Discovering Color Styles from Fine Art Images of Impressionism

Man-Kwan Shan mkshan@cs.nccu.edu.tw
Department of Computer Science
National Chengchi University
Taipei, 116, Taiwan

Abstract

Content-based image retrieval (CBIR) has attracted much interest since the last decade. Finding painting styles from fine art images is useful for CBIR. However, little research has been done on the painting style mining. In this paper, we investigated the color style mining technique for fine art of Impressionism. Three design issues for the color style mining are the feature extraction, the feature representation, and the style mining algorithm. For the feature extraction and presentation, dominate colors, adjacent color combinations and some MPEG-7 color descriptors, are utilized to represent the color features. Above all, we utilize the spatial data structure, 2D string, to represent color layout descriptor. For the style mining algorithms, we proposed a two-stage color style mining scheme. The first stage discovers the common properties of paintings of the same style. The second stage discovers the discriminative properties among styles. The experiment on the art work of European Impressionist was conducted. The performance of effectiveness is measured by the classification accuracy of the proposed style mining scheme. The classification accuracy ranges from 70% to 90%.

Keywords: Image mining, Painting Style, Associative Classification, Spatial Co-orientation Patterns.

1. INTRODUCTION

Content-based image retrieval (CBIR) has attracted more and more attention, as the development of multimedia technology. In recent years, several studies have been conducted on image data mining techniques which are useful for CBIR, such as indoor/outdoor[20, 21], city/ landscapes [20] and medical image classification [23]. However, little work has been done on the painting style mining on fine art images [19].

A painting style relates to the painting techniques which the artist uses to create the painting. On the other hand, the style may refer to the human perception. The periods and styles of 19th century western painting consist of Romanticism, Naturalism, Realism and Impressionism. The main characteristic of impressionist style is the concentration on the general impression produced by an object or scene, the use of small touches of pure color, rather than broader strokes, and painting out of doors to represent color and light. In this work, we concentrate on the impressionist style.

Generally speaking, there are two types of descriptions for the painting styles by art critics. One is the colors used frequently by the artist. For example, “the shimmering color and flickering light” for Pierre Auguste Renoir and “lush and brilliant pure colors” for Paul Gauguin, using bold, unrealistic colors and large flat areas. The other is the description of human perception. For instance, “joyful, vivid and spontaneous scenes” for Pierre Auguste Renoir and “happy and filled with nature” for Claude Monet. Because Pierre Auguste Renoir often used large reds and oranges with thick brush strokes on his works and Claude Monet used white lead, cadmium yellow, vermilion, madder, cobalt blue, chrome green.

Can we discover the feelings from the paintings? A number of studies in the field of cognitive psychology and industrial design have been done on the human perception of colors. For example, reddish orange stands for warmth while green stands for peace and relaxation. Moreover, adjacent color combinations, area of colors, and thickness and slope of the line also affect the feelings of humans. A popular example is the color combination of red and green which is related to Christmas in Western culture.

Consequently, both two types of painting style descriptions can be analyzed by low-level image features such as color and spatial relation. The objective of this research is to investigate the data mining techniques to find out the color styles of impressionists and to represent the style in a quantitative way. The feature extraction and representation for color style mining are investigated. The style mining algorithms which discover both the common characteristics and discriminative characteristics are presented.

This paper is organized as follows. Section 2 gives a brief review of previous work related to painting style mining. In section 3, we present the proposed painting style mining techniques. The effectiveness of the proposed techniques is analyzed in Section 4. Section 5 concludes this paper.

2. RELATED WORK

Little work has been done on the painting style mining. Sablatnig et al. proposed a hierarchically structured classification scheme according to stroke for artist’s painting style classification [19]. The hierarchical classification includes grouping portrait miniatures by mean RGB value of the image, face shape classification and stroke classification. The stroke was classified by grouping similar curvature and orientation. More recently, artistic concepts like art period, artist name and style were investigated. Marchenko et al. [13] took color usages as cues to analyze paintings. Gunsel et al. [6] extracted statistic features and performed classification via SVM for art movements. By eight given brushwork classes at block level and manually constructed decision hierarchy, Leslie et al. [9] employed transductive inference of concepts to annotate paintings. In [10], we investigated the approach to explore the affective space for Impressionism paintings. A new meta-level feature, color harmony, which encodes affective information, was proposed. Moreover, to discover the correlation between emotions and painting features, multiple-type latent semantic analysis is utilized to capture these underlying interrelated correlations [10].

