DOCTORAATSPROEFSCHRIFT

2011 | Faculteit Wetenschappen

A novel method for content-based retrieval in art image collections utilizing color semantics

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, te verdedigen door:

Krassimira IVANOVA

Promotor: prof. dr. Koen VANHOOF (UHasselt)
Copromotor: prof. dr. Peter STANCHEV
(Institute of Mathematics and Informatics, Bulgarije)

universiteit hasselt

DOCTORAATSPROEFSCHRIFT

2011 | Faculteit Wetenschappen

A novel method for content-based retrieval in art image collections utilizing color semantics

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, te verdedigen door:

Krassimira IVANOVA

Promotor: prof. dr. Koen VANHOOF (UHasselt)
Copromotor: prof. dr. Peter STANCHEV
(Institute of Mathematics and Informatics, Bulgarije)

D/2011/2451/49

universiteit hasselt

Hasselt University

Faculty of Science

November 2011

A Novel Method for Content-Based Image Retrieval in Art Image Collections Utilizing Colour Semantics

A thesis submitted for the degree of Doctor of Science by:

Krassimira Ivanova

Promoter: Prof. Dr. Koen Vanhoof

Co-promoter: Prof. Dr. Peter Stanchev

universiteit hasselt

Institute of Mathematics and Informatics

BULGARIAN ACADEMY OF SCIENCES

Abstract

The field of image retrieval has to overcome a major challenge: it needs to accommodate the obvious difference between the human vision system, which has evolved genetically over millenniums, and the digital technologies, which are limited within pixels capture and analysis. We have the hard task to develop appropriate machine algorithms to analyze the image. These algorithms are based on completely different logic and "instruments" compared to the human process of perception, but would give similar results in interpreting the input image. In the context of this thesis the challenges are even bigger because we focus our efforts on image analysis of the aesthetic and semantic content of art images. Naturally, the interpretation of what we – humans – see is hard to characterize, and even harder to teach to a machine. Yet, over the past decade, considerable progress has been made to make computers learn to understand, index and annotate pictures representing a wide range of concepts.

Convenient image capture techniques, inexpensive storage, and widely available dissemination methods have made digital images a widespread information format. This increased availability of images is accompanied by a need for indexing and retrieval tools [Chen et al, 2005a]. The digital repositories of cultural heritage objects can employ techniques similar to the ones used in any general-purpose image processing environment in order to provide standard functionality for searching objects. However, cultural heritage objects are rich in content describing events, monuments, places, people; and they are distributed across different locations. The users can formulate queries using different modalities such as free text, similarity matching, or metadata; one important current trend is the use of linked data [Gradmann, 2010]. To be addressed properly, the specificity of these objects suggests that support should be provided for some additional tasks which need to be reflected in the image analysis methods and techniques used. In the area of art paintings retrieval the sensory, semantic and abstraction gaps are a predominant problem.

Colour perception influences multiple aspects in image analysis of art colour ranging from the physical nature of light, to physiological specifics the human vision system, psychological peculiarities and socio-cultural ground, in which the artwork was created as well as the one of the receiver of the message which the artist expressed in his work.

In the process of perception, figure-ground separation is the first cognitive step. Colour plays an important but secondary role; however colour responses are more connected to human emotions than to rational mind. This property itself makes the colours' influence on human perception pivotal. The presence of one or more colours in different proportions conveys different messages, which can increase or suppress the perception of the observed objects. In the field of image retrieval the ways of perceiving colours and colour combinations as similar or not similar is crucial when one has to extract images based on a criterion reflecting the level of emotional perception, or to search for any specific characteristics of the artist's expressiveness.

The colour impact on people depends on multiple factors with physical laws and physiology being only the beginning. Further along this process psychological perception plays an important role; with both the particular psychological state and the socio-cultural environment in which a character of a person is composed playing a role. Perception of colour brings the whole emotional and mental identity of the artist as well as of the observer, joining their intelligence, memory, ideology, ethics, aesthetics and other sensations.

The primary goal of this work is to make an analysis of a range of theories on the existing interconnections in colour combinations and to formalize them in order to allow for extracting them from digitized artworks. We use Itten's colour theory s a basis of our research.

Global low-level features reflecting the quantitative presence of quantized colour characteristics are suggested as a helpful instrument for the formal definitions of harmonies/contrast descriptors.

We propose a classification of colour harmonies and contrasts, which is consistent with the human perception of visual expression and conforms to the possibilities of automatic extraction of visual information from digitised copies of art images. The classification combines the three main colour characteristics which are closest to human perception – hue, saturation and lightness. We provide the formal definition of extracting such descriptors from images.

The third group of descriptors, based on vector quantization of MPEG-7 descriptors over the partitioned images, are introduced in order to analyze what are the possibilities of capturing more detailed information for semantic and abstraction content of art images based on the MPEG-7 descriptors with significant dimensionality reduction.

Further, we propose architecture of an experimental CBIR system, where the extracted visual features are combined with the extraction of textual metadata on the examined art images. The textual metadata are necessary in order to apply supervised learning methods. A designated software system "Art Painting Image Colour Aesthetics and Semantics" (APICAS) was developed as an appropriate environment for implementing the algorithms suggested and for conducting experiments.

As a testing collection, we have created an experimental dataset of 600 paintings by 18 representative artists from different movements of the West-European fine arts from Renaissance, Baroque, Romanticism, Impressionism, Cubism, and Modern Art and Eastern Medieval Icons.

Using the functionalities of APICAS we have conducted several kinds of experiments.

The first set of experiments focused on the analysis of the colour distribution characteristics. We discover some specifics of given movements, which distinct them from the others. The examination of the features for all paintings revealed some common trends for art images. These results had been used later as a normalization factor in the process of defining the colour harmonies and contrast descriptors.

The second set of experiments is focused on results of automatic annotation of the images with harmonies' and contrasts' descriptors. These high level features can be used not only in the processes of categorization focused on cultural influences and specific techniques, but also for extracting images with specific characteristics closely connected with the emotional responses evoked by an image. These features refer to corresponding elements of the abstract space of the image content.

The third set of experiments is focused on examining the significance of local features, extracted by the proposed method, and more specifically on the type of underlying MPEG-7 descriptors; position of tiles; as well as number of clusters. They aim to decrease further examined features in order to accelerate the process without significant loss of accuracy in the classification tasks.

A further set of experiments focuses on the evaluation of classification accuracy achieved by different types of classifiers. For the purposes of this work we conducted experiments with PGN classifier developed by our team as a specialised classification tool. The goal was to confirm the hypothesis that the PGN classifier has good representation for such kind of features compared with the other well known classifiers, such as OneR, JRip, and J48.

To illustrate our hypothesis, the last group of experiments seeks to show the results of classification by PGN on the basis of painting periods in Goya's creative works.

The work conducted in this thesis demonstrates the possibilities of narrowing the semantic gap using an appropriate set of defined features in combination with machine learning algorithms for upgrading the concepts. This paves the way for future research which may explore two directions:

- Expanding the potential of the scope of examined visual features for defining higher-level concepts;
- Widening the range of knowledge acquisition methods, which can be used for extracting such concepts.

Work in these areas will be helpful in the transition from Web 2.0 to Web 3.0. In the era of Web 3.0 bridging the semantic gap is crucial. Finding appropriate combination of retrieval methods and techniques, which can lead to high quality image discovery, is a core problem in this domain and we hope that our work contributes to address part of these issues especially in the case of vast collections of digital images.

Acknowledgements

This work was supported by Hasselt University under the Project R-1875 "Search in Art Image Collections Based on Colour Semantics" and by the Bulgarian National Science Fund under the Project D002-308 "Automated Metadata Generating for e-Documents Specifications and Standards".

I would like to express my gratitude to the Institute of Mathematics and Informatics, Bulgaria and Hasselt University, Belgium for providing good conditions for establishing this work, and especially to the promoters of this study Prof. Koen Vanhoof from Hasselt University, Belgium and Prof. Peter Stanchev from Institute of Mathematics and Informatics, Bulgaria.

I much appreciate the continuous support for my work and the help in the field of digitisation of cultural heritage of my colleague and friend Milena Dobreva from the University of Strathclyde, United Kingdom.

My special thanks go to the artist Krassimir Zhelev, who patiently introduced me to painting techniques and helped me to choose the representative experimental collection.

I would also like to mention several colleagues who encouraged and supported my research: Iliya Mitov, PhD student at the Institute of Mathematics and Informatics, Bulgaria, with whom we developed the data mining environment PaGaNe used to test the classification accuracy of the selected attributes in this work; Eugenia Velikova from Sofia University, Bulgaria for the helpful collaboration in statistical analysis; Levon Aslanyan and Hasmik Sahakyan from the Institute for Informatics and Automation Problems of NAS, Armenia for fruitful discussions and collaboration in studying, establishing and implementing of data mining techniques; and other colleagues from the academic community, which supported not only my specific work but also my personal development as a researcher as well as opened my eyes for the colourfulness of the world of knowledge as much as of the world surrounding us.

I am very much in debt to my godparents Prof. Victor Gladun and Dina Gladun, which constantly encourage and support me to choose the right directions in my scientific and personal life.

And finally, I would like to express my deep gratitude to my mother and sister whose love and support during all my way had been outstanding, and to the sacred memory of my father who believed in me. I carry fire in my heart, my sister brought me the emollient force of the water, my mother taught me how to stay firmly on the Earth, and my father gave me the wings to fly.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAT	Art and Architecture Thesaurus
ACQUINE	AesthetiC QUality InfereNce Engine
ALIP	Automatic Linguistic Indexing of Pictures
ALIPR	ALIP real-time engine
APICAS	Art Painting Image Colour Aesthetics and Semantics
ART	Angular Radial Transform
Caliph	Common and Lightweight PHoto Annotation
CAR	Class Association Rules
CBIR	Content-Based Image Retrieval
CDWA	Categories for the Description of Works of Art
CIMWOS	Combined Image and Word Spotting
CL	Colour Layout
CLUE	CLuster-based Image rEtrieval
CMY	Cyan-Magenta-Yellow
CMYK	Cyan-Magenta-Yellow-blacK
CONA	Cultural Objects Name Authority
CS	Colour Structure
CSS3	Cascading Style Sheets 3
DC	Dominant Colour
DCT	Discrete Cosine Transform
EH	Edge Histogram
EMD	Earth Mover Distance
Emir	Experimental Metadata-based Image Retrieval
EOF	Empirical Orthogonal Function
GLOH	Gradient Location and Orientation Histogram
GNU	acronym for "GNU's Not Unix"
GoF	Group of Frames
GoP	Group of Pictures
HMMD	Hue-Max-Min-Diff
HSL	Hue-Saturation-Lightness
HSV	Hue-Saturation-Volume
HT	Homogeneous Texture
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization

IST	Information Society Technologies
JTC	Joint Technical Committee
k-NN	k-Nearest Neighbours
LESH	Local Energy based Shape Histogram
LIRE	Lucene Image REtrieval
LLE	Locally Linear Embedding
M4ART	Multimedia for Art ReTrieval
MDS	Multidimensional Scaling
Mirror	MPEG-7 Image Retrieval Refinement based On Relevance feedback
MPEG	Moving Picture Expert Group
MUFIN	Multi Feature Indexing Network
PCA	Principal Component Analysis
PGN	Pyramidal Growing Networks
PicSOM	Picture Self-Organizing Map
QBIC	Query by Image Contents
RDFS	Resource Description Framework Schema
RGB	Red-Green-Blue
RYB	Red-Yellow-Blue
SC	Scalable Colour
SIFT	Scale-Invariant Feature Transform
SIMPLIcity	Semantics-sensitive Integrated Matching for Picture LIbraries
SURF	Speeded Up Robust Features
SVD	Singular Value Decomposition
TGN	Thesaurus of Geographic Names
ULAN	Union List of Artist Names
VizIR	Visual Information Retrieval
VQ	Vector Quantization
VRA	Visual Resource Association
W3C	World Wide Web Consortium
Weka	Waikato Environment for Knowledge Analysis
YCrCb	YCrCb model –Y: Luma component, Cr and Cb: colour components

1 Introduction

1.1 The Art and the Digital Space

Digitalized art collections allow the user to immerse in an ocean of accumulated cultural artefacts. In previous times 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.

Chen et al. argue that "Research on significant cultural and historical materials is important not only for preserving them but for preserving an interest in and respect for them. Research on digital imagery for significant cultural and historical materials is imperative because of (1) its role in promoting cultural understanding, (2) its relevance to education at all levels, (3) its potential impact on many related sciences and engineering, (4) its potential to improve consolidation of such works of art, and (5) its contribution to improving our understanding of works of art themselves by knowing more about them." [Chen et al, 2005a].

Modern digital technologies made a reality large virtual exhibitions of works from multiple cultures. One of the great advantages of digitization is the improved accessibility of cultural artefacts, including images of art.

Several huge digitization projects are focused on this. For instance, Google Books signed contracts with several of the biggest USA libraries and with the National Archives in Washington (Figure 1)[1] (downloaded from) in order to digitize and store fractions of their huge collection of materials. Another similar example is the Yorck Project which provides access to 10 000 scanned images from public domain paintings (all of the artists passed away over 70 years ago) and thus expands the portfolio of pictures in Wikipedia Commons. In the recent years, Europeana project which is a showcase digital library supported by the

[1] http://www.nytimes.com/imagepages/2007/03/10/business/ 11archive.chart.ready.html

European Commission aims to unite the efforts of European institutions building on Europe's rich heritage, combining multicultural and multilingual environments with technological advances and new business models.

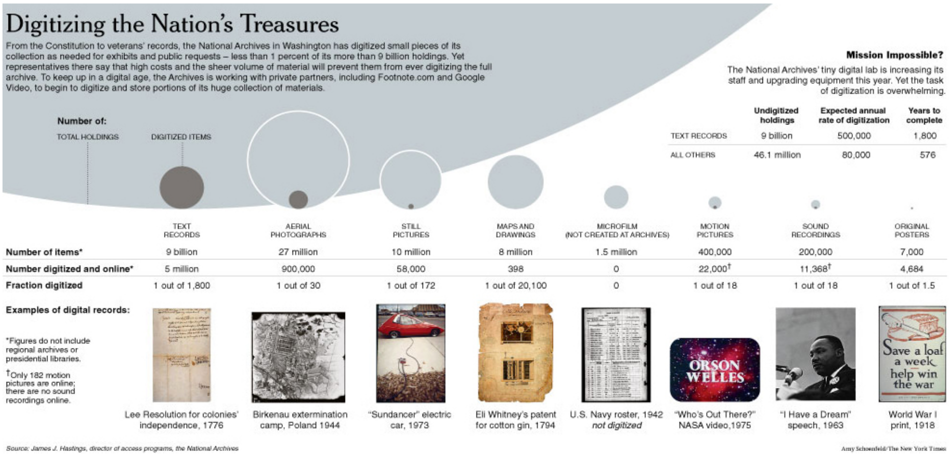


Figure 1. The Process of Digitization of Cultural Heritage

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 is 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). However, the typical digital library contains thousands of images. Very recent publicly-available systems, such as Tiltomo[2] or Alipr[3], claim to index hundreds of thousands of images. There is a high discrepancy between these numbers and the volumes of images available on the current Web.

1.2 Content Based Image Retrieval

Content-based image retrieval (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

[2] www.tiltomo.com
[3] www.alipr.com

images stored in the digital resource; current CBIR systems evolved and currently are helpful assistants when searching content by visual as well as by semantic similarity thus expanding the earlier work on image retrieval.

After decades of research, many efforts are aimed at converging text-based and content-based search technologies in real-world image retrieval. For example, text encoding techniques could be used to provide semantic description of content. Automated search can begin with user-defined constraints on the search domain. It can also accommodate definitions of semantic structures as relationships of concepts allowing high-level content-based retrieval, which can be integrated with existing retrieval techniques to better facilitate user access. Automated image analysis can also been used to enrich metadata providing additional text annotation on features, extracted from the visual content. The availability of a huge amount of digital material of cultural heritage requires the investigation of new cost-saving and effective methods for annotation and retrieval that are easy to use for most users. Image analysis processing techniques are a powerful means of extracting useful information from content. The current efforts seek to use extracted information on different levels, including low-level information, such as colour and texture; or primitives, such as salient points, corners and shapes; as well as higher-level information, such as objects, scene content, subject description, for constructing higher levels for understanding the meaning of the images under different aspects [Hurtut, 2010]. Content annotation based on pixels can be used to perform search operations from objective measures and descriptors of the visual content. However, effective descriptors which would be identical to the human perception and feeling are still required; and particular attention should be paid to the computer science "semantics" of images and scenes

1.3 Objectives of the Dissertation

The primary goal of this dissertation is to provide a detailed analysis of the colour theories about existing interconnections in successful colour combinations and to formalize them in order to make possible theirs extraction from digitized artworks. As a basis of the research we use Itten's colour theory.

The global low-level features that reflect quantitative presence of quantized colour characteristics are examined in order to be used as a basis in formal definitions of harmonies/contrast descriptors. We argue that analysis of such characteristics will outline some common trends for art images. Such results can be used at a later image analysis stage as a normalization factor in the process of defining colour harmonies and contrast descriptors.

In order to select descriptors that reflect Itten's harmonies and contrast laws we start with making a classification of such descriptors; it is consistent with the human perception of visual expression and is fine-tuned to the requirements for automatic extraction of visual information from digitized art images. We also provide a formal workflow of extracting such descriptors from images. To evaluate this approach, we also conducted a series of experiments in order to examine the possibilities of using these high level features in the process of image retrieval in connection with the emotional responses evoked by images.

The descriptors we suggest to use are based on the vector quantization of MPEG-7 descriptors over the partitioned images. We further analyze the possibilities of capturing more detailed information on the semantic and abstraction content of art images based on the MPEG-7 descriptors with significant dimensionality reduction.

Further, we analyze the use of class association rule algorithms, respectively PGN classifier, to extract rules that can be used as a profile of given class, such as movement, artist, artist period, genre, scene type, etc.

Finally, we present the software implementation of the proposed descriptors in a system called APICAS which had been used in our experiments.

1.4 Outline of the Dissertation

The dissertation is structured in eight chapters and appendix as follows:

1. Introduction
2. CBIR in Art Image Collections – the State of the Art
3. Real and Digitized Colour World
4. Some Examples of CBIR Systems
5. Proposed Set of Visual Attributes
6. The System APICAS
7. Experimental Results
8. Conclusions and Future Work

A brief overview of the content is given below.

Chapter 2 provides a literature overview of current work on image retrieval. The chapter discusses Content-Based Image Retrieval (CBIR) as alternative and extension of Text-Based Retrieval.

This overview is followed by introduction into current efforts in the area of resource discovery for art images. Digitized art images are important part of the online cultural heritage; however in the last decades the primary target of digitization and online accessibility projects were materials from libraries, museums, archives and audiovisual archives. Digital art images present an important challenge for the resource discovery of images since the methods applied in this domain should be a combination of queries executed on metadata and image retrieval methods.

We also discuss the existing gaps, which occur in different stages. The very first one is introduced in the process of digitization of real objects (Sensory Gap), the next one caused by lack of coincidence between semantic concepts, used by human users to express the requests and low-level visual data exploited by the systems (Semantic Gap), and concluding with more complicated gaps when user cannot express his/her requests because they are expressed in a way which is still not supported by the retrieval system (Subjective Gaps) or because the interpretation of the perceived information is highly subjective view on Arts and depends on particular user in a given situation (Aesthetic Gap). The abstract aspects, which are specific to art images and differ from the semantic challenge, also need to be mentioned. As a summary of these user-related issues in CBIR, we suggest a taxonomy of art image content as a source for satisfying different aspects of user needs and expectations.

Further in this chapter, we discuss the typical process in a CBIR workflow. A feature design is used as a mathematical description of an image for the retrieval purposes, hence as its signature. Visual features, which can be obtained and are organized in different types of data structures, usually include multiple attributes. This process inevitably raises the "curse of dimensionality" [Bellman, 1961], which impose applying different types of data reduction including dimensionality reduction as well as numerosity reduction. Regardless of the specific details of the task addressed – search by association, target search or category search [Stanchev et al, 2006] the process aims to measuring similarity between objects. While working on the similarity measures between two images, several processes of finding similarities on different levels and data types had been tried. A brief overview of different types of similarity measures, popular in image retrieval applications is also presented.

Some techniques, used for improving image retrieval, are discussed. *Clustering and categorization* are data mining techniques, which increasingly accelerate the searching process in image databases. *Unsupervised clustering techniques* are most helpful when handling large, unstructured image repositories such as the Web. *Image categorization* (classification) is advantageous when the image database is well specified, and labelled training samples are available. *Relevance feedback* is very important step in image retrieval, because it provides a compromise between a fully automated, unsupervised system and one based on subjective user needs. It is a query modification technique which attempts to capture the user's precise needs through iterative feedback and query refinement. Multimodal fusion is linked to the integration of information in human-machine interactive systems where several communication modalities are proposed to the user. The user can best describe his/her queries by a combination of media possibilities. This motivates the need for multimodal fusion as a technique for satisfying such user queries. Finally, ontologies provide useful apparatus for describing interconnections between concepts in a given area and such way to allow facilitating of response of the user requests.

Chapter 3 presents a brief review of colour theories from different points of view, which became a ground for our further study. Art paintings are connected with the highly complex area of colour perception. Physiological ground of the colour perception is discussed as a starting point for focusing search in art painting images. A brief historical overview of previous work on searching for colour interconnections and mutual influences is provided. Those colour models which are most suitable for representing the colours from human point of view are presented.

Chapter 4 discusses some existing art image analysis systems. Most of them aim to analyse high-resolution multi-spectral digital copies of the images in order to cover automatic artwork analysis techniques for different applications such as virtual restoration, image retrieval, studies on artistic praxis, authentication etc.

In addition some image retrieval systems using MPEG-7 descriptors are discussed. Despite the generic purpose of some of them they can be successfully implemented in the field of digitization of cultural heritage in order to fill some parts of the semantic gap; they also can contribute to reduce cultural and technical aspects of the abstraction gap.

Chapter 5 presents in detail the proposed features, which help to narrow the semantic and abstraction gap between low-level automatic visual extraction and high-level human expression. We propose three types of visual features:

1. Visual features, which represent colour distribution in art images.
2. Global higher-level features that reflect colour harmonies and contrasts in art images – we propose the classification of harmonies and contrasts in accordance to the Ittens' theory from the point of view of three main characteristics of colour – hue, saturation and luminance and give the formal description of defined harmonies and contrasts.
3. Local colour and texture features, based on vector quantization of MPEG-7 descriptors. We present the method for extracting local features that capture local colour and texture information, based on tiling the image and applying vector quantization of MPEG-7 descriptors, calculated for the tiles of the image.

Chapter 6 presents the implemented software APICAS ("Art Painting Image Colour Aesthetics and Semantics"). APICAS was developed in order to supply appropriate environment for realizing examined algorithms and conducted experiments.

The main structure of folders and on files, used in APICAS is provided. The functionality of the system is also presented. The main functions in APICAS can be grouped as follows:

1. Functions for choosing the input collection and setting up parameters.
2. Functions that serve textual metadata extraction and building ontology space.
3. Functions for extraction of proposed visual features.
4. Different kinds of functions for analyzing and visualizing the extracted features and preparing data for further data mining analysis

Chapter 7 presents experiments with the developed tools. A dataset of 600 paintings by 18 representative artists from the Renaissance, Baroque, Romanticism, Impressionism, Cubism, and Modern Art and Eastern Medieval Icons was used as a test collection. Several types of experiments were made using APICAS.

The analysis of the colour distribution characteristics showed that specific characteristics of some movements make them distinguishable from the others. The examination of the features for all paintings revealed some common trends for art images, which we have used later as a normalization factor in the process of defining the colour harmonies and contrast descriptors.

A further group of experiments is focused on the outcomes of automatic annotation of the images with harmonies' and contrasts' descriptors. These high level features can be used not only in the processes of categorization where cultural influences and specific techniques are revealed, but also by extracting the images with given characteristics, which are closely connected with the emotional responses evoked by an image. These features refer to the corresponding elements of the abstract space of the image content.

Another group of experiments examines the significance of local features extracted by the proposed method with regard to the type of underlying MPEG-7 descriptors; position of tiles; as well as number of clusters.

In addition, a designated group of experiments evaluates the classification accuracy by different types of used classifiers. The goal was to confirm the hypothesis that PGN classifier has good representation for such kind of features compared with the other well-known classifiers. This gives the opportunity to use the rules produced by the PGN classifier as a profile of a given class, such as movement, artist, artist period, genre, scene type, etc.

As an example of this hypothesis, the last group of experiments demonstrates the results of classification by PGN on the base of paintings periods in Goya's creative works.

Finally, **chapter 8** contains conclusions and an overview of topics for future research.

In the Appendix the confusion matrices of different classifiers and datasets with examined features are given.

The work contains 22 tables, 81 figures, and 154 references.

2 CBIR in Art Image Collections – the State of the Art

Abstract

This chapter presents an overview of the image retrieval and especially Content-Based Image Retrieval (CBIR) as an alternative and supplement of Text-Based Retrieval. A little tour of some existing digital repositories of art images is made. The existing gaps between human perception and computer interpretation of existing data as well as the taxonomy of art image content from the point of view of the messages that artworks "send" to the viewer are outlined. A short overview of existing techniques, which are employed in the main processes in CBIR, is given.

2.1 The Process of Image Retrieval

Information retrieval is the science of searching for digital items, based both on their content and the metadata about them. Information retrieval can be done on different levels, from personal digital collections to world repositories in the WWW. It is interdisciplinary and attracts the interest of wide range of researchers and developers from a number of domains - computer science, mathematics, library science, information science, information architecture, cognitive psychology, linguistics, statistics, physics, etc. Image retrieval is part of it; it focuses on the processes of browsing, searching and retrieving images from large collections of digital images. There are two basic methods in image retrieval: text-based retrieval and content-based image retrieval (CBIR), which are used separately or together.

Traditional text-based indexing uses controlled vocabulary or natural language to document what an image is or what it is about. Newly developed content-based techniques rely on a pixel-level interpretation of the data content of the image. The upper stage of indexing techniques – concept-based indexing is based on mixing of simple text-based and content-based tools taking into account additional information for interconnections between perceived information from the main player of this process – "the user".

2.1.1 Text-Based Retrieval

Search systems based on textual information contain metadata about the images such as captioning, keywords, or descriptions of the images; the retrieval is performed over the annotated words. These methods are easily implemented using already existing technologies, but require manual data input for each image in the system. Manual image annotation is time-consuming, laborious and expensive and is a potential bottleneck because the speed of manual description and data entry is lower than the speed of digitisation. This is unpractical for the huge collections or automatically-generated images. Usually text-based descriptions are not considered accurate and precise and are often incomplete. Another problem with text annotation is that it often does not conform any defined vocabulary in a particular domain and may not describe the relations of the objects in the images – besides the subjectivity of judgements of different people who entry data. Another inconvenience comes also from the lack of universal solutions for dealing with synonyms in the language, and difference of the users' languages. Additionally, the increase in social web applications and the semantic web have inspired the development of several web-based image annotation tools as crowdsourcing solutions. In the frame of text-based retrieval we can put also context-based technique, where retrieval is based on the analysis of free textual information, which became a context of the image [Hung et al, 2007].

The current efforts in the area of structuring information in digital repositories are focused mainly in two directions:

- Assistance in the processes of ordering and classifying the meta-information (such as Getty's AAT, ULAN, TGN, CONA). The use of these ontological structures in image retrieval processing leads to a decreasing metadata amount and expands the research scope utilizing defined interconnections between concepts;
- Development of metadata schemas and structures to classify image information (for instance Dublin Core, VRA Core, CIDOC CRM). They provide conceptual models intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information.

As example for successful project aimed to use the advantages of such directions is developed by the team of Prof. Radoslav Pavlov "Bulgarian Iconographical Digital Library (BIDL)" [Pavlov et al, 2010]. A tree-based annotation model had been developed and implemented for the semantic description of the iconographical objects. It provides options for auto-completion, reuse of values, bilingual data entry, automated media watermarking, resizing and conversing. A designated ontological model, describing the knowledge of East Christian Iconographical Art is implemented in BIDL; it assists in the annotation and semantic indexing of iconographical artefacts [Pavlova-Draganova et al, 2010].

2.1.2 Content-Based Image Retrieval (CBIR)

Content-based image retrieval, as we see it today, is any technology that in principle helps to organize digital images based on their content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls into the range of CBIR. This characterization of CBIR as a field of study places it at a unique juncture within the scientific community. While we witness continued effort in solving the fundamental open problem of robust image understanding, we also see specialists from different fields, such as, computer vision, machine learning, information retrieval, human-computer interaction, database systems, Web and data mining, information theory, statistics, and psychology contributing and becoming part of the CBIR community [Wang et al, 2006].

