{X} .

3.2 Previous Sparse Coding Algorithm for Painting Analysis

Mathematically, the basic sparse coding algorithm amounts to solve the following optimization problem (as Eq.(1)) or its tightest convex surrogate (as Eq.(2)):

$$\min_{\mathbf{D}, \mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2 \quad \text{s.t.} \quad \forall 1 \leq i \leq N, \|\mathbf{x}_i\|_0 \leq k \quad (1)$$

$$\min_{\mathbf{D}, \mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2 + \lambda \|\mathbf{X}\|_1 \quad (2)$$

where $\mathbf{Y} \in \mathbb{R}^{d \times N}$ contains all the available image patches $\{\mathbf{y}\}_{i=1}^N$ as columns, and similarly, $\mathbf{X} \in \mathbb{R}^{n \times N}$ contains all the sparse representation vectors $\{\mathbf{x}\}_{i=1}^N$. Sparse representation aims to find both the dictionary $\mathbf{D} \in \mathbb{R}^{d \times n}$ and the representations \mathbf{X} . In Eq.(1), k is a given constant controlling the sparsity, usually $k \ll d$ to keep \mathbf{X} sparse and $d < n$ to ensure the over-completeness of \mathbf{D} . In Eq.(2), λ is the regularization parameter to balance the ℓ_2 representation fidelity term and the ℓ_1 sparsity penalty term. In fact, extensive efforts have been devoted to the algorithms to find the global optimal solution of \mathbf{D} and \mathbf{X} efficiently, and lots of methods have been proposed. For more knowledge in this regard, readers are advised to refer to such monograph [7] and references therein.

The typical framework of using sparse coding for artistic authentication can be summarized briefly as this: \mathbf{D}_{tr} is firstly trained on the given authentic paintings, then the representation matrices of two testing samples, the authentic one and a forgery, are computed based on \mathbf{D}_{tr} . As [9, 16] did, kurtosis is applied to measure the sparseness of matrix. The one whose representation matrix is more sparse, namely the kurtosis value is larger, is determined as authentic. Obviously, such methods rely on the comparison of kurtosis of multiple query paintings without taking their difference of content complexity into account. The limitation of [9] will be demonstrated in the experiment part, Section 4.1.

3.3 Discriminative Patches Extraction

To facilitate more perceivable characteristics of the authentic paintings being learned by sparse coding model, here we propose to extract discriminative patches according to the statistics of their representation on the DCT basis.

Given a painting, we firstly estimate its representation coefficient matrix on the predefined DCT basis by replacing \mathbf{D} in Eq.(2) with \mathbf{D}_{dct} , as below:

$$\min_{\mathbf{X}_{dct}} \|\mathbf{Y} - \mathbf{D}_{dct}\mathbf{X}_{dct}\|_2 + \lambda \|\mathbf{X}_{dct}\|_1 \quad (3)$$

where the image patches $\{\mathbf{y}\}_{i=1}^N$ are of size 12×12 pixels, extracted from the whole image with half size sliding; the number of DCT atoms is 512, namely \mathbf{D}_{dct} is of dimensions 144×512 ; and the recommended value of λ is 0.6. As \mathbf{D}_{dct} is predefined and known, we apply the ℓ_1 -ls algorithm [12] to estimate \mathbf{X}_{dct} . Here, we clarify that the default patch size is 12×12 pixels extracted with half size sliding, the default atom number of dictionary, either predefined DCT or learned on training data, is 512. Such default parameters are used in this paper, except where indicated.