3. COLOR STYLE MINING

A painting style refers to the common properties of an artist’s works. On the other hand, it implies the artist’s characteristics which are different from others’. Consequently, the proposed painting style mining consists of the following three steps:
(1) feature extraction and representation.
(2) frequent pattern mining for finding out the common properties.
(3) painting style classification for discovering the discriminative characteristics.

3.1 Feature Extraction and Representation

The common features used in CBIR include color, shape, texture and spatial relationship [8]. In this work, the *dominant color* feature and the *adjacent color combination* features are extracted as the low-level image feature for color style mining. Moreover, MPEG-7 specifies a standard set of descriptors to describe the multimedia contents. Consequently, in addition to the features of dominant colors and adjacent color combination extracted from an image, MPEG-7 descriptors are utilized for the color style mining in the proposed approaches. MPEG-7 standard provides a set of standardized tools for multimedia content description [3, 15]. Tools for feature extraction and multimedia search using various algorithms are included in MPEG-7 eXperimental Model (XM) [14]. MPEG-7 visual part includes the basic structure and descriptors which cover the basic visual features: color, texture, shape, localization, etc. In this work, the following color descriptors are utilized: the scalable color descriptor, the color structure descriptor, and the color layout descriptor. The algorithms described in MPEG-7 XM are adopted to extract these three color descriptor.

One of the fundamental issues for the color feature extraction is the color space model. We choose *HSV* (Hue, Saturation, Value) color space because it corresponds to human vision. Hue is the color type, such as yellow, blue. Saturation is the vibrancy of the color. In other words, the amount of white was mixed into the color. Lower saturation indicates decreasing the contrast of the color. Value is brightness of the color. In this work, a color is represented by (H, S, V), where H is hue, $0 \le H < 360$, S is saturation and V is value, $0 \le S$, $V \le 1$.

Moreover, *LSLM* (Luminance, Red-Green Channel, Yellow-Blue Channel) color space is also considered. This comes from the fact that impressionists preferred to use opponent color to represent the variations of colors under lights. For example, in Claude Monet's Impression Sunrise, the boat's shadow under the orange sunrise has some strokes of green painted into it to increase its vitality. *LSLM* color space is a linear transformation of RGB based on the opponent signals of the cones:

$$\begin{cases} L = 0.209(R-0.5)+0.715(G-0.5)+0.076(B-0.5) \\ S = 0.209(R-0.5)+0.715(G-0.5)-0.924(B-0.5) \\ LM = 3.148(R-0.5)-2.799(G-0.5)-0.349(B-0.5) \end{cases}$$

We represent a color by (L, S, LM), where L is luminance, $0 \le L \le 1$, S is red-green channel and LM is yellow-blue channel, $0 \le S$, $LM \le 1$.

Another important issue is the color depth (the number of colors). More number of colors brings higher color precision. On the other hand, precise colors are too sensitive to discover color styles. Consequently, color quantization is performed on each image. Uniform quantization is performed for H, S and V respectively. Moreover, human beings are more sensitive to hue, so we divide H into more segments to get more representative values. For example, to quantize the image to 256 colors, for the *HSV* color space, H is divided into 16 levels while S and V are divided into 4 levels. And for the *LSLM* color space, L is divided into 4 levels; S and LM are divided into 8 levels.

- Dominant Color

For the dominant color feature, we generated the color histogram, which contains the number of pixels of representative colors. The color with pixel count less than one percent of the total image is discarded. Each image is therefore associated with a set of dominant colors. Although a color cab be represented as a three dimensional vector in a color space, in our approach, while each dominant color is assigned with a unique item number, each image is represented as a set of items.