Eakins and Graham affirm that Content-Based Image Retrieval is a term first used in 1992 by Toshikazu Kato [Eakins and Graham, 1999], when explaining his experiments on automatic extraction of colour and shape of paintings stored in a database [Kato, 1992]. Since then the term was used to describe the process of image retrieval from large collections using features extracted from the content of the image based on their visual similarity with a query image or image features supplied by an end user.

Before designing and constructing CBIR system one very important step is selecting the domain where the system will be used. Different domains would be addressed by specific functional and non-functional requirements, which have to be covered by the system. During the years a wide spectrum of areas refers to CBIR system, such as medical diagnostic, geographical information and remote sensing systems, crime prevention, the military, intellectual property, photograph archives, architectural and engineering design, art collections, etc. From the point of view of the application area the images can represent different type of sensor-related data, projected or directly received in digital formats. The digital imagery includes colour and black-and-white photographs, infra-red photographs, video snapshots, radar screens, synthetic aperture radar formats, seismographs records, ultrasound, electrocardiographic, electroencephalographic, magnetic resonance and others.

Typically, a content-based image retrieval system consists of three components:

1. Feature design.
2. Indexing.
3. Retrieval.

The feature design component extracts the visual feature(s) information from the images in the image database. The feature indexing component organizes the visual feature information to speed up the query or processing. The retrieval engine processes the user query and provides a user interface. During this process the central issue is to define a proper feature representation and similarity metrics. CBIR systems extract visual features from the images automatically. Similarities between two images are measured in terms of the differences between the corresponding features. To take into account the subjectivity of human perception and bridge the gap between the high-level concepts and the low-level features, relevance feedback is used as a means to enhance the retrieval performance.

All these steps are highly dependent of the domain where CBIR technology is applied. For instance, in the fields such as aerial image retrieval and medicine the goal is exactly defined, the searched objects in the images has homogeneous specifics, the received results usually do not need communication with the user to refine the queries. Absolutely different is the situation in the areas that are connected with the creative side of the human beings, such as art, architecture and design. The different kinds of users also stamp different requirements into specifics of CBIR systems.

In our work we are focused in the area of feature design for art images, presented in the web. Other components are touched only as an integral part of the experimental CBIR system, which is created in order to supply the appropriate environment for conducting the experiments with examined features.

2.2 Art Image Repositories

Since its first edition published in 1962 Janson's History of Art [Janson, 2004] is one of the most valuable sources of information spanning the spectrum of Western art history from the Stone Age to the 20th century. It became a major introduction to art for kids and a reference tool for adults trying to remember the identity of some embarrassingly familiar image. The colourful design and vast range of extraordinarily high-quality illustrations does not only present "dry" information, but also evokes deep emotional fulfilment by the touch to masterpieces.

However, nowadays online search engines have increased the appetite of web surfers for context and information, and there are numerous digital collections offering easy access to digital items. They present the colourfulness of art history as well as relevant metadata, provide additional information from purely technical details, ranging from the way of creating the artefacts to deeply personal details from the life of the creators, which help the observers to understand the original message in the masterpieces.

Digitised art images are important part of the online cultural heritage; however in the last decades the emphasis of digitisation and online accessibility was on materials from libraries, museums, archives and audiovisual archives. Digital art images pose an important challenge for the resource discovery of images since the methods applied in this domain should be a combination of queries executed on metadata and image retrieval methods. Here we provide an overview of the evolution of designated repositories for digital art.

In 2004 David Mattison, named as a master of the online archive universe published a series of lectures in Searcher magazine focusing on state-of-the-art of available Web resources and image databases, current techniques for image retrieval, concluding with a survey of national collections that document art history of Western civilization from medieval times to the 19th century. The creators of this wealth of image databases, art collections and guides usually are academic, librarian, commercial, and private art museums and galleries, amateur and professional art historians, artist sites, commercial image agencies, auction houses (usually on a temporary basis), etc.

Several different applications in the field of Fine Arts have led to specialized digital image processing. Numerous successful projects in the field of processing very large high quality colour images have been funded by a number of research programmes over the last decades, including but not limited to Esprit, Impact, Raphael, and IST programmes. They provide the most conventional image processing for the museums, such as geometric correction, registration, mosaicing, etc. [Maitre et al, 2001]. The European Union has funded numerous digital culture research and development projects. The EU's CORDIS (Community Research & Development Information Service)[4] is the primary resource to learn about past and current R&D projects involving art.

Some of the projects, such as *Vasari*[5] (1989-1992) and *Marc*[6] (1995-1996) focus on digital acquisition, storage and handling of colorimetric high-definition images of paintings (up to 2GB per image) for a range of galleries and museums in the European Union. The *Crisatel*[7] project (2001-2005) developed equipment for the direct fast capture of paintings, with a new ultra-high definition multi-spectral scanner in order to make spectrometric analysis of varnish layers to allow the effect of an aged varnish to be subtracted from an image of a painting. The *FingArtPrint*[8] project (2005-2008) aimed to combine 3D surface scanning and multispectral imaging in order to create a unique data record of the object which can be compared to check its authenticity.

Other projects and initiatives are aimed at establishing image repositories. One of the first projects in this domain was *NARCISSE*[9] (1990-1992), which created a very high-quality digitized image bank, supervised by a multilingual text database (in German, French, Italian and Portuguese). The objective of the project *Artiste*[10] (2000-2002) was to develop and prove the value of an integrated art analysis and navigation environment aimed at supporting the work of professional users in the fine arts. The environment has exploited

[4] http://www.cordis.lu/
[5] http://users.ecs.soton.ac.uk/km/projs/vasari/
[6] http://users.ecs.soton.ac.uk/km/projs/marc/
[7] http://www.ics.forth.gr/isl/projects/projects_individual.jsp?ProjectID=41
[8] http://users.ecs.soton.ac.uk/km/projs/fingartprint/
[9] http://www.culture.gouv.fr/culture/conservation/fr/laborato/narci.htm

advanced image content analysis techniques, distributed hyperlink-based navigation methods, and object-oriented relational database technologies. *Artiste* has integrated art collections virtually while allowing the owners of each collection to maintain ownership and control of their data, using the concept of distributed linking.

In more recent years several projects and initiatives focused on harmonizing activities carried out in digitization of cultural and scientific content in order to create a common platform for cultural heritage. Such project is *MINERVA+*[11], sponsored by FP6 of the EC, which enlarged the existing thematic network of European Ministries of Culture addressing this direction. Since 2005 the Netherlands' Organization for Scientific Research supports the research program *CATCH*[12] (Continuous Access to Cultural Heritage) that finances teams focusing on the improvement of cross-fertilization between scientific research and cultural heritage. In the light of transferability and interoperability, the research teams work on their research at the heritage institutions.

The latest and biggest project in this domain is *Europeana*[13], funded by the European Commission and the member states. The idea of *Europeana* was born in 2005, when the European Commission announced its strategy to promote and support the creation of a European digital library as a strategic goal within the European Information Society i2010 Initiative, which aims to foster growth and jobs in the information society and media industries. The European Commission's goal for *Europeana* is to make European information resources easier to use online. It builds on Europe's rich heritage, combining multicultural and multilingual environments with technological advances and new business models. Europeana.eu went live on 20 November 2008. Till now more than 15 million digital items (Images: paintings, drawings, maps, photos and pictures of museum objects; Texts: books, newspapers, letters, diaries and archival papers; Sounds: music and spoken word from cylinders, tapes, discs and radio broadcasts; Videos: films, newsreels and TV broadcasts) are available.

Numerous museums currently offer online galleries, supplying access to the museum's collections. The search engines of some of them are based only on the metadata information using categories such as artist, title of work, subject, chronology and reference number (for instance online gallery of Museo del Prado[14]). In other sites implementations of CBIR techniques during the search process are integrated (such as in the site of Hermitage museum[15]).

The first steps in the direction of supplying aggregated environments providing access to cultural heritage were to establish portals to existed sites of museums and online galleries. The simple hierarchical structure links to

---

[10] http://users.ecs.soton.ac.uk/km/projs/artiste/
[11] http://www.minervaeurope.org/whatis/minervaplus.htm
[12] http://www.nwo.nl/nwohome.nsf/pages/NWOP_66EUM7_Eng
[13] http://www.europeana.eu/portal/
[14] http://www.museodelprado.es/en/the-collection/online-gallery/
[15] http://www.hermitagemuseum.org/ fcgi-bin/db2www/qbicSearch.mac/qbic?selLang=English

preferred place. Some examples of such portals are *Museumland*[16] and *Zeroland*[17]*. Museumland* is the worldwide portal to museums and cultural heritage with more than 11 000 links accessible through interfaces in English and Italian. *Zeroland* is a systematic and comprehensive overview of the best, most authoritative and inspiring (and often entertaining) arts web pages from around the world. This style of presenting the information is typical for Web 1.0 [Agarwal, 2009].

During the years, the ability of processing the information as well as expanding the ways of data exchange increased in parallel. The development of computing and communication capacities allows to place the user in the centre of the process of information exchange and to afford him/her to use the overall power of the intellectualized tools for satisfying his/her needs and expectations. In the recent years as a result of this growth, the virtual museums change towards more compact and systematic presenting the information with abilities of common interoperable search between different collections, which are primary sources of the artworks.

Multiple additional issues arise in the web space, caused by the need to create a framework that includes the image, the delivery system, and the users [Jörgensen, 2001]. Numerous repositories are already available, and many services for reaching the content are implemented and used in the practice. In the current years, the problems of semantic, syntactic and profile interoperability and constructing reference layers become pivotal [DCMI, 2009]; additional area to explore is linked data which allow contextualizing objects in the cultural heritage domain. Today interoperability is considered a key-step to move from isolated digital repositories towards a common information space that allow users to browse through different resources within a single integrated environment [Vullo et al, 2010].

Some of the most successful digital virtual museums on a global scale are created independently, as a result of collaboration between different stakeholders – national or public institutions, or by supporting a specific company. Funds for establishing and supporting such initiatives are most commonly donated; or a shared between the budgets of the institutions; or by allocating funds from the profit of companies gained with other activities (for instance Artchive is supported by art.com, which is a company for selling digital posters of fine art and decor items).

Below we present some examples of such virtual art collections.

Artcyclopedia[18] is an online database of museum-quality fine art founded by the Canadian John Malyon. The site is the leading guide to museum-quality fine art images on the Internet. The Artcyclopedia is a form of Internet search engine and deals with art that can be viewed online, and indexes 2 600 art sites (from

[16] http://www.museumland.com/
[17] http://www.zeroland.co.nz/
[18] http://www.artcyclopedia.com/

museums and galleries), with links to around 140 000 artworks by 9 000 renowned artists. The site has also started to compile a list of art galleries and auction houses.

Artchive[19] is a virtual art gallery website, established by Mark Harden. It displays historic artworks with a convenient viewer that allows the size of the image to be set easily as required. The site is a leading example of an independently established collection of high-quality pictures important in the history of art. Works include art paintings from various periods, such as Abstract Expressionism, Baroque, Impressionism, Renaissance, Romanticism, Rococo, Surrealism and more.

OCAIW[20] (Orazio Centaro's Art Images on the Web) is an educational and non-profit site for art-lovers, teachers, students, artists and collectors. The catalogue consists of information about painters, sculptors, architects and photographers. Every section of the catalogue includes a listing of the greatest artists in Art History. The painters index lists about 1 500 artists from medieval times to the present.

WebMuseum[21] is one of the earliest examples of a virtual museum. Starting in 1994 as WebLouvre, now it has multiple mirror sites are established throughout the world. It provides an excellent archival and educational resource of good quality art images and information. WebMuseum is part of one of the largest "collections of collections" on the Internet ibiblio[22] a conservancy of freely available information, including software, music, literature, art, history, science, politics, and cultural studies.

Web Gallery of Art[23] is a virtual museum and searchable database of European painting and sculpture of the Romanesque, Gothic, Renaissance, Baroque, Neoclassicism, Romanticism periods (1000-1850), currently containing over 23 000 reproductions. Picture commentaries, artist biographies are also available there. Guided tours, period music, catalogue, free postcard and other services are provided.

Backtoclassics[24] is a new virtual art gallery (opened in 2009), created by the Italian division of Microsystems MS Lab, which gives a look at the creations of artists past and present. The paintings are classified not only by movements and artists, but also thematically (for instance Rembrandt's paintings are grouped into following series: Portraits; Biblical Themes; Various Paintings; Self-Portraits; Etchings; Drawings; Landscapes).

Olga's Gallery[25] is one of the largest online painting museums, containing works and biographies of most of the world's best known artists. Olga's Gallery

[19] http://www.artchive.com/
[20] http://www.ocaiw.com/
[21] http://www.ibiblio.org/wm/
[22] http://www.ibiblio.org/
[23] http://www.wga.hu/
[24] http://www.Backtoclassics.com/
[25] http://www.abcgallery.com/

was started by two sisters-in-law – Olga and Helen in 1999 as a fine art-themed website under the domain name abcgallery.com (where abc is part of the motto of the site "The abc of art"). Now Olga's Gallery contains over than 12 000 works of art by more than 300 painters and receives over 30 000 visitors and 1 000 000 page views daily. The artists can be searched by name, country or genre, and the site also includes brief biographies of the significant artists, with excellent links to more intensive sites, making it a significant research resource for young students.

Eastern Iconography borrows a special place in the European culture. Developed in parallel, and even earlier from Western Fine Art, the East European art of painting icons and frescoes followed more canonical tradition in scene representation, composition, and used colours. The main cradles of this rich heritage were Byzantium, Russia and the Balkans. Some of the famous representatives of the icons are presented in the examined portals. In parallel, the Orthodox Church has the initiatives such as "Orthodox Christianity on the web" for representing the cultural heritage and history of this region. One example is the site *Christian Art*[26], created as independent non-profit project of this initiative, where the artworks (icons, mosaics, frescoes, miniatures) are categorized by topics; schools; and geographical indices.

In our work we use these open access sites for receiving examples of artworks, representative for examined periods and artists as well as for additional textual information, needed during the process of testing the results.

The new technologies in the area of digitization of cultural heritage bring new ways of representation of art works in the digital space. From 1 February 2011 Google presented the *Google Art Project*[27]. Seventeen galleries and museums were included in the launch of the project. The 1061 high-resolution images (by 486 different artists) are shown in 385 virtual gallery rooms, with 6000 Street View-style panoramas. Each institute contributed one item of gigapixel artwork for free access.

Digital repositories for cultural heritage images can employ similar techniques as the ones exploited in generic systems in order to solve standard questions for searching objects. The cultural heritage objects are rich in content, describing events, monuments, places, people etc. and distributed across different locations. The users can formulate queries using different modalities such as free text, similarity matching, or metadata. The specifics of the observed objects profile some additional tasks, which are also of interest. In the area of art paintings retrieval the sensitive, semantic, and abstraction gaps are still major problems. Bridging the sensitive gap depends on the technical capacities for receiving good representation of presented object and establishing different kinds of formats for storing captured information. Bridging the semantic and

[26] http://www.icon-art.info/
[27] www.googleartproject.com

abstraction gaps is the main motivator of the intelligent feature design to which our work also contributes.

2.3 The Gaps

One of the most felicitous analogies for presenting the existing semantic gap in area of Content-Based Image Retrieval can be found in "The Hitch-Hiker's Guide to Galaxy" by Douglas Adams. In this story, a group of hyper-intelligent pan-dimensional beings demand to learn the "Answer to Life, the Universe, and Everything" from the supercomputer Deep Thought, specially built for this purpose. It takes Deep Thought 7½ million years to compute and check the answer, which turns out to be "42". The efforts of covering the semantic gap in CBIR are turned to avoid these misunderstanding between human perceiving and the ways of communications and computer manner of low-level representations [Ivanova and Stanchev, 2009].

Search in the context of content-based retrieval analyzes the actual contents of the image. The term content might refer to colours, shapes, textures, or any other information that can be derived from the image itself. Acknowledging the need for providing image analysis at semantic level, research efforts set focus on the automatic extraction of image descriptions matching human perception. The ultimate goal characterizing such efforts is to bridge the so called semantic gap between low-level visual features that can be automatically extracted from the visual content and the high-level concepts capturing the conveyed meaning [Dasiapoulou et al, 2007]. The semantic gap is not a unique cause of difficulties in the process of information retrieval where issues can arise on the whole range starting from the primary object' complexity and ending with end-user subjectivity. Currently different gaps are being discussed in the research literature: sensory, semantic, subjective, aesthetic and abstraction.

2.3.1 Sensory Gap

The sensory gap is "the gap between the object in the world and the information in a (computational) description derived from a recording of that scene" [Smeulders et al, 2000].

Figure 2. Detail of the photo of the Goya's painting "A Hunter and his dog" from Prado Museum

The sensory gap exists in the multimedia world as the gap between an object and the machine's capability to capture and define that object. Digitalization has a big challenge when applied to art-works and this is to develop techniques for creating digital objects, which allow to capture the paintings in good quality. Circumstances, such as the condition of the pictures, the lighting, the capabilities of used photo-cameras or scanners, the chosen resolution, etc., play a major role in this process. One example of this sensitiveness is illustrated in Figure 2, where a detail of the photo of the Goya's painting "A Hunter and his dog" from Prado museum is shown. In contrast to the others pictures of the same collection here one can clearly see the reflections of the texture.

The sensory gap in this area inevitably results in the impossibility to present real sizes of the pictures or to present all pictures in one proportional scale. One can only see the proportion of the height and length. For instance Picasso's Guernica is 3.59 m x 7.76 m, while the miniatures of Isaac and Peter Oliver not exceed 2.5 cm x 2.5 cm. This sensory gap can be omitted only with additional metadata, taken from the camera or manually added to the picture.

The granularity of digitalized sources of artworks is in accordance with their usage. Figure 3 summarizes the connections between different kinds of users and the amount and quality of corresponding digital sources.

For the purposes of professional analysis in museums, special kinds of images, received from different photographic processes such as multi-colour banding, x-rays and infra-red imaging are used. For the purposes of professional printing industry very high definition and quality images are needed [Maitre et al, 2001]. The royalties and copyright restrictions from one side [Mattison, 2004], the necessity of high speed delivery on the Internet from the other side, and the limitations of visual devices (monitors) from the third side, impose restrictions of the sizes and resolutions of digital images.

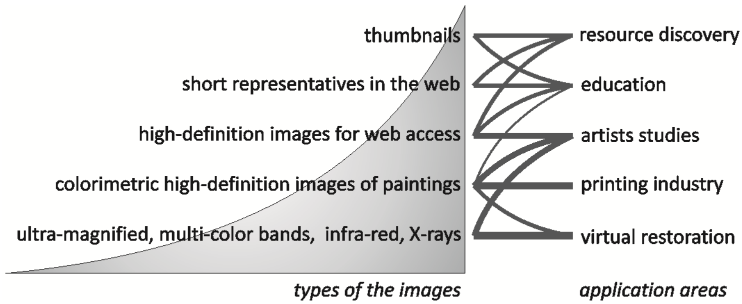


Figure 3. Digitised art images: quality and usage

Usually in the Internet space the presentations of the art paintings vary from:

- About 100x100 pixels for front presentation of the paintings as thumbnails;
- Image surrogates designed for presenting the painting on the screen, supplemented by additional text information, concerning the picture – author, sizes, techniques, locality, history of the creation, subject comment, etc.;

- High-definition images for web access, usually up to 1500 pixels by width or height;
- Finally, up to 4000 pixels by width or height, often watermarked items. The access to high-resolution and ultra-magnified images usually is defined by policies which set sets of use restrictions.

The fact that a digitized work of art is not the work itself but an image (instance) of this work, acquired at a certain time under specific conditions, makes semantic-based indexing and retrieval an absolute necessity in this area. For example, a query on "Mona Lisa" should retrieve all images of the painting regardless of their size, view angle, restoration procedures applied on the painting, etc. [Chen et al, 2005a].

2.3.2 Semantic Gap

The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data has for a user in a given situation [Smeulders et al, 2000]. The semantic gap is larger in visual arts images than in natural images since artworks are most often not perfectly realistic.

In simple terms, the semantic gap in content-based retrieval stems from the fact that multimedia data is captured by devices in a format, which are optimized for storage and very simple retrieval, and cannot be used to understand what the object "means". In addition to this, user queries are based on semantic similarity, but the computer usually processes similarity based on low-level feature similarity.

In early systems low-level representation was considered most reliable. Later, to bridge this gap, text annotations by humans were used in conjunction with low-level features of the objects. To extend the annotation list, different ontology systems have been used for further improving the results. It is applied generic ontologies as WordNet [Fellbaum, 1998], specialized international standard structures as Dublin Core Element Set[28] and VRA Core Categories[29], or special ontologies designed for description of artefacts, such as IconClass[30] and Categories for the Description of Works of Art (CDWA)[31]. The annotations are used to group images into a certain number of concept bins, but the mapping from image space to concept space is not one-to-one, since it is possible to describe an image using a large number of words. The labelling of images is made not only for the whole image, but also for separate parts of the image [Enser et al, 2006].

[28] http://dublincore.org/
[29] http://www.vraweb.org/vracore3.htm
[30] http://www.iconclass.nl
[31] http://www.getty.edu/research/conducting_research/standards/cdwa/

The semantic gap is very critical to content-based multimedia retrieval techniques. As Smeulders et al. states, "The aim of content-based retrieval systems should be to provide maximum support in bridging the semantic gap between the simplicity of available visual features and the richness of the user semantics." [Smeulders et al, 2000].

2.3.3 Subjective Gap

The subjective gap exists due to users' needs and the descriptions of these needs. It may be difficult for a user to express what he wants from a multimedia retrieval system [Agrawal, 2009]. The subjective gap also exists due to the non-availability of any features which user wants to express. Some authors [Castelli and Bergman, 2002] identify this as an *intermediate level* of extraction from visual content, connected with emotional perceiving of the images, which usually is difficult to express in rational and textual terms. In the recent years the term "*Emotional Semantic Image Retrieval*" enjoys growing popularity in scientific publications. The visual art is an area, where these features play significant role. Typical features in this level are colour contrasts, because one of the goals of the painting is to produce specific psychological effects in the observer, which are achieved with different arrangements of colours.

The subjective gap is similar to the semantic gap; it refers to the lack of ability of the user to describe his needs (queries) to a retrieval system. To bridge this gap, instead of defining user's requirements at a very fine granularity level, a higher level concept can be used. This higher level concept is used to return an initial set of results which helps the user to decide which images are subjectively similar to his/her ideas and annotate them in an appropriate way. The relevance feedback technique combined by neural network or fuzzy systems can bridge this subjective gap to some extent [Grosky et al, 2008].

2.3.4 Aesthetic Gap

In addition to the above three, the aesthetic gap has been introduced in [Croft, 1995] as "the lack of coincidence between the information that one can extract from low-level visual data and the interpretation of emotions that the visual data may arouse in a particular user in a given situation". Aesthetics is similar to quality as perceived by a viewer and is highly subjective. Modelling aesthetics of images will evolve in near future. The thesis of Rittendra Datta, presented in 2009 [Datta, 2009] is focused just to the semantic and aesthetic inference for image search, using statistical learning approaches.

Emotional abstraction relates to emotional responses evoked by an image. These issues are addressed in the research domain called affective computing which enjoys widespread attention among computer scientists beyond those working on cultural heritage. In principal artworks by their nature are images that naturally evoke affective effects. Due to their implicit stylization, one does not look at artistic images with the same kind of attention and expectation as for natural images.

Many approaches try to bridge the gap between selected low-level features and several emotions expressed with pairs of words, e.g. warm-cool, action-relaxation, joy-uneasiness [Colombo et al, 1999]. Wei-ning et al. argued that emotional expression of the image is closely connected with such low-level characteristics as colour and luminance distributions, saturation and contrast information as well as edge sharpness in images [Wei-ning et al, 2006].

2.3.5 Abstraction Gap

The abstract aspects are specific to art images and differ from the semantic challenge. There are two major directions in this area, the first one addressing cultural specifics and the second one addressing technical differences.

Cultural abstraction relates to information inferred from cultural knowledge [Hurtut, 2010]. Artistic style analysis belongs to that category. According to the American Dictionary [Pickett et al, 2000] style is "the combination of distinctive features of artistic expression, execution or performance characterizing a particular person, group, school or era." According to [Hurtut, 2010] style and semantic depiction share the same visual atomic primitives (lines, dots, surfaces, textures). Manual style recognition is a very difficult task, requiring the knowledge of numerous art historians and experts. An artwork from the blue period of Picasso for instance is recognizable not only because of its blue tonality, but because of many semantic features and iconographical cues.

Technical abstraction deals with questions about real artist of the artwork (artwork authentication), but can be focused into artistic praxis such as perspective rendering, pigment identification, searching for preliminary sketches in underlying layers, pentimenti, engraving tools, etc. These aspects can be analyzed using different imaging techniques such as X-rays, UV and infrared imaging.

2.4 User Interaction

The main focus in the creation of digital art resources has to be user-centred rather than system-centred since most of the issues around this content are related to making it accessible and usable for the real users [Dobreva and Chowdhury, 2010].

✓ ***Complexity of the queries***

Bridging the gaps is closely connected with user interaction. This is the place where the user and the system communicate. In the image retrieval systems, an important parameter to measure user-system interaction level is the complexity of queries supported by the system. The queries can use different modalities, such as:

- *Direct entry of values* of the desired features (query by percentage of properties). This method is not usually used in current systems because it is not particularly convenient for the users;
- *Image, also known as query by example:* Here, the user searches for images similar to a query image. Using an example image is perhaps the most popular way of querying a CBIR system in the absence of reliable metadata;
- *Graphics, or query by sketch:* This consists of a hand-drawn or computer-generated picture, where graphics are used as a query.
- *Keywords or free-text:* This is a search in which the user makes a query in the form of a word or group of words, selected from previously defined set or in free form. This is currently the most popular way in web image search engines like Google, Yahoo! and Flickr. Usually this search is based on manually attached metadata or context driven information. Numerous current efforts are directed to finding the methods for automated labelling of the images – a challenge for the CBIR systems in present days;
- *Composite:* These are methods that combine one or more of the aforesaid modalities for querying a system. This also covers interactive querying such as the one implemented in relevance feedback systems.

Exploring user needs and behaviour is a basic and important phase of system development and is very informative when done as a front-end activity to system development. Currently users are mostly involved in usability studies when a set of digital resources has already been created and is being tested (for an overview on usability evaluation methods in the library domain see [George, 2008]). It would be really helpful to involve users on early stages of design and planning the functionality of the product which is being developed.

✓ ***Relevance Feedback***

Relevance feedback is a very important step in image retrieval, because it defines the goals and the means to achieve them. Relevance feedback provides a compromise between a fully automated, unsupervised system and one based on subjective user needs. It is a query modification technique which attempts to capture the user's precise needs through iterative feedback and query refinement. It can be thought of as an alternative search paradigm to other paradigms such as keyword-based search. In the absence of a reliable framework for modelling high-level image semantics and subjectivity of perception, the user's feedback provides a way to learn case-specific query semantics. A comprehensive review can be found in [Zhou and Huang, 2003] and [Crucianu et al, 2004]. The goal in relevance feedback is to optimise the amount of interaction with the user during a session. It is important to use all the available information to improve the retrieval results. Based on the user's relevance feedback, learning-based approaches are typically used to appropriately modify the feature set or similarity measure. In practice, learning instances are very small number. This circumstance has generated interest in novel machine-learning techniques to solve the problem, such as *one-class* learning, *active* learning, and *manifold* learning. Usually, classical relevance feedback consists of multiple rounds, which leads to loosing the patience in the user. Recent developments are directed to find techniques for minimizing the rounds. One decision is to use information of earlier user logs in the system. Another approach is presented in [Yang et al, 2005], where a novel feedback solution for semantic retrieval is proposed: *semantic feedback*, which allows the system to interact with users directly at the semantic level. This approach is closely neighboured to the new relevance feedback paradigms aimed to help users by providing the user with cues and hints for more specific query formulation.

2.5 Web 3.0

Not much time passed before the idea of "Web 3.0" appeared. Amit Agarwal suggests that Web 3.0 is about semantics (or the meaning of data), personalization (e.g. iGoogle), intelligent search and behavioural advertising among other things [Agarwal, 2009]. While Web 2.0 uses the Internet to make connections between people, Web 3.0 will use the Internet to make connections with information. The intelligent browsers will analyze the complex requests of the users made in natural language, will search the Internet for all possible answers, and then will organize the results. The adaptation to user specifics and aptitudes (personalisation) will be based on capturing the historical information thorough searching the Web. Many of the experts believe that the Web 3.0 browser will act like a personal assistant. The computer and the environment will become artificial subjects, which will pretend to communicate in real manner as real humans. A core problem in this domain continues to be finding appropriate combination of retrieval methods and techniques, which can lead to high quality image discovery. In the era of Web 3.0 bridging the semantic gap stands crucial.

