Local Features in APICAS

Analyzing of Added Value of the Descriptors Based on MPEG-7 Vector Quantization

Krassimira Ivanova[1, a], Iliya Mitov[1, b], Peter Stanchev[1, 2, c], Evgeniya Velikova[3, d],
Koen Vanhoof[4, e], Benoit Depaire[4, f], Rajkumar Kannan[5, g]

[1]Institute of Mathematics and Informatics – Bulgarian Academy of Sciences, Sofia, Bulgaria
[2]Kettering University, Flint, MI, 48504, USA
[3]Faculty of Mathematics and Informatics, Sofia University, Sofia, Bulgaria
[4]IMOB, Hasselt University, Hasselt, Belgium
[5]Bishop Heber College, India

[a]kivanova@math.bas.bg; [b]imitov@math.bas.bg; [c]pstanche@kettering.edu; [d]velikova@fmi.uni-sofia.bg;
[e]koen.vanhoof@uhasselt.be; [f]benoit.depaire@uhasselt.be; [g]rajkumarkannan@yahoo.co.in

Abstract- **An approach for extracting higher-level visual features for art painting classification based on MPEG-7 descriptors was implemented in the system "Art Painting Image Colour Aesthetics and Semantics" (APICAS). The approach consists of the following steps: (1) tiling images into non-overlapping rectangles in order to capture more detailed local information; (2) the tiles of the images are clustered for each MPEG-7 descriptor; (3) vector quantization is used to assign a unique value to each tile, which corresponds to the number of the cluster where the tile belongs, in order to reduce the dimensionality of the data. The distribution of significance of the attributes, the importance of the underlying MPEG-7 descriptors as well as analysis of spatial granularity for class prediction in this domain are analyzed.**

Keywords- Content-Based Image Retrieval (CBIR); Multimedia Semantics; MPEG-7 Descriptors; Vector Quantization; Categorization

I. INTRODUCTION

Digitalized art collections allow the user to immerse in an ocean of accumulated cultural artifacts. Just a few years ago we could only dream of seeing some masterpieces. Now our computers move us to every chosen place and time. The growth of available digital resources increases the users' expectations for easy resource discovery by different criteria. While one user could be interested in art paintings from a specific movement or artist, others would search for images with particular theme or composition, while others would be attracted by the purely aesthetic influence of the paintings. Modern digital technologies made reality large virtual exhibitions of works from multiple cultures. One of the great advantages of digitization is the improved accessibility of cultural artifacts, including images of art.

Current search engines headed by Google are paving the way in our information age. Google answers daily more than 200 million queries against over 30 billion items. However, the search power of these engines is typically limited to text and its similarity. Since less than 1% of the Web data are in textual form, the rest being of multimedia/streaming nature, there is a clear need to extend next-generation search to accommodate these heterogeneous media. The new search engines must combine search according to textual information or other attributes associated with the files with abilities of extracting information from the content, which is scope of action of Content-Based Image Retrieval (CBIR).

CBIR is an area for which we can say that uses baron Münchhausen methods for self-extracting knowledge from the image content of the digital picture archives. The earlier CBIR systems provided search based on simple visual similarity between given picture or a sketch and the images stored in the digital resource. Current CBIR systems are helpful assistants when searching content by visual as well as by semantic similarity.

One standard for describing multimedia content data is MPEG-7. MPEG-7 is an ISO/IEC standard developed by Moving Picture Experts Group. MPEG-7 is not aimed at any application in particular. The visual descriptors in the MPEG-7 standard describe different aspects of the image content such as dominant colors, edginess, texture, etc. The MPEG-7 descriptors are often used in the process of image-to-image matching, searching for similarities, sketch queries, etc.

In the paper we apply a vector quantization (VQ) method on the MPEG-7 descriptors as a possible approach for dimensionality reduction. These descriptors were firstly introduced in [1]. Here we focused on the analysis of the added value received by the defined descriptors. The proposed algorithm is implemented in the experimental software system "Art Painting Image Colour Aesthetics and Semantics" (APICAS), created as an environment for testing applicability of different visual features in image retrieval.

The rest of the paper is organized in the following way. In Section 2 existing systems in the field of art painting image analysis are described. Section 3 presents our methodology for extracting visual features from MPEG-7 descriptors based on vector quantization. Section 4 describes the developed software. Section 5 presents the results. Finally, some conclusions and future work directions are highlighted.

II. SOME EXISTING ART IMAGE ANALYZING SYSTEMS

In the last 20 years, numerous research and development efforts addressed the image retrieval problem, adopting the similarity-based paradigm. The technical report made by Veltkamp and Tanase in 2000 [2] still remains a precise and comprehensive review of industrial systems developed since the beginning of the work on CBIR until the end of the previous century. While earlier developments were focused on extracting visual signatures, the recent ones address mostly the efficient pre-processing of visual data as a means to improve the performance of neural networks and other learning algorithms when dealing with content-based classification tasks. Given the high dimensionality and redundancy of visual data, the primary goal of pre-processing is to transfer the original data to a low-dimensional representation that preserves the information relevant for the classification. The performance of the techniques is assessed on a difficult painting-classification task that requires painter-specific features to be retained in the low-dimensional representation.

In the field of art-painting image retrieval, several systems propose a variety of approaches for defining high level visual characteristics. Some aspects of image content are used with an existing ontology of art space.

S. Kröner and A. Lattner [3] trained a naive Bayes classifier to distinguish among the artists. They use five measures: three of them measure the ratio of black and white pixels and two take the stroke direction. The method is evaluated on a data set of 60 drawings, containing 41 drawings by Delacroix and 19 drawings from other artists. They achieved correct classification for about 87% of the drawings.

A study of the personal style of artists in the database of the Austrian National Library is presented in [4]. In this research, a computer-aided classification and recognition system for portrait miniatures has been developed. It enables a semi-automatic classification based on brush strokes and the hierarchical structure of the classification steps allows a top-down classification. Especially for face recognition, a bottom-up approach is used. The final goal is to reduce the number of potential candidate artists to a minimum. The extension of the system, presented in [5], uses a hierarchical database indexing method based on Principal Component Analysis. The description incorporates the eyes as the most salient region in the portraits. 586 portrait miniatures have been tested in order to determining the orientation of eyes. The overall accuracy was about 87%.

