APICAS – An Example of Incorporating Colour Semantic Descriptors into Digital Art Image Repositories

Krassimira Ivanova
Institute of Mathematics and Informatics – BAS, Sofia, Bulgaria
kivanova@math.bas.bg

Abstract. The paper presents the architecture of experimental Content-Based Image Retrieval (CBIR) system APICAS ("Art Painting Image Colour Aesthetics and Semantics"). This system has been developed within a doctoral thesis which aims to provide a suite of specialized tools for CBIR within a digital library of art images. The main functions in APICAS are: data ingest; visual features extraction; and data delivery. The high-level architecture suggested in this work adds a designated layer allowing CBIR functions to be used both within ingest and access to the digital objects in a digital library.

Keywords: CBIR, colour semantics, digital art image repository

1 Introduction

The development of specialized digital libraries (DL) for art images has to combine the traditional DL functionality for storage, resource discovery and access with specialized image processing tools. Such tools can be used at ingest of digitized art objects as a means to enhance their metadata in automated way, or for access if the users would like to benefit from content-based image retrieval (CBIR) or other semantic-oriented tools.

In this paper we are presenting the architecture for a specialized art image DL which integrates general digital library functionality with designated CBIR tools. The suggested architecture had been implemented and the experience from this implementation informed this work.

2 APICAS Architecture

The software system APICAS was developed in order to supply appropriate environment for testing several kinds of visual and higher level features, connected with the colour presence and interaction between colours within art images [1][2][3][4]. In Figure 1 the architecture of proposed system is shown. The main functions in APICAS are:

- data entry – establishing connections with image sources as well as supplying controlling textual metadata;
- feature extraction – such functions produces automated metadata for image labelling;