2.6 Feature Design

The process of feature design is achieved to make mathematical description of an image for the retrieval purposes as its signature. Most CBIR systems perform feature design as a pre-processing step. Once obtained, visual features act as inputs to subsequent image analysis tasks, such as similarity estimation, concept detection, or annotation.

The process of feature design is achieved to make mathematical description of an image for the retrieval purposes, as its *signature*. The extraction of signatures and the calculation of image similarity cannot be cleanly separated. One the one hand, the formulation of signatures determines the necessity of finding new definitions of similarity measures. On the other hand, intuitions are often the early motivating factors for designing similarity measures in a certain way, which puts requirements on the construction of signatures. In terms of methodology development, a strong trend which has emerged in the recent years is the employment of statistical and machine learning techniques in various aspects of the CBIR technology. Automatic learning, mainly clustering and classification, is used to form either fixed or adaptive signatures, to tune similarity measures, and even to serve as the technical core of certain searching schemes, for example, relevance feedback. The fixed set of visual features may not work equally well to characterize different types of images. The signatures can be tuned either based on images alone (when some property does not characterize the image – than signatures vary according to the classification of images) or by learning from user feedback (when the user is not interested in a particular feature).

In contrast with early years, where global feature representations for images, such as colour histograms and global shape descriptors were used, currently the focus shifts towards using local features and descriptors, such as salient points, region-based features, spatial model features, and robust local shape characterizations.

2.6.1 Taxonomy of Art Image Content

Johannes Itten [Itten, 1961] has given very good formulation of messages that one artwork sends to the viewer. He points three basic directions of evincing colour aesthetics:

- Impression (visually);
- Expression (emotionally);
- Construction (symbolically).

These characteristics are mutually connected and cannot live of full value alone: symbolism without visual accuracy and without emotional force would be merely an anaemic formalism; visually impressive effect without symbolic verity and emotional power would be a banal imitative naturalism; emotional effect without constructive symbolic content or visual strength would be limited to the plane of sentimental expression. Each artist works according to his temperament, and must emphasize one or another of these aspects [Itten, 1961].

Different styles in art paintings are connected with used techniques from one side and aesthetic expression of the artist from other side. The process of forming artist style is a very complicated one, where current fashion painting styles, social background and personal character of the artist play significant role. All these factors lead to forming some common trends in art movements and some specific features, which distinguish one movement to another, one artist style to another, one artist period to another, etc. On the other hand the theme of the paintings also stamps specifics and can be taken into account. The compositions in different types of images (portraits, landscapes, town views, mythological and religious scenes, or everyday scenes) also set some rules, aesthetically imposed for some period.

When humans interpret images, they analyze image content. Computers are able to extract low-level image features like colour distribution, shapes and texture. Humans, on the other hand, have abilities that go beyond those of computers. The humans draw own subjective conclusions. They place emphasis on different parts of images, identify objects and scenes stamping theirs subjective vision and experience. The emotion that one person gets from seeing an image, and therefore associates with it, may differ from another person's point of view.

Trying to define some useful grounds for bridging the gaps between interpreting the information from human and from computers several taxonomies of image content as extracted by the viewer of an image had been suggested.

Jaimes and Chang [Jaimes and Chang, 2002] focus on two aspects of image content – the received visual *percepts* from the observed images and underlying abstract idea, which corresponds to *concepts*, connected with the image content.

In his brilliant survey for 2D artistic images analysis Tomas Hurtut [Hurtut, 2010] expands the taxonomy suggested by Burford, Briggs and Eakins [Burford et al, 2003]. He gives profiling of extraction primitives and concepts accounting the specific of artworks, splitting image categories into three groups: *image space*, *object space* and *abstract space*.

For the purposes of this study we adopted Hurtut's proposition adjusting the distribution of features in the groups. We examine *image space*, *semantic space* and *abstract space*.

- *Image space* contains visual features, needed to record an image through visual perception. Image space includes perceptual primitives (colour, textures, local edges), geometric features (strokes, contours, shapes) and design constructions (spatial arrangement, composition);
- *Semantic space* is related to the meaning of the elements, their potential for semantic interpretation. It consists of semantic units (objects), 3D relationship between them (scene, perspective, depth cues) and context (illumination, shadow);
- *Abstract space* is connected with the aspects that are specific to art images and reflect cultural influences, specific techniques as well as emotional responses evoked by an image form the abstract space.

In Figure 4 we give our vision for classifying feature percepts and concepts (they slightly differ from Hurtuts' proposition) while pointing examples of used techniques for extracting visual primitives as well as some of closer relationships between concepts from defined spaces [Ivanova et al, 2010/MCIS].

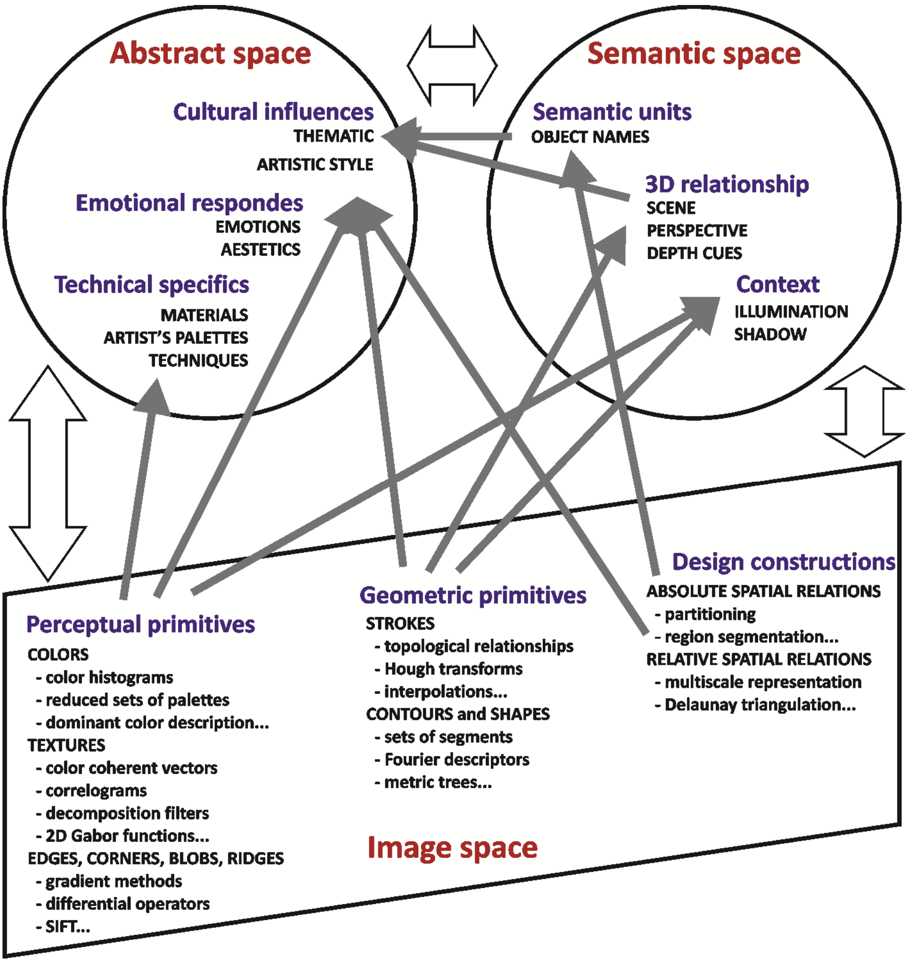


Figure 4. A taxonomy of art image content, inspired from [Burford et al, 2003] and [Hurtut, 2010]

Of course, all concepts are mutually connected – for instance emotional abstractions depends on specific expressive power of the artists (which is closely connected with visual perception primitives), thematic of the painting (concerning objective semantics of the paintings) as well as with the viewpoint of observer with his/her cultural and psychological peculiarities. Resolving these questions come up against the problem that in real-world contexts, it is in fact dynamic in nature. The information that one can extract from the visual data for a one-time trained image recognition model does not change, but on the other hand, the interpretation that the same data have for a user in a given situation changes across users as well as situations.

2.6.2 Visual Features

According to Figure 4, main features, extracted from image space, are:

- *Perceptual features*, especially colour, texture and interesting points features;
- *Geometric features*, where the main focus for art image analysis is on contours and shapes as a source for further semantic interpretation and on strokes, which together with the previous ones are the source for extracting technical abstractions;
- *Design constructions*, connected with absolute or relative spatial relations.

Image features can be extracted at *a global level* to represent the entire image or the image can be split into parts and then features are computed *in a local level* from each part.

The most commonly used features include those reflecting *colour*, *texture*, *shape*, and *salient points* in an image. On a global level, features are computed to capture the overall characteristics of an image. The advantage is high speed for both processes: construction of signatures and computing similarity. The processing on the *local level* increases the robustness to spatial transformation of the images and gives more detailed representation of specific features of the image. Both approaches have their advantages: global features help to build an integral overview of the image as well as local ones can capture more detailed specifics.

➢ ***Colour Features***

Colour features are focused on the summarization of colours in an image. A set of colour descriptors are included in the MPEG-7 standard, which reflect different aspects of colour presence in an image. *Dominant Colour descriptor* presents the percentage of each quantized colour over the observed area. *Scalable Colour descriptor* builds a colour histogram, encoded by a Haar transformation. *Colour Layout descriptor* effectively represents the spatial distribution of colour in an image or in an arbitrary shaped region in a very compact form and is resolution-invariant. *Colour Structure descriptor* captures both colour content (similar to a colour histogram) and information about the structure of this content. Usually the exploration of colour features is attended with conversing colour representation to other colour spaces, which are more comprehensive for the human vision and in this way facilitate the choice of appropriate distance measures.

➢ ***Texture Features***

Texture features are intended to capture the granularity and repetitive patterns of surfaces within in a picture. Their role in domain-specific image retrieval is particularly vital due to their close relation to the underlying semantics. In image processing, a popular way to form texture features is by using the coefficients of a certain transformation on the original pixel values, or by statistics computed from such coefficients. Such descriptors encode significant, general visual characteristics into standard numerical formats that can be used for various higher-level tasks. In many application areas, for example in aerial image retrieval and in medicine, thesauri for texture have been built. A thesaurus of brushwork terms concerning the annotation of paintings covering the period from Medieval to Modern art, which includes terms as "shading", "glazing", "mezzapasta", "grattage", "scumbling", "impasto", "pointillism", and "divisionism" had been proposed in the field of art image retrieval [Marchenko et al, 2007]. The brushwork is defined as a combination of colour presence and contrast features with texture features, such as directional ("impasto"), non-directional ("pointillism"), contrasting ("divisionism") and smooth ("mezzapasta"), or in case of spatial homogeneity they can be grouped into homogeneous ("mezzapasta" and "pointillism"), weakly homogeneous ("divisionism") and inhomogeneous ("scumbling", "shading" and "glazing").

➢ ***Salient Point Features***

Feature detection is a low-level image processing operation. That is, it is usually performed as the first operation on an image, and examines every pixel to see if there is a feature present at that pixel. If this is part of a larger algorithm, then the algorithm will typically only examine the image in the region of the features. There are very large number of feature detectors, which vary widely in the kinds of feature detected, the computational complexity and the repeatability. The salient point feature detectors can be divided into: *edges*, *corners*, *blobs* and *ridges* (with some overlap). A detailed review of salient point features is showed in [Ivanova et al-TR, 2010].

"*Edges*" are points where there is a boundary between two image regions. In practice, edges are usually defined as sets of points in the image which have a strong gradient magnitude. Furthermore, some common algorithms chain high gradient points together to form a more complete description of an edge. These algorithms usually place some constraints on the properties of an edge, such as shape, smoothness, and gradient value. Locally, edges have one dimensional structure.

The terms "*Corners*" or "*Interesting points*" refer to point-like features in an image, which have a local two dimensional structure. The name "corner" arose since early algorithms first performed edge detection, and then analyzed the edges to find rapid changes in direction (corners). These algorithms were then developed so that explicit edge detection was no longer required, for instance by looking for high levels of curvature in the image gradient. It was then noticed that the so-called corners were also being detected on parts of the image which were not corners in the traditional sense (for instance a small bright spot on a dark background may be detected).

"*Blobs*" provide a complementary description of image structures in terms of regions, as opposed to corners that are more point-like. Blob descriptors often contain a preferred point (a local maximum of an operator response or a gravity centre), which means that many blob detectors may also be regarded as "point of interest" operators. Blob detectors can detect areas in an image which are too smooth to be detected by a corner detector.

The concept "*Ridges*" is a natural tool for elongated objects. A ridge descriptor computed from a grey-level image can be seen as a generalization of a medial axis. From a practical viewpoint, a ridge can be thought of as a one-dimensional curve that represents an axis of symmetry, and in addition has an attribute of local ridge width associated with each ridge point. It is algorithmically harder to extract ridge features from general classes of grey-level images than edge-, corner- or blob features. The ridge descriptors are frequently used for road location in aerial images and for extracting blood vessels in medical images.

Multiple scale-invariant features extraction algorithms such as SIFT, GLOH, SURF, LESH exist; they are widely used in current object recognition. They transform an image into a large collection of feature vectors, each of which is invariant to image translation, scaling, and rotation, partially invariant to illumination changes and robust to local geometric distortion (an overview is provided in [Ivanova et al-TR, 2010]).

- ***Shape Features***

Shape is a key attribute of segmented image regions, and its efficient and robust representation plays an important role in retrieval. Shape representations are closely connected with the particular forms of shape similarities used in each case. The current state of the art of this area is described in detail in [Data et al, 2008]. They have marked the shift from global shape representations which was dominant in early research to the use of more local descriptors in the last years. In MPEG-7 standard also are included Region Shape, Contour Shape and Shape 3D descriptors. The Region Shape descriptor utilizes a set of ART (Angular Radial Transform) coefficients. The Contour Shape descriptor is based on the Curvature Scale Space representation of the contour. The Shape 3D descriptor specifies an intrinsic shape description for 3D mesh models, which exploits some local attributes of the 3D surface [ISO/IEC 15938-3]. Shape features can play very significant role in semantic retrieval.

- ***Spatial Relations***

Representing spatial relations among local image entities plays a very important role in the process of preparing visual signatures. Several kinds of indexing methods are used for representing absolute or relative spatial relations.

✓ ***Partitioning***

Partitioning can be defined as data-independent grouping [Data et al, 2008]. This method is not closely connected with representing absolute spatial relations, but allows a simple way for receiving more local information for the examined images. There are different methods for *partitioning* the image depending on the type of application. The simplest method is to divide image into non-overlapping tiles. In [Gong et al, 1996] the image is split into nine equal sub-images. [Striker and Dimai, 1997] have split image into oval central region and four corners. These methods have low computational cost and can be used for deriving more precisely (than from the whole image) low-level characteristics. They are not suitable if the goal is object segmentation.

✓ ***Segmentation***

Segmentation is opposite of partitioning and is characterized as data-driven grouping. [Estrada, 2005] formulates *segmentation* as "the problem of defining a similarity measure between image elements that can be evaluated using image data, and the development of an algorithm that will group similar image elements into connected regions, according to some grouping criterion. The image elements can be pixels, small local neighbourhoods, or image regions produced by an earlier stage of processing, or by a previous step of an iterative segmentation procedure. The similarity function can use one or many of the available image cues (such as image intensity, colour, texture, and various filter responses), or be defined as a proximity measure on a suitable feature space that captures interesting image structure."

A great variety of segmentation techniques exists. Some applied approaches used either agglomerative (by merging) or divisive (by splitting) hierarchical clustering with different similarity functions (based on the entropy or statistic distance) and stopping criteria (like minimum description length, chi-square, etc.). Agglomerative algorithms are usually more frequently used than the divisive ones. An excellent example for agglomerative clustering is the algorithm where "normalized cut criterion" measures both the total dissimilarity between the different groups as well as the total similarity within the groups [Shi and Malik, 2000].

Other algorithms are not hierarchical. The simplest and widely used segmentation approach is based on *k*-means clustering. This basic approach enjoys a speed advantage, but is not as refined as some recently developed methods. Another disadvantage is that the number of clusters is an external parameter. The mean-shift algorithm is nonparametric clustering technique; it does not require prior knowledge of the number of clusters and the algorithm recursively moves to the kernel smoothed centroid for every data point looking for the point with highest density of data distribution [Comaniciu and Meer, 1999].

Amongst other approaches it is worth mentioning the multi-resolution segmentation of low-depth-of-field images [Wang et al, 2001], a Bayesian framework-based segmentation involving the Markov chain Monte Carlo technique [Tu and Zhu 2002], and the EM-algorithm-based segmentation using a Gaussian mixture model [Carson et al, 2002], forming *blobs* suitable for image querying and retrieval. A sequential segmentation approach that starts with texture features and refines segmentation using colour features is explored in [Chen et al, 2001]. An unsupervised approach for segmentation of images containing homogeneous colour/texture regions has been proposed in [Deng and Manjunath, 2001].

Yet another group of algorithms are the so called model-based segmentation algorithms. The central assumption is that structures of interest have a repetitive form of geometry. These algorithms work well when the segmented image contains the search object and are widely used in medicine and radiological image retrieval.

✓ ***Relative Relationships***

Considering that homogeneous regions or symbolic objects have already been extracted, the relative relationships try to model or characterize the spatial relations between them, for instance "object A is under and on the left of object B" [Freeman, 1975].

Another convenient way of representing local spatial relations is Delaunay triangulation. This method was invented by Boris Delaunay in 1934 for the case of Euclidean space. In this space the Delaunay triangulation is the dual structure of the Voronoi diagram. Several algorithms can be used for computing Delaunay triangulation, such as flipping, incremental, gift wrap, divide and conquer, sweep-line, sweep-hull, etc. [de Berg et al, 2000].

2.7 Data Reduction

Data reduction techniques can be applied to obtain a reduced representation of the data set that is much smaller in volume, yet closely maintains the integrity of the original data.

That is, mining on the reduced data set should be more efficient yet produce the same (or almost the same) analytical results. Strategies for data reduction include *dimensionality reduction*, where encoding mechanisms are used to reduce the data set size and *numerosity reduction*, where the data are replaced or estimated by alternative, smaller data representations. In Figure 5 a hierarchy of some data reduction techniques is presented.

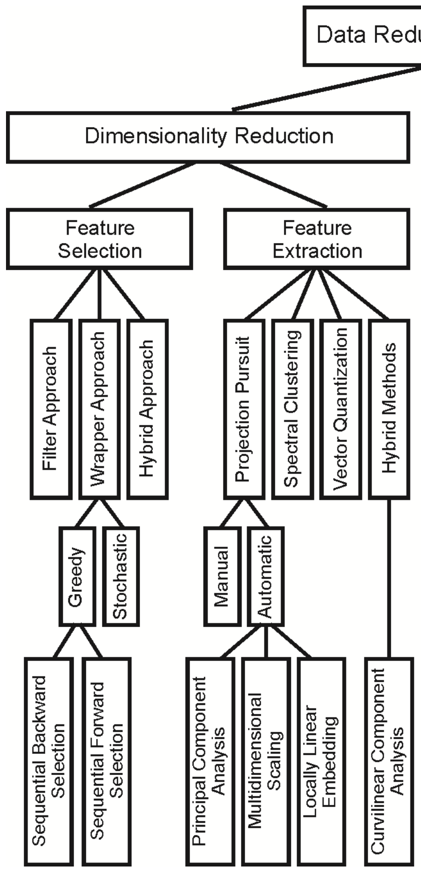


Figure 5. Data Reduction Techniques

2.7.1 Dimensionality Reduction

The "curse of dimensionality", is a term coined by Bellman [Bellman, 1961] to describe the problem caused by the exponential increase in volume associated with adding extra dimensions to a feature space. In image clustering and retrieval applications, the feature vectors tend to use high dimensional data space and in such case to fall into "curse of dimensionality" since the search space grows exponentially with the dimensions. In image databases, the volume of the data is very large and the amount of time needed to access the feature vectors on storage devices usually dominates the time needed for a search. This problem is further complicated when the search is to be performed multiple times and in an interactive environment. Thus high dimensionality of data causes increased time and space complexity, and as a result decreases performance in searching, clustering, and indexing.

When the attribute space is not high-dimensional, the standard method is representing the features as points in a feature space and using distance metrics for similarity search. The problem with this method is that with the increasing of data dimension, the maximum and minimum distances to a given query point in the high dimensional space are almost the same under a wide range of distance metrics and data distributions. All points converge to the same distance from the query point in high dimensions, and the concept of nearest neighbours become meaningless.

There are two main ways to overcome the curse of dimensionality in image search and retrieval. The first is to search the approximate results of a multimedia query, and the second is to reduce the high dimensional input data to a low dimensional representation.

The dimensionality reduction techniques are based on either *feature selection* (also named *attribute subset selection*) or *feature extraction* methods.

- ***Feature Selection (Attributive Subset Selection)***

In feature selection, an appropriate subset of the original features is found to represent the data. This method is useful when the data available limited amount, but is represented with a large number of features [Agrawal, 2009]. It is crucial to determine a small set of relevant variables to estimate reliable parameter. The advantage of selecting a small set of features is that you need to use few values in the calculations.

Data sets for analysis may contain attributes, which may be irrelevant or redundant to the mining task. Leaving out relevant attributes or keeping irrelevant attributes may aggravate data mining process. Attribute subset selection reduces the data set size by removing irrelevant or redundant attributes (or dimensions). The goal of attribute subset selection is to find a minimum set of attributes such that the resulting probability distribution of the data classes is as close as possible to the original distribution obtained using all attributes. Finding on optimal subset is a hard computational process. Therefore, heuristic methods that explore a reduced search space are commonly used for attribute subset selection. Optimal feature subset selection techniques can be divided to: filter, wrapper and hybrid [Gheyas and Smith, 2010].

  - ✓ ***Filter Approaches***

In filter approaches, features are scored and ranked based on certain statistical criteria and the features with highest ranking values are selected. Usually as filter methods t-test, chi-square test, Wilcoxon Mann-Whitney test, mutual information, Pearson correlation coefficients and principal component analysis are used. Filter methods are fast but lack robustness against interactions among features and feature redundancy. In addition, it is not clear how to determine the cut-off point for rankings to select only truly important features and exclude noise.

✓ ***Wrapper Approaches***

In the wrapper approaches, feature selection is "wrapped" in a learning algorithm. The learning algorithm is applied to subsets of features and tested on a hold-out set, and prediction accuracy is used to determine the feature set quality. Generally, wrapper methods are more effective than filter methods. Since exhaustive search is not computationally feasible, wrapper methods must employ a designated algorithm to search for an optimal subset of features. Wrapper methods can broadly be classified into two categories based on the search strategy: (1) greedy and (2) randomized/stochastic.

(1) *Greedy wrapper methods* use less computer time than other wrapper approaches. Two most commonly applied wrapper methods that use a greedy hill-climbing search strategy are:

- Sequential backward selection, in which features are sequentially removed from a full candidate set until the removal of further features increase the criterion;
- Sequential forward selection, in which features are sequentially added to an empty candidate set until the addition of further features does not decrease the criterion.

The problem with sequentially adding or removing features is that the utility of an individual feature is often not apparent on its own, but only in combinations including just the right other features.

(2) *Stochastic algorithms*, developed for solving large scale combinatorial problems such as ant colony optimization, genetic algorithm, particle swarm optimization and simulated annealing are used as feature subset selection approaches. These algorithms efficiently capture feature redundancy and interaction, but are computationally expensive.

✓ ***Hybrid Approaches***

The idea behind the hybrid method is that filter methods are first applied to select a feature pool and then the wrapper method is applied to find the optimal subset of features from the selected feature pool. This makes feature selection faster since the filter method rapidly reduces the effective number of features under consideration [Gheyas and Smith, 2010].

➢ ***Feature Extraction***

In feature extraction, new features are found using the original features without losing any important information. Feature extraction methods can be divided into linear and non-linear techniques, depending of the choice of objective function. Some of the most popular dimensionality reduction techniques are:

✓ ***Projection Pursuit***

Projection pursuit is a method, which finds the most "interesting" possible projections of multidimensional data. A good review of projection pursuit can be found in [Huber, 1985]. The projection index defines the "interestingness" of a direction. The task is to optimize this index. A projection is considered interesting if it has a structure in the form of trends, clusters, hyper-surfaces, or anomalies. These structures can be analyzed using manual or automatic methods. The scatter-plot is one such manual method, which can be used to understand data characteristics over two selected dimensions at a time. There are many methods to automate this task.

✓ ***Principal Component Analysis (PCA)***

One often used and simple projection pursuit method is the Principal Component Analysis, which calculates the eigenvalues and eigenvectors of the covariance or correlation matrix, and projects the data orthogonally into space spanned by the eigenvectors belonging to the largest eigenvalues. PCA is also called the discrete Karhunen-Loève method (K-L method), the Hotelling transform, singular value decomposition (SVD), or empirical orthogonal function (EOF) method. A good tutorial on PCA can be found in [Smith, 2002]. PCA searches for k n-dimensional orthogonal vectors that can best be used to represent the data, where $k \leq n$. The original data are thus projected onto a much smaller space, resulting in dimensionality reduction. PCA transforms the data to a new coordinate system such that the first coordinate (also called the first principal component) is the projection of the data exhibiting the greatest variance, the second coordinate (also called the second principal component) exhibits the second greatest variance, and so on. In this way, the "most important" aspects of the data are retained in the lower-order principal components.

PCA is computationally inexpensive, can be applied to ordered and unordered attributes, and can handle sparse data and skewed data. Principal components may be used as inputs to multiple regression and cluster analysis.

✓ ***Multidimensional Scaling (MDS)***

Multidimensional scaling (MDS) is used to analyze subjective evaluations of pairwise similarities of entities. In general, the goal of the analysis is to detect meaningful underlying dimensions that allow the researcher to explain observed similarities or dissimilarities (distances) between the investigated objects. In PCA, the similarities between objects are expressed in the covariance or correlation matrix. MDS allows analyzing any kind of similarity or dissimilarity matrix, in addition to correlation matrices.

Assume, there are p items in n-dimensional space and a $p \times p$ matrix of proximity measures, MDS produces a k-dimensional representation ($k \leq n$) of the original data items. The distance in the new k-space reflects the proximities in the data. If two items are more similar, this distance will be smaller. The distance measures can be Euclidean distance, Manhattan distance, maximum norm, or other. MDS is typically used to visualize data in two or three dimensions, to uncover underlying hidden structure.

Any dataset can be perfectly represented using $n-1$ dimensions, where n is the number of items scaled. As the number of dimensions used goes down, the stress must either come up or stay the same. When the dimensionality is insufficient the non-zero stress values occur. It means that chosen dimension k cannot perfectly represent the input data. Of course, it is not necessary that an MDS map has zero stress in order to be useful. A certain amount of distortion is tolerable. Different people have different standards regarding the amount of stress to tolerate. The rule of thumb is that anything under 0.1 is excellent and anything over 0.15 is unacceptable.

Both PCA and MDS are eigenvector methods designed to model linear variability in high dimensional data. In PCA, one computes the linear projections of greatest variance from the top eigenvectors of the data covariance matrix. Classical MDS computes the low dimensional embedding that best preserves pair-wise distances between data points. If these distances correspond to Euclidean distances, the results of metric MDS are equivalent to PCA.

✓ ***Locally Linear Embedding (LLE)***

Locally Linear Embedding (LLE) is also an eigenvector method that computes low dimensional, neighbourhood preserving embeddings of high dimensional data. LLE attempts to discover nonlinear structure in high dimensional data by exploiting the local symmetries of linear reconstructions. Notably, LLE maps its inputs into a single global coordinate system of lower dimensionality, and its optimizations – though capable of generating highly nonlinear embeddings – do not involve local minima. Like PCA and MDS, LLE is simple to implement, and its optimizations do not involve local minima. At the same time, it is capable of generating highly nonlinear embeddings [Saul and Roweis, 2000].

✓ ***Spectral Clustering***

The main tools for spectral clustering are graph Laplacian matrices. The technique is based on two main steps: first embedding the data points in a space in which clusters are more "obvious" (using the eigenvectors of a Gram matrix) and then applying an algorithm to separate the clusters, such as K-means. A good tutorial for Spectral Clustering can be found in [Luxburg, 2006]. Sometimes called Diffusion Maps or Laplacian Eigenmaps, these manifold-learning techniques are based on graph-theoretic approach. A low dimensional representation of the data set is computed using the Laplacian of the graph that optimally preserves local neighbourhood information. The vertices, or nodes, represent the data points, and the edges connected the vertices, represent the similarities between adjacent nodes. After representing the graph with a matrix, the spectral properties of this matrix are used to embed the data points into a lower dimensional space, and gain insight into the geometry of the dataset. Though these methods perform exceptionally well with clean, well-sampled data, problems arise with the addition of noise, or when multiple sub-manifolds exist in the data.