As well known that the DCT basis is usually used for image compression, an image patch with more complex content requires more DCT basis functions to represent, and vice versa. To measure the content complexity of an image patch of size $n \times n$ pixels, the activity measure is proposed in [7], a higher value indicating the presence of more complex content, as below:

$$Activity(I) = \sum_{i=2}^n \sum_{j=1}^n |I[i, j] - I[i-1, j]| + \sum_{i=1}^n \sum_{j=2}^n |I[i, j] - I[i, j-1]| \quad (4)$$

It is worth mentioning that the outline in the painting is mainly decided by the object itself, and includes little characteristics of the artist. Therefore we eliminate such pixels first based on the segmentation results of K-means segmentation. In our work, the DCT basis (or say dictionary) which is popularly applied as initialization for dictionary learning in sparse coding is applied as a baseline to represent a general painting style. Therefore, patches which do not conform to such criteria are defined as discriminative, which contain representative characteristics of specific artist. Mathematically, discriminative patches are those patches with lower kurtosis (namely, more atoms used for representation) and relatively lower activity (namely, content is simpler). Instead of using the values directly, we perform discriminative patches extraction based on the sorting of values of both kurtosis and activity, as summarized in Algorithm 1.

In Fig.1, the result of discriminative patches extraction is shown. Moreover, the intermediate result of outline detection of K-means segmentation is also included. Through careful observation, we find that such extracted patches cover most of the area with obvious brushstrokes, especially, areas with impasto are extracted successfully, which are supposed to be the most representative characteristics of the artist [25]. In contrast to those brushstroke detection methods which are limited for oil painting, our discriminative patch extraction method is more general, and can be easily applied on other kinds of artworks. In fact, we have tried other bases including wavelet and coutourlet for this part, and find that the DCT basis works well to meet our requirement.

3.4 Adaptive Sparse Coding Algorithm

With discriminative patches P_i^d ($i = 1, \dots, K$) in hand, whose discriminative level (represented as dl_i) is the descending order, we propose the following adap-

Algorithm 1 Discriminative Patches Extraction.

Input: A painting image IM ;
Output: Discriminative patches: P_i^d , $i = 1, \dots, K$;

- 1: Get patches of size 12×12 pixels from IM with half-size sliding: P_i , $i = 1, \dots, N$;
- 2: Get the boundary map IM_e by performing the K-means segmentation on IM ;
- 3: Based on IM_e , obtain patches without boundary pixels: P_i , $i = 1, \dots, M$;
- 4: Set $K = M \times 0.2$; //20% of the patches will be extracted.
- 5: Using Eq.(3), compute coefficient vector \mathbf{x}_i for patch P_i , $i = 1, \dots, M$, compute the kurtosis for each coefficient vector and sort in ascending order, get kurArray ;
- 6: Compute activities for the patches P_i , $i = 1, \dots, M$, and sort the values in ascending order, get actArray ;
- 7: Set $nCounter = 0$; //to count how many discriminative patches have been found.
- 8: **for** $j=1$; $j \leq 20$; $j++$
- 9: **for** $k=1$; $k \leq j$; $k++$
- 10: **if** $nCounter < K$
- 11: Find patches whose kurtosis value is the beginning $j \times 5\%$ in the kurArray ;
- 12: Find patches whose activity value is the beginning $k \times 5\%$ in the actArray ;
- 13: Find intersection of patches found in step 11, 12, save to the discriminative patches, P_i^d , $i = 1, \dots, p$;
- 14: $nCounter = nCounter + p$;
- 15: **end for**
- 16: **end for**
- 17: In the discriminative patches list, the first K patches are exported.

tive sparse coding model, as below:

$$\min_{\mathbf{D}, \mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_1 + \sum_{i=1}^K \lambda_i \|\mathbf{x}_i\|_1, \quad \lambda_i = \lambda_{min} + \frac{dl_i - dl_{min}}{dl_{max} - dl_{min}} (\lambda_{max} - \lambda_{min}) \quad (5)$$

where the adaptive regularization parameter λ_i , instead of a constant, is applied to balance the representation fidelity term and the sparsity penalty term, and $\lambda_{max} = 2.5$, $\lambda_{min} = 0.5$. Obviously, patch with higher dl value will be assigned with larger λ value. In other words, if the patch with more discriminative information is enforced to be represented with less dictionary atoms. Crafted in this



Fig. 1. Result of discriminative patches extraction. Left: edge detected by K-means. Right: extracted discriminative patches.