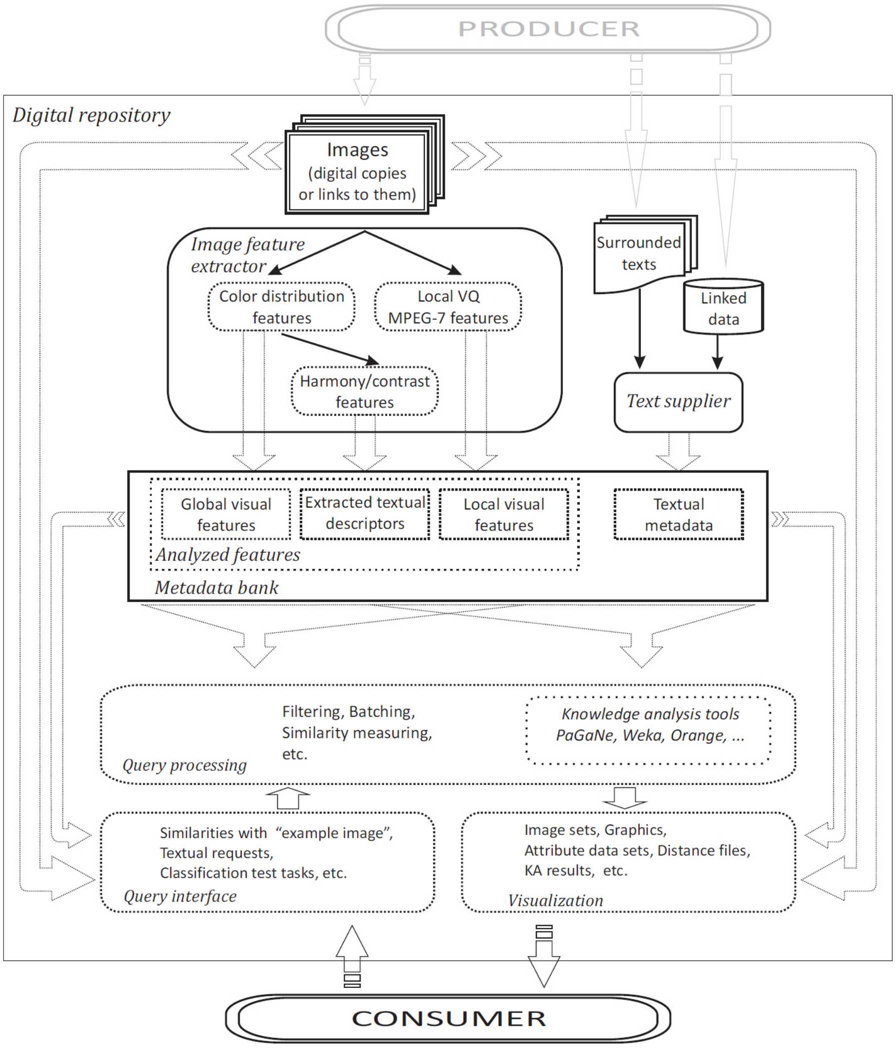

Fig. 1. APICAS architecture

- query interface – part of user-interface functions, connected with receiving of the tasks from the consumer. Here an image bank is used in order to select "an example" for searching images with greatest similarity to the selected image. The metadata bank is used for constructing a "controlled vocabulary", from which users can select desired feature(s);
- query processing – analysis of extracted metadata, their potential to meet user query for receiving images with specified colour harmonies or contrast or to be used for building artist practice profile or movement description;
- visualization – the other part of user-interface functions, connected with visualizing of received results. A variety of tools is used, such as image sets (whole

images or patches), attribute data sets, distance files, graphics, knowledge analysis results, etc.

The main goals of APICAS are two-fold:

- to analyze the possibilities of defined harmonies and contrast features for narrowing the semantic gap;
- to investigate possibilities for finding regularities between these features that can be used as semantic profile of the art paintings.

The analysis of the significance of the received characteristics and finding regularities between them can be used as discriminating semantic profile of the art paintings. It can predict several characteristics such as: the artists' names, movements, themes, techniques, etc. In this way the high level visual concepts, formed by combination of the features, can be used for narrowing the semantic gap between low-level automatic visual extraction and high-level human expression. We use data mining analysis environment PaGaNe [5][6], developed in the Institute of Mathematics and Informatics, which supplied statistical and attribute analyzing tools as well as specially designed Class-Association Rule classifier PGN.

3 APICAS Ground

The system is realized using CodeGear Delphi 2007 for Win32. As metadata storage space Arm 32, property of FOI Creative Ltd., is used. For obtaining the MPEG-7 descriptors APICAS refers to the Multimedia Content Management System MILOS [7]. For obtaining the results of multidimensional scaling we used the open component-based data mining and machine learning software suite Orange [8]. As clustering algorithm the program "vcluster", which is part of the CLUTO open source software package [9] is implemented in the system. As knowledge analysis and testing environment, we used the data mining analysis environment PaGaNe [5][6]. We use the PGN classifier, the ArmSquare association rule miner and the implemented statistical analyzing tools for checking up our results and extracting regularities for artists' and movements' styles based on the extracted attributes. For comparing received results of PGN classifier we used Waikato Environment for Knowledge Analysis (Weka) [10].

4 APICAS Functionality

The system is formed as an environment for carrying out different types of experiments. Program modules developed gradually over few years, are now combined in a common background.

Receiving the colour distributions for further establishing colour harmonies/contrast descriptors may differ by used colour models or the way of calculating presence of dominant colours in the images using:

- exact quantization of the HSL colour model for extracting colour distribution [1];
- MPEG-7 Dominant Colour descriptor as a source for determining colour distribution in the image and calculating of harmonies and contrasts features [2];

- fuzzy calculations of the HSL-artists colour model for establishing belonging of the colour characteristics into quantized bins [3].

4.1 Functions that Serve Data Entry

Choosing the Collection
The system operates with images in JPEG-format. Images, stored in one directory, form a collection. The user can choose the specific collection, changing the working folder. The system automatically scans new collections and forms database. The system allows rescanning and searching for the new images added in the collection.

Setting up Quantization Parameters and Boundaries
A special screen allows changing the settings of some parameters and boundaries, concerning quantization of hue as well as of saturation or lightness. Also access to user defined minimal thresholds, used in definition of different kinds of harmonies and contrast descriptors, are given.

Setting up Parameters for Vector Quantization
The system allows flexible apparatus for defining parameters, used in the functions, which calculate local features, connected with vector quantization of MPEG-7 descriptors. These parameters are used also in the process of visualizing the results and forming the data for knowledge analysis. The user can choose which of MPEG-7 descriptors to be put under the process of preparing the datasets for knowledge analysis. The examined tiles can be given from all surface, or only from left or right half of the image. From other side all tiles from chosen surface can be given in chess-board order starting from first tile (odd start) or from second tile (even start). The numbers of tiling by width and by height as well as the number of resulting clusters are given as parameters.