✓ ***Vector Quantization (VQ)***

Vector Quantization (VQ) is used to represent not individual values but (usually small) arrays of them. In vector quantization, the basic idea is to replace the values from a multidimensional vector space with values from a lower dimensional discrete subspace. A vector quantizer maps k*-dimensional* vectors in the vector space R^k into a finite set of vectors $Y = \{y_i : i = 1,...,n\}$. The vector y_i *is* called a *code vector* or a *codeword* and the set of all the codewords Y is called a *codebook*. Unfortunately, designing a codebook that best represents the set of input vectors is NP-hard. There are different algorithms, which try to overcome this problem. A review of vector quantization techniques used for encoding digital images is presented in [Nasrabadi and King, 1988]. VQ can be used for any large data sets, when adjacent data values are related in some way. VQ has been used in image, video, and audio compression.

✓ ***Curvilinear Component Analysis***

The principle of Curvilinear Component Analysis is a self-organized neural network performing two tasks: vector quantization of the submanifold in the data set (input space); and nonlinear projection of these quantizing vectors toward an output space, providing a revealing unfolding of the submanifold. After learning, the network has the ability to continuously map any new point from one space into another: forward mapping of new points in the input space, or backward mapping of an arbitrary position in the output space [Demartines and Herault, 1997].

2.7.2 Numerosity Reduction

In numerosity reduction data is replaced or estimated by alternative, smaller data representations. These techniques may be parametric or nonparametric. For *parametric methods*, a model is used to estimate the data, so that typically only the data parameters need to be stored, instead of the actual data. For comprehensive data representation outliers may also be stored. Regression and Log-linear models, which estimate discrete multidimensional probability distributions, are two examples. *Nonparametric methods* for storing reduced representations of the data include histograms, clustering, and sampling. Data discretization is a form of numerosity reduction that is very useful for the automatic generation of concept hierarchies. Discretization and concept hierarchy generation are powerful tools for data mining, in that they allow the mining of data at multiple levels of abstraction.

✓ ***Regression and Log-Linear Models***

Regression and Log-Linear models can be used to approximate the given data. They are typical examples of parametric methods [Han and Kamber, 2006].

In *simple linear regression*, the data are modelled to fit a straight line. A random variable, y (called a *response variable*), can be modelled as a linear function of another random variable, x (called a *predictor variable*), with the equation $y = wx + b$, where the variance of y is assumed to be constant. In the context of data mining, x and y are both numerical attributes. The coefficients, w and b (called *regression coefficients*), specify the slope of the line and the y-intercept, respectively. These coefficients can be solved by the *method of least squares*, which minimizes the error between the actual line separating the data and the estimate of the line.

Multiple linear regression is an extension of simple linear regression, which allows a response variable y to be modelled as a linear function of two or more predictor variables.

Log-linear models approximate discrete multidimensional probability distributions using logarithmic transformations. Given a set of tuples in n dimensions (e.g., described by n attributes), we can consider each tuple as a point in a n-dimensional space. Log-linear models can be used to estimate the probability of each point in a multidimensional space for a set of discretized attributes, based on a smaller subset of dimensional combinations. This allows a higher-dimensional data space to be constructed from lower dimensional spaces. Log-linear models are therefore also useful for dimensionality reduction (since the lower-dimensional points together typically occupy less space than the original data points) and data smoothing (since aggregate estimates in the lower-dimensional space are less subject to sampling variations than the estimates in the higher-dimensional space).

Regression and log-linear models can both be used on sparse data, although their application may be limited. While both methods can handle skewed data, regression does it exceptionally well. Regression can be computationally intensive when applied to high dimensional data, whereas log-linear models show good scalability for up to 10 or so dimensions.

✓ ***Discrete Wavelet Transforms***

The Discrete Wavelet Transform (DWT) is a linear signal processing technique that transforms input vector to another vector with same length, but elements are wavelet coefficients. A wavelet is a mathematical function used to divide a given function into different scale components. A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as "daughter wavelets") of a finite-length or fast-decaying oscillating waveform (known as the "mother wavelet"). The first DWT was invented by the Hungarian mathematician Alfred Haar in 1909. The most commonly used set of discrete wavelet transforms was formulated by the Belgian mathematician Ingrid Daubechies in 1988. Haar wavelet is the first one of the family of Daubechies wavelets. The Daubechies wavelets are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (also called father wavelet) which generates an orthogonal multi-resolution analysis [Daubechies, 1988].

✓ ***Histograms***

Histograms use binning to approximate data distributions and are a popular form of data reduction [Han and Kamber, 2006]. A histogram for an attribute partitions the data distribution of the attribute into disjoint subsets, or *buckets*. There are several partitioning rules, including *Equal-width* (where the width of each bucket range is uniform), *Equal-frequency* (each bucket contains roughly the same number of contiguous data samples), *V-Optimal* (the histogram with the least variance) and *MaxDiff* (where a bucket boundary is established between each pair for pairs having the $b-1$ largest differences, where b is user-specified number of buckets). V-Optimal and MaxDiff histograms tend to be the most accurate and practical. Histograms are highly effective at approximating both sparse and dense data, as well as highly skewed and uniform data.

✓ ***Clustering***

In data reduction, the cluster representations of the data are used to replace the actual data. The effectiveness of this technique depends on the nature of the data. It is much more effective for data that can be organized into distinct clusters than for smeared data. In database systems, multidimensional index trees are primarily used for providing fast data access. They can also be used for hierarchical data reduction, providing a multi-resolution clustering of the data. This can be used to provide approximate answers to queries. An index tree can store aggregate and detail data at varying levels of abstraction. It provides a hierarchy of clustering of the data set, where each cluster has a label that holds for the data contained in the cluster. If we consider each child of a parent node as a bucket, then an index tree can be considered as a *hierarchical histogram*. The use of multidimensional index trees as a form of data reduction relies on an ordering of the attribute values in each dimension. Multidimensional index trees include R-trees, quad-trees, and their variations.

Special cases of clustering are data discretization techniques, which can be used to reduce the number of values for a given continuous attribute by dividing the range of the attribute into intervals. Interval labels can then be used to replace actual data values. Replacing numerous values of a continuous attribute by a small number of interval labels thereby reduces and simplifies the original data. From point of view of using class information in the discretization process the methods are *supervised* or *unsupervised*. Supervised discretizators usually fall into the following categories: if the process starts by finding one or a few points (called *split points* or *cut points*) to split the entire attribute range, and then repeats this recursively on the resulting intervals, it is called *top-down discretization* or *splitting*. In contrast, *bottom-up discretization* or *merging* starts by considering all continuous values as potential split-points, removes some by merging neighbourhood values to form intervals, and then recursively applies this process to the resulting intervals. Discretization can be performed recursively on an attribute to provide a hierarchical or multi-resolution partitioning of the attribute values, known as a concept hierarchy. Concept hierarchies are useful for mining at multiple levels of abstraction. We have made a brief overview of discretization techniques in [Mitov et al, 2009b].

✓ ***Sampling***

Sampling allows a large data set to be represented by a much smaller random sample (or subset) of the data. The most common ways for receiving data reduction, using sampling, according to [Han and Kamber, 2006] are:

- *Simple random sample without replacement* (SRSWOR), where all tuples are equally likely to be sampled;
- *Simple random sample with replacement* (SRSWR), where after a tuple is drawn, it is placed back into the primary set, so that it may be drawn again;
- *Cluster sample*, where all tuples are grouped into mutually disjoint "clusters", then a simple random sample can be obtained;
- *Stratified sample*, where if source set is divided into mutually disjoint parts called strata, a stratified sample is generated by obtaining a simple random sampling at each stratum. This helps ensure a representative sample, especially when the data are skewed.

An advantage of sampling for data reduction is that the cost of obtaining a sample is proportional to the size of the sample when applied to data reduction, sampling is most commonly used to estimate the answer to an aggregate query.

2.8 Indexing

The second component of CBIR-systems is indexing. Efficient indexing is critical for building and functioning of very large text-based databases and search engines. Research on efficient ways to index images by content has been largely overshadowed by research on efficient visual representation and similarity measures.

In [Markov et al, 2008] we provide an expanded survey of different spatial access methods, based on the earlier analyses of [Ooi et al, 1993] and [Gaede and Günther, 1998]. The access methods are classified in several categories:

one-dimensional; multidimensional spatial; metric; and high dimensional access methods (Figure 6). The article [Markov et al, 2008] includes more detailed description of interconnections between access methods as well as the references of sources, where these methods are described. Here we provide a summary of the methods.

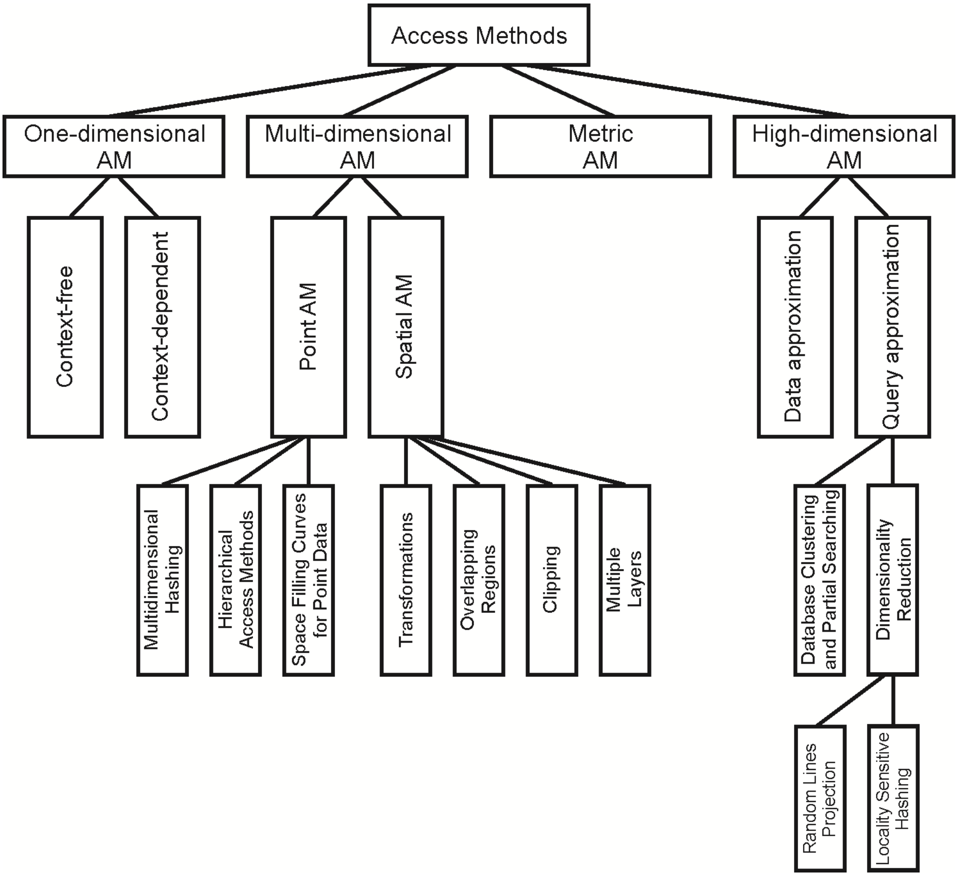


Figure 6. Taxonomy of the Spatial Access Methods

Multidimensional Spatial Access Methods are developed to serve information about spatial objects, approximated with points, segments, polygons, polyhedrons, etc. From the point of view of the spatial databases can be split in two main classes of access methods – Point Access Methods and Spatial Access Methods [Gaede and Günther, 1998].

Point Access Methods are used for organizing multidimensional point objects. Typical instance are traditional records, where one dimension corresponds to every attribute of the relation. These methods can be clustered in three basic groups: (1) Multidimensional Hashing; (2) Hierarchical Access Methods; (3) Space Filling Curves for Point Data.

Spatial Access Methods are used for work with objects which have arbitrary form. The main idea of the spatial indexing of non-point objects is using of the approximation of the geometry of the examined objects to more simple forms, using approximations as minimum bounding rectangles, spheres or other polytopes, as well as their combinations. The usual problem when one operates with spatial objects is their overlapping. There are different techniques to avoid this problem. From the point of view of the techniques for organization of the spatial objects Spatial Access Methods form four main groups: (1) *Transformation* – this technique uses transformation of spatial objects to points in the space with more or less dimensions. Most of them spread out the space using space filling curves and then use some of point access method upon the transformed data set; (2) *Overlapping Regions* – here the data set are separated in groups; different groups can occupy the same part of the space, but every space object associates with only one of the groups. The access methods of this category operate with data in their primary space (without any transformations) eventually in overlapping segments; (3) *Clipping* – this technique use eventually clipping of one object to several sub-objects. The main goal is to escape overlapping regions. But this advantage can lead tearing of the objects, extending of the resource expenses and decreasing of the productivity of the method; (4) *Multiple Layers* – this technique is a variant of the technique of Overlapping Regions, because the regions from different layers can also overlap. However, there are some important differences: first, the layers are organizing hierarchically; second, every layer splits the primary space in different way; third, the regions of one layer never overlap; fourth, the data regions are separated from space extensions of the objects.

Metric Access Methods deal with relative distances of data points to chosen points, named anchor points, vantage points or pivots [Moënne-Loccoz, 2005]. These methods are designed to limit the number of distance computation, calculating first distances to anchors, and then finding the point searched for in the narrowed region. These methods are preferred when the distance is highly computational, as e.g. for the dynamic time warping distance between time series. Metric Access Methods are employed to accelerate the processing of similarity queries, such as the range and the k-nearest neighbour queries too [Chavez et al, 2001].

High Dimensional Access Methods are created to overcome the bottleneck problem, which appears with increasing of dimensionality. These methods are based on the *data approximation* and *query approximation* in sequential scan. For *query approximation* two strategies can be used: (1) examine only a part of the database, which is more probably to contain a resulting set – as a rule these methods are based on the clustering of the database; (2) splitting the database to several spaces with fewer dimensions and searching in each of them, using Random Lines Projection or Locality Sensitive Hashing.

2.9 Retrieval Process

The third component of CBIR systems is served by retrieval engines. The retrieval engines build the bridge between the internal space of the system and the user making requests which need to be satisfied. Looking at the system side, the design of these engines is closely connected with the chosen feature representation and indexing schemes as well as the selected similarity metrics. From the user point of view, in order to take into account the subjectivity of human perception and bridge the gap between the high-level concepts and the low-level features, relevance feedback has been proposed to enhance the retrieval performance. The other direction for facilitating that process is examining the image retrieval in more general frame of multimedia retrieval process, where content-based, model-based, and text-based searching can be combined.

2.10 Similarity

In the process of image retrieval, choosing the features as well as indexing the database are closely connected with the used similarity measures for establishing nearness between queries and images or between the images in a given digital resource in the processes of categorization.

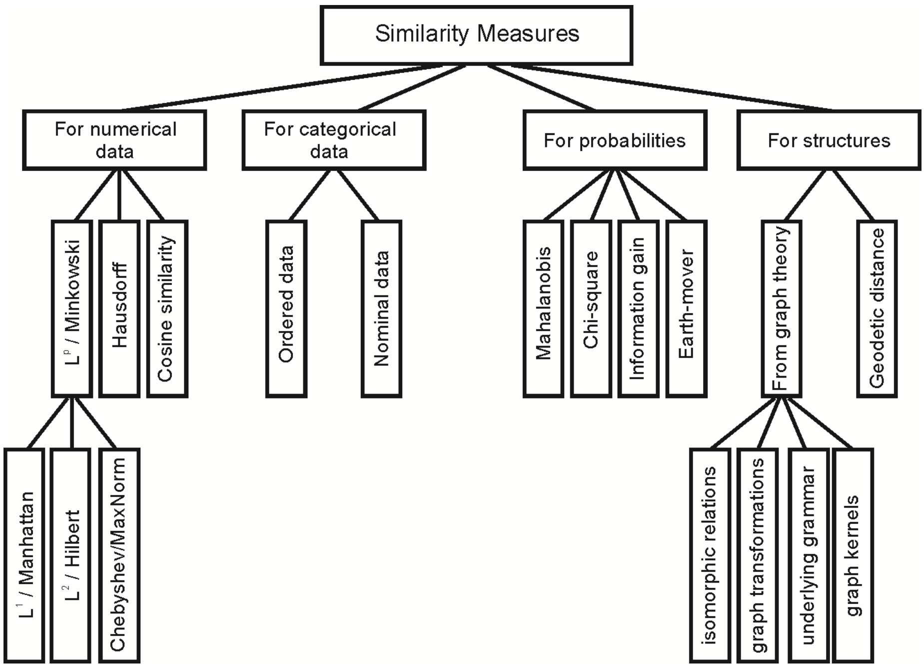


Figure 7. Different kinds of similarity measures

The concept of similarity is very complex and is almost a whole scientific area itself. Similarity measures are aimed to give answer how much one object is close to another one. In the process of obtaining the similarity between two images several processes of finding similarity on different levels and data types need to be resolved. Image signature is a weighted set of feature vectors. In the case when a region-based signature is represented as such set of vectors, each of them would be represented in this way. A natural approach to defining a region-based similarity measure is to match every two corresponded vectors and then to combine the distances between these vectors as a distance between sets of vectors. For every level different similarities may be used. In Figure 7, similarity measures used in CBIR for different features types are presented.

2.10.1 Distance-Based Similarity Measures

The most popular similarity measures are distance measures. They can be applied in each area which meets the conditions to be a metric space for equal self-similarity, minimality, symmetry and triangle inequality[32]. In mathematical notation the distance $d(X,Y)$ between two vectors X and Y is a function for which $d(X,Y) \geq 0$; if $d(X,Y)=0$ then $X=Y$; $d(X,Y)=d(Y,X)$. Fulfilment of triangle inequality $d(X,Y) \leq d(X,Z)+d(Z,Y)$ defines distance $d(X,Y)$ as a metric. Replacing for triangle inequality with the condition $d(X,Y) \leq \max\{d(X,Z),d(Z,Y)\}$ defines an ultra-metric, which plays an important role for the hierarchical cluster analysis.

[Perner, 2003] shows a classification of some distance metrics explaining their interconnections. For two vectors $X=(x_1,x_2,...,x_n)$ and $Y=(y_1,y_2,...,y_n)$ we can use different metrics:

- The L^p-metric also called *Minkowski metric* is defined by the following formula: $d_{L^p}(X,Y)=\left[\sum_{i=1}^{n}\left|x_i-y_i\right|^p\right]^{1/p}$, where the choice of the parameter p depends on the importance of the differences in the summation;
- L^1-metric, also known as rectilinear, taxi-cab, city-block or Manhattan metric is received for $p=1$: $d_{L^1}(X,Y)=\sum_{i=1}^{n}\left|x_i-y_i\right|$. This measure, however, is insensible to outlier since big and small difference are equally treated;
- In the case of $p=2$ the resulting spaces are the so called *Hilbert spaces.* One of most popular of them are *Euclidean spaces*, where the distance is

[32] http://www.britannica.com/EBchecked/topic/378781/metric-space

calculated as: $d_{Euclidean}(X,Y)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2}$. This metric gives special emphasis to big differences in the observations and is invariant to translations and orthogonal linear transformations (rotation and reflection). In image retrieval weighted Euclidean distance is also used [Wang et al, 2001];

- In the case $p=\infty$ L^{∞}*-metric*, which can be also called *Chebyshev* or *Max Norm* metric is obtained: $d_{Chebyshev}(X,Y)=\max_{i=1}^{n}\left|x_i-y_i\right|$. This measure is useful if only the maximal distance between two variables among a set of variables is of importance whereas the other distances do not contribute to the overall similarity.

In image matching often *Hausdorff measure* is used. The Hausdorff distance measures how far two subsets of a metric space are from each other. It turns the set of non-empty compact subsets of a metric space into a metric space in its own right. Informally, two sets are close in the Hausdorff distance if every point of either set is close to some point of the other set. The Hausdorff distance is the longest distance you can be forced to travel by an adversary who chooses a point in one of the two sets, from where you then must travel to the other set. The Hausdorff distance is symmetricized by computing in addition the distance with the role of X and Y reversed and choosing the larger of the two distances:

$$d_{Hausdorf}(X,Y)=\max(\max_{i=1}^{n}\min_{j=1}^{n}d(x_i,y_j),\max_{j=1}^{n}\min_{i=1}^{n}d(y_j,x_i)) .$$

Often techniques, used in text matching, are applied in image retrieval too. One such example is *cosine similarity*, which is a measure of similarity between two vectors of n dimensions by finding the cosine of the angle between them, often used to compare documents in text mining. Given two vectors of attributes, X and Y , the cosine similarity is represented using a dot product and magnitude as: $d_{cosine}(X,Y)=\frac{X\ \cdot\ Y}{\|X\|\ \|Y\|}=\frac{\sum_{i=1}^{n}x_i*y_i}{\sqrt{\sum_{i=1}^{n}x_i^2}\sqrt{\sum_{i=1}^{n}y_i^2}}$.

2.10.2 Distance Measures for Categorical Data

Categorical data are two types – ordered and nominal. The analysis of symbolic data has led to a new branch of Data Analysis called Symbolic Data Analysis (SDA) [Esposito et al, 2002].

The degree of dissimilarity can be defined by assigning levels of dissimilarity to all the different combinations between attribute values. The mapping can be made to discrete space [0, 1] or more complex discrete linear spaces. Special distance coefficients have been designed for nominal attributes. The basis for the calculation of these distance coefficients is a contingency table, where as columns and rows either the status "not present" (0) or "present" (1) of the property are placed. The cells of the tables contain the frequency of observations that do not share the property (N_{00}), either only one object contains the property (N_{01} or N_{10}), or both of them share the property (N_{11}).

Given that distance coefficients for nominal data can be calculated as variants of generalized formula [Nieddu and Rizzi, 2003]:

$$\frac{N_{11} + tN_{00}}{N_{11} + v(N_{10} + N_{01}) + wN_{00}}, \ t = \{0,1\}, v = \{0,1,2\}, w = \{0,1\}.$$

Several coefficients are subordinated of this formula, such as Jaccard coefficients ($t = 0, v = 1, w = 0$), Russel-Rao coefficients ($t = 0, v = 1, w = 1$), Sokal-Sneath coefficients ($t = 0, v = 2, w = 0$), Sokal-Michener coefficients ($t = 1, v = 1, w = 1$), Roger-Tanimoto coefficients ($t = 1, v = 2, w = 1$), etc.

Other similarity measures that do not fit in the previous class are considered respectively as arithmetic and geometric mean of the quantities $N_{11}/(N_{11} + N_{10})$ and $N_{11}/(N_{11} + N_{01})$ that represent the proportional of agreements on the marginal distributions: $\frac{1}{2}\left(\frac{N_{11}}{N_{11} + N_{10}} + \frac{N_{11}}{N_{11} + N_{01}}\right)$ (Kulczynski) and $\frac{N_{11}}{\sqrt{(N_{11} + N_{10})(N_{11} + N_{01})}}$ (Occhiai-Driver-Kroeber).

More sophisticated similarity measures, concerning recent data mining techniques, take into account the distribution of combinations of examined attributes as presented in [Boriah et al, 2008].

2.10.3 Probability Distance Measures

The disadvantage of the metric measures is that they require the independence of the attributes. A high correlation between attributes can be considered as a multiple measurement for an attribute. That means the measures described above give this feature more weight as an uncorrelated attribute. Some examples of the used distances of such type are shown below.

The *Mahalanobis distance* (or "generalized squared interpoint distance") is defined as: $d_i = \sqrt{(\overline{x_i} - \overline{y_i})S^{-1}(x_i - y_i)}$ and takes into account the covariance matrix S of the attributes.

The most familiar measure of dependence between two quantities is *Pearson's correlation*, which is obtained by dividing the covariance of the two variables $\mathrm{cov}(X,Y)$ by the product of their standard deviations σ_X and σ_Y: $\rho(X,Y)=\dfrac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y}$.

Some similarity measures are defined in statistics. *Chi-square* is a quantitative measure used to determine whether a relationship exists between two categorical variables.

Other similarities come from the information theory. One example is the *Kullback-Leibler distance* also called *information divergence*, *information gain* or *relative entropy*. It is defined for discrete distributions of compared objects X and Y, which have probability functions x_i and y_i: $d(X,Y)=\sum_{i=1}^{n} x_i \log_2\left(\dfrac{x_i}{y_i}\right)$. Although the information divergence is not a true metric because $d(X,Y)\neq d(Y,X)$, it satisfies many useful properties, and is used to measure the disparity between distributions.

An example of probability measure based approach is the *Earth Mover's Distance* (EMD) [Rubner et al, 1998]. EMD is a measure, which can be used for signatures in the form of sets of vectors. The concept was first introduced by Gaspard Monge in 1781. It is a mathematical measure of the distance between two distributions. Informally, if the distributions are interpreted as two different ways of piling up a certain amount of dirt over the region, the EMD is the minimum cost of turning one pile into the other. The cost is assumed to be amount of dirt moved times the distance by which it is moved. A typical signature consists of list of pairs ((x_1,m_1), … , (x_n,m_n)), where each x_i is a certain "feature" (e.g., colour, luminance, etc.), and m_i is "mass" (how many times that feature occurs). Alternatively, x_i may be the centroid of a data cluster, and m_i – the number of entities in that cluster. To compare two such signatures with the EMD, one must define a distance between features, which is interpreted as the cost of turning a unit mass of one feature into a unit mass of the other. The EMD between two signatures is then the minimum cost of turning one of them into the other. EMD can be computed by solving an instance transportation problem using the so-called Hungarian algorithm [Kuhn, 1955]. The EMD is widely used to compute distances between colour histograms of two digital images. The same technique is used for any other quantitative pixel attribute, such as luminance, gradient, etc. Several attempts are focused on proposing fast algorithms for calculating EMD. For instance a fast algorithm for angular type of histograms, which make good representation of hue or gradient distribution, is suggested in [Cha et al, 1999].

2.10.4 Structural Similarity Measures

Structural similarity is involved in a variety of pattern recognition problems when considered from an abstract perspective. The abstraction refers to measurements and observations whose specifics are ignored. One class of such problems is encountered in image processing, where a set of features or objects with topological interrelations is detected in several scenes. Whenever these are presumed to be similar according to position, proximity or else, the degree of similarity is of interest.

Structures are represented throughout by labelled graphs such as image graphs. In image graphs, vertices represent image edges, corners or regions of interest such as regions of constant intensity or homogenous texture. Graph edges represent relations such as neighbourhoods or concept hierarchies. Edge labels represent distances, degrees of association or else. Special branch of measures observed similarities in graph theory. Examples of such measures are given in [Dehmer et al, 2006].

Most classical methods are *based on isomorphic and sub-graph relations*. For large graphs these measures are faced with the complexity of the sub-graph isomorphism problem. Other measures are *based on graph transformations*. The graph edit distance is defined as minimum cost of transformations (deletion, substitutions, insertions) of vertices and edges, which need to transform one graph into another one. The idea of *finding underlying graph grammar*, in which both of graphs belong, is used to define some further measures. The application of such measures is very complex, because the underlying grammar is difficult to define. *Graph kernels* take the structure of the graph into account. They work by counting the number of common random walks between two graphs. Even though the number of common random walks could potentially be exponential, polynomial time algorithms exist for computing these kernels.

From graph theory in image retrieval, often *Geodesic distances* are used to measure the similarities between images. The geodesic distance is defined as the shortest path between two vertexes of a graph [Deza and Deza, 2009].

2.11 Techniques for Improving Image Retrieval

Single similarity measure is not sufficient to produce robust, perceptually meaningful ranking of images. The results achieved with the classical content based approaches are often unsatisfactory for the user. As an alternative, learning-based techniques such as clustering and classification are used for speeding-up image retrieval, improving accuracy, or for performing automatic image annotation. Including the relevance feedback in this process allows the user to refine t query-specific semantics. Bridging the gaps between primitive feature levels, which are produced in the classic CBIR systems and higher levels, which are convenient for the user, can be made with examining the image retrieval process in more general frame of multimedia retrieval process. It needs to integrate multimedia semantics-based searching with other search techniques (speech, text, metadata, audio-visual features, etc.) and to combine content-based, model-based, and text-based searching. It can be made in two main ways:

- Creating the semantic space through statistical pattern recognition and machine learning techniques;
- Creating semantic concepts, which can be incorporated in an already built semantic space.

Several techniques in these directions are used.