- Adjacent Color Combination

The rationale of the adjacent color combination feature comes from the color harmony. According to the theory of the color harmony, different combinations of colors would bring different feelings.

For example, the combination of blue and green appears in many Monet's paintings. It brings humans the feelings of coolness and peace.

(a) (b)

FIGURE 1: An example of JSEG segmentation. (a) The original painting of Yang San-Lang, (b) Regions of similar colors.

To capture the adjacent color combination feature, first image segmentation technique, JSEG, is employed to segment an image into homogeneous regions. JSEG is an unsupervised segmentation technique based on colors and textures. It consists of two stages: color quantization and spatial segmentation. The first stage, color quantization, is to find out several representative colors which are suitable to discriminate regions in an image. Each representative color is assigned to a color label, and an image is transformed into a labeled image by replacing the colors by labels. The second stage is the spatial segmentation based on region growing method. For more detail, please refer to [5]. Figure 1 shows an example of JSEG segmentation; Figure 1.(a) shows the image before segmentation while Figure 1.(b) shows the regions after JSEG segmentation.

After image segmentation, each region is associated with a representative color. Each image is represented as a set of color pairs. Each color pair is made up of the representative colors of two *adjacent regions*.

- Scalable Color Descriptor

The scalable color descriptor is a color histogram, encoded by the Haar transform, in the *HSV* color space. In other words, in this work, the scalable color descriptor is represented as a tuple of numeric attributes.

- Color Structure Descriptor

The color structure descriptor represents the information of colors and corresponding spatial arrangement. It is a histogram that counts the occurrences of colors appeared in an 8x8 window sliding over the rows and columns of the image. Color values in color structure descriptor are represented in the *HMMD* color space, which is quantized non-uniformly into 32, 64, 128 or 256 bins. Therefore, in this work, the color structure descriptors of images in 32, 64, 128, or 256 *HMMD* color space are represented as tuples of 32, 64, 128, or 256 numeric attributes respectively.

- Color Layout Descriptor

The color layout descriptor specifies the spatial distribution of colors in *YCbCr* color space. To obtain the color layout descriptor, an image is segmented into 8 x 8 grids. Then 2D 8 x 8 Discrete Cosine Transform (DCT) is performed on the average colors of these 8 x 8 grids. The descriptor is a series of 64 nonlinear quantized DCT coefficients.

In this work, instead of performing Discrete Cosine Transform, the 8 x 8 average colors of grids are represented as a compact spatial data structure, 2D string, to preserve the spatial relationships among colors of grids. 2D string was originally proposed by Chang et al. for iconic indexing of image retrieval [4]. Some definitions concerning 2D string are given in the following.

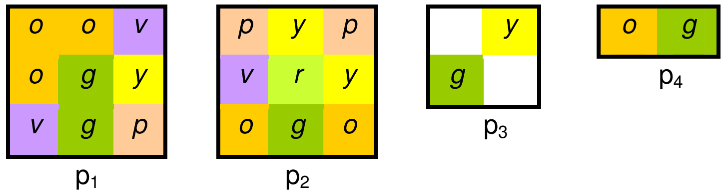

FIGURE 2: Examples of grid images for 2D string representation.

[Definition 1] A 1D string S, over a set of colors C, is represented as $S = c_1c_2\ldots c_m$ where $c_i \in C$ for $1 \le i \le m$, and m is the length of S.

[Definition 2] Let C be a set of colors and R be the set {"=", "<"} which is used to specify the relative direction among colors of grids. The symbol "=" denotes the "at the same spatial location", the symbol "<" denotes the "left-right" or "below-above" spatial relationship. A 2D string (S_x, S_y) over C is defined as $(c_1r_{1x}c_2r_{2x}\ldots r_{(n-1)x}c_n\ ,\ c_{p(1)}r_{1y}c_{p(2)}r_{2y}\ldots r_{(n-1)y}c_{p(n)})$, where $c_1c_2\ldots c_n$ and $c_{p(1)}c_{p(2)}\ldots c_{p(n)}$ are 1D strings over C, p is a permutation function from {1,…,n} to {1,…,n}, $r_{1x}r_{2x}\ldots r_{(n-1)x}$ and $r_{1y}r_{2y}\ldots r_{(n-1)y}$ are both 1D strings over R and n is the length of (S_x, S_y). A 2D string with n objects is called the size-n 2D string [4].