Fig. 2. Authentication on query samples from van Gogh (left) and Monet (right).

way, the dictionary can better represent the stylistic features. Compared with Eq.(2), another significant difference is that the ℓ_1 -norm based representation fidelity is applied, since it is less vulnerable to outliers, as proved in [24, 6]. In this way, our learnt dictionary is robust to the outliers generated from the flaws of previous operations.

To tackle the optimization problem of Eq.(5), we separate it into two sub-problems, sparse coding for training samples and dictionary updating. Briefly, the objective function of each sub-problem is defined as below:

$$\min_{\mathbf{x}_i} \sum_{i=1}^K \|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_1 + \lambda_i \|\mathbf{x}_i\|_1, \quad i = 1, \dots, K \quad (6)$$

Eq.(6) is to estimate the sparse coefficients of each training patch, when \mathbf{D} is given. We can rewrite Eq.(6) into an equivalent ℓ_1 approximation problem:

$$\min_{\mathbf{x}_i} \sum_{i=1}^K \left\| \begin{pmatrix} \mathbf{y}_i \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \lambda_i \mathbf{I} \end{pmatrix} \mathbf{x}_i \right\|_1, \quad i = 1, \dots, K \quad (7)$$

With \mathbf{X} fixed, the dictionary \mathbf{D} can be updated as follows:

$$\min_{\mathbf{D}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_1 \quad (8)$$

All the sub-problems in Eq.(7) and (8) are standard ℓ_1 -regression problems. We resort to the iterative reweighted least squares (IRLS) [3] for solutions, and the algorithm is summarized as below:

In Algorithm 2, δ is a small positive value ($\delta = 0.001$ in our experiments). Each iteration of IRLS involves minimizing a quadratic objective function. The global optimum can be reached by taking derivatives and setting them to zeros.

An intuitive yet important question on adaptive sparse coding model is that why not apply much larger λ value to make the dictionary more adaptive to the patches. However, this is not the case. If λ is out of certain range, although the solution of \mathbf{D} , \mathbf{X} still can be obtained, but results in a poor representation fidelity measure. This observation again justifies that besides kurtosis, the representation fidelity measure must be taken into account to evaluate how well a dictionary is learned, or how well the signal is consistent with the dictionary. Here, we further exploit a well proved conclusion in sparse coding that the dictionary learned from the data itself is more adapted to the data, resulting in significantly better denoising performance than applying predefined dictionaries [8], therefore, the denoising performance, namely the PSNR (peak signal-noise ratio) value is also incorporated to indicate the reconstructive power of the dictionary. In summary, on the same testing sample, if the learnt dictionary can achieve higher PSNR and higher kurtosis than that of applying DCT, the testing is deemed as much consistent with the dictionary. Here, the importance of estimating a reasonable range for the dynamic λ must be highlighted. The value of $\lambda_{max} = 2.5$, $\lambda_{min} = 0.5$ applied in our work is based on extensive experiments.

Algorithm 2 Adaptive Sparse Coding Algorithm.

Input:Image patches matrix $\mathbf{Y} \in \mathbb{R}^{d \times K}$, λ_{max} , λ_{min} , and the atom numbers of \mathbf{D} is n ;**Output:**Solution $\mathbf{D} \in \mathbb{R}^{d \times n}$, $\mathbf{X} \in \mathbb{R}^{n \times K}$ of Eq.(5);

- 1: Initialization: set \mathbf{D} as overcomplete DCT dictionary;
 - 2: **while** not converged **do**
 - 3: //Lines 4 is the sparse coding stage to solve Eq. 7.
 - 4: apply the IRLS method to solve \mathbf{X} :
the key objective function for IRLS is:

$$\min_{\mathbf{x}_i} \sum_{j=1}^n \omega_i^j (\mathbf{y}_i[j] - \mathbf{D}(j, :)\mathbf{x}_i)^2 + \sum_{j=1}^n \eta^j (\lambda_i \mathbf{x}_i[j])^2;$$