2.11.1 Unsupervised Clustering

Unsupervised clustering techniques are a natural fit when handling large, unstructured image repositories such as the Web. Clustering methods fall roughly into three main types: pair-wise-distance-based, optimization of an overall clustering quality measure, and statistical modelling. The pair-wise distance-based methods (e.g., linkage clustering and spectral graph partitioning) are of general applicability, since the mathematical representation of the instances becomes irrelevant. One disadvantage is the high computational cost. Clustering based on the optimization of an overall measure of clustering quality is a fundamental approach used in pattern recognition. The general idea in statistical modelling is to treat every cluster as a pattern characterized by a relatively restrictive distribution, and the overall dataset is thus a mixture of these distributions. For continuous vector data, the most commonly used distribution of individual vectors is the Gaussian distribution.

2.11.2 Image Categorization (Classification)

Image categorization (classification) is advantageous when the image database is well specified, and labelled training samples are available.

There are a great variety of classification algorithms, mainly grouped into: Bayesian methods, Support Vector Machines (SVM), Decision Trees, Decision Rules, Class Association Rules, Lazy Learners, Neural Networks and Genetic Algorithms [Maimon and Rokach, 2005]. Here we are interested mainly from Decision Trees, Decision Rules and Class Association Rules because for these groups of classifiers [Zaiane and Antonie, 2005]:

- The training process is very efficient;
- The classification is very fast;
- The produced classification models are easily understandable by humans.

The decision tree algorithms are based on a divide-and-conquer approach to the classification problem. They work top-down, seeking at each stage an attribute to split on that best separates the classes; then recursively processing the sub-problems that result from the split. This strategy generates a decision tree, which can be converted into a set of classification rules if necessary [Witten and Frank, 2005]. The typical example of decision tree classifier is Quinlan's C 4.5 [Quinlan, 1993].

An alternative way is to apply the covering approach, used in the decision rules algorithms. It takes each class in turn and seeks for a way of covering all instances in the class, at the same time excluding instances which are not in the class. The representative of this group is RIPPER (Repeated Incremental Pruning to Produce Error Reduction), proposed by William Cohen [Cohen, 1995], which used a heuristic based on the minimum description length principle as stopping criterion for adding more conditions to a rule.

The simplest representative of decision rules is OneR [Holte, 1993]. OneR takes as input a set of examples, each with several attributes and a class. The OneR algorithm chooses the most informative single attribute and bases the rule on this attribute alone. Shortly algorithm consists of creating the rules with antecedent each possible value of each attribute and consequent corresponded class label, after that for each class label find the rule with maximal accuracy.

In [Witten and Frank, 2005] the differences between rules and trees are mentioned in two main directions:

- On the one hand the rules can be symmetric whereas trees must select one attribute to split on first, and this can lead to trees that are much larger than an equivalent set of rules;
- On the other hand in the multiclass case, a decision tree split takes all classes into account, trying to maximize the purity of the split, whereas the rule-generating method concentrates on one class at a time, disregarding what happens to the other classes.

Association rules are similar to classification rules. The associative rule mining algorithms, which had been suggested in the field of market basket analysis for discovering interesting rules from large collections of data [Agrawal et al, 1993] quickly got expanded as Class Association Rules algorithms for modelling relationships between class labels and features from a training set [Bayardo, 1997].

In 2009 we have proposed a CAR algorithm, named PGN, which strives for maximal accuracy of produced rules and use a coverage measure in the classification stage [Mitov et al, 2009a, 2009b]. Being a CAR algorithm it has the already mentioned advantages, from which in the present work the most important one is the creation of easily understandable set of rules, which can be used as a discriminative profile of the examined concept. The use of coverage measure in the classification stage improved the classification accuracy and achieved values comparable to the outcomes of instance-based algorithms. These advantages informed our decision to use PGN in the classification tasks in our experimental system.

Conclusion

We have made an overview of the area of resource discovery in the case of art images. Digital art images provide an important challenge for the resource discovery of images since the methods applied in this domain should be a combination of queries executed on metadata and image retrieval methods. An overview of the evolution of designated repositories for digital art is presented.

The taxonomy of art image content as a source for satisfying different aspects of user wishes is presented.

The feature design is an essential pre-processing step in the CBIR workflow. According to the proposed taxonomy of image content, main features, extracted from image space, are analyzed.

The systematic overview of data reduction strategies, which are crucial in CBIR due to the huge volume of received primary data, is made.

The concept of similarity is very complex and is merely a scientific area in its own right. In the process of obtaining the similarity between two images several processes of finding similarity on different levels and data types had to be addressed. A classification of used similarity measures, depending from the types of extracted features had been outlined.

Some techniques for improving image retrieval such as clustering and categorization had been succinctly introduced.

3 Real and Digitized Colour World

Abstract

In this chapter we offer a succinct review of colour theory from different points of view. The rationale for that is the strong connection of any work on art paintings with the complex area of colour perception. Physiological grounds of this phenomenon are taken as a starting point for focusing the search within art painting images. A brief historical overview of attempts to define colour interconnections and mutual colour influences is made. Those colour models which are best suited to present colours from human point of view, are summarised.

3.1 Colour – Physiology and Psychology

From all the senses that connect us to the world – vision, hearing, taste, smell, and touch – vision is the most important. More than 80% of our sensory experiences are visual [Holtschue, 2006]. When the brain receives light stimulus, it first recognizes shapes and objects and separates the objects from their surrounding environment. Figure-ground separation or pattern recognition is the first cognitive step in the process of perception. In this process, colour plays an important but secondary role. Colour responses are tied stronger to human emotions than to intellectual judgement. Even this property on its own illustrates why colours have such a powerful influence on human perception. The presence of one or more colours in different proportions conveys different messages, which can increase or suppress the perception of the observed objects.

Jointly with shape, colour is one of the fundamental building blocks of visual symbols. It is also closely associated with mental and emotional states, and can affect them profoundly [O'Connel et al, 2009].

Colours play a major role in the field of image retrieval. Within this context it is not the colour itself but the perception of colours and colour combinations as similar or dissimilar what is crucial when one has to extract images by some criterion related to the level of emotional perception, or to search for the specifics of expressiveness of the artist. All these tasks fall in already discussed abstraction aspects of image content. For instance, different painting techniques reflect technical abstractions, as well as the use of colour combinations as particular expressive means grounded on emotional abstractions.

Here we make an overview of the existing qualitative descriptions of the phenomena, which we will use later to propose an appropriate transition to quantitative formal description of successful colour combinations already defined by artists and art researchers.

3.2 Colour Theory

The nature of colour is a subject of study in various sciences. Physics studied electromagnetic structure of light waves, physiology is interested in the perception of light waves as colours, psychology explores the problems of colour perception and its impact on intelligence, mathematics constructs techniques for structuring colour spaces and their measurement.

It appears that the basic laws for the establishment of harmony, colour, different ways to use contrast, the ratio of colour components with other forms of art such as line, plastic, lights, etc., which in theory and practice of painting are born intuitively, have scientific explanations in different disciplines (which does not mean of course that creating of the masterpiece is a simple process of following any schemes blindly).

The problems of colour can be examined from several aspects.

The physicist studies the nature of the electromagnetic energy vibrations and particles involved in the phenomenon of light, the several origins of colour phenomena such as prismatic dispersion of white light, and the problems of pigmentation. He investigates mixtures of chromatic light, spectra of the elements, frequencies and wave lengths of coloured light rays. Measurement and classification of colours are also a topic in physical research.

The chemist studies the molecular structure of dyes and pigments, problems of colour fastness, vehicles, as well as preparation of synthetic dyes. Colour chemistry today embraces an extraordinarily wide field of industrial research and production.

The physiologist investigates the various effects of light and colours on our visual apparatus – eye and brain – and their anatomical relationships and functions. Research on light and dark adaptation and on chromatic colour vision occupies an important place. The phenomenon of afterimages is another physiological topic.

The psychologist is interested in the influence of colour radiation on human mind and spirit. Colour symbolism, and the subjective perception and discrimination of colours, are important psychological problems. Expressive colour effects – what Goethe called ethico-aestetic values of colours – likewise fall within the psychologist's research [Itten, 1961].

Cultural studies and semiology are both concerned with the meaning and interpretation of colours in different cultures.

Engineering is investigating what are the best ways of generating high quality colours in different devices – starting from television and ending with small portable devices.

The artist, finally, is interested in colour effects from their aesthetic aspect, and needs both physiological and psychological information.

Discovery of relationships, mediated by the eye and brain, between colour agents and colour effects in man, is a major concern of the artist. Visual, mental and spiritual phenomena are multiply interrelated in the realm of colour and the colour art [Itten, 1961].

3.3 Physiological Ground of the Colour Perceiving

The colour of the physical point of view is part of the electromagnetic spectrum with a wavelength of 380 nm to 780 nm (usually rounded or between 400 nm and 700 nm). The colour of an object depends on both the physics of the object in its environment and the characteristics of the perceiving eye and brain. Two complementary theories of colour vision are the trichromatic theory and the opponent process theory.

- ***Trichromatic Theory***

Great importance for the development of the colour theory is the Newton discovery in 1666 that white light is a mix of all colours of the spectrum. In 1801, Thomas Jung suggests the hypothesis that mixing only three primary colours can produce all colours. Later Hermann von Helmholtz elaborated on this theory with the assumption that in the retina of the human eye has receptors responding to the three primary colours and all colours are obtained by mixing these three colours with different intensities. The trichromatic theory has been confirmed experimentally in 1960, when three types of receptors were identified in the retina, preferentially sensitive to red, green and blue light waves.

- ***Opponent Theory***

The idea of opponent theory emerged in the studies of Leonardo da Vinci about 1500. Similar views were expressed by Arthur Schopenhauer , but the first integral presentation of this theory had been proposed in the works of Ewald Hering in 1872 [Hering, 1964]. The theory suggests that there are three channels opposing each other: red against green, blue against yellow, and black against white (the latter channel is achromatic and carries information about variations of lightness). Responses to one colour of an opponent channel are antagonistic to those of the other colour. To put it in another way, there are certain pairs of colours one never sees together at the same place and at the same time. One does not see reddish greens or yellowish blues but does see yellowish greens, bluish reds, yellowish reds, etc.

One practical example for this theory is the so called after-image phenomenon: if one looks at a unique red patch for about a minute and then switches the gaze to a homogeneous white area he would see a greenish patch on the white area. In other words after-image will produce such colours that in combination with the first colour are neutral.

The opponent theory was confirmed in the 1950s, when opposing colour signals were found in optical connections between the eye and brain. At that time a pair of visual scientists working at Eastman Kodak conceived a method for quantitatively measuring the opponent processes responses. Leo Hurvich and Dorothea Jameson invented the hue cancellation method to psychophysically evaluate the opponent processing nature of colour vision [Hurvich and Jameson, 1957].

Modern theories combine these two theories: the process starts by light entering the eye, which stimulates the trichromatics cones in the retina, and is further processed into three opponent signals on their way to the brain. More recent developments are in the Retinex theory, proposed by Edwin Land. Experiments show that people have a considerable amount of colour constancy (i.e. colours are perceived the same even under different illumination) [Gevers, 2001].

➢ ***Colour Perception***

Colour perception is not an independent process and is influenced by conditions in which this act takes place. On the one hand a physical interference of waves leads to the perception of two colours as another one, which had been used by impressionists after the invention of new techniques of laying the paints. On the other hand, the perception of colour provokes mutual induction of the nerve processes; according to Pavlov, the law of mutual induction of nerve processes is one of the fundamental laws of the nerve physiology [Raychev, 2005]. Mutual induction in the perception of colour leads to a change in the perception of a given colour, depending on the stimuli in another part of the retina (simultaneous contrast) or stimuli applied earlier on the same spot of the retina (consecutive contrast).

Contrasting colour changes, resulting from the simultaneous operation of different colours, can be analysed through three main features characterizing the colour – hue, brightness and saturation. Exceptions are achromatic and monochromatic images that use only the contrast of brightness. The perception of colour depends largely on the background. Under its influence, colours are seen in other tints and shades.

The perception of achromatic colours, placed among chromatic ones, is also changing. Gray colour on a red background is perceived with a greenish hue, on a yellow background – with a bluish one, on a green – with a pinkish one, on a blue – with a yellowish, i.e. the colour hue of the object acquires the additional background colour.

The same principles are valid to consecutive contrast, but in this case cannot refer to the background and the object, but to preceding visual stimuli, which affect the next colour, laid in the same position of sight. This is a result of the colour eye fatigue. The result is a coherent image, which remains different depending on the length of preceding visual stimulus as well as on its colour composition.

3.4 Image Harmonies and Contrasts

The contrasts are experienced when we establish differences between two observed effects. When these differences reach maximum values we talk about diametrical contrast. Our senses perceive only through comparison. For instance an object is perceived as being short when it is near to a long object and vice versa. In a similar way colour effects become stronger or weaker thorough contrasts.

Multiple scholars observed and examined the influence of colours on each other. Aristotle in his "De meteorologica" formulated questions about the difference of violet near white or black wool [Gage, 1993].

In 1772 – the same year that Johann Heinrich Lambert constructed his colour pyramid and demonstrated for the first time that the completeness of colours can only be reproduced within a three dimensional system [Spillmann, 1992], another colour circle was published in Vienna by Ignaz Schiffermüller. He was one of the first who arranged the complementary colours opposite each other: blue opposite orange; yellow opposite violet; red opposite green [Gage, 1993].

Leonardo da Vinci noticed that when observed adjacent to each other, colours are influencing the perception. Goethe, however, was the first to specifically draw attention to these associated contrasts.

Johann von Wolfgang Goethe in his book Theory of Colours, published in 1810, studied the emotion and psychological influence of colours. His six-hue spectrum of colours remains the standard for artists even nowadays [Birren, 1981].

Michel Eugène Chevreul (1786-1889) had contributed to the study of contrast establishing the law of *simultaneous contrast* in 1839 [Gage, 1993]. When colours interact, they are capable of change in appearance, depending on particular relationships with adjacent or surrounding colours. Simultaneous contrast is strongly tied to the phenomenon of afterimage, also known as *successive contrast*, when the eye spontaneously generates the complementary colour even when the hue is absent. The explanation of successive contrast is given in opponent colour vision theory. Successive and simultaneous contrasts suggest that the human eye is satisfied, or in equilibrium, only when the complementary colour relation is established.

Research on the mutual influences of colours had strongly manifested in the studies of Georges Seurat who suggested the optical fusion theory, also called Pointillism or Illusionism. The theory behind this optical mixture was set out as early as in the 2nd century by Ptolemy who identified two ways of achieving optical fusion; one by distance where "the angle of vision formed by rays of light from the very small patches of colour was too small for them to be identified separately by the eye, hence many points of different colours seemed together to be the same colour" [Gage, 1993]. The other related to after images and moving objects. The use of this theory lays in the established new painting technique, firstly showed by Seurat in his painting "Sunday Afternoon on the Island of La Grande Jatte" in 1886. He called this phenomenon "Chromoluminarisme" or "Peinture Optique". The Pointillist technique consists of "placing a quantity of small dots of two colours very near each other, and allowing them to be blended by the eye at the proper distance" [Birren, 1981].

Adolf Hoelzel suggested seven contrast groups, based on his own understanding of the colour wheels. Every contrast marks some quality of colour perception. His contrasts are: (1) *Contrast of the hue*; (2) *Light-Dark*; (3) *Cold-Warm*; (4) *Complementary*; (5) *Gloss-Mat*; (6) *Much-Little*; (7) *Colour-Achromatic* [Gage, 1993].

The great contribution in revealing effects of colour interactions was made by Josef Albers (1888-1976). His book "The Interaction of Colour" [Albers, 1963] became quintessential in understanding colour relationships and human perception. Albers stated that one colour could have many "readings", dependent both on lighting and the context in which it is placed. He felt that the comprehension of colour relationships and interactions was the key to gaining an eye for colour. According to Albers, we rarely see a colour that is not affected by other colours. Even when a colour is placed against a pure neutral of black, white, or gray, it is influenced by that neutral ground. Colours interact and are modified in appearance by other colours in accordance with three guiding rules: *Light/dark value contrast*, *Complementary reaction*, and *Subtraction*.

Johannes Itten (1888-1967) expanded the theories of Hoelzel and Albers. He defined and identified strategies for successful colour combinations [Itten, 1961]. Through his research he devised seven methodologies for coordinating colours utilizing the hue's contrasting properties. These contrasts add other variations with respect to the intensity of the respective hues; i.e. contrasts may be obtained due to light, moderate, or dark value. He defined the following types of contrasts:

- *Contrast of hue*: the contrast is formed by the juxtaposition of different hues. The greater the distance between hues on a colour wheel, the greater the contrast;
- *Light-dark contrast*: the contrast is formed by the juxtaposition of light and dark values. This could be a monochromatic composition;

- *Cold-warm contrast*: the contrast is formed by the juxtaposition of hues considered "warm" or "cold";
- *Complementary contrast*: the contrast is formed by the juxtaposition of colour wheel or perceptual opposites;
- *Simultaneous contrast*: the contrast is formed when the boundaries between colours perceptually vibrate. Some interesting illusions are accomplished with this contrast;
- *Contrast of saturation*: the contrast is formed by the juxtaposition of light and dark values and their relative saturation;
- *Contrast of extension* (also known as the *Contrast of proportion*): the contrast is formed by assigning proportional field sizes in relation to the visual weight of a colour.

3.5 Psychological Colour Aspects

The colour impact on people depends on many factors, where physical laws and physiology are only the beginning. Psychological perception plays an important role in this process which is influenced by the particular psychological state on the one hand, and by socio-cultural environment in which the character of a person is composed on the other hand. Perception of colour brings the whole emotional and mental identity of the observer, his/her intelligence, memory, ideology, ethics, aesthetic feelings and other sensations. These feelings as well as philosophical, religious and other aspects of the categories of colour perception are essential to its nature and create relative symbolic aspects of colour impression.

Using colour as a symbol dates back to antiquity. Simple natural feeling, caused by the colour, gradually had been canonized in a system of secular or religious symbols. Thus a deep stratification of religious, social, historical, moral, ethical, psychological, etc. symbolism occurs, which sometimes leads to impossibility to detect what is the primary affective value of a particular colour. Heraldry is a typical example of such formal system in which every colour and composition conditionally acquired some symbolic meaning [Raychev, 2005]. Such orderly system of colour symbols is built in within the liturgical system of the Catholic, Orthodox and Protestant churches in which each liturgical colour carries some message and importance and may be used only under certain circumstances. The white colour for example symbolizes innocence, purity and joy. The red colour symbolizes fire and blood, sacrifice and martyrdom. The green colour brings hope and life. The purple colour is associated with relaxation, contemplation and repentance. The pink colour marks moments of joy during periods of penance and fasting. The black colour is associated with sorrow and sadness. Gold is allowed during the holidays and is associated with praise and high mood [Goldhammer, 1981].

Such symbolic colour systems had been developed in almost all nations. For example, in China five colours: white, black, blue, yellow and red symbolize certain concepts and attributes of the objects of the world around us as the cardinal directions, seasons, weather events, taste, character features and others [Raychev, 2005].

There are general principles of psychological elements of colour perception regardless of the formation sources of conditional symbolic systems in different cultures. All phenomena in the field of psychological perception of colour, which cannot be explained as a direct result of the visual impression of colour, can find an explanation by way of association. Associations are built between factors which coexist constantly or frequently. For instance binding of green with the concept of hope is the result of constantly repeated relationship between green plants and hope for future good harvest. Associations can be strictly individual, for example the relationship between blue and the mother can arise only for a person whose mother is blue-eyed or wears blue most of the time, but by no means it can be a common association.

Adopted secular and religious symbolism of colour is reflected in art because of its social function. For example there is no reason to be surprised that in West European Renaissance works of art before Constable (1776-1837) missed green tones. The academic style of painting at that time has presented a green with brown tones with scarce presence of green. Red robes as a symbol of the martyrdom of Jesus seemed natural and in line with religious symbolism in the painting of El Greco "The disrobing of Christ (1583)".

3.6 Colour Models

Attempts to refer colours to each other and establish correlations between them have been made from Antiquity. The ancient models were coupled with more mystical links between colours and primary forces (the belief in ancient Greece that colours arose from the struggle between light and darkness), or musical tones, or planets, etc. Colour models established later gradually focused on the relationships of colours each other, searching for basic characteristics as well as exploring which combinations can cover the whole spectrum of visible colours. A detailed survey of colour models was made by the team of Urs Baumann[33]. A particularly comprehensive overview is provided also in the remarkable book of John Gage "Color and Culture" [Gage, 1993]. We will not present in details all steps toward revealing the colour aspects. In Figure 8 we only show a timeline of colour systems and/or contributors to the development of colour knowledge.

[33] ©Echo Productions, http://www.colorsystem.com/

Antiquity Middle Ages 1600 ... ... 1750 1775 1800 1825 1850 1875 1900 1925 1950 1975 2000

Pythagoras
Aristotle
Plato
(1220-1235) Robert Grosseteste (England)
(1435) Leon Battista Alberti (Italy)
(1510) Leonardo da Vinci (Italy)
(1611) Aron Sigfrid Forsius (Finland)
(1613) Franciscus Aguilonius (Belgium)
(1629-31) Robert Fludd (England)
(1646) Athanasius Kircher (Germany)
(1766-1770) Moses Harris (England)
(1772) Ignaz Schiffermueller (Austria)
(1772) Johann Heinrich Lambert (Germany)
(1809) James Sowerby (England)
(1810) Johann Wolfgang von Goethe (Germany)
(1810) Philipp Otto Runge (Germany)
(1826) Charles Hayter (England)
(1839) Michel Eugene Chevreul (France)
(1846) George Field (England)
(1855-1860) RGB - James Clerck Maxwell (England)
(1860) Hermann von Helmholtz (Germany)
(1868) William Benson (England)
(1874) Wilhelm von Bezold (Germany)
(1874, 1893) Wilhelm Wundt (Germany)
(1878) Ewald Hering (Austria)
(1879) Charles Blanc (France)
(1879) Nicholas Odgen Rood (USA)
(1883-1897) Alois Hoefler (Austria)
(1890) Charles Lacouture (France)
(1902) Hermann Ebbinghaus (Germany)
(1905-1916) Albert Henry Munsell (USA)
(1912) Robert Ridgway (USA)
(1916-1917) Wilhelm Ostwald (Germany)
(1923) Michel Jacobs (Canada)
(1924) Max Becke (Austria)
(1927-1928) R. Luther; N. Nyberg (Germany)
(1928) CIE - S. Roesch (Germany)
(1929) Arthur Pope (USA)
(1929) Edwin Boring (USA)
(1931) CIE 1931 System
(1934) Faber Birren (USA)
(1937-1939) Tryggve Johansson (Sweden)
(1941-1953) DIN System (Germany)
(1944) CIE - D.L. MacAdam (USA)
(1946) CIE - Walter Stiles (USA)
(1952) Alfred Hickethier (Germany)
(1953) Sven Hesselgren (Sweden)
(1955) ISCC-NBS System (USA)
(1960) OSA System (USA)
(1962-1974) Coloroid-System (Hungary)
(1965) Aemilius Mueller II (Switzerland)
(1968-1969) NCS-System (Sweden)
(1975) J. Frans Gerritsen (Holland)
(1976) CIE L*A*B* System
(1978) ACC System (Netherlands)
(1983) Michel Albert-Vanel (France)
(1986) CMN System (Italy)
(1978) HSL, HSV Systems - A.R.Smith (USA)
(2003) HMMD - H. Kim and J. Lee (USA)

Figure 8. History of researching in the colour phenomenon

A colour model is an abstract mathematical model for describing the colour as a numerical vector, usually with three or four values, which are called colour components. Colour space is another concept used to represent colours.

The difference between these two terms is that a colour model is intended to present the interconnections between colours, but is not exactly fixed to real colours. When a mapping function to the real colours is associated to a colour model, a part of the visible spectrum, known as a gamut, represents this colour space. In the most generic sense of the definition above, colour spaces can be defined without the use of a colour model. These spaces, such as Pantone, are in effect a given set of names or numbers which are defined by the existence of a corresponding set of physical colour swatches. As it could be seen, colour space is a more specific term. For a certain combination of a colour model plus a colour mapping function, the term "colour space" tends to be used to identify colour models, since identifying a colour space automatically identifies the associated colour model.

Informally, the two terms are often used interchangeably, though this is incorrect. For example, although several specific colour spaces are based on the RGB model (Adobe RGB, sRGB), there is no such thing as the RGB colour space.

Different models and spaces serve various domains – from physics and colorimetry; through painting, architecture, and design; to digital coding for printers, monitors and TV. The history and practice show that a perfect colour model cannot be created: one is suitable to supply compact coding and transmitting of the colour characteristics, another is easy perceived from humans, etc. Colour spaces used in various processes such as image display, video information transmission, printing, storing image and video files are aimed at maximum density and speed of perception of the device. In most cases, the methods of the colour coding in the technological system are hard to understand for humans. A typical example of such a space is the compact model YCrCb, used in television transmission and storage of JPG-files containing information about luminance (Y – Luma component) and two colour components (Cr and Cb), showing variations of red and blue respectively.

From human point of view, it is most easy to define the colour as a composition of three components – hue, saturation and lightness. Hue means the name of the colour – red, orange, etc. Black, different shades of gray and white are called achromatic. Saturation measures the hue intensity or brilliance of a sample, its dullness or vividness. Lightness refers to relative light and dark in a sample [Holtzschue, 2006]. Such point of view on the colour facilitates the structuring of colour contrasts and harmonies are evinced in art images.

Now we will discuss some of presented colour models, which are connected with human perception of the colour and with the representation of visual digital items.

- ***Artist's Colour Wheel***

In his experiments with light, Isaac Newton recognized that colours could be created by mixing colour primaries. In his *Opticks*, Newton published a colour wheel to show the geometric relationship between these primaries (Figure 9) [Gage, 1993]. This circular diagram became the model for many colour systems of the 18th and 19th centuries.

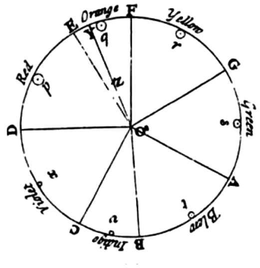


Figure 9. Newton's Colour Circle

Another approach was suggested by the German painter Phillip Otto Runge who aimed to establish the complete world of colours resulting from the mixture of the three (yellow, red, and blue), among themselves and together with white and black (Figure 10). Runge presented his views on colour order in a sketch of a mixture circle, with the three primary colours forming an equilateral triangle and together with their pair-wise mixtures - a hexagon. He arrived at the concept of the colour sphere sometime in 1807 by expanding the hue circle into a sphere, with white and black forming the two opposing poles. This proposal had been overshadowed shortly after by Michel Eugène Chevreul's hemispherical system of 1839. A spherical colour order system was patented in 1900 by Albert Henry Munsell, soon replaced with an irregular form of the solid.

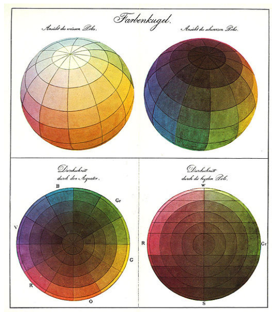


Figure 10. Runge's Colour Sphere

The RYB (Red-Yellow-Blue) model is used by artists (Figure 11). It had been developed before physics examination of the lengths of light waves and the discovery that the primary colours are not actually such. These are colours that artists are accustomed to accept as sufficient on their palettes to obtain the other colours. Actually it is a slightly shifted analogue of CMY, and its proximity to the RGB-model is somewhat confusing. RYB make up the primary colour triad in a standard artist's colour wheel. The secondary colours VOG (violet-orange-green) also make up another triad. Triads are formed by 3 equidistant colours on a particular colour wheel. In the 18th century, the RYB primary colours became the foundation of theories of colour vision, as the fundamental

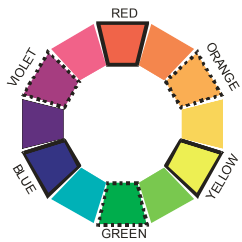


Figure 11. The Artists' Colour Wheel

sensory qualities that are blended in the perception of all physical colours and equally in the physical mixture of pigments or dyes. These theories were enhanced by 18th-century investigations of a variety of purely psychological colour effects, in particular the contrast between "complementary" or opposing hues that are produced by colour afterimages and in the contrasting shadows in coloured light. These ideas and many personal colour observations were summarized in two founding documents in colour theory: the Theory of Colours (1810) by Johann Wolfgang von Goethe, and The Law of Simultaneous Colour Contrast (1839) by Michel-Eugène Chevreul.