D. Keren [6] proposed a framework for classification of paintings based on local features derived from coefficients of a discrete cosine transform. After calculating the local features, each pixel is classified and the overall classification of the image is determined from a majority vote of the pixel values. The experiments were made over the works of Rembrandt, Van Gogh, Picasso, Magritte, and Dali. The training set contained ten paintings from each painter and the test set consisted of twenty to thirty paintings for each painter. For the five painters tested, a "tournament scheme" classifier was implemented and the rate of success was 86%.

The Art Historian system [7] utilizes an automatic extraction of features of paintings' art movements. It is shown that the feature set enables one to highlight art movements efficiently.

In the classifier design, a statistical pattern recognition approach is exploited using Bayesian, k-NN and SVM classifiers. It has been shown that the art movements of paintings can be indexed by exploiting the following six measures: (1) percentage of dark colors; (2) gradient coefficient calculated from the gradient map of the painting image; (3) number of local and global maxima in the luminance histogram; (4) color range that the peak point of the luminance histogram corresponds; (5) the deviation of average grey level acquired within each block from the average grey level acquired within the entire image after the partition of the painting image into identical blocks; (6) the deviation of grey level distribution from the Gauss distribution. In the given experiments, the training set was constructed with 27 original paintings that belong to classic, cubic or impressionist movements, 9 from each class. The six features were extracted for every image in each movement. In order to observe efficiency of the proposed feature set and robustness to changes in the illumination and resolution, a test set was created with lighting effects and resolution changes. The test set included 107 paintings from the different movements taken into account. The overall classification accuracy varied from 81.66% to 91.66% depending of used dimensionality and classification method.

In Lightweight Image Retrieval System for Paintings [8], the image indexing features are divided into the following three groups: (1) canvas features: max, min, mean, median, and standard deviation from each of the red, green, and blue color channels; (2) color features: intensity mean (measures the global brightness of a grayscale image), color frequency distribution (measures the degree of disorder found in the frequency distribution of colors in a painting); (3) edge characteristics: line count (uses the Sobel edge detector to identify lines in the image). The tests for identifying the work of individual painters were run over set of 100 training images, which contained 10 images from the corpus of 10 artists and a test set of 90 images, randomly chosen from the work of these artists. In 81% of test cases, the system retrieved at least one painting by the same artist in the top ten closest matches, suggesting that the model is effective for interactive classification of paintings by the artist. Another test was based on a database of 500 images from the Web Museum. Although the overall retrieval rate was only 49.2%, the system performed particularly well with respect to certain artists, distinguishing 87.5% for Aertsen and Velazquez. An analysis of the classification errors reveals that the system is effectively classifying artistic style even when it fails to classify the artist correctly.

The team, headed by R. Jain [9] uses annotation of paintings based on brushwork, where brushwork is modeled as part of the annotation of high-level artistic concepts such as the movements of artists using defined ontology of so cold "brushworks", based on specific combination of low-level feature values of gradient, texture, hue and light contrast, and homogeneity, which characterize some style of the painting. The experiments with 4880 patches from 30 paintings of Renaissance, Fauvism, Impressionism, Post-Impressionism, Expressionism and Pointillism painting styles, where 75% are used for training and 25% for testing, show accuracies between 80% and 95% in variants, where multiple experts are used.

III. PROPOSED APPROACH FOR FEATURE EXTRACTION

MPEG-7 descriptors are complex descriptors, which provide a good presentation of different types of the visual features. The formal description of the structure of MPEG-7 descriptors and algorithms for their use is given in [10].

These complex structures need specific processing and similarity measures and cannot be put directly into generic classification algorithms.

A. Used MPEG-7 Descriptors

Here we give a brief explanation of each MPEG-7 descriptor examined in our work and which values we use in further processing:

- *Scalable Color (SC)* represents the color histogram in the HSV color space, encoded by a Haar transform. For presenting the image or a selected part, *Scalable Color* needs a vector with 64 attributes;
- *Color Layout (CL)* specifies the spatial distribution of colors using *YCbCr* color space. We use the first quantized *DCT* coefficient of the *Y*, *Cb* and *Cr* components, the next five successive quantized *DCT* coefficients of the *Y* component and the next two quantized *DCT* coefficients of the *Cb* and *Cr* component. Thus, the *Color Layout* vector has 12 attributes;
- *Color Structure (CS)*, which specifies both color content and the structure of the content. The descriptor expresses local color structure in an image by means of a structuring element that is composed of several image samples. We use a vector with 64 attributes for representing the *Color Structure*;
- *Dominant Color (DC)*. We reconfigured the presentation of this descriptor as three vectors, representing distribution of quantized hue, saturation and luminance. Such method is already precisely described and used by us in [11]. After this quantization we receive a vector with 23 attributes (13 for hue + 5 for saturation + 5 for luminance);
- *Edge Histogram (EH)* specifies the spatial distribution of five types of edges in local image regions (four directional edges – vertical, horizontal, 45 degree, 135 degree and one non-directional). *Edge Histogram* descriptor produces a vector with 80 attributes;
- *Homogeneous Texture (HT)* characterizes the region texture using the energy and energy deviation in a set of frequency channels. A vector with 60 attributes is used, presenting *Energy* and *Energy Deviation*.

As a result of using these descriptors, we obtain a vector, which contains altogether more than 300 attributes. From other side each descriptor needs specific similarity measure. Some of the MPEG-7 descriptors are alternative. Scalable Color, Color Layout, Color Structure and Dominant Color concern different aspects of the same phenomenon, i.e. distribution of the color within the image or region. It means that not all descriptors have to be used in the classification process. One of our tasks is to examine which are most for extracting simple visual features for the purposes of class recognition.

Visual attributes can represent global characteristics concerning whole images, or they can be extracted over the part of the images (specific region or tile of the image). Both approaches have their strengths: global attributes deliver integral temper of the image. Local attributes can capture more detailed information, which characterize the artists' styles or movements' specifics but introduce redundancy for the classifier.

For reducing the computational weight and redundancy one possibility is to choose only part of the tiles – only chess ordered tiles, starting from the first tile (1, 1) or from the second one (1, 2) as well as taking into account only left sided or right sided tiles. In this paper we use all of the tiles. The second goal of this research is to examine the trade of granularity.