For example, the size-9 2D string representations for the 3 x 3 grid images p_1, p_2 in Figure 2 is $(S_{1x}, S_{1y}) = (v{=}o{=}o{<}g{=}g{=}o{<}p{=}y{=}v,\ v{=}g{=}p{<}o{=}g{=}y{<}o{=}o{=}y)$, $(S_{2x}, S_{2y}) = (o{=}v{=}p{<}g{=}r{=}y{<}o{=}y{=}p,\ o{=}g{=}o{<}v{=}r{=}y{<}p{=}y{=}p)$, respectively. Note that, to ensure the unique 2D string representation (S_x, S_y) of a grid image, if two colors of grids c_i, c_j are at the same spatial location along the vertical axis, the relative orders of c_i, c_j in S_y should be of the same as those in S_x.

3.2 Frequent Pattern Mining

In order to obtain the interesting hidden relationships between color features and painting styles, different approaches are utilized for the five types of color features mentioned in Sec. 3.1. Table 1 lists the summary of color features, color representations, and approaches to discover common characteristics in this paper.

3.2.1 Frequent Itemset Mining

For the features of dominate colors and adjacent color combinations which are represented as sets of items, frequent itemset mining [1] developed in the field of data mining is employed to discover frequented used colors and adjacent color combinations. The techniques of frequent itemset mining originated from the market basket analysis which analyzes customer buying behaviors by discovering associations between the items bought by most customers. Given a transaction database where each transaction is a set of items, frequent itemset mining finds the items that are frequently purchased together. The percentage of transactions in the transaction database that contain the itemset (the set of items purchased together) is called the *support* of this itemset. In our work, the discovered frequent itemset specifies the set of colors (or the set of adjacent adjacent color combinations) frequently used together by an artist.

Moreover, there are similarities between colors. For example, both of the two colors, (72, 0.6, 0.8) and (72, 0.8, 0.8) in HSV color space, are green with slight different saturations. Some artists preferred green colors with different saturations and/or luminance while some others preferred high saturation colors in spite of hue and luminance. Therefore, the concept of multi-level association rule mining is applied into the frequent itemset mining of the dominant color feature and the adjacent color combination feature. Multi-level association rules involve concepts at multiple levels of abstraction [7]. Figure 3 illustrate an example of multi-level frequent itemset mining for the dominant color feature. In Figure 3, a color can be generalized by replacing the colors in lower-level by their higher level color. In our work, the multi-level mining with reduced support is adopted. In other word, the lower the level of abstraction, the smaller the corresponding minimum support threshold is.

Color Features	Representation	Common Characteristics
Dominant Color, Adjacent color combination	A set of items	Frequent Itemset Mining
Scalable Color Descriptor, Color Structure Descriptor	A tuple of numeric attributes	Gain Ratio of C4.5
Color Layout Descriptor	A 2D string	Frequent Spatial Co-orientation Pattern Mining

TABLE 1: Summary of color features, representations and approaches to discover common characteristics.

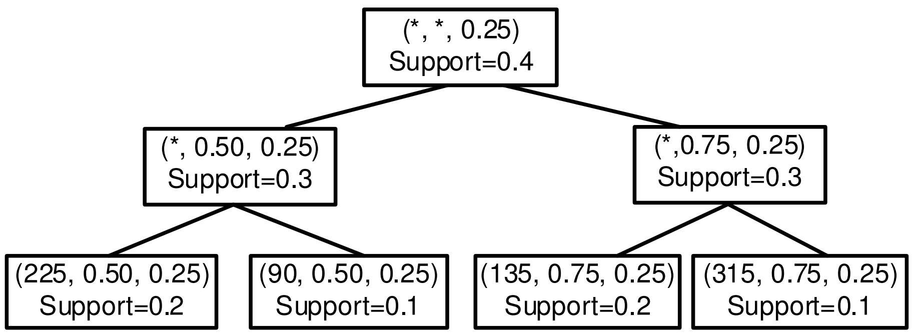

FIGURE 3: An example of multi-level frequent itemset mining for the dominant color feature.