Subsequently, German and English scientists established in the late 19th century that colour perception is best described in terms of a different set of primary colours – red, green and blue (RGB) – modelled through the additive, rather than subtractive, mixture of three monochromatic lights.

Painters have long used more than three primary colours in their palettes – and at one point considered red, yellow, blue, and green to be the four primaries. Red, yellow, blue, and green are still widely considered the four psychological primary colours, with black and white added as additional primary colours.

- ***Munsell Colour System***

In the first decade of 20th century Professor Albert Munsell suggested a colour system, where colours are specified according to three colour dimensions: hue, lightness, and chroma [Munsell, 1969]. This is the first model were colour is separated into perceptually uniform and independent dimensions, and the first, which systematically illustrates the colours in three dimensional space.

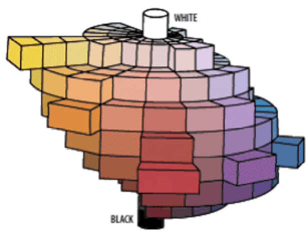


Figure 12. Munsel colour model

Munsell defined hue as "the quality by which we distinguish one colour from another". He selected five principle colours: red, yellow, green, blue, and purple; and five intermediate colours: yellow-red, green-yellow, blue-green, purple-blue, and red-purple. He arranged these colours in a wheel measured in units of 100 compass points: Munsell's arrangement of colours in this way was also important for his concept of colour harmony, or balance (Figure 12). Munsell was a conservative artist with strict views on the aesthetics of painting. He wanted his system to serve not only as a guide for notating colours, but also as a guide for choosing complimentary colours for artistic work.

Value was defined by Munsell as "the quality by which we distinguish a light colour from a dark one". Value is a neutral axis that refers to the grey level of the colour.

Chroma is "the quality that distinguishes the difference from a pure hue to a gray shade". Very specific feature of this colour model is that chroma is not uniform for every hue at every value. Munsell saw that full chroma for individual hues might be achieved at very different places in the colour sphere. In the Munsell system, reds, blues, and purples tend to be stronger hues that average higher chroma values at full saturation, while yellows and greens are weaker hues that average fullest chroma saturation relatively close to the neutral axis. Reds, blues, and purples reach fullest saturation at mid-levels on the value scale, while yellows and greens reach it at higher values.

As a result of these differences, the three-dimensional solid representation of Munsell's system is radically asymmetrical. Munsell's system, although dating near the 19th century and devised more by intuition than exact science, is still an internationally accepted, leading colour system.

- ***RGB (Red-Green-Blue) Colour Model***

The great part of the colours in the visible spectrum can be obtained by mixing of light waves corresponding to red, green and blue with different intensity. RGB is an additive system used to describe colours in light sources (e.g. monitors). The colours represented in RGB coordinates are hardly comprehensible by humans but most of the digital visual items are rendered in this format.

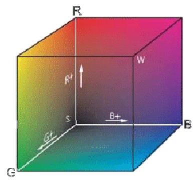


Figure 13. RGB Colour Model

Typically, digital display devices represent each component of an RGB colour coordinate as an n-bit integer in the range of 0 to 2^n-1 (Figure 13). Each displayable colour is an RGB coordinate triple (R, G, B) of n-bit numbers yielding a palette containing 2^{3n} total colours. In the current systems usually n=8; therefore each colour coordinate can range from 0 to 255, and the total palette contains 2^{24} colours. Each component contains the presence of corresponded colour in the range from 0 to 255. The value 0 means that this colour component does not take part in the construction of a chosen colour, while the value 255 (FF hexadecimal) corresponds to maximum presence of this colour component. For instance, (0, 0, 0) encodes black colour; (FF, FF, FF) encodes white. More generally, equal values of each component encode the grayscale from black to white. (FF, 0, 0) encodes spectral red colour. (FF, FF, 0) stands for spectral yellow colour (as a mixture of red and green colours).

- ***CMY (Cyan-Magenta-Yellow) Colour Model***

The CMY and CMYK (Cyan-Magenta-Yellow-blacK) colour models are the opposite of the RGB colour model and are used in printing processes. Black is added as a colour because its creation as a mixture of other pigments is impractical and inefficient. This system is subtractive, because it produces objects that reflect light. In the present work we are focused on the direct observation of a digital object, therefore this colour model will not be used. However, we mention it here due to its connection with the RYB colour model.

➢ ***HSL (Hue-Saturation-Lightness), HSV (Hue-Saturation-Value) Colour Models***

HSL (Hue-Saturation-Lightness) and HSV (Hue-Saturation-Value) are two related representations of points in an RGB colour model that attempt to describe perceptual colour relationships more accurately than RGB colour model, while remaining computationally simple (Figure 14). Both models were first formally described in 1978 by Alvy Ray Smith [Smith, 1978]. We present both models, because HSV is used in MPEG-7 descriptors and HSL we used in some of our experiments.

Figure 14. Hue in HSL and HSV Colour Wheel

In HSL and HSV colour models the colour is presented by three independent characteristics (attributes):

– Hue (colour tone): shows the specific combination of colours, it brings the name of the colour – red, green, etc. The black, gray and white colours do not have a colour tone and are called achromatic. Hue is represented as an angle on the colour circle (i.e. the rainbow represented in a circle). This angle is measured in degrees. By definition red is 0° or 360°. The other colours are spread around the circle. For instance: Green is located at 120°, Blue at 240°, etc. Achromatic colours are coded by "-1";

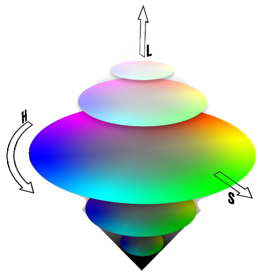


Figure 15. HSL colour model

– Saturation: indicates the purity of the colour and may vary from clear to cloudy. The saturation is presented as a value in the range [0, 1]: 1 means full saturation, 0 is a shade of grey;

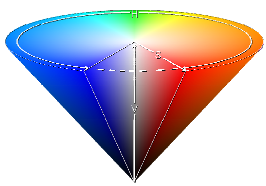


Figure 16. HSV colour model

– Lightness/Value: presents how dark or light is a particular colour in the range from black to white. Lightness/Value also are expressed in values in the range [0, 1], but the function of correspondence of lightness and value are different. Lightness 0 is black, 1 is white, and 0.5 is "normal" lightness (Figure 15). Value component 0 also means black, but 1 represents the plane where colours are completely saturated (Figure 16).

Some examples for presenting different colours in both models follow:

- In HSL colour model: (H=0°, S=1, V=0.5) means spectral red; (H=180°, S=1, V=0.25) means dark cyan; (H=120°, S=1, V=0.75) encodes light green;
- In HSV colour model: (H=0°, S=1, V=1) means spectral red; (H=180°, S=1, V=0.5) means dark cyan; (H=-1, S=0, V=v) codes the gray colours from v=0: black to v=1: white.

Based on the definitions in [Smith, 1978], taking into account the chosen measurements of different coordinates, the conversions from RGB colour model to HSV and HSL colour models is defined as follows:

If $R,G,B \in [0,255]$ are the red, green, and blue coordinates in RGB colour model. Let max be the greatest of R, G, and B, and min the least.

The hue angle in HSL and HSV colour models $H \in [0,360]$ is calculated as:

$$H = \begin{cases} 0 & \text{if } max = min \\ (60^\circ \times \frac{G-B}{max-min} + 360^\circ) \bmod 360^\circ & \text{if } max = R \\ 60^\circ \times \frac{B-R}{max-min} + 120^\circ & \text{if } max = G \\ 60^\circ \times \frac{R-G}{max-min} + 240^\circ & \text{if } max = B \end{cases} .$$

The saturation in HSL colour model $S \in [0,1]$ is calculated as:

$$S = \begin{cases} 0 & \text{if } max{=}min \\ \frac{max-min}{max+min} & \text{if } max+min \le 1/2 \\ \frac{max-min}{2-(max+min)} & \text{if } max+min > 1/2 \end{cases} .$$

The luminance in HSL colour model $L \in [0,1]$ is calculated as:

$$L = \frac{max+min}{2*255} .$$

The saturation in HSV colour model $S \in [0,1]$ is calculated as:

$$S = \begin{cases} 0 & \text{if } max=0 \\ 1-\dfrac{min}{max} & \text{if } max>0 \end{cases}.$$

The value in HSV colour model $V \in [0,1]$ is calculated as:

$$V = \frac{max}{255}.$$

The CSS3 specification from the W3C states, "Advantages of HSL colour model are that it is symmetrical to lightness and darkness (which is not the case with HSV colour model)" [W3C, 2011]. From one side HSL colour model better reflects the intuitive notion of "saturation" and "lightness" as two independent parameters, but from other side its definition of saturation is not very clear, because, for example very pastel, almost white colour can be defined as fully saturated in the HSL colour model. It might be controversial, though, whether HSV colour model or HSL colour model is more suitable for use in human user interfaces. Another problem with the definition of lightness/value is that they have little to do with colour science definitions of the terms. In photometry, concepts such as lightness and luminance relate to a weighted spectrum, in which green is much more luminous than blue and red is in between. In HSL and HSV colour models, the "L" and "V" dimensions treat the three colour channels equally. These disadvantages are compensated of simplicity of the transforming formulas from/to RGB colour model.

- ***YCbCr Colour Model***

YCbCr is the second primary colour model used to represent digital component video, the other one being RGB colour model. The difference between these two models is that YCbCr represents colour as brightness and two colour difference signals, while RGB represents colour as red, green and blue. In YCbCr colour model, *Y* is the brightness (luma), *Cb* is blue minus luma and *Cr* is red minus luma. YCbCr is a digital MPEG compression, which is used in DVDs, digital TV and Video CDs, and digital camcorders (MiniDV, DV, Digital Betacam, etc.) output YCbCr over a digital link such as FireWire or SDI.

If normalized R, G, B components are $r, g, b \in [0,1]$, then according to [ISO/IEC 15938-3]:

$$Y = 0.299 * r + 0.587 * g + 0.114 * b$$

$$Cb = -0.169 * r - 0.331 * g + 0.500 * b$$

$$Cr = 0.500 * r - 0.419 * g - 0.081 * b$$

It follows that the *Y* component values are in the range [0, 1] and *Cb* and *Cr* component values are in the range [-0.5, 0.5].

➢ *HSL-artists Colour Model*

HSL colour model presents perceptual colour dimensions accurately from human point of view and is computationally simple, but it suffers from some disadvantages. HSL colour model does not present Hue as it is adopted in artists' practices. On the other hand Lightness does not take into account the dependence of Hue. Spectral yellow and spectral indigo have the same lightness value 0.5, but indigo is vastly darker then yellow.

However, the YCbCr colour model is very useful and widely adopted in high-speed video transition. In this model the luma component (Y) represents very accurately the brightness of the colour, but the other two components (Cb and Cr) are not convenient for human perception of the nature of the colour.

In order to overcome these disadvantages, in our work we use one colour model, which has also three dimensions – Hue, Saturation and Luma. Hue is based on the definition of Hue in the HSL colour model. But later, when we use quantization of the Hue, we take into account the misplacement of artists' colour wheel (Figure 11). In this way we establish the basis for easy definition of harmonies and contrast relationships as functions of the angles between hues. Saturation is calculated in the same way as in HSL colour model. The luma component is defined in YCbCr colour model as $Luma = \frac{0.299 * R + 0.587 * G + 0.114 * B}{255}$ where $R, G, B \in [0,255]$ are the red, green, and blue coordinates in the RGB colour model. This formula is widely used for converting the colour images into grayscale images in many products, for instance in MathLab[34].

➢ *HMMD (Hue-Max-Min-Diff) Colour Model*

HMMD colour model [Kim and Lee, 2003], is supported in MPEG-7. The *Hue* has the same meaning as in the HSV and HSL models. *Max* and *min* are the maximum and minimum among the R, G, and B values, respectively. *Max* indicates how much black colour it has, giving a flavour of the shade or blackness. *Min* indicates how much white colour it has, giving a flavour of the tint or whiteness. The *Diff* component is defined as the difference between max and min. *Diff* indicates how much gray it contains and how close to the pure colour, giving a flavour of the tone or colourfulness. The last component *Sum* is calculated as a mean

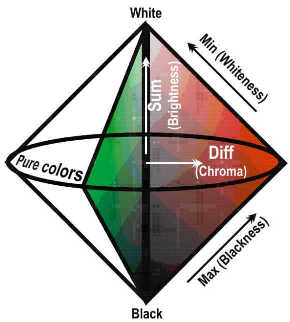


Figure 17. HMMD colour model

[34] http://www.mathworks.com/access/helpdesk/help/toolbox/images/rgb2gray.html

value between *max* and *min* and simulates the brightness of the colour. *Hue* component takes values in the range of [0, 360]. *Max*, *Min*, *Sum* and *Diff* components have values in the range of [0, 1]. Three of the components are sufficient to describe the HMMD colour model – (*Hue*, *Max*, *Min*) or (*Hue*, *Diff*, *Sum*). This colour model can be depicted using the double cone structure as shown in the Figure 17. In the MPEG-7 core experiments for image retrieval it was observed that the HMMD colour model is very effective and compared favourably to the HSV colour model.

The real world is complicated indeed. One phenomenon, depending of different points of view and different purposes is captured by multiple models, which integrate the specific features concerning only these aspects, which are most significant for the goal a particular model seeks to achieve. The variety of colour models does not go beyond these frameworks. Some of them are easy to comprehend by the human user such as the Artist's colour wheel, Munsell system, HSL and HSV colour models. Others are specifically designed for fast processing and serve different media, for instance RGB for monitors, CMYK for printers. Other models, such as YCbCr and HMMD, offer appropriate structures for fast metadata extraction.

3.7 MPEG-7 Standard

The Moving Picture Experts Group (MPEG) [ISO/IEC JTC1/SC29 WG11] was formed by the ISO in 1988 to set standards for audio and video compression and transmission. A series of currently widely spread standards such as the standard for audio recording MP3 (MPEG-1 Layer 3, ISO/IEC 11172), the standards for transmission for over the air digital television, digital satellite TV services, digital cable television, DVD video and Blue-ray (MPEG-2, ISO/IEC 13818 and MPEG-4, ISO/IEC 14496) are outcomes of this group. In addition to the above standards, the group deals with different standards to describe content and environment. Here our interest is focused on the MPEG-7 standard ISO/IEC 15938, named "Multimedia Content Description Interface" [ISO/IEC 15938-3], which provides standardized core technologies allowing the description of audiovisual data content in multimedia environments. Audiovisual data content that has MPEG-7 descriptions associated with it may include: still pictures, graphics, 3D models, audio, speech, video, and composition information about how these elements are combined in a multimedia presentation (scenarios).

The MPEG-7 descriptions of visual content are separated into three groups:

- *Colour Descriptors:* Colour Space, Colour Quantization, Dominant Colours, Scalable Colour, Colour Layout, Colour Structure, and GoF/GoP Colour;
- *Texture Descriptors:* Homogeneous Texture, Edge Histogram, and Texture Browsing;
- *Shape Descriptors:* Region Shape, Contour Shape, and Shape 3D.

Colour descriptors used in MPEG-7 are as listed below:

- *Colour Space:* The feature colour space is used in other colour based descriptions. In the current version of the standard the following colour models are supported: RGB, YCrCb, HSV, HMMD, Linear transformation matrix with reference to RGB and Monochrome;
- *Colour Quantization:* This descriptor defines a uniform quantization of a colour space. The number of bins which the quantizer produces is configurable; this allows for great flexibility within a wide range of applications. For a meaningful application in the context of MPEG-7, this descriptor has to be combined with Dominant Colour descriptors, e.g. to express the meaning of the values of dominant colours;
- *Dominant Colour(s):* This colour descriptor is most suitable for representing local (object or image region) features where a small number of colours are enough to characterize the colour information in the region of interest. Whole images are also applicable, for example, flag images or colour trademark images. Colour quantization is used to extract a small number of representative colours in each region/image. Correspondingly, the percentage of each quantized colour in the region is calculated. A spatial coherency on the entire descriptor is also defined, and is used in similarity retrieval. The specific presentation of this descriptor allows for the variety of possibilities of using different kind of similarity measures. The Earth mover distance [Wang et al, 2003] is the most convenient for such kind of features. Other types of similarity measures are used in [Yang et al, 2008];
- *Scalable Colour:* The descriptor specifies a colour histogram in HSV colour space, which is encoded by a Haar transformation. Its binary representation is scalable in terms of bin numbers and bit representation accuracy over a broad range of data rates. The Scalable Colour descriptor is useful for image-to-image matching and retrieval based on colour feature. Retrieval accuracy increases with the number of bits used in the representation. The sum of absolute difference of coefficients can be used (L_1 metric) as a distance measure;
- *Colour Layout:* This descriptor effectively represents the spatial distribution of colour of visual signals in a very compact form. This compactness allows visual signal matching functionality with high retrieval efficiency at very small computational costs. It provides image-to-image matching without dependency on image format, resolutions, and bit-depths. It can be also applied both to a whole image and to any connected or unconnected parts of an image with arbitrary shapes. It also provides very friendly user interface using hand-written sketch queries since this descriptor captures the layout information of colour feature. The sketch queries are not supported in other colour descriptors. The colour Layout

descriptor uses the YCbCr colour space with 8 bits quantization. The elements of colour Layout specify the integer arrays that hold a series of zigzag-scanned DCT coefficient values. The DCT coefficients of each colour component are derived from the corresponding component of local representative colours. For similarity measure can be used standard L_1 or L_2 metrics as well as specific functions, which takes into account the significance of the order of coefficients [Herrmann, 2002];

– *Colour Structure:* This is a colour feature descriptor that captures both colour content (similar to a colour histogram) and information about the structure of this content. Its main functionality is image-to-image matching and its intended use is for still-image retrieval, where an image may consist of either a single rectangular frame or arbitrarily shaped, possibly disconnected, regions. The extraction method embeds colour structure information into the descriptor by taking into account all colours in a structuring element of 8x8 pixels that slides over the image, instead of considering each pixel separately. Unlike the colour histogram, this descriptor can distinguish between two images in which a given colour is present in identical amounts but where the structure of the groups of pixels having that colour is different in both images. Colour values are represented in the double-coned HMMD colour space, which is quantized non-uniformly into 32, 64, 128 or 256 bins. Each bin amplitude value is represented by an 8-bit code. The Colour Structure descriptor provides additional functionality and improved similarity-based image retrieval performance for natural images compared to the ordinary colour histogram. The descriptor expresses local colour structure in an image by means of a structuring element that is composed of several image samples. The semantics of the descriptor, though related to those of a colour histogram, is distinguishable in the following way. Instead of characterizing the relative frequency of individual image samples with a particular colour, this descriptor characterizes the relative frequency of structuring elements that contain an image sample with a particular colour. Hence, unlike the colour histogram, this descriptor can distinguish between two images in which a given colour is present in identical amounts but where the structure of the groups of pixels having that colour is different in the two images. Usually the sum of absolute normalized difference of coefficients is used (L_1 metric) as a distance;
– *GoF/GoP Colour:* The Group of Frames/Group of Pictures Colour descriptor extends the Scalable Colour descriptor that is defined for a still image to colour description of a video segment or a collection of still images. The same similarity/distance measures that are used to compare Scalable Colour descriptions can be employed to compare GoF/GoP Colour descriptors.

From texture descriptors we will stop our attention on the following ones:

- *Edge Histogram:* This descriptor 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 in each local region called a sub-image. The sub-image is a part of the original image and each sub-image is defined by dividing the image space into 4x4 non-overlapping blocks, linearized by raster scan order. For each sub-image a local edge histogram with 5 bins is generated. As a result 16*5=80 histogram bins forms an Edge Histogram descriptor array. Each sub-image if divided into image-blocks. The value for each histogram bin in is related to the total number of image blocks with the corresponding edge type for each sub-image. These bin values are normalized by the total number of image blocks in the sub-image and are non-linearly quantized by quantization tables, defined in MPEG-7 standard. For this descriptor can be used each similarity measure function for histograms. [Won et al, 2002] suggests an extension to this descriptor in order to capture not only the local edge distribution information but also semi-global and global ones;
- *Homogeneous Texture:* This descriptor characterizes the region texture using the energy and energy deviation in a set of frequency channels. This is applicable for similarity based search and retrieval applications. The frequency space from which the texture features in the image are extracted is partitioned with equal angles of 30 degrees in the angular direction and with an octave division in the radial direction. As a result of applying of 2D Gabor function for feature channels and consequent quantization and coding average, standard deviation, energy and energy deviation are extracted.

The main issue with the MPEG-7 standard is that it focuses on the representation of descriptions and their encoding rather than on the practical methods on the extraction of descriptors. The creation and application of MPEG-7 descriptors are outside the scope of the MPEG-7 standard. For example, description schemes used in MPEG-7 specify complex structures and semantics groupings descriptors and other description schemes such as segments and regions which require a segmentation of visual data. MPEG-7 does not specify how to automatically segment still images and videos in regions and segments; likewise, it does not recommend how to segment objects at the semantic level [Tremeau et al, 2008].

MPEG-7 does not strictly standardize the distance functions to be used and sometimes does not propose a dissimilarity function leaving the developers the flexibility to implement their own dissimilarity/distance functions. A few techniques can be found in the MPEG-7 eXperimentation Model (XM) [MPEG-7:4062, 2001]. Apart from that, there are many general purpose distances that may be applied in order to simplify some complex distance function or even to improve the performance [Eidenberger, 2003]. A large number of successful distance measures from different areas (statistics, psychology, medicine, social and economic sciences, etc.) can be applied on MPEG-7 data vectors [Dasiapoulou et al, 2007].

MPEG-7 is not aimed at any particular application. The elements that MPEG-7 standardizes support as broad a range of applications as possible. The MPEG-7 descriptors are often used in the processes of image-to-image matching, searching of similarities, sketch queries, etc. [Stanchev et al, 2006].

Conclusion

This chapter presented a brief state-of-the art review of colour theory from different points of view; this theory is used as a basis for the study presented in this dissertation. Physiological grounds of colour perception as a starting point for focusing the search in art painting images had been outlined. A brief historical overview of attempts to define colour interconnections and mutual influences had shown that there are multiple models available suit for the purposes of different domains. Some colour models which are most suitable for representing the colours from human point of view are presented. Finally, the MPEG-7 standard, which provides standardized core technologies allowing the description of audiovisual data content in multimedia environments, is presented. Its colour and texture descriptors will be used in the subsequent chapters as the potential basis for constructing high-level visual features.

4 Some Examples of CBIR Systems

Abstract:

Numerous experimental as well as commercial systems had been developed in the field of image retrieval over the last 20 years.

More specifically, in the area of art painting image analysis the efforts were focused in several directions. Some of them aimed to create new digital objects, based on the source one with specific features, such as: image enhancement, expressive rendering, watermarking, virtual restoration of artworks, image-based 3D reconstruction, etc. Another trend of work was focused on the content based retrieval of artworks in digital collections.

The efforts to enhance the process of automatic metadata generation for artworks collections, as well as the analysis of specific artists' and art history studies informed the development of multiple systems which combine image analysis with machine learning techniques.

In this chapter we present some existing art image analyzing systems as well as some image retrieval systems, based on MPEG-7 descriptors.

The correlation between these systems and their possibilities to bridge semantic and abstraction gaps is examined.

4.1 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 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 [Veltkamp and Tanase, 2000]. 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.

Most existing specialized systems focus on artworks analysis techniques over digital imagery, e.g. virtual restoration of artworks (inpainting, fading colour enhancement, crack removal, etc.). Other issues addressed are aspects requiring specific technical expertise (authentication, cracks, and art forgeries).

However, our primary area of interest in this dissertation is the content based retrieval of artworks in databases, as well as specific artists' studies (colour palette statistics, creative processes, etc.) and art history investigation. The emotional and aesthetic charge is inextricably bound up with other components that create the whole presentation of the artwork. Hence, emotional based image retrieval has also its place in the review.

We will stop our attention on some systems from the area of digital cultural heritage:

QBIC [Flickner et al, 1995] is the pioneer in CBIR; QBIC stands for "Query by Image Contents". This product was developed by the IBM and it allows to make queries of large image databases based on visual image content using properties such as colour percentages, colour layout, and textures occurring in the images. The queries use the visual properties of images, and the user can match colours, textures and their positions without describing them in words. Content based queries are often combined with text and keyword predicates to get powerful retrieval methods for image and multimedia databases. The QBIC system was successfully implemented as a search machine in the site of the Hermitage museum, which, in the words of the QBIC-team[35], had been recently voted the best in Russia.

PICASSO system [Del Bimbo and Pala, 1997] was developed in the University of Florence, Italy. The system hurdles multiple descriptions of each data set, each one covering a different level of precision. Images are analyzed at several levels of resolution in order to obtain a pyramidal segmentation of colour patches. Each region at level *n* is obtained by clustering adjacent regions at level *n*-1. Region energy measure is associated to each region. This energy is obtained as a weighted sum of three entries: the area, the colour uniformity and the colour contrast. The further elaboration of this team [Corridoni et al, 1999] took into account the semantics of colour usage. On the one hand, low-level properties directly reflect numerical features and concepts tied to the machine representation of colour information. On the other hand, high-level properties address concepts such as the perceptual quality of colours and the sensations that they convey in humans. Colour-induced sensations include *warmth*, *accordance* or *contrast*, *harmony*, *excitement*, *depression*, *anguish*, etc. In particular, paintings are an example where the message is contained more in

[35] http://wwwqbic.almaden.ibm.com/

the high-level colour qualities and spatial arrangements than in the physical properties of colours. The system translates the Itten theory into a formal language that expresses the semantics associated with the combination of chromatic properties of colour images. It exploits a competitive learning technique to segment images into regions with homogeneous colours. Fuzzy sets theory is used to represent low-level region properties such as hue, saturation, luminance, warmth, size and position. A formal language and a set of model-checking rules are implemented to define semantic clauses and to verify the degree of truth by which they hold over an image.

Analysis of the pictorial portrait database of miniatures of the Austrian National Library [Sablatnig et al, 1998] was developed in order to study the personal style of the artist, or the "structural signature" based on brush strokes in particular in portrait miniatures. A computer-aided classification and recognition system for portrait miniatures had been developed; it provided for a semi-automatic classification based on brush strokes. The hierarchical structure of the classification steps allowed to apply a top-down classification. First, the colour impression is used for a rough classification. Within this subset of all possible artists a more detailed classification is further performed on the basis of shape features within certain face regions which reduces the set of possible candidates once again. Within the regions of the face under examination the stroke classification introduces a bottom-up approach. The final goal is the greatest possible reduction of the number of potential candidate artists. An extension of the system, presented in [Saraceno et al, 1999] uses a hierarchical database indexing method based on Principal Component Analysis. The description incorporates the eyes being the most salient region in the portraits. The algorithm has been tested on 600 portrait miniatures from the Austrian National Library.

Painting classification system [Keren, 2002] suggested a framework for classification of paintings based on the 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 testing set comprises works of Rembrandt, Van Gogh, Picasso, Magritte, and Dali. Ten paintings from each painter consisted the training set, and the test set consisted of twenty to thirty paintings for each painter. The rate of success was 89%.

Art historian system [Icoglu et al, 2004] utilized an automatic extraction of features of paintings' art movements. The authors of this system demonstrated that the feature set enables one to highlight art movements efficiently. In the classifier design, statistical pattern recognition approach using Bayesian, k-NN and SVM classifiers had been implemented. This study showed that art movements of paintings can be indexed by exploiting the following six measures: (1) Percentage of dark colours; (2) Gradient coefficient calculated from the gradient map of the painting image; (3) Number of local and global maxima in the luminance histogram; (4) Colour range corresponding to the peak point of the luminance histogram; (5) Deviation of average grey level acquired within each block from the average grey level acquired within entire image after the partition of the painting image into identical blocks; (6) Deviation of grey level distribution from Gauss distribution. A training set was constructed with the 27 original paintings that belong to classical paintings, Cubism or Impressionism, 9 from each class. The training samples included 107 paintings (31 classical, 34 from Cubism, 42 from Impressionism). The accuracy varied from 80% to 90% for different used classifiers.

A lightweight image retrieval system for paintings [Lombardi et al, 2005] utilizes image indexing features grouped as follows: (1) Canvas features: max, min, mean, median, and standard deviation from each of the red, green, and blue colour channels; (2) Colour features: intensity mean (measures the global brightness of a greyscale image), colour frequency distribution (measures the degree of disorder found in the frequency distribution of colours in a painting); and (3) Edge characteristics: line count (the Sobel edge detector is used to identify lines in the image). The experiments with the nearest neighbour classifier achieved to reliably distinguished between two painters (Picasso and Van Gogh) shows the accuracies from 83 to 94%. In the experiments made on the bigger datasets with more artists (10 images from each of fifty artists) the overall retrieval rate stand around 49%.