B. Vector Quantization Approach

The flowchart of the process of extracting VQ-MPEG7 descriptors is given in Figure 1.

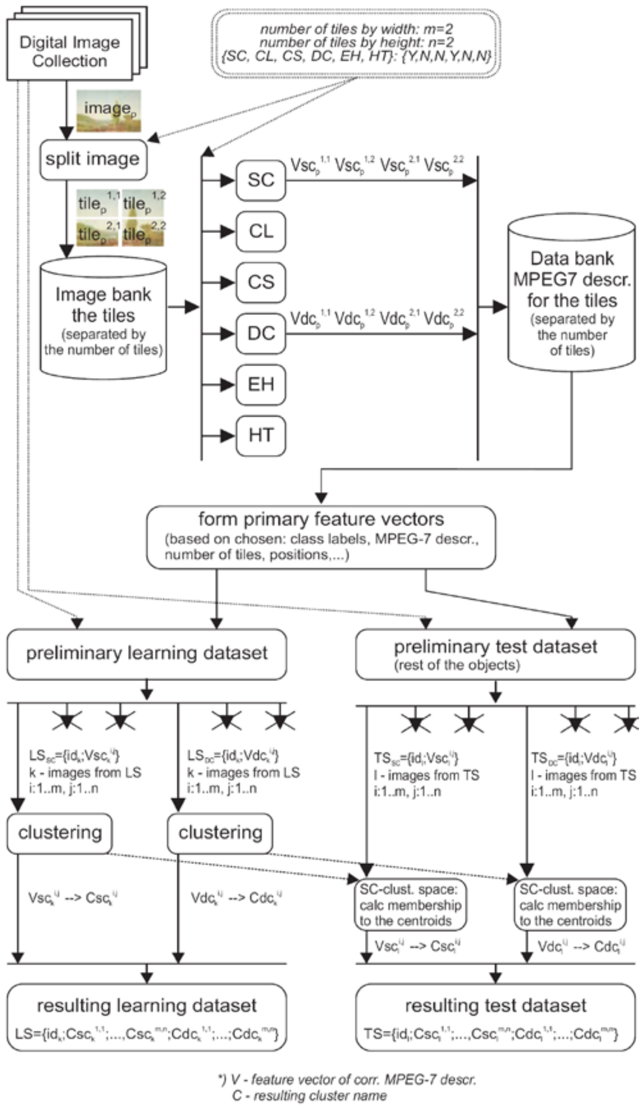

Figure 1 Flowchart of the process

In the proposed approach we split the images into $m \times n$ non-overlapping rectangles (tiles). The tiles are marked as (i, j), where $i \in 1...m$ and $j \in 1...n$. Index i increases from the left tile to the right tile and index j increases from the top tile to the bottom tile of the image.

Let $I = \{I_p \mid p = 1...r\}$ be the observed set of r images. Each image is divided into non-overlapping $m \times n$ tiles (sub-images) and as a result a set of $k \times m \times n$ tiles $T = \{I_p^{ij} \mid p = 1...r, i = 1...m, j = 1...n\}$ is produced.

For each of chosen MPEG-7 descriptors $X \in \{SC, CL, CS, DC, EH, HT\}$ the corresponded feature vector is calculated for each of formed tiles $\{(V_X)_p^{i,j} \mid X \in \{SC, CL, CS, DC, EH, HT\}, i = 1..m, j = 1..n\}$.

From the observed set, a subset of images is chosen, used as learning set. The rest of the images are used for test set.

Clustering procedure for received vectors of the tiles from learning set is applied. The number of the clusters can be different for different MPEG-7 descriptors, but in this realization an equal number α is used for all descriptors.

Each cluster receives as a name its corresponded number of cluster from clustering procedure. These names are used as attribute values in the resulting dataset. In this way the complex feature vectors in primary learning dataset is simplified with replacing with one corresponded value.

The centroids of clusters are calculated. For the tiles from the test dataset the membership to the centroids is calculated, the number of the corresponding cluster is assigned as an attribute value.

As a result, each image is represented with a feature vector with $x \times m \times n$ attributes, where x is the number of examined MPEG-7 descriptors and m and n are the numbers of tiles by width and height. For instance in case of using all MPEG-7 descriptors for 3×3 tiling, the number of attributes in this vector is $6 \times 3 \times 3 = 36$. In case of selecting only a subset of the available tiles (chess order or left/right side), the number of features reduces.

IV. SOFTWARE REALIZATION

The proposed algorithm is implemented in the experimental system "Art Painting Image Color Aesthetic and Semantic" (APICAS) [12].

APICAS aims to provide a suite of specialized tools for CBIR within digital collections of art images. The main functions of APICAS can be grouped into following groups:

- data entry support: choosing the collection; setting up parameters; selecting the samples of the learning set; supplying with test textual metadata;
- visual characteristics extraction, where the algorithm for extracting local VQ-MPEG7 features is implemented;
- data delivery: descriptive analysis; visualizing; statistical and DM analysis.

A comprehensive description of the architecture and implemented functions in APICAS is given in [13]. Here we stop our attention in some specific function concerning the implementation of the proposed algorithm.

The system is created on module principle and allows incorporating of external modules for serving some processes. In this case, for obtaining the MPEG-7 descriptors APICAS refers to the Multimedia Content Management System MILOS [14] and as a clustering algorithm the program "vcluster", which is a part of the CLUTO open source software package [15], is used in the system.

The descriptive analysis of received characteristics is conducted in the Waikato environment for Knowledge Analysis (WEKA) [16] under the procedures of attribute subset selection.

The examination of the accuracy for predicting the artists' names or movements is made with representatives of decision rules (OneR [17] and JRip as WEKA implementation of RIPPER [18]), decision trees (J48 as WEKA implementation of C4.5 [19]), and associative classifiers (PGN [20]). These classification schemes have similar behavior of creating the task model. The associative classifier PGN is realized in the data mining analysis environment PaGaNe [21] by members of the authors collective. Its main feature is focusing primarily on the confidence of the association rules and only in a later stage on the support of the rules. The comparison of PGN with the other mentioned classifiers showed very good results in the case of multiclass data sets [22].