3.2.2 Discretization
For the features of scalable color descriptor and color structure descriptor which are represented as tuples of numeric attributes, discretization techniques on continuous numeric attributes are utilized. Discretization of numeric attributes is essential for the classification task which discovers the discriminative characteristics of painting styles performed after frequent pattern mining. There exist several techniques for discretization. In this paper, the discretization technique of C4.5 inspired by the Minimum Description Length principle and gain ratio is utilized. For detail, please refer to [18].

3.2.3 Frequent Spatial Co-orientation Pattern Mining
For the feature of color layout descriptor which is represented as the spatial data structure, 2D string, the technique of frequent spatial co-orientation pattern mining is utilized. Spatial co-orientation patterns refer to the spatial objects that occur frequently and collocate with the same orientation among each other. For example, in Figure 2, both p_3 and p4 are the spatial co-orientation patterns of grid images p_3 and p_4. In other words, the discovered frequent spatial co-orientation patterns specifies the color layouts frequent used by an artist.

Given a 2D string database, the problem of mining spatial co-orientation patterns is to find the complete set of maximal 2D subsequences among all 2D strings that satisfies a user-specified minimum support threshold. We have developed an Apriori-based algorithm to discover the spatial co-orientation patterns. For more detail, please refer to [22].

3.3 Painting Style Classification
Common properties of images with the same style are discovered by frequent pattern mining. However, it is not enough to discriminate one style from others only by the frequent patterns. In generally, people recognize a painting style not only by the common characteristics of itself, but

also by the discrimination between this style and the others. In this work, we modified the associative classification algorithm [11] and proposed two improved algorithms to discriminate common properties of one style from those of the others. Moreover, bagging classification was utilized to improve the accuracy of discrimination.

In associative classification, a classifier is an ordered set of rules. Each rule is of the form $l \Rightarrow y$, where $l \in \bigcup_k L_k$, l is a frequent pattern and y is a category. The format of a classifier is $< r_1, r_2, ..., r_n, default\_class >$, where each rule r_i is ordered by its confidence and support. The confidence of a rule is the percentage of training samples that satisfy l and belong to class y. A training sample conforms to l if its feature f contains l. Test data is classified by the first rule which cover it. If there are no rules satisfying the test data, the test data is classified to the $default\_class$.

To discover the discriminative characteristics between the style of one artist and that of the other one, first, for each artist, the frequent patterns of fine art images are discovered as stated in section 3.2. The frequent patterns include the frequent itemsets for the dominant color feature, the frequent itemsets for the adjacent color combination feature, the classification rules generated by C4.5 [2, 17, 18] for the scalable color descriptor and the color structure descriptor, and the spatial co-oriental patterns for the color layout descriptors. Then, a classifier is constructed from the discovered patterns over the training data.

3.3.1 Single Feature Classification

In the original algorithm of associative classification, the minimum support threshold for each class is the same. However, the uniform support is not appropriate for all cases in the classification. For instance, an artist who used colors diversely may have more rules but with lower supports. Consequently, given the same minimum support, fewer numbers of rules would be discovered for artists who preferred colors more diversely. Furthermore, minimum support threshold should be determined manually by experts in previous algorithm. We presented the Single Feature Classification algorithm (SFC) which is modified from the concept of msCBA algorithm to solve these problems [12].

To determine the appropriate value of minimum support for each category, the proposed algorithm builds the classifier iteratively with possible values of minimum support and then selects the most effective one among the built classifiers. Five-fold cross-validation is employed to avoid over-fitting. The training set is divided into five disjoint subsets of equal size. The algorithm trains five times on four of these five subsets (training set) and tests on the one left out (validation set). After finishing five times of training, the classifier with the highest average accuracy of testing on validation set is selected.