Selecting the Samples of Learning Set
The files that contain samples of learning or examining set can be made manually. The names of these files are arbitrary, which allows keeping different variants of learning and examining set during the experimental process. The system APICAS helps the creation of these files by marking which image to be included in the corresponding set. The paintings, which belong to either the learning set or the examining set, are extracted from a file selected by the user. The system is checking the sample for availability in the current collection and marks them as participants of the learning set. In this way selected paintings can be used in different collections.

4.2 Function for Supplying Textual Metadata

Textual metadata, which describe different aspects of the image content, are necessary in order to use learning algorithms and also for testing. This information can be received by different ways – filled manually or derived from the context. In the Web space this information can be extracted from the Internet page, which contains examined image. This process taken alone is separate field for investigation. Here we use simple ways for supplying the process with needed metadata using the names of the images as a source for the names of the artist and the picture, and eventually – the year of painting. A simple ontology contains the information for the movements and sub-movements and artists. The ontology contains the connections between described concepts, which allow, using extracted from the filename information to receive all additional information that can be attached to the examined paintings. Other kinds of metadata, such as theme of the paintings, genre, used techniques, etc., can be added manually.

4.3 Functions for Calculating Visual Characteristics

Calculating Colour Distribution

A special function calculates three dimensional array containing coefficients of participation of colours with correspondingly measured characteristics of the image. The function is used in the process of examining colour distribution as well as part of the process of defining harmonies and contrast descriptors. It gives each pixel from observed area, convert the colour value from RGB-colour model to colour coordinates in HSV or HSL colour model and as result of applying selected quantization increase the presence of colour with quantized coordinates. In the case of fuzzy quantization, the increasing catches neighbour coordinates with corresponded value. Finally the normalization of the values in the array is made. The output is three dimensional array which contain colour distribution by selected dimensions [1].

Estimating Harmonies' and Contrasts' Descriptors

Special functions for calculating defined harmonies' and contrast' descriptors are realized in the system. The exact algorithms for estimating these descriptors are explained in [2] and [3].

Establishing Local Features, Based on Vector Quantization of MPEG-7 Descriptors over the Tiles of the Image

This algorithm is presented in [4]. Several functions are connected with this process:

- *choosing learning samples*: the function reads text file that contains learning samples, check the images for existing in current collection, and writes correct samples in a database;
- *clustering*: this function passes into several steps: (1) creating tiles from images of the learning set with given parameters (numbers of tiles by width and by height); (2) calculating MPEG-7 descriptors for these tiles using MILOS system; (3) for each MPEG-7 descriptor building a dataset, which contains corresponding feature vector for each chosen tile from the learning set; (4) executing clustering procedure with selected number of clusters; (5) calculating the centroids of each cluster; (6) assigning the corresponded number for each tile and writing in a database;
- *finding most similar tiles to the centroids*: it is not used in the straight process of finding local features. It is connected with visualizing function of representatives of cluster values for corresponded MPEG-7 descriptor. The system finds the tile from the image base, which is closest to the centroids of examined descriptor;
- *defining corresponded features for the rest of the images*: for tiles of the images, which were not in the learning set, the membership of their centroids is calculated

and the number of the corresponding cluster is assigned as a value of the tile. The result is written in the same way as for the images from the learning set.

For each MPEG-7 descriptor two types of similarity measures are realized: first is based on L_1-metric; second is based on L_2-metric, but for some of descriptors specific similarity measure can be used. For instance for *Scalable Colour* function takes into account the significance of the order of coefficients [11]. For *Edge Histogram* [12] proposed one extension in order to capture not only the local edge distribution information but also semi-global and global ones.

4.4 Functions, Connected with Output Information

Examining Colour Distribution

One class of functions is directed to carry out the analysis of distribution of colour characteristics in the images – hue, saturation or luminance, or combination of them. These functions are firstly introduced in [1], where the analysis is made on the base of HSL colour space. Further developments of these and additional features has shown that quantization of hue in respect of artists colour wheel is more appropriate, because of this additional possibility to make analyses based on constructed by us HSL-artist colour model is added.