In order to facilitate the processing, the results of some previously made operations are kept and used by next experiments. Such operations are:

- *Creation of the tiles*. They are stored as separate JPG-files in specific directory (subfolder of the main folder with name starting with "_tmp", followed by m and n). The name of the tile starts with the position of the tile by width i, $i = 1,...,m$, the position by height j, $j = 1,...,n$, followed with the name of the image. This way, already processed tiles can be used over again directly;
- *Calculation of MPEG-7 descriptors*. They are stored in corresponding XML files. If the tiles were already created and MPEG-7 descriptors were calculated by previous running of the functions, the system passes over this step. This significantly accelerates the conducting of experiments with varying number of clusters.

The system allows viewing the results of the clustering procedure. It can show the tiles, which belong to the selected number of cluster for specific MPEG-7 descriptor. For instance, part of the tiles, which are in cluster No: 2 of Colour Structure Descriptor for 5×5 tiling, is seen in Figure 2.

Another function allows displaying the tiles from the learning set, which are closest to the centroids of a given clustering for a chosen MPEG-7 descriptor.

Figure 3 shows the tiles, closest to the centroid of Colour Structure Descriptor with tiling 4×4 and 20 clusters. The idea of this presentation is that these tiles can be used later as elements in a visual lexicon for representing specifics of some image profiles.

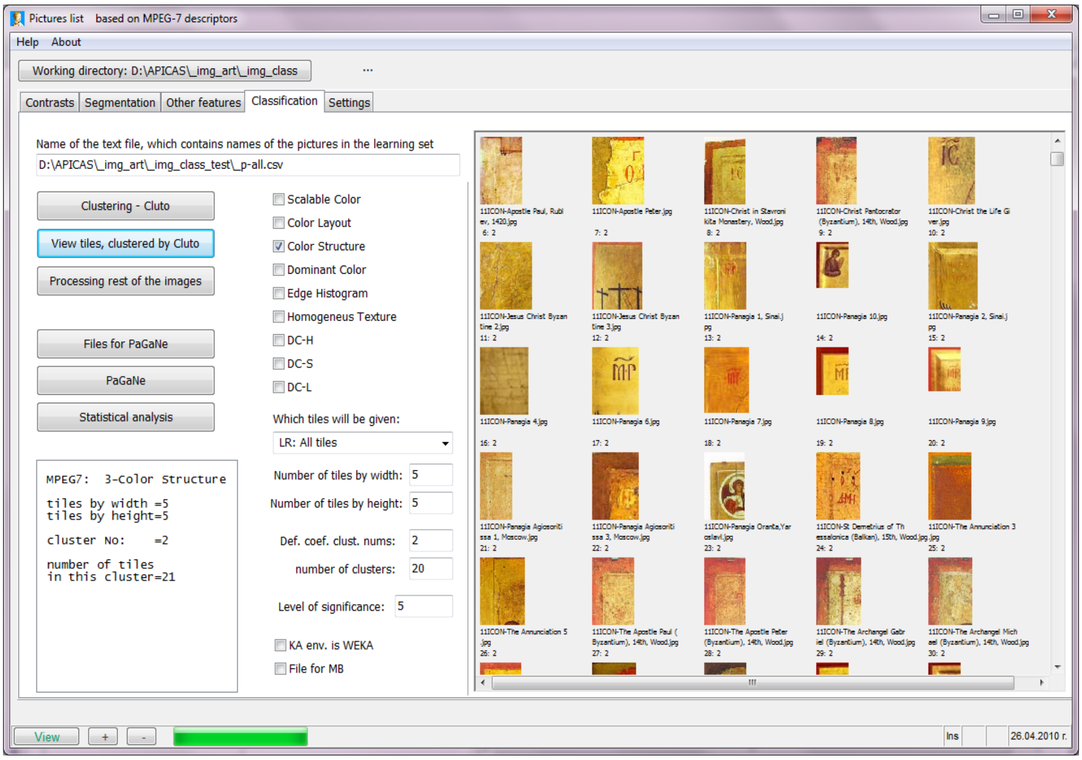

Figure 2 The screenshot of viewing 5×5 tiles, belonging to cluster No:2 of Colour Structure Descriptor

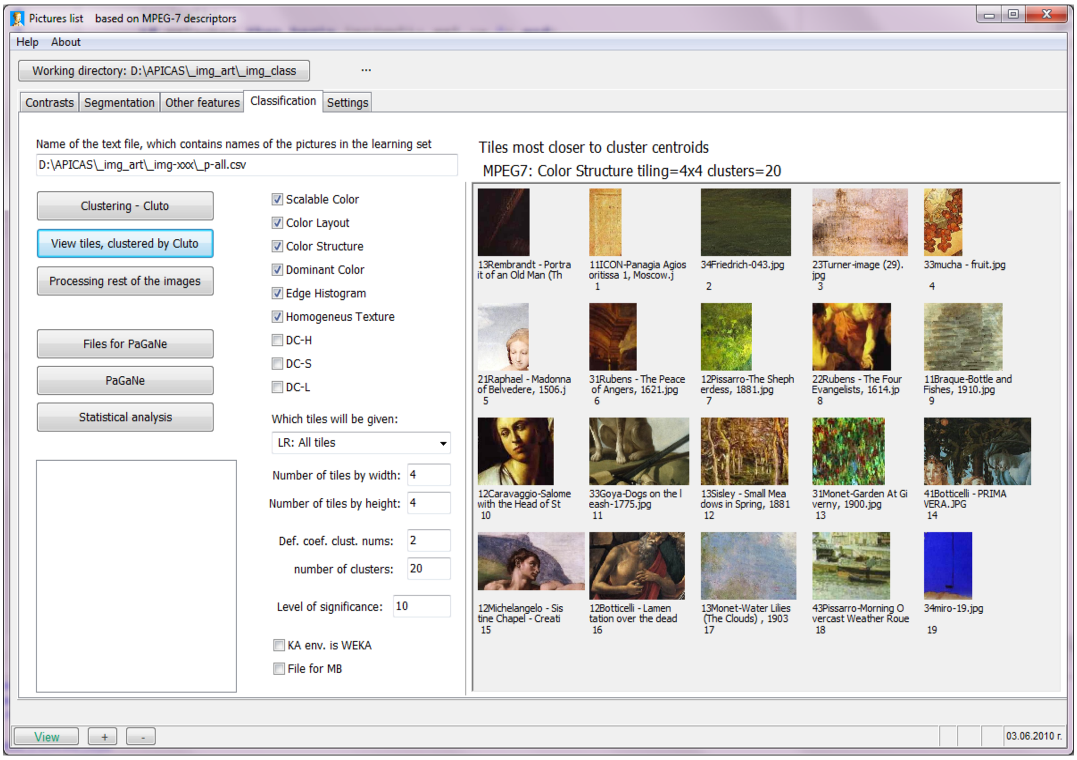

Figure 3 The tiles, closest to the centroid of Colour Structure Descriptor (tiling 4×4 with 20 clusters)