3.3.2 Multiple Feature Classification

SFC algorithm considers the difference of consistency of intra-category style and builds a classifier for single type of feature. However, the appropriate feature for each category may differ from other categories. For example, an artist often used red and the colors with low hue and low value, e.g. (*, 0.25, 0.25), in one painting. In addition, blue and white are often adjacent in his work. Both dominant color and adjacency color should be considered to represent this artist's style. We extend the SFC algorithm to build a classifier that contains rules of different types of features, which is called Multiple Feature Classification (MFC).

MFC classifier consists of rules of multiple features. Figure 4 is an example of a two-way classifier constructed by MFC algorithm for the color styles between Vincent van Gogh and Paul Gauguin. There are five rules and three types of patterns in the classifier.

MFC algorithm trains for each combination of features with various corresponding minimum supports. It first mines all frequent patterns of the categories, and trains the classifier with these patterns. For each combination of patterns and minimum supports, we take five-fold cross-validation which is the same as SFC algorithm to evaluate the classifiers and select the highest score one.

(0.45, *, 0.25)→van Gogh	(dominant colors)
(*,0.75,0.25)|(*,0.25,0.25) → Gauguin	(adjacent color combinations)
(0, *, 0.75)^(0.45,*,0.25) → van Gogh	(dominant colors)
<(0.3,*,0.8)<(*,0.25,*), (*,0.25,*)<(0.3,*,0.8)> → van Gogh	(color layout)
default class: Gauguin	

FIGURE 4: An example of the classifier discovered between works of Gauguin and van Gogh by multiple feature classification.

4. EXPERIMENTS

We collected images of impressionists' works from the Internet. The categories include Paul Gauguin, Claude Monet, Pierre Auguste Renoir, and Vincent van Gogh. Our data set consists of 126, 182, 154 and 201 images of Gauguin, Monet, Renoir and van Gogh respectively. The sizes of images range from 366×400 to 984×840. We first transform the color space of each image from *RGB* to *HSV* and *LSLM*. Then, each images is quantized to 128 (*H*:*S*:*V* is 8:4:4, *L*:*S*:*LM* is 4:8:8).

Two series of experiments were performed to evaluate the performance of our proposed approaches. The performance of the two-way classification is measured by the accuracy, which is defined as the percentage that the test images are classified correctly. Five-fold cross-validation was performed to obtain accuracy of the classification method. In each time, one of the folds is selected as the test set while the other four folds are collected to derive the classifier. The accuracy is therefore measured as the average accuracy over the five tests.

The first part of experiments is to compare the performance of classification for each color feature. For each color feature, SFC algorithm is used for classification. Note that SFC algorithm generates and tests all combinations of minimum support thresholds while the original associative classification only considers a subset of SFC. Consequently, it is expected the accuracy of the original associative classification is not higher than that of SFC. Therefore, only the result of SFC algorithm is listed.

Table 2 shows the result of the first experiment. The min_sup columns show the corresponding minimum support threshold for each artist selected by SFC algorithm. The overall accuracy achieved 53% to 93%. The average accuracy shows that the dominant color feature and the adjacent color combination feature perform slightly better than MPEG-7 color descriptor. It is observed that the pairs of Gauguin versus Renoir, and Monet versus Renoir are less discriminating between each other. Actually, Gauguin and Renoir used higher contrast for adjacency colors and reds/oranges. According to adjacent color combination, the pair of Gauguin and Van Gogh is better discriminating because Gauguin often used higher contrast adjacent color combination, {(*,0.75,0.25), (*,0.25,0.25)}, {(*,0.75,0.5) , (*,0.25,0.5)} and {(*,0.75,0.75) , (*,0.25,0)}, and the painting style of Van Gogh is lack of adjacent color combinations.