The analysis can be conducted over the whole array (all three dimensions); a simple projection of selected characteristics; or projection of two characteristics (for instance, Hue and Luminance). The functions can be executed for:

- all pictures in the collection;
- all movements or for concrete movement, presented in the collection;
- all sub-movements or for selected sub-movement;
- all artists or for chosen artist in the collection.

The function makes selecting and/or sectioning of the images in the collection. For receiving colour distribution of given image it refers to already discussed function for calculation colour distribution. If the calculations are already made, the function can overcome this step and only visualize features using stored information.

Visualizing Extracted Colour Harmonies and Contrast Features

The extracted descriptors (from the content and from the context) can be observed in a grid. The user can sort it by any selected feature. Pointing on the exact image, the user can see the extracted metadata, connected to this image – an example is given in Figure 2.

The user can set different conditions on the extracted descriptors and receive the images that satisfy these conditions. The results can be obtained in thumbnail form, where the images can be seen or in a file, where selected images can be additionally batched using other features, chosen by the user.

The system allows searching within a collection of images, which has

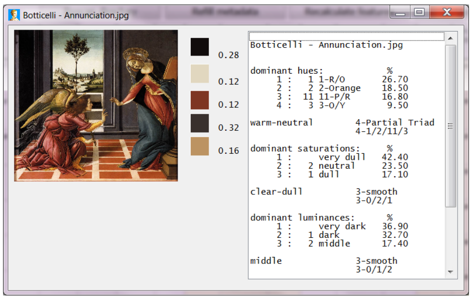

Fig. 2. Harmonies/contrasts
for the painting "Annunciation" by Botticelli

specific combination of the colours, defined by some harmony or contrast.

Another branch of the system allows creating a datasets, containing extracted attributes or selected part of them labelled with chosen profile such as artist name, movement, scene-type. These datasets are used for further analysis by data mining tools for searching typical combinations of characteristics, which form profiles of artists or movements, or reveal visual specifics, connected to the presented thematic in the images.

Visualizing the Results of Clustering

The system allows viewing of the results of clustering procedure showing all tiles, which belongs to selected number of cluster for specific MPEG-7 descriptor.

Another function allows showing of the tiles from the learning set, which are closest to the centroids of given clustering for chosen MPEG-7 descriptor. The function uses the results of the already discussed function for finding most similar tiles to the centroids. The idea is that these tiles can be used later as elements in a visual lexicon for representing specifics of some image profiles.

Preparing Data for Multidimensional Scaling

Multidimensional scaling is realized in the system Orange. In order to implement special distances the system needs to receive not only primary data, but also distance matrix between the instances. The functions in APICAS are aimed to prepare corresponded data for multidimensional scaling. The distance matrices are calculated using Earth Mover Distance (a kind of transport task algorithm) with special defining the distances of underlying data as follows:

· *For Colour Distribution Features*

If $I = (h_i, s_i, l_i)$ and $J = (h_j, s_j, l_j)$ are two points, where h_i and h_j are their hue values, s_i and s_j are their saturations, and l_i and l_j are their luminances.

The distance between I and J is calculated as Manhattan distance (L^1-metric) between distances of theirs characteristics: $d(I,J) = d_h(h_i,h_j) + d_s(s_i,s_j) + d_l(l_i,l_j)$.

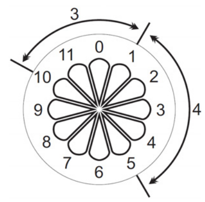

Fig. 3. Angular distance

Because of the angular type of hue characteristic, the distance is calculated as (Figure 3): $d_h(h_i,h_j) = \min(|h_i - h_j|, maxh - |h_i - h_j|)$.

For saturation and luminance distances the L^1-metric are used.

· *For Harmonies/Contrast Descriptors*

In spite of the categorical nature of these descriptors, their values can be partially ordered – for instance "Monochromatic" is closer to "Analogous" than to "Complementary". Special matrices that describe the distances between each two values of given descriptor are implemented.

· *For Local VQ MPEG-7 Features*

The resulting features of this process are strictly categorical and cannot be ordered in any manner. We use Jaccard coefficients as a ground for establishing similarity measures.