$$\omega_i^j = \frac{1}{\sqrt{(\mathbf{y}_i[j] - \mathbf{D}(j, :)\mathbf{x}_i)^2 + \delta}}, \eta^j = \frac{1}{\sqrt{(0 - \mathbf{I}(j, :)\mathbf{x}_i)^2 + \delta}}$$
 - 5: //Lines 6 is the dictionary update stage to solve Eq. 8.
 - 6: apply the IRLS method to update \mathbf{D} :

$$\min_{\mathbf{D}(i, :)} \left\| (\mathbf{Y}(i, :))^T - (\mathbf{D}(i, :)\mathbf{X})^T \right\|_1, i = 1, \dots, d;$$
the key objective function for IRLS is:

$$\min_{\mathbf{D}(i, :)} \sum_{j=1}^N \phi_i^j (\mathbf{Y}[i, j] - \mathbf{D}(i, :)\mathbf{x}_j)^2;$$

$$\phi_i^j = \frac{1}{\sqrt{(\mathbf{Y}[i, j] - \mathbf{D}(i, :)\mathbf{x}_j)^2 + \delta}}$$
normalize each column of \mathbf{D} after all rows are updated;
 - 7: **end while**
 - 8: Output \mathbf{D} , \mathbf{X} .
-

4 Experiments And Analysis

4.1 Authentication Experiment

The direct comparison strategy of [9] is questionable, since the decision relying on the single statistic of sparse representation could be unfair (or biased). The bias arises because the difference between query samples in terms of the content complexity has not been taken into account.

To demonstrate the effect of content complexity, let us take an authentication example shown in Fig.2, in which the given set of authentic paintings are from van Gogh, and the objective is to determine the authenticity of two query paintings, from van Gogh and Monet respectively. A dictionary is trained from van Gogh's painting, and then the kurtosis values of the sparse coefficients of the two query paintings associated with the van Gogh dictionary are computed. As a result, the kurtosis of van Gogh's painting is 43.00, whereas that of the Monet's painting is 61.24. This clearly contradicts the assumption drawn in [9] that the painting from the same style has a higher kurtosis.

Previous method fails in this case where the content of the authentic van Gogh's painting is more complex than Monet's, thus resulting in dense coefficients, which will lead to the wrong conclusion that Monet's painting is authentic to a van Gogh's dictionary.

In our method, we estimate the decision boundary of a single query sample via leveraging the DCT basis as baseline, where the margin is obtained by the

original kurtosis subtracting off that associated with the DCT basis. The motivation behind the introduction of a baseline is because the DCT basis, being unbiased and often used for dictionary initialization, can represent the general style. It then follows that a decision boundary can be estimated by comparing the sparse representation of a test painting to this baseline. Instead of using the original kurtosis values, the margin is perceived as a more suitable measure. Consequently, for the same paintings used previously, the margin for van Gogh’s painting is 29.56, while that for Monet’s is 29.72. The effect of content complexity is thus mitigated to certain extent.

Besides kurtosis, which shows the representativeness of the dictionary, the denoising performance is another important measure indicative of the reconstructive power of the dictionary. Here the denoising performance is also compared against the DCT basis. Using the dictionary trained from the same space, the PSNR margin for van Gogh’s painting is 1.46 dB compared with -1.72 dB for Monet’s. The positive PSNR margin for van Gogh’s painting demonstrates that the dictionary trained from the same image space is indeed better than DCT basis in terms of the reconstruction ability. On the contrary, DCT basis denoises Monet’s painting better than the dictionary trained from van Gogh.

In summary, the lack of a decision boundary in the previous methods can be addressed by introducing DCT basis as a baseline, and then the sparse representation and denoising performance can be compared against it. Sparser representation and better denoising performance indicate that the query painting is authentic, belonging to the same image space with the dictionary.

4.2 Style Diversity Experiment

To analyze the style diversity in the context of decision boundary, we follow the experimental settings in [15]. The objective of this experiment is to determine the similarity of van Gogh’s paintings in different periods.

As mentioned before, the collection of data set for painting analysis is quite challenging, since the high quality reproductions owned by the museums are rarely publicly available. Hence we employ an alternative way to acquire the large amount of paintings needed for the experiments, by using a python script to collect paintings from the web, following the same strategy of [4]. The paintings collected are mainly from WIKIART, which provides artworks in public domain.