Table 3 shows the classification accuracy with different multiple feature set. In Table 3, FS1 denotes the feature set including the dominant color feature, and the adjacent color combination feature. FS2 denotes the feature set including the dominant color feature, the adjacent color combination feature and the color layout descriptor. Besides the features included in FS2, FS3 includes the scalable color descriptor and the color structure descriptor. The min_sup columns show the minimum support threshold for each artist selected by MFC algorithm. The overall accuracy achieved 83% to 93%. The average accuracy shows that MFC performs better than SFC. Moreover, the consideration of color layout descriptor improves the average performance slightly, but performs much better for the discrimination between Gauguin and van Gogh. This phenomenon also occurs in the consideration of the color structure descriptor and the scalable color descriptor for the discrimination between Renoir and van Gogh. At last, we have performed the experiments for the style classification among these four artists (4-way classification). The classification accuracy is 75.14%.

Feature / Artists	Dominant Color		Adjacency Color Combination		Color Structure Descriptor	Scalable Color Descriptor	Color Layout Descriptor	
	Accuracy	Min_Sup	Accuracy	Min_Sup	Accuracy	Accuracy	Accuracy	Min_Sup
G - M	79.17%	0.3/0.1	79.17%	0.3/0.1	77.24%	73.06%	59.31%	0.2/0.3
G - R	83.46%	0.1/0.1	74.60%	0.3/0.2	66.86%	72.70%	61.85%	0.2/0.3
G – V	87.02%	0.1/0.1	88.63%	0.2/0.2	86.54%	92.92%	67.98%	0.2/0.3
M - R	82.91%	0.1/0.1	68.50%	0.3/0.3	70.34%	66.72%	68.14%	0.3/0.2
M – V	87.12%	0.2/0.2	84.66%	0.1/0.2	88.20%	84.54%	70.95%	0.3/0.3
R – V	87.23%	0.2/0.3	87.23%	0.2/0.3	87.78%	89.78%	66.78%	0.2/0.3
Average Accuracy	84.49%		80.47%		79.49%	79.95%	65.84%	

TABLE 2: Classification accuracy with different features.
(G : Gauguin, M : Monet, R : Renoir, V : van Gogh)

Features / Artists	FS1		FS2		FS3	
	Accuracy	Min_Sup	Accuracy	Min_Sup	Accuracy	Min_Sup
G - M	89.30%	0.3/0.1	89.34%	0.2/0.2	89.80%	0.2/0.2
G - R	87.69%	0.2/0.3	88.02%	0.2/0.2	88.94%	0.2/0.2
G – V	89.73%	0.1/0.1	92.41%	0.3/0.3	93.08%	0.3/0.3
M - R	83.64%	0.2/0.3	83.71%	0.2/0.3	83.99%	0.2/0.3
M – V	89.79%	0.1/0.3	89.84%	0.3/0.3	90.74%	0.3/0.3
R – V	90.83%	0.2/0.2	91.78%	0.3/0.3	93.39%	0.3/0.3
Average Accuracy	88.50%		89.18%		89.99%	

TABLE 3: Classification accuracy with different multiple features.
(G:Gauguin, M:Monet, R:Renoir, V:van Gogh)

5. CONCLUSIONS & FUTURE WORK

In this paper, the techniques to discover color styles of fine art images of Impressionism are investigated. The dominant color feature, the adjacent color combination feature, and three of MPEG-7 color descriptors are extracted. Frequent pattern mining techniques are employed to discover the common characteristics of an artist's works. To discover the color styles in terms of distinguished common characteristics, the classification techniques are developed. The experiments show that the proposed style mining approaches are satisfied to discover the color styles of Impressionism paintings. Future works include the style mining of brush strokes, the consideration of quality for the frequent itemset mining [16], the fuzzy logic approach for painting style mining, and the consideration of other images features.

6. ACKNOWLEDGEMENTS

The author wish to thank Mr. Chin-Nan Liu and Miss Ling-Yin Wei for their implementation of the algorithms.

7. REFERENCES

1. R. Agrawal and R. Srikant. "*Fast Algorithms for Mining Association Rules*". In Proceedings of International Conference on Very Large Data Bases, 1994.