V. EXPERIMENTS

For our experiments we have used datasets that include 600 paintings of 18 artists from different movements of West-European fine arts and one group, which represent Orthodox Iconographic Style from Eastern Medieval Culture (Table 1).

The paintings were chosen by an art expert reviewer. He has included in the collection the representative artists for the movements and most valuable paintings for each artist. The digitized artworks were obtained from different web-museum sources using ArtCyclopedia as a gate to the museum-quality fine art on the Internet as well as from different Eastern public virtual art galleries and museums for extracting Icons. The images are free available, designed for presenting the paintings on the screen for the purposes of the resource discovery and educational use.

TABLE I LIST OF THE ARTISTS, WHICH PAINTINGS WERE USED IN EXPERIMENTS, GROUPED BY MOVEMENTS

Movement	Artist
Icons (60)	Icons (60)
Renaissance (90)	Botticelli (30); Michelangelo (30); Raphael (30)
Baroque (90)	Caravaggio (30); Rembrandt (30); Rubens (30)
Romanticism (90)	Friedrich (30); Goya (30); Turner (30)
Impressionism (90)	Monet (30); Pissarro (30); Sisley (30)
Cubism (90)	Braque (30); Gris (30); Leger (30)
Modern Art (90)	Klimt (30); Miro (30); Mucha (30)

We have made the experiments with different parameters for the type of descriptors; position of tiles; number of clusters.

The obtained characteristics were analyzed mainly in two directions:

- analysis of the *distribution of significance* of observed characteristics;
- predictive analysis using *accuracy* and *confusion matrices*.

For obtaining the distribution of significance of observed characteristics we used the attribute selection. These methods allow reducing the dimensionality of the data by deleting unsuitable attributes. This way the learning algorithms speed up and yield more compact, more easily interpretable representation of the target concept, focusing the attention on the most relevant variables. We used filter methods that make independent assessment based on general characteristics of the data scoring terms according to predetermined numerical functions (such as Chi square or Information Gain) that measure the “importance” of the attributes.

The accuracy presents the number of correct answers over the total number of the test instances as percentage.

The confusion matrix is usually applied as a basis for analyzing the results of the classifiers. The confusion matrix is $m \times m$ matrix, where m is the number of class labels $C_i, i = 1..m$. The rows indicate the class where the test query actually belongs to. The columns show the class label assigned to the query by the classifier. The cell (i, j) contains the percentage of queries, which are actually of class C_i but were predicted as C_j . The percentages of correctly recognized instances are represented on the diagonal.

We have processed the datasets under the procedures of attribute selection with a view to receive the order of distribution of attributes significance for prediction. We have implemented the Chi-square evaluation method. As datasets we have used different numbers of tiling from 3×3 to 7×7 tiling. We have made experiments with different numbers of clusters – 20, 40 and 60 in order to study the convenient number of used clusters.

As class value we have used “movements” and “artists’ names”. We have summarized the obtained order of attributes by different points of view – types of descriptors; positions of the tiles by width; positions of the tiles by height.

A. Evaluation of the Attributes in Respect to the Type of Underlying MPEG-7 Descriptor

Figure 4 shows the distribution of significance of MPEG-7 descriptors for class prediction. As it is shown, the Color Structure (CS) descriptor is the most informative for our datasets.

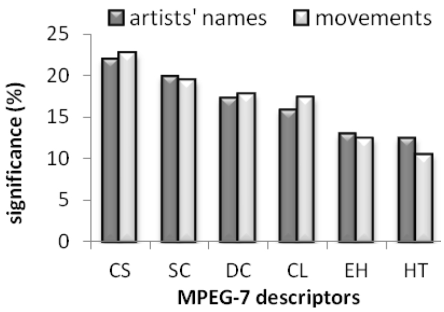

Figure 4 Average distribution of significance by MPEG-7 descriptors over datasets (different variants of tiling and clustering) using Chi-square evaluation method

The dominance of the features based on the color descriptors (CS, SC, DC, CL) leads to the following assumptions:

- The artists’ palettes, which are captured in color descriptors, are a powerful tool for creating the profiles of art painting images;
- Using this approach, texture descriptors (EH, HT) cannot produce sufficient quality attributes to present the specifics of the brushwork of the artists.

B. Optimize Spatial Granularity

We have made the analysis of the distribution of significance of the attributes that represent left/right side, respectively up/down part of the image. We have made from 3×3 to 7×7 tiling and average the results giving a half of centre tiles for odd tiling to participate on both parts.

Figure 5 shows the distributions of significance of left side and right side of the images. The construction of many classical paintings is based on central symmetry. A little superiority of the right part of the image confirms the results from psychological theories for understanding human perception [23]. We intend to use this fact in further investigation with analyzing the tiles only of the right half of the image.

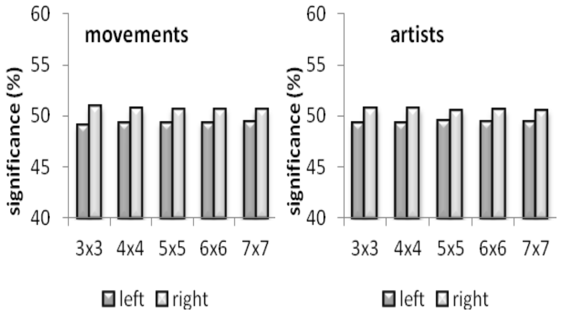

Figure 5 Distributions of significance of left side and right side of the images with different tiling

Figure 6 shows the distributions of significance of upper and lower zone of the images with different tiling. Based on these results we can conclude that upper part of the images is more informative than the lower one.