The given training set comprises of paintings by van Gogh produced in four periods, namely Paris Period, Arles Period, Saint-Remy Period and Auvers-sur-Oise Period. 20 paintings are chosen from each period, thus totally 80. Among these paintings, 5 in each period are randomly selected to form a group of 20 paintings to represent the overall style of van Gogh, which is used for training the dictionary. 20 paintings by Monet are also included as a group of outliers.

The sparsity and denoising performance are plotted in Fig.3. The horizontal axis shows the PSNR margin between the van Gogh dictionary and DCT basis, whereas the vertical axis reveals the kurtosis margin. These values are normalized such that their magnitudes are smaller than 1. The hollow markers represent the original data and the filled markers indicate the position of the mean values.

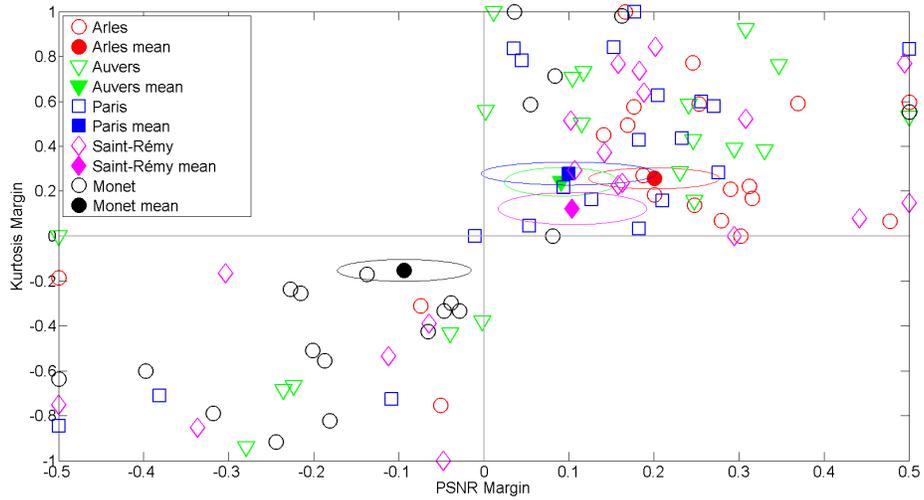


Fig. 3. Style similarity for van Goghs painting from four periods and Monets painting. Hollow marker is for each painting, filled marker is for the mean values of each group.

The plot demonstrates that as outliers, most of Monet’s paintings lie in the third quadrant and have negative kurtosis and PSNR margin on average, suggesting that they are from a different image space. For the majority of van Gogh’s paintings located in the first quadrant, those from Paris Period have a higher mean kurtosis, which indicates that this period is more similar to overall van Gogh style than other three periods, meaning that van Gogh’s paintings can be represented better by the Paris Period. These observations are consistent with the conclusion drawn in [15].

In addition to the similarity analysis in [15], we also measure the spread of the paintings in each period. This is achieved by computing the variation of the kurtosis margin and PSNR margin, which are set to be the minor and major axis of the ellipses centered at the mean. It could be observed that the Paris Period has the longest major axis, revealing that the stylistic variation of this period is the greatest among the four periods. This observation agrees with the fact that the painter constantly varies his style and technique during his stay in Paris.

To summarize, the Paris Period is the most similar one to van Gogh’s overall style. The consistency of our finding with the previous paper demonstrates the efficiency of our method. Furthermore, our algorithm finds that his painting techniques vary most frequently during Paris Period as well, which shed some new lights on style diversity analysis. The experiment also shows that the DCT basis is capable of providing a decision boundary for authentication, and successfully identifying the outliers.

4.3 Classification of van Gogh’s Painting from Four Periods

To illustrate the usefulness and accuracy of the painting authentication based on a decision boundary, a novel experiment is designed to classify van Gogh’s paintings produced in the four periods. The dataset is the same with the experiment in Section 4.2, except that only 10 paintings are chosen from each period to reduce computational time. The details of these paintings are listed in the supplementary material.