Collage [Ward et al, 2005] was developed by the Institute for Image Data Research at the University of Northumbria at Newcastle and iBase Image Systems; they collaborated with the London Guildhall Library to introduce CBIR technology for image retrieval on the Collage web site. At the time, Collage was the largest digital imaging project in Europe and faced many technical and curatorial challenges. These were successfully solved, leading the way for other organizations across the UK heritage sector to adopt similar systems. Each image has three data fields in Collage: the identification number, feature vector length, and feature vector data. The search queries can be divided into 3 categories: (1) retrieval by features – colour, texture, and shape or the spatial location; (2) retrieval by derived features, known as logical features, that involve some degree of logical inference about the identity of the objects depicted in the image; and (3) retrieval by abstract attributes involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted.

Brushwork identification [Marchenko et al, 2006] had been developed by a team, headed by R. Jain and used annotation of paintings based on brushwork. It is modelled as a part of the annotation of high-level artistic concepts such as the movements or artists using defined ontology of different types of brushworks, based on specific combinations of low-level feature values of gradient, texture, hue and light contrast, as well as homogeneity, which altogether characterize some style of painting. They employ the ontology of artistic concepts that includes visual, abstract and application-level concepts. Visual level concepts include various pictorial attributes of paintings such as colour palettes, brushwork techniques, colour contrasts and others. Abstract level concepts include concepts defined by artistic theories for the art experts such as *expressive*, *rational* and *gestural*. Application-level concepts present the widely used concepts for retrieval by novice users in online galleries such as the artists' names, painting styles, periods of art and others. This ontology is based on the Getty's AAT and ULAN ontologies. It has several advantages. First, the explicit assignment of visual and abstract concepts offers more flexibility for paintings' annotation and retrieval. For example, queries such as paintings in *warm* colours, paintings with *temperature* contrast and *impasto* brushwork class are possible. Second, the use of domain-specific ontology within the proposed framework facilitates concept disambiguation and propagation. Third, a large variety of ontology concepts facilitates retrieval by both expert and novice user groups. For the experiments, they extracted 4880 patches of size 32x32 from 30 paintings of Renaissance, Fauvism, Impressionism, Post-Impressionism, Expressionism and Pointillism painting styles (randomly selecting 75% of the dataset for training and 25% for testing). Different feature selecting models gave overall performance of brushwork recognizing between 80% and 95%.

M4ART [Broek et al, 2006], Multimedia for Art ReTrieval system, is an entry point to the digitized collection of the Rijksmuseum, the National Gallery of the Netherlands. The global colour distribution of the sample query image, and a set of texture features are extracted and compared with those of the images in the collection. Thus, the collection can be queried based on text as well as content-based features. Moreover, the matching process of M4ART can be scrutinized. With this feature, M4ART not only integrates the means to give experts and general user's equal access to the system, but it also helps users understand the system's inner logic. These qualities make M4ART unique in its ability to let the user access, enhance, and retrieve the knowledge available in digitized art collections.

MECOCO [Berezhnoy et al, 2007] expands traditional research avenues where the analysis of visual arts is performed by human art experts only. The availability of advanced artificial intelligence techniques makes it possible to support art experts in their judgment of visual art. The goals of the MECOCO team were: (1) to determine the usage of complementary colours in the work of Vincent Van Gogh and (2) whether this characteristic may make his paintings identifiable in time. It is commonly acknowledged that, especially in his French period, Van Gogh started employing complementary colours to emphasize contours of objects or parts of scenes. MECOCO offers a new method for measuring complementary-colour usage in a painting by combining an opponent-colour space representation with Gabor filtering. To achieve the aim it first defines a novel measure called "opponency value" that quantifies the usage of complementary-colour transitions in a painting and then applies the measure to the specific Van Gogh's painting style. In addition, MECOCO also provides an objective and quantifiable way to support the analysis of colours in individual paintings.

ALIPR [Li and Wang, 2008] is a machine-assisted image tagging and searching service developed at Penn State University. The research in content-based image retrieval and automatic learning-based linguistic indexing had been made by the group of Prof. James Wang and Prof. Jia Li (since 1995 in the Stanford University). This group made several successful applications in the field of image retrieval and digital imagery:

- *SIMPLIcity* [Wang et al, 2001] is Semantics-sensitive Integrated Matching for Picture LIbraries – Visual similarity search engine, developed from 1999 to 2000, which uses semantics classification methods, a wavelet-based approach for feature extraction, and integrated region matching based upon image segmentation;
- *Advancing Digital Imagery Technologies for Art and Cultural Heritages* [Wang et al, 2002] was a collaborative research project, started in August 2002which aimed at advancing information technologies related to the preservation, retrieval, and dissemination of digital imagery for Asian art and cultural heritage;
- *CLUE* [Chen et al, 2005b] is a Cluster-based image retrieval, developed in 2002-2003 which experimented with the use of unsupervised learning;
- *AUTHENTIC* [Berezhnoy et al, 2005] investigated new image analysis techniques which could be applicable for the assisted analysis of visual art forgeries. The texture-analysis technique revealed two main clusters of brushstroke shapes used by Van Gogh;
- *ACQUINE* [Datta et al, 2006], Aesthetic Quality Inference Engine, was developed in 2005-2006. It is a machine-learning based online system of computer-based prediction of aesthetic quality for colour natural photographic images.

ALIPR is a real-time engine based on the system ALIP (Automatic Linguistic Indexing of Pictures). The system was inspired by the fact that a human being can recognize objects and concepts by matching a visual scene with the knowledge structures stored in the brain. The system builds a knowledge base on different concepts automatically from training images. Statistical models are created about individual concepts by analyzing a set of features extracted; the algorithm uses wavelet transform [Li and Wang, 2004]. A dictionary of these models is stored in the memory of the computer system and used in the recognition process. A mixture of 2-D multiresolution hidden Markov models is developed and used to capture different styles in the paintings. The team has implement their results on various collections including The Van Gogh Museum in Amsterdam and rich image resources of the Emperor and the Chinese Memory Net projects by Ching-chih Chen.

4.2 Some Image Retrieval Systems, which Use MPEG-7 Descriptors

Development of MPEG-7 content-based indexing systems is amongst the active research topics in image retrieval. Many demos, prototypes and commercial systems have been developed or are currently under development. The MPEG-7 & MPEG-21 community[36] aims at bringing together experts from research and industry in the area of multimedia metadata interoperability for collaborative working environments. By establishing a community of professionals it is intended to bridge the gap between academic research and industrial scale development of innovative products for natural collaboration.

Some systems worth noting are listed below:

PicSOM[37] (Picture Self-Organizing Map) [Koskela et al, 2002] is a neural-network-based image retrieval system developed at the Helsinki University of Technology, Finland in 2001. It uses self-organizing maps to retrieve relevant images from the database and MPEG-7 visual descriptors for the image description space. The results of comparing PicSOM with a reference system, presented in [Koskela et al, 2002], show that that the MPEG-7 defined content descriptors can be used as such in the PicSOM system even though Euclidean distance calculation, inherently used in the system, is not optimal for all of them. Also, the results indicate that the PicSOM technique is a bit slower than the reference system in starting to find relevant images, but retrieval precision exceeds that of the reference system when its relevance feedback mechanism of begins to function.

CIMWOS[38] [Papageorgiou and Protopapas, 2003] is a multimedia, multimodal and multilingual system supporting content-based indexing, archiving, retrieval, and on-demand delivery of audiovisual content. It was developed within a Greek-Belgian project entitled "CIMWOS – Combined Image and Word Spotting" funded under the IST program in 2000-2003. The system uses a multifaceted approach to locate important segments within multimedia material employing state-of-the-art algorithms for text, speech and image processing. It is a powerful tool for any organization that produces, markets and/or broadcasts video and audio programs, facilitating common procedures of retrieving audio-visual material during a research, a production of a documentary, etc.

VizIR[39] (a framework for Visual Information Retrieval) [VizIR, 2003] was a project of the Vienna University of Technology, supported by the Austrian Scientific Research Fund in 2003-2006, aimed implementing MPEG-7 visual descriptors in Content-based retrieval of images and video. The VizIR group was

[36] http://www.multimedia-metadata.info/
[37] http://www.cis.hut.fi/picsom/
[38] www.xanthi.ilsp.gr/cimwos
[39] http://cbvr.ims.tuwien.ac.at/

member of the EU DELOS network of excellence in Digital Libraries and an affiliated member of the EU SCHEMA project. The VizIR source code is free under GNU Public License.

Mirror[40] (MPEG-7 Image Retrieval Refinement based On Relevance feedback) was developed in 2005 in Hong Kong to evaluate MPEG-7 visual descriptors in retrieval algorithms. The system core is based on MPEG-7 Experimentation Mode (XM) with web-based user interface for query by image example. A new Merged Colour Palette approach for DCD similarity measure and relevance feedback had also been developed in the system. The system is highly modularized and new algorithms, new ground truth set, and even new image database can be added easily. MIRROR mainly supports the following image retrieval functions: (1) Hierarchical image browsing; (2) Content based image retrieval by example image; (3) Relevance feedback retrieval; (4) Automatic performance measure using ANMRR; (5) Random browsing.

MARVEL[41] [Natsev et al, 2007] had been developed within The Intelligent Information Management Department at IBM Research. It is a multimedia analysis and retrieval system which supports the management of large and growing amounts of multimedia data (e.g., video, images, audio) by using machine learning techniques to automatically label its content. The system won the Wall Street Journal Innovation Award in the multimedia category in 2004. MARVEL was designed to support the emerging MPEG-7 multimedia content description standard by providing automated meta-tagging for all XML standard content, including images, video, audio, and text. MARVEL aims to replace costly, time-consuming, and error-prone processes of authoring metadata with a semantics machine learning approach. It works by building statistical models from visual features using the training examples and applies the models for automatically annotating large repositories. The MARVEL process reduces the total cost of creating metadata, reduces annotation errors and allows more effective search and retrieval. The MARVEL system consists of two components: the MARVEL multimedia analysis engine and the MARVEL multimedia search engine. The MARVEL multimedia analysis engine applies machine learning techniques to model semantic concepts in video from automatically extracted audio, speech, and visual content. It automatically assigns labels (with associated confidence scores) to new video data to reduce manual annotation load and improve searching and organizes semantic concepts using ontologies that exploit semantic relationships for improving detection performance. The MARVEL multimedia retrieval engine integrates multimedia semantics-based searching with other search techniques (speech, text, metadata, audio-visual features, etc.). It also combines content-based, model-based, and text-based searching for video searching. The IBM Marvel system has achieved the top

[40] http://abacus.ee.cityu.edu.hk/
[41] http://mp7.watson.ibm.com/marvel

performance on the TRECVID semantic concept detection evaluation in 2003 and 2004.

MPEG-7 Library in JOANNEUM Research[42] [Schallauer et al, 2006] has acquired know-how in using MPEG-7 and developed tools for working with MPEG-7 descriptions. It had been partially funded by the European Commission and by national research programs. Its basic features are:

- C++-implementation of the full MPEG-7 standard ISO/IEC 15938. To promote the uptake of MPEG-7, JOANNEUM RESEARCH releases its MPEG-7 library under an open source license;
- Detailed Audiovisual Profile: a MPEG-7 profile for detailed description of single image, audio and video media items, designed to allow interoperability between MPEG-7 based systems;
- MPEG-7 extensions for the description of quality and defects of audiovisual material.

With this library application developers are able to create multimedia content descriptions, manipulate them, serialize to XML and de-serialize from XML. One major design goal was to simplify extending single classes to allow the developer to enrich interface functionality for certain descriptors. The library contains data structures for representing, accessing, and modifying descriptors and description schemes and provides tools for parsing and serializing descriptions.

Caliph (Common and Lightweight PHoto Annotation) and ***Emir*** (Experimental Metadata-based Image Retrieval) [Lux, 2009] are two programs starting in the IICM of the University of Technology Graz, Austria[43]. Further development was done by adapting it to a project of the Know-Center, in Graz, Austria, by Mathias Lux and Werner Klieber. *Caliph* & *Emir* are MPEG-7 based Java prototypes for digital photo and image annotation and retrieval supporting graph like annotation for semantic metadata and CBIR using MPEG-7 descriptors. Besides using text information which already exists in digital photographs, and transformation of this information to MPEG-7, *Caliph* supports the creation of new metadata. Semantic information about the image is presented as directed graph, where the nodes reflect semantic objects, locations, agents, states, times or concepts and the edges define the relations between these semantic entities. The MPEG-7 description consists of the following parts: metadata description, creation information, media information, textual annotation, semantics, visual descriptors. The experimental metadata based image retrieval tool *Emir* supports retrieval in file system based photo repositories created with *Caliph*. Different types of retrieval mechanisms are supported: CBIR using the MPEG-7 descriptors Colour Layout, Edge Histogram and Scalable Colour; Keyword based searching using the Lucene search engine; Graph based retrieval supporting wildcards for semantic relations and semantic objects.

[42] http://iiss039.joanneum.at/

[43] http://www.semanticmetadata.net/features/

LIRE[44] (Lucene Image REtrieval) library [Lux and Chatzichristofis, 2008] provides a simple way to retrieve images and photos based on their colour and texture characteristics. LIRE create a Lucene index of image features for CBIR. Three of used features are taken from the MPEG-7 Standard: Scalable Colour, Colour Layout and Edge Histogram; a fourth one, the Auto Colour Correlogram has been implemented based on recent research results. The LIRE library is part of the Caliph & Emir project and aims to provide the CBIR features of Caliph & Emir to other Java projects in an easy and lightweight way.

Fedora Repository [Burgi and Monbaron, 2008] integrated CBIR built on the MPEG-7 standard into their repository; this had been announced at the conference of Fedora Commons in April 2008. The idea developed from necessity of migration of the bank of images with their associated metadata collected in the University of Geneva. Beside textual metadata, the MPEG-7 documents also contain content descriptors that are extracted using image processing methods (Caliph and Emir tools). The second part of the project was about building a user interface for image retrieval. Content similarity involves distance measurements performed on the MPEG-7 descriptors represented as data vectors built from the textual data stored in Lucene. The outcome of this matching process is a set of images URLs, which are included in the result web page so that the user web agent can retrieve them from Fedora.

MUFIN Search[45] (Multi Feature Indexing Network) [Batko et al, 2009] aims to develop a general purpose technology solution for the problem of searching in various and very large databases. The extensibility is ensured by adopting the metric space to model the similarity; thus MUFIN can evaluate queries over a wide variety of data domains compared by metric distance functions. The scalability is achieved by utilizing the paradigm of structured peer-to-peer networks, where the computational workload of query execution is distributed over multiple independent peers that can work in parallel. These unique capabilities of MUFIN are demonstrated on a database of 100 million images indexed according to a combination of five MPEG-7 descriptors (Colour Structure, Colour Layout, Scalable Colour, Edge Histogram, Homogeneous Texture).

The examined systems have not addressed directly the case of art image retrieval. However, they clearly demonstrate the state of the art of current repositories. The analysis shows that MPEG-7 as standard is already supported by implementations in several image banks. Appropriate metrics and similarity measures for MPEG-7 descriptors are already established. The techniques for enhancing image indexing based on MPEG-7 descriptors show good initial results. This evidence support the suggestion that is worth using MPEG-7 descriptors as a ground in the field of art painting analysis.

---

[44] http://www.semanticmetadata.net/lire/
[45] http://mufin.fi.muni.cz/search/

4.3 Systems and Their Connections with Gaps Tasks

Although all presented systems have, generally speaking, a common task, they differ in their goals, depending on the specific needs for which they were created. Some systems (such as QBIC, Collage, M4ART, etc.) are industrial systems which support the complete digital object lifecycle in the overall process of image retrieval. Such systems have to support a wide range of functions, starting from reaching the data from repository through feature extraction, creating and keeping metadata, building appropriate indexing techniques for easy access, providing a friendly user interface and accelerating relevance feedback, etc. Others are experimental systems, aimed to study the ability of some features and/or algorithmic techniques to enhance image retrieval or to solve exact classification task. Such systems do not claim to be so comprehensive, but being at the frontier of the research they are more focused of studying particular concept or technique and in this way are of definite interest.

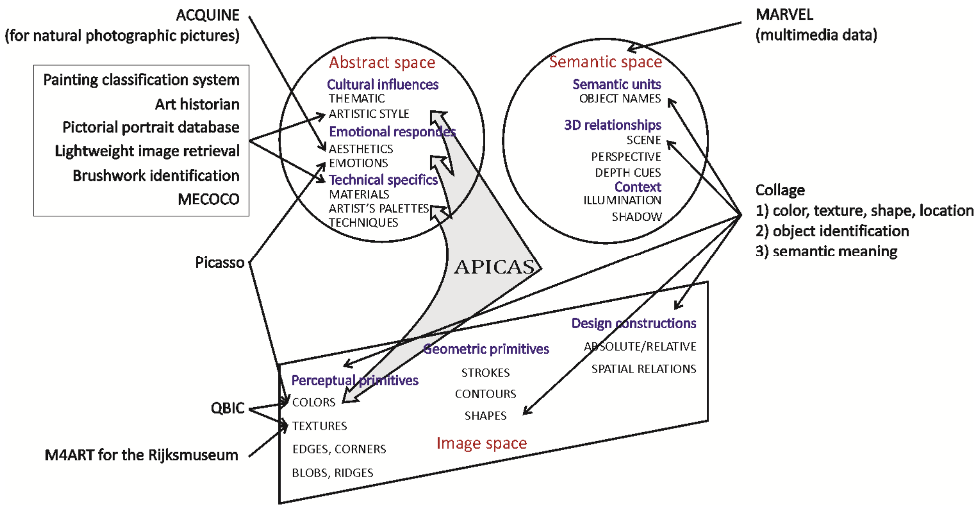


Figure 18. The systems and their connection with the taxonomy of art image content

Returning to our goal – studying the ways for closing different kinds of gaps – on Figure 18 we show our vision of connection of some of reviewed systems with the taxonomy of art image content. Here we do not stop our attention of used features and algorithms within the process of analysis made by the systems. We are focused on resulting features or concepts, produced by the systems. For instance, all systems in the rectangle engage with the field of artistic practice studies and art history investigation, notwithstanding that all of them use different kinds of visual primitives.

In the figure we also incorporate our vision which shows which aspects will be addressed by our proposal, APICAS. APICAS is intended mostly to study the ability of colour analysis for covering mainly the abstraction gap. The analysis that we already made shows that colour plays a significant role in all three parts of abstraction space. The colour always brings some symbolism, generated by cultural environment. On the other side each artist builds his own colour vision, which expresses his/her emotional and aesthetic feelings. Technical aspects are also narrowed from current existence. For instance, the new painting direction in the landscapes of Impressionists for studying the air and sun influence arose after the industrial production of paints in tubes started; it allowed artists to work easily outside of their studios.

The scientific efforts are gradually progressing and step by step we can immerse into the processes of resolving the problems of classifying paintings of different styles or of different painters. New search paradigms such as mental image search with visual words enhance the user expression of visual target without a starting example. Alternatively, image fingerprinting, which deals with extracting unique image identifiers that are robust to image deformations (cropping, resizing, illumination changes, rotations etc.), might be used along with query-by-example techniques to partially deal with this task. This area is still in its infancy of research and development.

Conclusion

In this chapter we presented some existing art image analyzing systems. Most of them address the analysis of high-resolution multi-spectral digital copies of the images in order to cover automatic artwork analysis techniques, and are used in different applications such as virtual restoration, image retrieval, studies on artistic praxis, authentication etc.

In addition, some pioneering image retrieval systems, which use MPEG-7 descriptors, were discussed. Although these systems do not have direct application in the field of art image retrieval yet, they clearly show the current trend to include MPEG-7 descriptors as metadata in current digital repositories. This fact confirms the conviction that it is worth using MPEG-7 descriptors as a ground in the field of art painting analysis. On the other hand, in spite of the generic character of some of them (such as Marvel) they can be successfully implemented in the field of digitization of cultural heritage in order to fill some parts of the semantic gap.

We showed the connection of reviewed systems with the taxonomy of art image content, stopping our attention on resulting features or concepts produced by the systems. We motivated our vision which parts need to be addressed by the proposed system APICAS, which is aimed mostly at studying the ability of colour analysis for covering mainly the abstraction gap.

5 Proposed Set of Visual Features

Abstract

In this chapter we describe the types of features suggested for the purposes of the APICAS system, as follows:

- *visual features, which represent colour distribution in the images;*
- *global features that reflect colour harmonies and contrasts in art images in accordance with Itten's theory;*
- *local colour and texture features, based on vector quantization of MPEG-7 descriptors.*

We use these features for:

- *analyzing the colour distribution in art images;*
- *searching the images by higher-level concepts, concerning harmonies and contrasts;*
- *analyzing how more detailed information on semantic and abstraction content of art images based on MPEG-7 descriptors with significantly dimensionality reduction can be captured;*
- *classifying the art images by different abstraction criteria, such as artists' names, periods or movements or semantic profile as a genre of the art image.*

The chapter includes the following parts:

- *proposing colour histograms as appropriate features for examining distribution of colour characteristics;*
- *producing one classification of harmonies and contrasts in accordance to the Ittens' theory from the point of view of three main characteristics of the colour – hue, saturation and luminance;*
- *providing the formal description of defined harmonies and contrasts;*
- *presenting the methods for extracting local features that capture local colour and texture information, based on tiling the image and applying vector quantization of MPEG-7 descriptors, calculated for the tiles of the image.*

5.1 Introduction

Digitalized art painting space provides the users with the opportunity to immerse in the ocean of accumulated culture. Earlier we could only dream to see some masterpieces as the originals saved in various public or private collections. Now our computer moves us to every chosen place and time. These abilities increase the 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; or would be attracted by some purely aesthetic influence of the image. Many efforts are aimed at combining text-based and content-based search technologies in real-world image retrieval.

In this way the high level visual concepts can be used for narrowing the semantic and abstraction gap between low-level automatic visual extraction and high-level human expression.

Figure 19. Contrast of Hue – Theotokos of the Passion (13th-15th c.), currently in the church of Sant'Alfonso di Liguori all'Esquilino, Rome

The analyses of used expressive techniques of art masters show the tendencies of using one or another artistic techniques, i.e. as is named by Itten "the subjective timbre" [Itten, 1961]. Here we will mention some examples, given by Itten in his book "The Art of Colour", by Margaret Walch and Augustine Hope in "Living Colours" [Walch and Hope, 1995] and by Becky Koenig in "Colour Workbook" [Koenig, 2010].

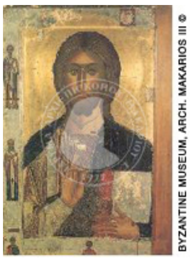


Figure 20. Blessing Christ, c.1192, currently in Byzantine museum of Cyprus

Many different colours are rendered in equal brilliance in the medieval manuscripts as well as in Orthodox icons and became typical examples for applying contrast of hues (Figure 19). The resulting vivid cold-warm effects were re-discovered years later in the Impressionists.

This tendency is observed from earliest stages of iconographical painting. Dr. Ioannis Elliades, director of the Byzantine museum and art gallery of the archbishop Makarios III Foundation in Cyprus, presenting the icon "Blessing Christ" (c. 1192), which is one of the rare brilliant remaining artefacts from this epoch, pays special attention on the use of different types of contrast – green against red, light against dark, and gold against silver (Figure 20).

In the first half of the 15th century, the brothers Hubert and Jan van Eyck began to construct patterns of composition around the local colours of the person or thing represented. Using variations of these local colours, through dull and bright, light and dark tones, they produced realistic images very closely approaching nature.

The painting "Deposition", by the Northern Renaissance painter Rogier Van der Weyden (1399-1464) is an altarpiece that uses the primary triad of red, yellow, and blue in a balanced and rhythmic manner (Figure 21).

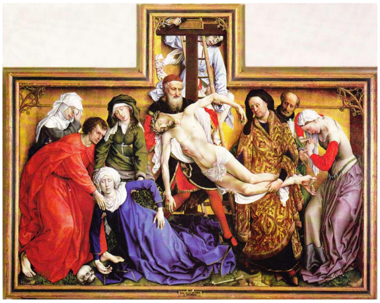


Figure 21. Primary triad of red, yellow, and blue – Rogier Van der Weyden: Deposition (1436) [Koenig, 2010]

Piero della Francesca (1410-1492) painted individuals in sharp outline and clearly expressive areas, with balancing complementary colours. In contrast, Leonardo da Vinci (1452-1519) rejected strong colouration. His drawings and paintings are virtuoso studies of light and shadow – chiaroscuro, as Italians called it. In fact, Leonardo developed his own method of modelling all the forms in his pictures with monochrome browns, blacks, and greys - tones that generally dominated the finished painting – before the application of any strong colour. He painted in infinitesimal tonal gradations; the endmost representatives of such approach are "St. Jerome" and "Adoration", which are composed entirely in sepia tones of light and shade.

El Greco (1545-1614) looked to action and colour for an expression of his experience. He developed objectively correct colour concords for each subject, leaving the treatment of form to subjective bias. Dramatic effects from cold but intense colours in "Purification of the Temple" heighten the anger of Christ as he drives out the moneychangers from the temple precinct (Figure 22).

Contrary to El Greco (1545-1614), which shades are always sharply and characteristically defined by gray and black tones, Grunewald (1475-1528) set colour against colour. In Rembrandt's works (1606-1669) pure light colours often shine like jewels in dark dull settings (Figure 23). Picasso's blue period was characterized by his dominant use of a monochromatic blue colour scheme to enhance the melancholy subjects of the paintings (Figure 24).

However, the psychologically expressive power of his colours, their symbolic verity, and their realistic signification – all these three potentialities of colour are, in a deeper sense, fused into unity. Visual, mental and spiritual phenomena are interrelated multiple times in the realm of colour and the colour art.

Figure 22. Cold Intense Colours – El Greco: Purification of the Temple (1600)

Figure 23. Light-Dark Contrast – Rembrandt: Portrait of Saskia (1633)

Figure 24. Monochromatic Colour Scheme – Pablo Picasso: Self-Portrait in Blue Period (1901)

Contrast effects and their classification are a proper starting point in the study of colour aesthetics. The problems of subjectively conditioned colour perception are especially pertinent to art education and scholarship, architecture and commercial design.

5.2 Proposed Features

We propose a set of visual features, with the aim to represent the human perception of colours. We try to formalize the qualitative achievements of Itten's theory of successful colour combinations. The Itten's investigation of the "subjective timbre" [Itten, 1961] shows that the existence of laws of combining the colours does not restrict the variety of used colour combinations by the artists. They are key to the identification of the individual's natural mode of thinking, being and doing.

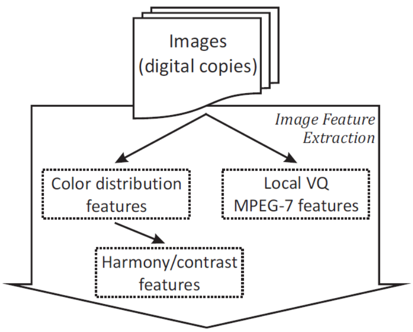


Figure 25. Proposed visual features

We use three different categories of visual features which represent the image (Figure 25).

The first class of features is a group of several global colour low-level attributes, which represent colour distribution in the images. The analysis of the distribution of colour features in art images is made in order to be used in tuning up the similarity measure functions.

The second one is based on an attempt to formulate high-level features which represent colour harmonies and contrasts, based on the three main characteristics of the colour, which are closest to the human perception – hue, saturation and lightness. Functions for automatic features extraction from digital images based on the defined low-level global colour distribution features, are defined.

The third method for obtaining visual features consists of observing the tiles of the images from chosen learning set. MPEG-7 descriptors for these tiles are extracted. For each descriptor after clustering a set of centroids is defined. The new images are putted under similar splitting and the calculation of MPEG-7 descriptors is attached to the closest centroid from the corresponding cluster set. In this way we overcome the complexity of MPEG-7 descriptors, which made good presentation of different types of visual features but need specific processing and cannot be putted directly into generic classification algorithms.

5.3 Colour Distribution Features

These features represent colour distribution in the images. One popular way is to use colour histograms, which are a statistic that can be viewed as an approximation of an underlying continuous distribution of colours' values.

We want to use these characteristics for two connected purposes:

- Analyzing the colour distribution in art images;
- Using these low-level features in the process of calculating higher-level colour harmonies and contrast features.

For representing colours we chose colour models that describe perceptual colour relationships and are computationally simple – such as HSV, HSL, HSL-artist colour models. HSV is used in MPEG-7 descriptors. HSL better reflects the intuitive notion of "saturation" and "lightness" as two independent parameters. HSL-artist colour model is a base for further development for defining colour harmonies and contrast characteristics. Later, except in the places, where is specially pointed, the term "lightness" is used as a collective concept for "Lightness" in HSL colour model, "Value" in HSV colour model or "Luma" in HSL-artist colour model depending on the colour model chosen.