2. M. N. Anyanwu and S. G. Shiva. "Comparative Analysis of Serial Decision Tree Classification Algorithms". International Journal of Computer Science and Security, 3(3) 2009.

3. S. F. Chang, T. Sikora, and A. Puri. "*Overview of MPEG-7 Standard*". IEEE Transactions on Circuits Systems for Video Technology, 11(6), 2001.

4. S. K. Chang, Q. Y. Shi, and C. W. Yan. "*Iconic Indexing by 2D Strings*", IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(3), 1987.

5. Y. Deng, and B. S. Manjunath. "*Unsupervised Segmentation of Color-Texture Regions in Images and Video*". IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(8), 2001.

6. B. Gunsel, S. Sariel, and O. Icoglu. "*Content Based Access to Art Paintings*". In Proceedings of IEEE International Conference on Image Processing, 2005.

7. J. Han, and Y. Fu. "*Discovery of Multiple-Level Association Rules from Large Databases*". IEEE Transactions on Knowledge and Data Engineering, 11(5), 1999.

8. P. S. Hiremath and J. Pujari. "Content Based Image Retrieval based on Color, Texture and Shape features using Image and Its Complement". International Journal of Computer Science and Security, 1(4), 2007.

9. L. Leslie, T. S. Chua and R. Jain. "*Annotation of Paintings with High-level Semantic Concepts Using Transductive Inference and Ontology-based Concept Disambiguation*". In Proceedings of ACM International Conference on Multimedia, 2007.

10. C. T. Li, and M. K. Shan. "*Affective Space Exploration for Impressionism Paintings*". In Proceedings of Pacific-Rim Conference on Multimedia, 2008.

11. B. Liu, W. Hsu, and Y. Ma. "*Integrating Classification and Association Rule Mining*". In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1998.

12. B. Liu, Y. Ma and C. K. Wong. "*Classification Using Association Rules: Weaknesses and Enhancements*". In Vipin Kumar, et al, (eds), Data Mining for Scientific Applications, 2001.

13. Y. Marchenko, T. S. Chua, I. Aristarkhova. "*Analysis and Retrieval of Painting Using Artistic Color Concepts*". In Proceedings of IEEE International Conference on Multimedia and Expo., 2005.

14. MPEG-7 Visual Experimentation Model (XM), Version 10.0, ISO/IEC/JTC1/SC29/WG11, Doc. N4063, 2001.

15. Overview of the MPEG-7 Standard, Version 5.0, Final Committee Draft, ISO/IEC JTC1/SC29/WG11, Doc. N4031, 2001.

16. K. Preetham and V. S. Ananthanarayana. "*Discovery of Frequent Itemsets Based on Minimum Quality and Support*". International Journal of Computer science and Security, 3(3), 2009.

17. J. R. Quinlan. "*C4.5: Programs for Machine Learning*", Morgan Kaufmann, San Mateo, CA (1993)

18. J. R. Quinlan. "*Improved Use of Continuous Attributes in* C4.5", Journal of Artificial Intelligence Research, 4, 1996.

19. R. Sablatnig, P. Kammerer and E. Zolda. "*Hierarchical Classification of Painted Portraits using Face and Brush Stroke Models*". In Proceedings of International Conference on Pattern Recognition, 1998.

20. A. Vailaya, A. T. Figueriedo, A. K. Jain and H. J. Zhang. "*Image Classification for Content-Based Indexing*". IEEE Transactions on Image Processing, 10(1), 2001.

21. J. Z. Wang, J. Li, and G. Wiederhold. "*SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries*". IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9), 2001.

22. L. Y. Wei, and M. K. Shan, "*Efficient Mining of Spatial Co-orientation Patterns from Image Databases*". In Proceedings of IEEE International Conference on Systems, Man and Cybernetics, 2006.

23. O. R. Zaïane, M. Antonie and A. Coman, "*Mammography Classification by an Association Rule-Based Classifier*". In Proceedings of International ACM SIGKDD Workshop on Multimedia Data Mining, 2002,