The experimental procedure is as follows. Firstly, a dictionary is trained from the 10 paintings randomly selected from each of the four periods. Secondly, the kurtosis margin and the PSNR margin of each painting associated with the four dictionaries are computed and then plotted respectively. Under the scenario where all the paintings are painted by the same painter, the ‘authentic’ paintings are defined as those produced in the same period with the dictionary. For instance, the paintings in the Arles Period are ‘authentic’ to the dictionary of this period, and the paintings in other three periods are regarded as ‘forgeries’. Having a sparser representation and better denoising performance by the trained dictionary in a specific period than the DCT, the paintings located in the first quadrant are deemed to be from the same image space of that dictionary and are authentic. Therefore, the first quadrant is the decision boundary to determine authenticity for each query painting.

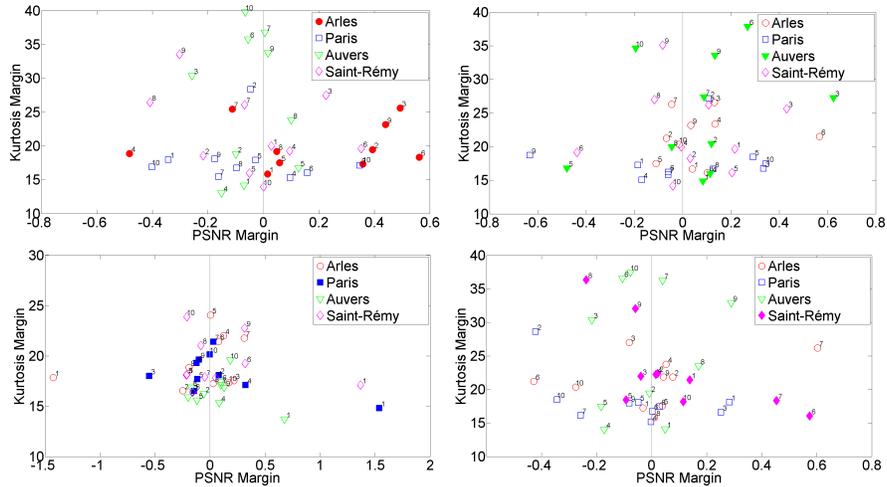


Fig. 4. Classification plots of the four periods. Filled markers: paintings in the same period with the dictionary used for testing. Hollow markers: paintings in other periods.

The classification results are plotted in Fig.4. The results demonstrate that our method classifies the paintings with acceptable accuracy. For instance, most of the paintings in the Arles Period, Saint-Remy Period and Auvers-sur-Oise

Period are correctly identified. This has corroborated the usefulness of the decision boundary since most of the authentic paintings indeed lie in the first quadrant. Nevertheless, there are some paintings from other periods misclassified as authentic. This may be explained by the high similarity of these paintings observed from the similarity experiment. After all they are all painted by the same artist.

As for Paris Period, only 5 authentic paintings are correctly identified. However, the low accuracy does not necessarily undermine the effectiveness of the decision boundary. From the experiment in Section 4.2, it is observed that the paintings in the Paris Period vary most significantly in styles. Hence it is not surprising to find that the dictionary trained from these paintings cannot represent each painting well. From this perspective, this result is consistent with our findings in the previous experiment.

To sum up, the decision boundary built upon kurtosis and PSNR margin is able to classify paintings with high accuracy even in the extremely challenging case where both the positive and negative samples are paintings of van Gogh. Furthermore, the classification accuracy can be increased by selecting the paintings with high stylistic consistency.

4.4 Dating Challenge

While the paintings used in the experiment in Section 4.3 can be easily dated by the connoisseurs as to which period they belong, the attribution for other paintings by van Gogh are more ambiguous. To address the real dating problems raised by experts, which has also been explored extensively in [14, 1], we examine three paintings which seem to share the traits of varied periods.