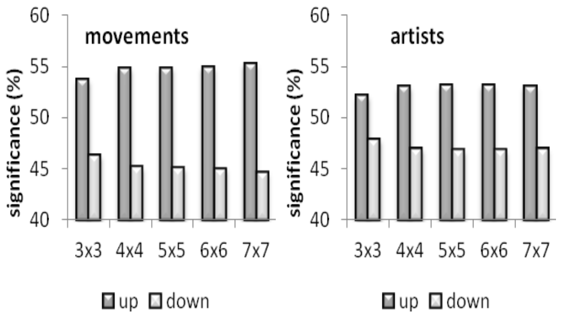

Figure 6 Distributions of significance of upper and lower zone of the images with different tiling

Similar analysis in respect to the vertical position of the tiles is shown on Figure 7. Here, it becomes clear that outer tiles (and especially border tiles) are more informative (more distinctive for different classes) than inner tiles (and especially centre tiles).

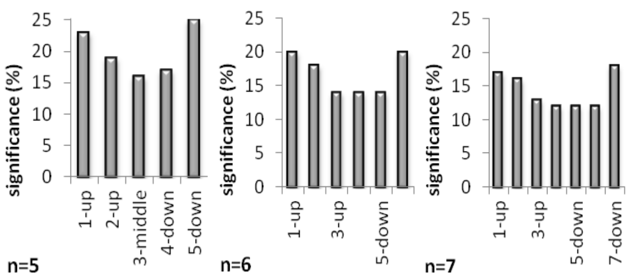

Figure 7 Distributions of significance of the tiles by position of height, $j \in 1...n$ (up to down)

This fact can also be explained with differences in the composition in different styles [23]. While the central part of the image brings objects or scene information, the borders are less burdened with this task. In order to supply the focus of the image, there are not usually specific objects found here, but only the ground patterns, which are specific for the artists or the school, in which the artists belong. These patterns capture the ground of the artists' palettes and brushworks.

Other experiments focused on establishing the appropriate number of clusters in order to receive good classification results with lower computational cost. We have run ten-fold cross-validation over the datasets.

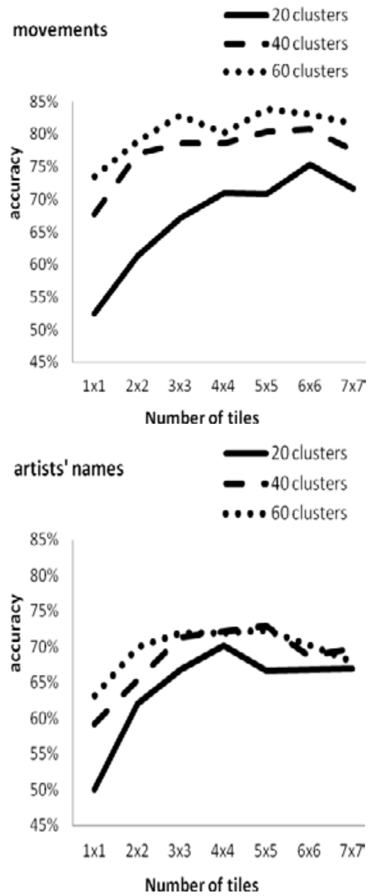

Figure 8 Classification accuracies for datasets using 20, 40, 60 clusters and different numbers of tiling

The results, displayed in Figure 8, show that using vector quantization on MPEG-7 descriptors for the entire image (i.e. 1×1 tiling) is not so informative. Better but non-sufficient results are obtained for 2×2 tiling. Tiling 3×3 is the first with relatively good results. This result is also conceptually validated by the fact that 3×3 tiling corresponds to a rough approximation of the golden ratio, which usually lies in the compositions of the art paintings. The last observed tiling 7×7 shows a decrease of accuracy, which can be explained by the fact that the pictures became too fragmented and the clusters fall not in the proper positions.

C. *Predictive Analysis of the VQ-MPEG Descriptors*

Table 2 and Figure 9 show the accuracies of different classifiers, based on VQ-MPEG descriptors with movements and artists' names as class labels. The experiments were produced with 4×4 tiles and 40 clusters using 5-fold cross validation.

TABLE II ACCURACIES OF DIFFERENT CLASSIFIERS USING VQ-MPEG7 DESCRIPTORS

Database	OneR	Jrip	J48	PGN
movements	51.33	52.33	51.33	66.00
artists' names	32.00	42.00	37.33	46.83

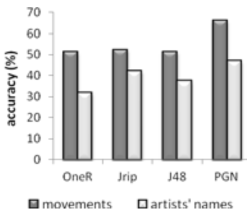

Figure 9 Accuracies of different classifiers using VQ-MPEG7 descriptors

Figures 10 and 11 visualize the confusion matrices of classification based on VQ-MPEG7 descriptors. Again we can see the best performance of PGN. Two facts immediately get attention when analyzing the movements' results. First, OneR is able to predict all movements - and independently of the classifier a black/grey downwards diagonal can be seen. These facts are in favor for the constructed VQ-MPEG descriptors. The obtained accuracies (66% for movement and 47% for artists) are at the level of accuracies (63% for movement and 49% for artists) obtained with visual features.

Analyzing the artist results the two mentioned patterns are confirmed and there is one vertical line (dark or light) and the presence of "movement" squares.