Preparing Data for Knowledge Analysis
Several functions for preparing datasets for further statistical and data analyses, which contain features calculated by proposed algorithms, are given. The functions allow user to choose which of attributes to be included in dataset. As class label artist's name, movement, sub-movement, scene-type can be given. The results are prepared in different formats, convenient for PaGaNe, Weka or Orange. An additional function allows creating dataset for associative rule miner ArmSquare, which makes frequency analysis over transactional datasets. The structure of the file differs from previous ones, which operate with rectangular datasets.

5 Conclusion

We have proposed architecture of an experimental CBIR lab-system, aimed at analyzing different types of visual features, which strive to narrow the semantic and abstraction gap between low-level automatic visual extraction and high-level human expression. We have explained the structure and functionality of the software system "Art Painting Image Colour Aesthetics and Semantics" (APICAS). All these functions we have realized and put into common environment APICAS.

The vividness of proposed features will open the door for indexing and searching in paintings repositories, according to such characteristics of their content. The proposed features can be used as a step in the transition from Web 2.0 to Web 3.0. Without a breakthrough technology, superior Web 3.0 tools will be more difficult to develop than their counterparts for Web 2.0. This will be part of creation of new tools which will offer society new greater sophistication, complexity, and functionality.

Acknowledgements

This work was supported in part by Hasselt University under the Project R-1875 and by the Bulgarian National Science Fund under the Project D002-308.

6 References

1. Ivanova, K., Stanchev, P., Dimitrov, B.: Analysis of the Distributions of Color Characteristics in Art Painting Images. Serdica Journal of Computing, 2/2, Sofia, 111-136 (2008)
2. Ivanova, K., Stanchev, P.: Color Harmonies and Contrasts Search in Art Image Collections. First International Conference on Advances in Multimedia (MMEDIA-2009), Colmar, France, 180-187 (2009)
3. Ivanova, K., Stanchev, P., Vanhoof, K.: Automatic Tagging of Art Images with Color Harmonies and Contrasts Characteristics in Art Image Collections. Int. Journal on Advances in Software, vol. 3, no.3&4, 474-484 (2010)
4. Ivanova, K., Stanchev, P., Velikova, E., Vanhoof, K., Depaire, B., Kannan, R., Mitov, I., Markov, K.: Features for art painting classification based on vector quantization of MPEG-7 descriptors. 2nd Int. Conf. on Data Engineering and Management ICDEM, India, LNCS, (2010)
5. Mitov, I., Ivanova, K., Markov, K., Velychko, V., Vanhoof, K., Stanchev, P.: PaGaNe – a Classification Machine Learning System Based on the Multidimensional Numbered Information Spaces. Fourth Int. Conf. Intelligent Systems and Knowledge Engineering

(ISKE-2009), Hasselt, Belgium, Printed in World Scientific Proceedings Series on Computer Engineering and Information Science, No:2, 279-286 (2009)
6. Mitov, I., Ivanova, K., Markov, K., Velychko, V., Stanchev, P., Vanhoof, K.: Comparison of Discretization Methods for Preprocessing Data for Pyramidal Growing Network Classification Method. In New Trends in Intelligent Technologies, Sofia, 31-39 (2009)
7. Amato, G., Gennaro, C., Rabitti, F., Savino, P.: Milos: a Multimedia Content Management System for Digital Library Applications. Research and Advanced Technology for Digital Libraries: Volume 3232/2004 in LNCS, Springer Berlin/Heidelberg, 14-25 (2004)
8. Demsar, J., Zupan, B., Leban, G., Curk, T.: Orange: From Experimental Machine Learning to Interactive Data Mining. White Paper (www.ailab.si/orange), Faculty of Computer and Information Science, University of Ljubljana (2004)
9. Karypis, G.: CLUTO: A Clustering Toolkit Release 2.1.1. University of Minnesota, Department of Computer Science, Minneapolis, Technical Report: #02-017 (2003)
10. Witten, I., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. 2nd Ed., Morgan Kaufmann, San Francisco (2005)
11. Herrmann, S.: MPEG-7 Reference Software. Munich University of Technology. http://www.lis.e-technik.tu-muenchen.de/research/bv/topics/mmdb/e_mpeg7.html
12. Won, Ch., Park, D., Park, S.: Efficient use of MPEG-7 edge histogram descriptor. ETRI Journal, 24/1, 23-30 (2002)