The paintings under examination are Still Life: Potatoes in a Yellow Dish (f386), Willows at Sunset (f572), and Crab on its Back (f605), as shown in Fig.5. The decisions for these paintings are still under debate among connoisseurs because some insist that they belong to Paris Period while others argue that they belong to Provence Period, corresponding to the Arles and Saint-Remy Period in the similarity and period classification experiments.

For the ground truth, two groups of artworks, each containing eight paintings produced in Paris Period and Provence Period respectively, are selected and we train two dictionaries from each group. Then the kurtosis margin and PSNR margin associated with the dictionaries are plotted. From Fig.5, it is clear that f386 should be dated to the Provence Period since it is located in the first quadrant when using the Provence dictionary. Similarly, f572 is dated to the Paris Period. One may wonder how to date f605 since both locations are in the second quadrant. In fact, this painting can be dated to the Provence Period because both kurtosis and PSNR margin are higher for the Provence dictionary. Moreover, in the previous works it is highlighted that compared with f386, f605 is less similar to the paintings in the Provence Period. This is also reflected in our results since the PSNR margin for f605 is negative when using the Provence dictionary, meaning that the reconstructive power of the Provence dictionary is less significant for f605 in comparison to f386. In short, the effectiveness of

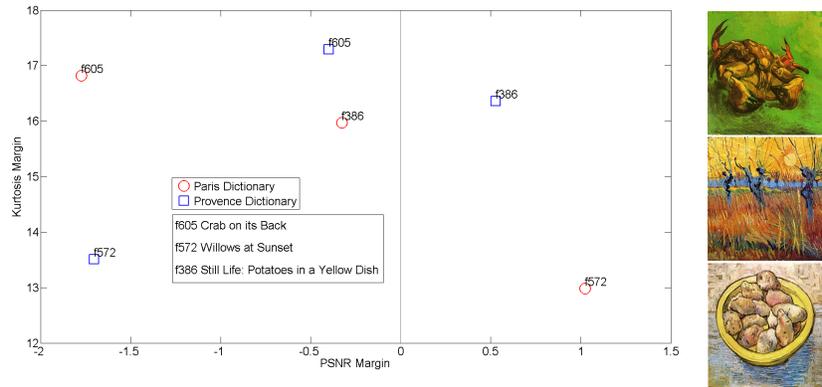


Fig. 5. Dating results (left) of three test paintings (right).

the painting classification based on decision boundary can also be demonstrated when applied to address the dating question. Which period a painting belongs to not only depends on the quadrant but also relates to the comparison of the margins. Although the correct attribution of the three paintings investigated is controversial in the art historical literature, the conclusion reached in our experiment is consistent with the examinations conducted by other state-of-the-art computerized painting analysis methods. Hence it may be concluded that the performance of our method is also superior.

5 Conclusion

In this paper we have developed a groundbreaking method for painting style analysis. Leveraging the proposed DCT baseline, both the discriminative patches extraction in the early stage and the decision making in the final stage are performed efficiently and successfully. In between, an adaptive sparse coding algorithm is proposed to learn the perceivable characteristics of the artist, provided a set of authentic data is given. A promising aspect of our method is that the authenticity of a single query painting can be determined based on a decision boundary. With this procedure we have performed extensive authentication and classification experiments where consistent conclusions are drawn. We have also shown that our novel approach of authentication based on a decision boundary sheds some new lights on style diversity analysis. Furthermore, our method can classify paintings with high stylistic similarity fairly accurately. We believe that with the advances in vision and machine learning, more art related questions can be addressed from a computational point of view.

Acknowledgement. This work is supported by these Grants: theory and methods of digital conservation for cultural heritage (2012CB725300), PSF Grant 1321202075, and the Singapore NRF under its IRC@SG Funding Initiative and administered by the IDMPO at the SeSaMe centre.