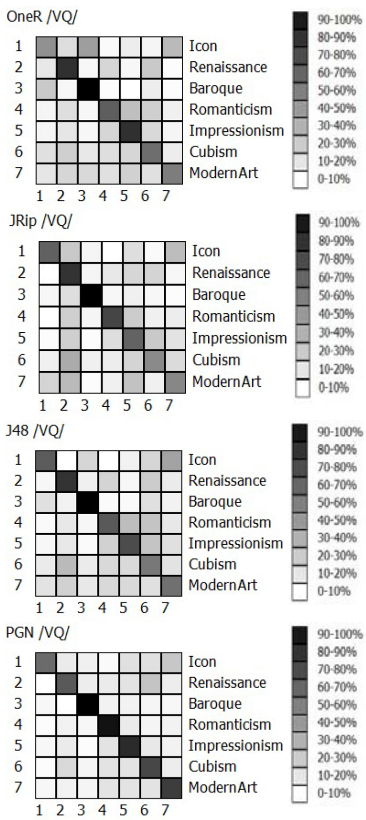

Figure 10 Confusion matrices for movements as class labels

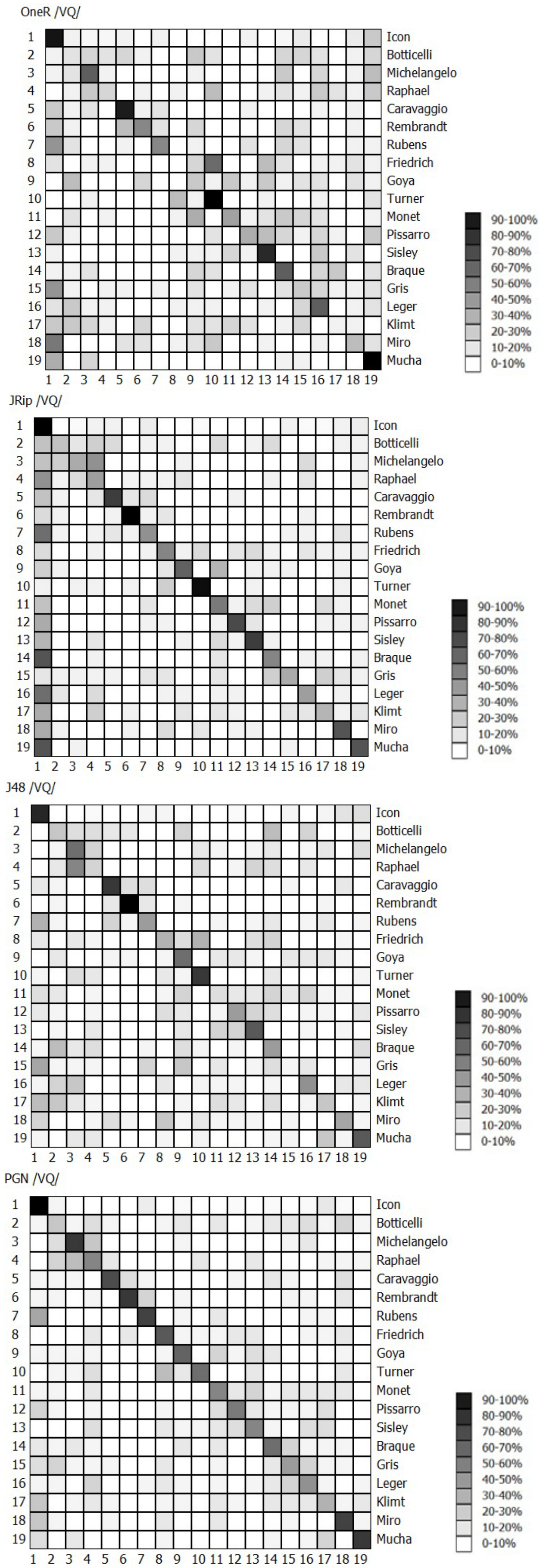

Figure 11 Confusion matrices for artists' names as class labels

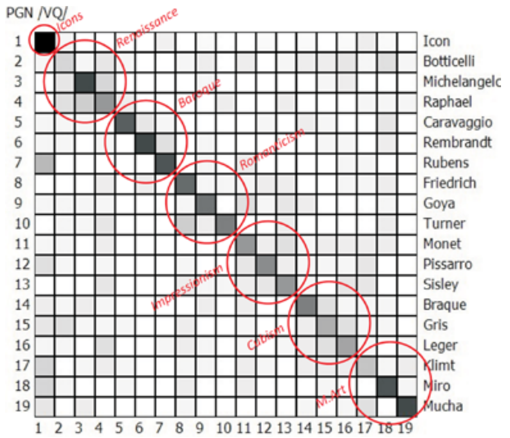

Figure 12 Visualization of confusion matrix of PGN-classifier – artists' names as class label with grouping by movements

Figure 12 shows the confusion matrix of PGN classifier for artists' names, with marks of the movements' groups. It is seen that local misclassifying within the frame of movements happens mainly for Renaissance and to some extent for Impressionism and Cubism; this confirms the proposition for similar existing features in the movements.

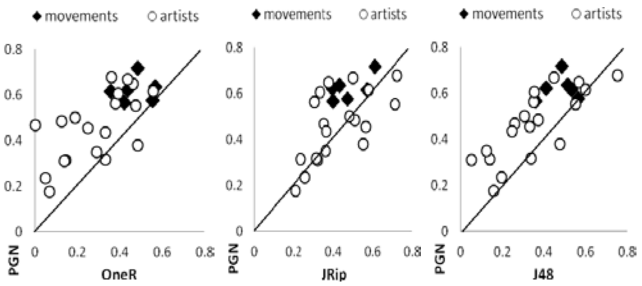

Figure 13 Scatterplots of F-measure of PGN against other classifiers

The scatterplots of F-measures of PGN against other classifiers (Figure 13) shows that PGN has not only the best accuracy, but has also better local behavior within class labels – most of the F-measures are in the upper zone of the graphics. That confirms our expectation that PGN would be well suited to predict multi-class datasets.

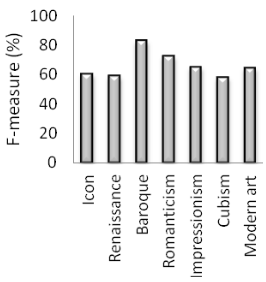

Figure 14 F-measures for PGN classifier - classes "movements"

The results (Figure 14) show that the F-measures for some movements (for instance, Baroque) are very good. The nature of a part of paintings from Romanticism, presenting landscapes [24], tends to confuse with similar paintings from Impressionism. Our expectation that Impressionism and Modern art with their different brushwork techniques will be distinctively separated by attributes, based on vector quantization of Homogeneous Texture or Edge Histogram descriptors was not confirmed, which means that for capturing brushwork specifics we have to use another approach or to find different set of features.