References

1. Abry, P., Wendt, H., Jaffard, S.: When Van Gogh meets Mandelbrot: Multifractal classification of painting's texture. *Signal Processing* **93** (2013) 554–572
2. Berezhnoy, I.E., Postma, E.O., van den Herik, H.J.: Automatic extraction of brush-stroke orientation from paintings. *Machine Vision and Applications* **20** (2009) 1–9
3. Bissantz, N., Dmbgen, L., Munk, A., Stratmann, B.: Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces. *SIAM Journal on Optimization* **19** (2009): 1828–1845
4. Blessing, A., Wen, K.: Using machine learning for identification of art paintings. Technical report, Stanford University, Stanford (2010)
5. Castrodad, A., Sapiro, G.: Sparse modeling of human actions from motion imagery. *IJCV* **100** (2012) 1–15
6. Cong, Z., Xiaogang, W., Wai-Kuen, C.: Background subtraction via robust dictionary learning. *EURASIP Journal on Image and Video Processing* **2011** (2011)
7. Elad, M.: *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer (2010)
8. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. on Image Processing* **15** (2006) 3736–3745
9. Hughes, J.M., Graham, D.J., Rockmore, D.N.: Quantification of artistic style through sparse coding analysis in the drawings of pieter bruegel the elder. In: *Proc. of the National Academy of Sciences* **107** (2010) 1279–1283
10. Jacobsen, R.: *Digital Painting Analysis: Authentication and Artistic Style from Digital Reproductions*. PhD thesis, Aalborg University, Aalborg (2012)
11. Johnson, C.R., Hendriks, E., Berezhnoy, I.J., Brevdo, E., Hughes, S.M., Daubechies, I., Li, J., Postma, E., Wang, J.Z.: Image processing for artist identification. *IEEE Signal Processing Magazine* **25** (2008) 37–48
12. Koh, K., Kim, S.J., Boyd, S.P.: An interior-point method for large-scale l1-regularized logistic regression. *Journal of Machine Learning Research* **8** (2007) 1519–1555
13. Li, J., Wang, J.Z.: Studying digital imagery of ancient paintings by mixtures of stochastic models. *IEEE Trans. on Image Processing* **13** (2004) 340–353
14. Li, J., Yao, L., Hendriks, E., Wang, J.Z.: Rhythmic brushstrokes distinguish van gogh from his contemporaries: findings via automated brushstroke extraction. *IEEE Trans. on PAMI* **34** (2012) 1159–1176
15. Liu, Y., Pu, Y., Xu, D.: Computer analysis for visual art style. *SIGGRAPH Asia 2013 Technical Briefs*, ACM (2013) 9
16. Mairal, J., Bach, F., Ponce, J.: Task-driven dictionary learning. *IEEE Trans. on PAMI* **34** (2012) 791–804
17. Melzer, T., Kammerer, P., Zolda, E.: Stroke detection of brush strokes in portrait miniatures using a semi-parametric and a model based approach. *International Conference on Pattern Recognition* **1** (1998) 474–476
18. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381** (1996) 607–609
19. Stork, D.G.: *Computer vision and computer graphics analysis of paintings and drawings: An introduction to the literature*. *Computer Analysis of Images and Patterns*, Springer (2009) 9–24
20. Stork, D.G., Coddington, J.: Computer image analysis in the study of art. In: *Proc. of SPIE* **6810** (2008)

21. Stork, D.G., Coddington, J., Bentkowska-Kafel, A.: Computer vision and image analysis of art II. In: Proc. of SPIE **7869** (2011)
22. Taylor, R.P., Micolich, A.P., Jonas, D.: Fractal analysis of pollock's drip paintings. *Nature* **399** (1999) 422–422
23. van den Herik, H.J., Postma, E.O.: Discovering the visual signature of painters. *Future Directions for Intelligent Systems and Information Sciences*. Springer (2000) 129–147
24. Wagner, A., Wright, J., Ganesh, A., Zhou, Z., Mobahi, H., Ma, Y.: Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *IEEE Trans. on PAMI* **34** (2012) 372–386
25. Yelizaveta, M., Tat-Seng, C., Ramesh, J.: Semi-supervised annotation of brushwork in paintings domain using serial combinations of multiple experts. In: Proc. of the 14th annual ACM international conference on Multimedia, ACM (2006) 529–538