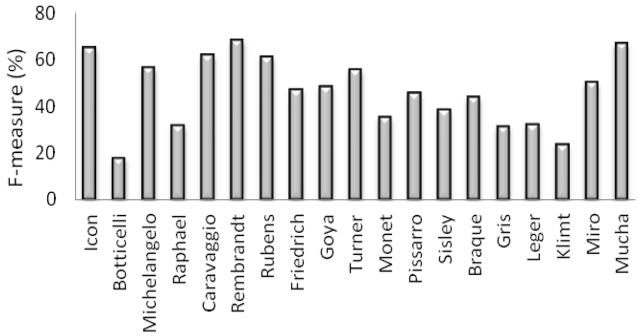

Figure 15 F-measures for PGN classifier - classes "artists' names"

Figure 15 presents the results for artists' names. It is seen the good measures for Mucha, Rembrandt, Turner, Caravaggio, and Icons. Raphael was mainly misplaced by Michelangelo, as well as Friedrich – with Sisley and Turner.

VI. CONCLUSIONS

In this article we have analyzed the added values of proposed descriptors based on vector quantization of MPEG-7 descriptors. The method allows capturing local information with a reduction of the data dimensionality. For the used data sets, 4×4 tiling and 40 clusters seemed optimal. From the analysis we learned that the artist's palettes, which are captured in color descriptors, are a powerful tool for creating the profiles of art painting images. The associative classifier PGN shows its powerfulness for categorizing multi-class datasets.

REFERENCES

[1] K. Ivanova, P. Stanchev, E. Velikova, K. Vanhoof, B. Depaire, R. Kannan, I. Mitov, and K. Markov, "Features for art painting classification based on vector quantization of MPEG-7 descriptors", *Second Int. Conf. on Data Engineering and Management ICDEM*, 2010, LNCS 6411, Springer, pp.146-153, 2012.

[2] R. Veltkamp and M. Tanase, *Content-Based Image Retrieval Systems: A Survey*. Technical Report UU-CS-2000-34, 2000.

[3] S. Kröner and A. Lattner, "Authentication of free hand drawings by pattern recognition methods", *Proc. of the 14th Int. Conf. on Pattern Recognition*, Brisbane, Australia, pp. 462-464, 1998.

[4] R. Sablatnig, P. Kammerer, and E. Zolda, "Hierarchical classification of paintings using face- and brush stroke models", *Proc. of the 14th International Conference on Pattern Recognition*, Vol. 1, IEEE CS, Washington, DC, p. 172, 1998.

[5] C. Saraceno, M. Reiter, R. Kammerer, E. Zolda, and W. Kropatsch, "Pictorial portrait indexing using view-based eigen-eyes", *Visual Information and Information Systems*, LNCS 1614, pp. 649-656, 1999.

[6] D. Keren, "Painter identification using local features and naive Bayes", *Proc. of the 16th Int. Conf. on Pattern Recognition*, Quebec, Canada, Vol. 2, pp. 474-477, 2002.

[7] O. Icoglu, G. Bilge, and S. Sanem, “Classification and indexing of paintings based on art movements”, *12th European Signal Processing Conference (Eusipco)*, Vienna Austria, pp. 749-752, 2004.

[8] T. Lombardi, C. Sung-Hyuk, and C. Tappert, “A lightweight image retrieval system for paintings”, *Proceedings of SPIE – Volume: Storage and Retrieval Methods and Applications for Multimedia*, 5682, pp. 236-246, 2005.

[9] Y. Marchenko, T. Chua, and R. Jain, *Semi-supervised Annotation of Brushwork in Painting Domain Using Serial Combinations of Multiple Experts*, Technical Report, NUS, Singapore, 2006.

[10] *International Standard ISO/IEC 15938-3 Multimedia Content Description Interface – Part 3: Visual*, http://www.iso.org/iso/ iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34230.

[11] K. Ivanova and P. Stanchev, “Color harmonies and contrasts search in art image collections”. *Proceedings of First Int. Conf. on Advances in Multimedia (MMEDIA)*, Colmar, France, pp. 180-187, 2009.

[12] K. Ivanova, “A Novel Method for Content-Based Image Retrieval in Art Image Collections Utilizing Color Semantics”, PhD Thesis, Hasselt University, Belgium, 2011.

[13] K. Ivanova, “Content-based image retrieval in digital libraries of art images utilizing colour semantics”, *Research and Advanced Technology for Digital Libraries*, S. Gradmann et al (Eds.), LNCS 6966, pp.515-518, 2011.

[14] G. Amato, C. Gennaro, F. Rabitti, and P. Savino, “Milos: a multimedia content management system for digital library applications”. *Research and Advanced Technology for Digital Libraries*, LNCS 3232, Springer, pp. 14-25, 2004.

[15] G. Karypis, *A Clustering Toolkit Release 2.1.1*. Tech. Report: #02-017, University of Minnesota, Department of Computer Science, Minneapolis, 2003.

[16] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[17] R. Holte, “Very simple classification rules perform well on most commonly used datasets”, *Machine Learning*, Vol. 11, pp.63-91, 1993.

[18] W. Cohen, “Fast effective rule induction”, *In Proc. of the 12th Int. Conf. on Machine Learning*, Morgan Kauffman, pp.115-123, 1995.

[19] R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, 1993.

[20] I. Mitov, “Class Association Rule Mining Using Multi-Dimensional Numbered Information Spaces”, PhD Thesis, Hasselt University, Belgium, 2011.

[21] I. Mitov, K. Ivanova, K. Markov, V. Velychko, K. Vanhoof, and P. Stanchev, “PaGaNe – a classification machine learning system based on the multidimensional numbered information spaces”, *World Scientific Proceedings Series on Computer Engineering and Information Science*, No:2, pp. 279-286, 2009.

[22] K. Ivanova, I. Mitov, P. Stanchev, M. Dobreva, K. Vanhoof, and B. Depaire, “Applying associative classifier PGN for digitised cultural heritage resource discovery", *Proc. of the 1st Int. Conf.” Digital Preservation and Presentation of Cultural Heritage"*, IMI-BAS, pp.117-126, 2011.

[23] R. Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*, University of California Press, Berkeley, 1974.

[24] K. Clark, *Civilisation – a Personal View*. British Broadcasting Corporation, London and John Murray, London, 1969.