Colour histograms represent the number of pixels that have colours in each of a fixed list of colour ranges that span the image's colour space, the set of all possible colours. The used colours in digital presentations of art paintings can receive almost all possible colour values; because of this we divide the dimensions of the colour model into appropriate numbers of ranges.

The pixels in the images are converted into one of the chosen colour model (preferably HSL-artist colour model). The quantization of Hue is made to 13-bins, $ih = -1,\ldots,NH-1$, $NH = 12$, where one value is used for achromatic colours ($ih = -1$) and twelve hues are used for fundamental colours ($ih = 0,\ldots,NH-1$). The quantization of Hue is linear equidistant when HSL colour model or HSV colour model is chosen. For HSL-artist colour model the quantization function is non-linear with respect to taking into account the misplacement of artists' colour wheel and Hue definition in HSL colour space. The quantization intervals are given in Figure 26.

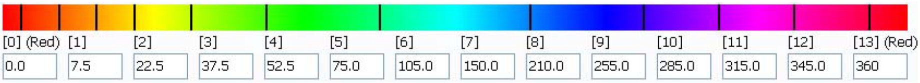


Figure 26. Quantization of Hue

The saturation and lightness in HSL colour model, respectively the saturation and values in HSV colour model, and saturation and luma in HSL-artist colour model are linearly quantized into NS-bins ($is = 0,\ldots,NS-1$), respectively NL-bins ($il = 0,\ldots,NL-1$).

The MPEG-7 Dominant Colour descriptor, which extracts a small number of representative colours and calculates the percentage of each quantized colour in the image, also can be used as a kind of colour distribution feature. In order to equalize further definitions we reconfigure extracted RGB-values of Dominant Colour descriptor into values in chosen quantized feature space and use the corresponded percentage of each quantized colour as a percentage in the defined three-dimensional array.

In [Ivanova and Stanchev, 2009] we have used an exact function of defining the belonging of the colour characteristic to quantizing segment. Further in [Ivanova et al, 2010/IJAS] we added the possibility to make the quantization of colour characteristics using fuzzy calculating of belonging of colour to a corresponded index (Figure 27). If the position of the examined value is in the inner part of one defined segment (more than one half from the left bound and less than three half from the right bound) the characteristic is considered to belong to this segment. In any other case (except the endmost parts for saturation and lightness), part of the characteristic is considered to belong to this segment and the rest part is considered to belong to the adjacent

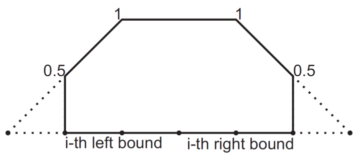


Figure 27. Fuzzy function for calculating quantization part of colour characteristic

segment. For receiving that part a linear function, which reflects the decrease of belonging of that characteristic to the segment, is used.

As a result, every picture is represented with three dimensional array containing coefficients of participation of colours with correspondingly measured characteristics of the picture.

$A = \{A(ih, is, il) \mid ih = \text{-}1, \ldots, NH\text{ -}1;\ is = 0, \ldots, NS\text{ -}1;\ il = 0, \ldots, NL\text{ -}1\}$.

Analysis of colour distribution can be made by three directions together or only by two or one of them. On the base of three dimensional array A , using summarizing over the discarded dimension(s), we can receive the corresponding projections:

- For examining two of the dimensions: $A_{HS} = \{A(ih, is, -)\}$, $A_{HL} = \{A(ih, -, il)\}$, $A_{LS} = \{A(-, is, il)\}$;
- For representing the distribution of one dimension: $A_H = \{A(ih, -, -)\}$, $A_S = \{A(-, is, -)\}$, $A_L = \{A(-, -, il)\}$

where $ih = \text{-}1, \ldots, NH\text{ -}1;\ is = 0, \ldots, NS\text{ -}1;\ il = 0, \ldots, NL\text{ -}1$.

5.4 Harmonies/Contrasts Features

Usually, in accordance of Johannes Itten proposition, the colour wheel which represents relations between hues is divided into twelve sectors. The centres of three equidistance sections correspond to primary colours. Secondary colours are located between them, which from one side are middle points of two primary colours, and from other side are complementary to the third colour. The quantization is expanded with the intermediate colours, which lays at the midpoint to adjacent primary and secondary hues.

In Figure 11 the position of the hues in standard artists' colour wheel is shown. This order and correlations between hues is described in RYB (Red-Yellow-Blue) colour model, used by the artists. Let us mention that this arrangement of hues differs from many contemporary colour models – RGB (Red-Green-Blue), CMY (Cyan-Magenta-Yellow), HSL (Hue-Saturation-Luminance), HSV (Hue-Saturation-Value), being based on the definition of colours as primary or secondary in accordance with the trichromatic theory [Colman, 2006]. But all classic theories, connected with definition of contrast are based on the opposition of the colours as they appear in the artists' colour wheel.

We use HSL-artist colour model, which is based on the advantages of HSL and YCbCr colour models and render an account of disposition of hues in RYB colour model. We present one classification of different types of harmonies and contrasts, from the point of view of the three main characteristics of the colour – hue, saturation and lightness.

➢ *Harmonies/Contrasts from the Hue Point of View*

✓ *Harmonies/Contrasts Based on the Hues Disposition*

The figures below show only relatively disposition of the colours, not the absolute meaning of the colour. Some of these combinations are discussed in [Holtzschue, 2006] and [Eiseman, 2006].

Monotone compositions: These compositions use one hue, and the image is built on the base of varying of lightness of colour (Figure 28). These images are used to suggest some kind of emotion since every hue bears specific psychological intensity.

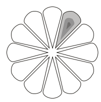


Figure 28. Monotone Composition

Analogous hues: Analogous hues can be defined as groups of colours that are adjacent on the colour wheel (Figure 29). They contain two, but never three primaries and have the same hue dominant in all samples.

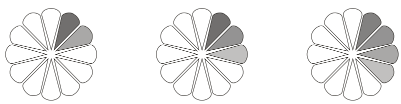


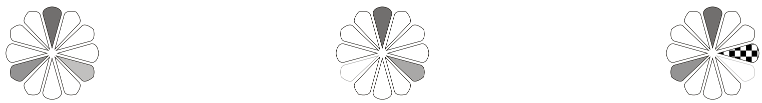
Figure 29. Different Variants of Analogous Composition

Complementary contrasts: Complementary colours are hues that are opposite one another on the colour wheel. When more than two colours take part in the composition the harmonic disposition suggests a combination between analogous and complementary hues (Figure 30).

a) Complementary; *b) Double Complementary;* *c) Split Complementary;* *d) Complementary;*

Figure 30. Variants of Complementary Contrast

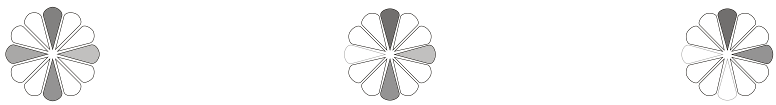
Triads: Three colours that are equidistant on the colour wheel form a triad. This means that all colours are primary or secondary, or intermediate. When we have analyzed art paintings also "partial" triads were observed in landscapes images, when two of the colours forming a triad were found most significant for the image. In some cases a tweaked form of triads is observed also, when two of the colours are in two of the possible positions of triad, but the third one is slightly skewed from the third position (Figure 31).



a) Triad; *b) Partial Triad;* *c) Tweaked Triad;*

Figure 31. Triads

Tetrads: The tetrad includes four colours in equidistance on the colour wheel. This contrast produces very complicated scheme and can lead to disharmony. Of course the presence of partial forms of the tetrads (Figure 32) can also be examined here. Usually this scheme is used only in more schematic pictures with precisely coordinated saturation and lightness values, for instance in heraldic signs, flags, etc.



a)Tetrad; *b)Partial Tertrad* *c)Partial Tetrad*

Figure 32. Tetrads

Achromatic compositions: As a special case, images composed by black, greys and white tones or contain colours with very small saturation.

✓ ***Warm-Cold Contrast***

Warm and cold are two opposing qualities of hue. Warm colours are hues around red and orange; cold colours are these around blue. The terms "warm" and "cold" are helpful for describing families of colours. Such kind of contrast is pointed by Adolf Hoelzel [Gage, 1993] and Johannes Itten [Itten, 1967].

The combination of colour families that reflect such kind of contrast can be defined as follows:

- *Warm:* The image is warm when the composition is built from family of warm colours;

- *Cold:* The image is cold when it is composed only (or predominantly) with cold colours;
- *Neutral:* The composition contains colours mainly from neutral zones;
- *Warm-cold:* The composition lays in this category when the percentage of cold family is in some proportion to the percentage of warm family;
- *Warm-neutral:* In such compositions there is proportion between warm colours and neutral ones;
- *Cold-neutral:* The image contains cold and neutral tones in some proportion.

➢ ***Harmonies/Contrasts from the Saturation Point of View***

Unlike hue, which is circular and continuous, saturation and lightness are linear. That difference determines different definitions of harmonies/contrasts for these characteristics. This harmony appears together with the hue ones. It is used to give different perception when the colour is changed. As a whole we can define three big groups of harmonies and contrasts:

- *Dull:* An image can be classified as dull when the composition is constructed mainly from unsaturated colours;
- *Clear:* Clear images have been build mostly from clear colours (spectral and near to spectral, respectively only with varying in lightness);
- *Different proportion of saturations:* Usually in composition of clear colours in combination of dull ones. Depending on the content of different saturation and on the distance between predominate quantities harmonies can be defined such as *smooth*, *contrary*, etc.

➢ ***Harmonies/Contrasts from the Lightness Point of View***

The whole effect of the lightness of the image as well as light-dark contrast is a very powerful tool in art mastering. In some cases an artwork may not contain light-dark contrast – at that case the image has one integral vibration of the lightness. In the other case sharp light-dark contrast is used to focus the attention in exact points of the image.

- *Dark:* Dark compositions are built mainly from dark colours;
- *Light:* Light images contain mostly colours near white;
- *Different proportion of lightness:* Light colours combined with dark ones compose the image. Depending on content of different lightness and of distance between predominate quantities contrasts can be defined as*:* *smooth*, *contrary*, etc.

5.5 Formal Description of Harmonies/Contrasts Features Using HSL-artist Colour Model

For defining colour harmonies/contrast features we use representation of the colour distribution as colour histograms, defined above:

$A = \{A(ih, is, il) \mid ih = \text{-}1,\ldots,NH\text{-}1;\ is = 0,\ldots,NS\text{-}1;\ il = 0,\ldots,NL\text{-}1\}$.

Here $NH = 12$ and corresponds to the number of quantized colours in Ittens' circle. "-1" index percentage of achromatic tones; "0" to "NH-1" points percentage of colours, ordered as it is shown on Figure 11, starting from reds and ending to purples.

We use $NS = 5$ for defining harmonies' and contrasts' descriptors. Index "0" holds percentage of greys and almost achromatic tones, and "4" contains percentage of pure (in particular – spectral) tones.

For indexing of luminance we use $NL = 5$. "0" holds percentage of very dark colours, and "4" contains percentage of very light colours.

To simplify further calculation up to three arrays, containing percentage values of corresponding characteristics in the picture is calculated on the basis of this array.

These arrays are:

- $H(h_{-1}, h_0, \ldots, h_{NH-1})$ for hues (the projection $A_H = \{A(ih,-,-)\}$, $ih = -1,\ldots NH-1$);
- $S(s_0, \ldots, s_{NS-1})$ for saturation (the projection $A_S = \{A(-,is,-)\}$, $is = 0,\ldots NS-1$);
- $L(l_0, \ldots, l_{NL-1})$ for lightness (the projection $A_L = \{A(-,-,il)\}$, $il = 0,\ldots NL-1$).

➢ ***Hue Order Vector***

Hue Order Vector contains number of dominant hues nh, and positions of dominant hues, ordered in decreasing percentage. nh can vary from zero for achromatic paintings to maximum values of defined dominant colours. For the purposes of defining hue harmonies maximum values of the dominant colours are restricted to 5. The value of nh is defined as the number of ordered hues, which sum of the percentages exceed some (expert-defined) value x when an image is not achromatic.

$(nh; p_1, p_2, \ldots, p_{nh})$, $nh \in \{0,\ldots,5\}$, $p_i \in \{-1,\ldots,NH-1\}$: $h_{p_i} \geq h_{p_{i+1}}$, $h \in H$, $i \in \{1,\ldots,nh-1\}$

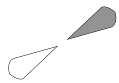
$$nh : \begin{cases} nh = 1 \ \ if \ \ h_{p_1} \geq x \\ nh = n \ \ if \ \ \sum_{i=1}^{n-1} h_{p_i} < x \ \ and \ \ \sum_{i=1}^{n} h_{p_i} \geq x \end{cases}$$

➢ ***Hue Harmony***

In order to define hue harmonies we first define functions which reflect the mutual disposition of two colours. Below we provide the mathematical formulation on the left side, and some graphical examples of defined disposition on the right. The dark leaf corresponds to the primary colour p, the light leaf shows relative disposition of the second colour, which is defined by corresponding function:

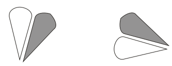
Colour, which lay opposite to the colour p:

$$opposite(p) = \begin{cases} p + NH\ div\ 2 & if \quad p \le NH\ div\ 2 \\ p - NH\ div\ 2 & if \quad p \ge NH\ div\ 2 \end{cases}$$



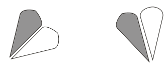
Colour, which is left neighbour of the colour p:

$$l\_neighbour(p) = \begin{cases} NH - 1 & if \quad p = 0 \\ p - 1 & if \quad p\ in\ \{1, \ldots, NH-1\} \end{cases}$$



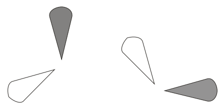
Colour, which is right neighbour of the colour p:

$$r\_neighbour(p) = \begin{cases} 0 & if \quad p = NH - 1 \\ p + 1 & if \quad p\ in\ \{0, \ldots, NH-2\} \end{cases}$$



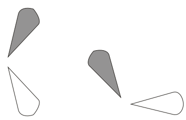
Colour, which can participate in triad with colour p, at its left side

$$l\_triad(p) = (NH + p - NH\ div\ 3) \bmod NH$$



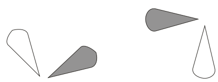
Colour, which can participate in triad with colour p, at its right side

$$r\_triad(p) = (p + NH\ div\ 3) \bmod NH$$



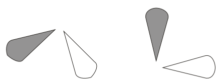
Colour, which can participate in tetrad with colour p, at its left side

$$l\_tetrad(p) = (NH + p - NH\ div\ 4) \bmod NH$$



Colour, which can participate in tetrad with colour p, at its right side

$$r\_tetrad(p) = (p + NH\ div\ 4) \bmod NH$$



The values of the hue harmony depend on the number of dominant hues nh:

✓ $nh = 0$**:**

Achromatic: the composition is constructed by black, white and grey tones. This construction can be examined as a special case of monochromatic harmony;

✓ $nh = 1$**:**

- *Monochromatic:* only one hue predominates in image;

✓ $nh = 2$**:**

- *Analogous*: when $p_2 = l\_neighbour(p_1)$ or $p_2 = r\_neighbour(p_1)$;
- *Complementary:* when $p_2 = opposite(p_1)$;
- *Partial Triad:* when $p_2 = l\_triad(p_1)$ or $p_2 = r\_triad(p_1)$;

✓ $nh = 3$**:**

- *Analogous:* if for one of dominant hues $p_i,\ i \in \{1,...,nh\}$ is fulfilled that the other two colours are $l\_neighbour(p_i)$ and $r\_neighbour(p_i)$ respectively;
- *Split complementary:* if for one of dominant hues $p_i,\ i \in \{1,...,nh\}$ is fulfilled that the other two colours are $l\_neighbour(opposite(p_i))$ and $r\_neighbour(opposite(p_i))$;
- *Triad:* if for one of dominant hues $p_i,\ i \in \{1,...,nh\}$ the other two colours are $l\_triad(p_i)$ and $r\_triad(p_i)$;

✓ $nh = 4$**:**

- *Analogous:* if for one of dominant hue $p_i,\ i \in \{1,...,nh\}$ is fulfilled that one of the other three colours $p_j, j \in \{1,...,nh\}, j \neq i$: $p_j = l\_neighbour(p_i)$ or $p_j = r\_neighbour(p_i)$ and other two colours are $l\_neighbour(p_j)$ and $r\_neighbour(p_j)$;
- *Double Complementary:* if for one of dominant hue $p_i,\ i \in \{1,...,nh\}$ is fulfilled that one of the other three colours $p_j, j \in \{1,...,nh\}, j \neq i$: $p_j = opposite(p_i)$ and other two colours are $l\_neighbour(p_i)$ and $l\_neighbour(p_j)$ or $r\_neighbour(p_i)$ and $r\_neighbour(p_j)$;
- *Split Complementary:* if for one of dominant hue $p_i,\ i \in \{1,...,nh\}$ is fulfilled that one of the other three colours $p_j, j \in \{1,...,nh\}, j \neq i$: $p_j = opposite(p_i)$ and other two colours are $l\_neighbour(p_j)$ and $r\_neighbour(p_j)$;
- *Tetrad:* if for first hue p_1 the other hues are $l\_tetrad(p_1)$, $opposite(p_1)$, $r\_tetrad(p_1)$ respectively;

✓ $nh = 5$:

– *Multicolour:* here the presence of defined combinations discarding the colour with smallest presence can be searched.

➢ ***Cold/Warm Contrast***

For defining cold/warm contrast we use the proportion of the percentage values of families of colours p_{warm}, p_{cold}, and $p_{achromatic}$. We have taken into account the fact of changing the type of a colour depending on its saturation and lightness [Koenig, 2010]. Because of this we calculate these values using the three-dimensional array A. The strongest contrasts points is the warmest "red-orange" ($ih = 1$) and the coolest "blue-green" ($ih = 7$). We use semi-linear function of including colours in warm, respectively cold family, with the following properties (Figure 33):

– All achromatic values ($ih \in \{-1\}$)are added to an achromatic family;
– All strongly unsaturated colours ($is \in \{0\}$) are added to an achromatic family;
– Increasing the lightness in unsaturated colours ($is \in \{1,2\}$) leads to increasing of coldness. For instance dark unsaturated colours are added in warm family from magenta to orange-yellow ($ih \in \{11,0,1,2,3\}$), but from light ones only red, red-orange, and orange are added ($ih \in \{0,1,2\}$). On the contrary, dark colours added to cool family are only these near blue-green ($ih \in \{6,7,8\}$); increasing the light expands the range and in cool family from lightest yellow-green to lightest blue-magenta ($ih \in \{5,6,7,8,9\}$) and halves of neighbours are included ($ih \in \{4,10\}$);
– Colours with middle saturation ($is \in \{3\}$) include stable families of warm colours ($ih \in \{0,1,2,3\}$) and cold colours ($ih \in \{6,7,8\}$);
– For saturated colours ($is \in \{4\}$) increasing the lightness cause expanding of both families of warm and cold colours. For instance for dark saturated colours in warm family belongs from magenta to orange-yellow ($ih \in \{11,0,1,2,3\}$), while in light spectrum half of their neighbours also are included ($ih \in \{4,10\}$). Cold colours also expands with half tones when the lightness increases.

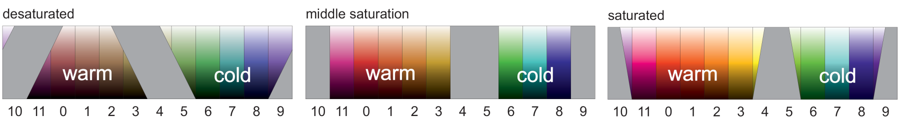


Figure 33. Cold/warm depending by saturation and lightness

An image is defined as *warm*, *cold*, or *neutral* if a corresponding value is greater than some threshold. If none of these values exceeds given parameters, the image is *warm-cold*, *warm-neutral*, *cold-neutral* according to order of decreasing of corresponded values.

- ***Saturation Order Vector***

The Saturation Order Vector contains number of dominant saturations ns, $ns \in \{1,...,NS\}$, and positions of dominant saturations, ordered in decreasing percentage. The value of ns is defined as the numbers of ordered saturations, which sum of the percentages exceeds some value y.

$(ns; p_1, p_2, ..., p_{ns})$, $ns \in \{1,...,5\}$, $p_i \in \{0,...,NS-1\}$: $s_{p_i} \geq s_{p_{i+1}}$, $s \in S$, $i \in \{1,...,ns-1\}$

$$ns : \begin{cases} ns = 1 \;\; if \;\; s_{p_1} \geq y \\ ns = n \;\; if \;\; \sum_{i=1}^{n-1} s_{p_i} < y \;\; and \;\; \sum_{i=1}^{n} s_{p_i} \geq y \end{cases}$$

- ***Saturation Combination***

If $ns = 1$ a picture is defined as *monointense*. If $ns > 1$ some combinations of presence of dominant saturations can be outlined. If p_0 and p_{NS-1} are dominant saturations, the image is defined as *contrary*; if saturations are adjoining – the feature is *smooth*, etc.

- ***Clear/Dull Contrast***

Depending of the global lightness of the image the saturation distribution of an image is possessed in another attribute, which can receive values as *soft* or *sharp* for light images, *ground* or *spectral* for images with medium lightness and *dull* or *clear* for dark images.

- ***Lightness Order Vector***

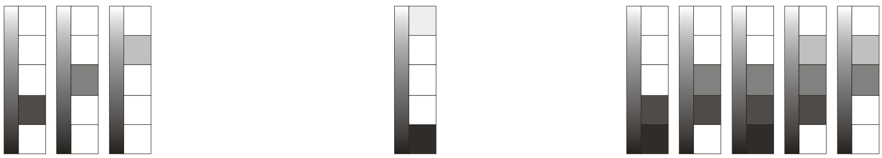
The Lightness Order Vector $(nl; p_1, p_2, ..., p_{nl})$ is defined in the same way as the saturation order vector. It contains number of dominant lighting values nl, $nl \in \{1,...,NL\}$, and their positions, ordered in decreasing percentage. The value of nl is defined as the numbers of ordered values of lightness, which sum of the percentages exceeds some value z.

$(nl; p_1, p_2, ..., p_{nl})$, $nl \in \{1,...,5\}$, $p_i \in \{0,...,NL-1\}$: $l_{p_i} \geq l_{p_{i+1}}$, $l \in L$, $i \in \{1,...,nl-1\}$

$$nl : \begin{cases} nl = 1 \ \ if \ \ l_{p_1} \geq z \\ nl = n \ \ if \ \ \sum_{i=1}^{n-1} l_{p_i} < z \ \ and \ \ \sum_{i=1}^{n} l_{p_i} \geq z \end{cases}$$

- ***Lightness Combination***

The Lightness Combinations are defined in the equal manner as saturation ones where the same function are used and only corresponding parameters are changed.



a) monointense *b) contrary* *c) smooth*

Figure 34. Variants of lightness combinations

When $nl = 1$ a picture is defined as *monointense* (Figure 34a). If $nl > 1$ some combinations of presence of dominant lightness can be outlined. If p_0 and p_{NL-1} are dominant lightnesses, the image is defined as *contrary* (Figure 34b); if the lightnesses are adjoining – the feature is *smooth* (Figure 34c), etc.

- ***Light/Dark Contrast***

The attribute, which receives values for light-dark contrast depends of user defined threshold of darkness and lightness. The images, which hold l_0 more than given dark threshold, are identified as *very dark*. *Dark* images are these for which $l_0 + l_1$ exceed this threshold. Similarly, the images with l_{NL-1} receive value *very light* and these for which $l_{NL-2} + l_{NL-1}$ exceed the threshold are *light*. Depending of distribution of lightness, images can be categorized as *dark-light*, *light-dark*, *middle*, etc.

5.6 Local Features, based on Vector Quantization of MPEG-7 Descriptors over Tiles

MPEG-7 descriptors are complex descriptors, which provide a good presentation of different types of visual features. The description of the structure of MPEG-7 descriptors and algorithms are given in [ISO/IEC 15938-3]. These complex structures need specific processing and cannot be put directly into generic classification algorithms. Here we give a brief explanation of each MPEG-7 descriptor examined in our work and which values we use in the further processing:

- *Scalable Colour (SC)* represents the colour histogram in the HSV colour space, encoded by a Haar transformation. For presenting the image or a selected part, *Scalable Colour* needs a vector with 64 attributes;
- *Colour Layout (CL)* specifies the spatial distribution of colours using *YCbCr* colour 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. These coefficients are used to form three vectors, which contain all extracted values of the *Y*, *Cb* and *Cr* components and thus, the *Colour Layout* vector has 12 attributes:

$$DY = \{DY_y; y = 1..6\} \quad DY_y = \begin{cases} YDCCoeff & y = 1 \\ YACCoeff_{y-1} & y \neq 1 \end{cases}$$

$$DCb = \{DCb_y; y = 1..3\} \quad DCb_y = \begin{cases} YDCbCoeff & y = 1 \\ YACbCoeff_{y-1} & y \neq 1 \end{cases}$$

$$DCr = \{DCr_y; y = 1..3\} \quad DCr_y = \begin{cases} YDCrCoeff & y = 1 \\ YACrCoeff_{y-1} & y \neq 1 \end{cases}$$

- *Colour Structure (CS)*, which specifies both colour content and the structure of the content. The descriptor expresses local colour 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 *Colour Structure*;
- *Dominant Colour (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 [Ivanova and Stanchev, 2009]. 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 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 Colour, Colour Layout, Colour Structure and Dominant Colour concern different aspects of the same phenomenon, i.e. distribution of the colour 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 a 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 our approach [Ivanova et al, 2010/ICDEM] 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 $j \in 1...n$. The index i increases from the left tile to the right tile and the index j increases from the top tile to the bottom tile of the image.

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

A subset $LI = \left\{ I_q \mid q = 1...l,\ l < k,\ I_q \in I \right\}$ of representative images is extracted from the observed set and is used as a learning set. The images of this subset produce a set of $l \times m \times n$ tiles $LT = \left\{ I_q^{ij} \mid q = 1...l, i = 1...m, j = 1...n \right\}$.

For each MPEG-7 descriptor $X \in \{SC, CL, CS, DC, EH, HT\}$ the algorithm consists of the following steps:

- For all tiles in LT, the MPEG-7 descriptor and the corresponding feature vector $X \in \{SC, CL, CS, DC, EH, HT\}$ is calculated;
- The tiles in LT are clustered. We used "repeated bisections". 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;
- The centroids of each cluster are calculated;

- For each tile a value, which corresponds to the number of the cluster where the tile belongs to, is assigned;
- For tiles which were not in the learning set $IT \setminus LT$, the membership of their centroids is calculated using L_1-metric, and the number of the corresponding cluster is assigned as a value of the tile.

As a result, each image of the observed set I is represented with a feature vector with $x \times m \times n$ numerical attributes, where x is the number of MPEG-7 descriptors. 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.

A specific feature of this approach is that the obtained attributes are nominal. The main purpose of the prepared datasets after implementing this approach is to examine the significance of the attributes and the local/global trade-off for class prediction.

Conclusion

We have proposed to extract visual features, which represent colour distribution in art images in colour histograms. We suggest using presentation of the colour in comprehensive manner using the human dimensions, which reflects the hue, the saturation and the lightness of the colour. As appropriate colour models we use HSV, HSL and HSL-artist colour models.

We have developed one classification of different types of harmonies and contrasts, based on the Itten's theory, from the point of view of the three main characteristics of the colour – hue, saturation and lightness. The classification consists of following types:

- *Harmonies/Contrasts Based on the Hues Disposition;*
- *Warm-Cold Contrast;*
- *Harmonies/Contrasts from the Saturation Point of View;*
- *Clear-Dull Contrast;*
- *Harmonies/Contrasts from the Lightness Point of View;*
- *Light-Dark Contrast.*

We have defined formal functions for extracting defined harmonies and contrasts, based on already received low-level features for colour distribution.

We also have suggested a method for local extraction of colour and texture characteristics, based on vector quantization of MPEG-7 descriptors for tiles of the images in order to analyze possibilities of capturing more detailed information for semantic and abstraction content of art images based on the MPEG-7 descriptors with significant dimensionality reduction.

6 The System APICAS

Abstract

We propose the architecture of an experimental CBIR system as an environment for testing applicability of suggested visual features for image retrieval.

The realization of the proposed architecture as a special software system, named "Art Painting Image Colour Aesthetics and Semantics" (APICAS) is presented.

The main functions in APICAS are:

- *data entry support: choosing the collection; setting up quantization parameters and boundaries; setting up parameters for vector quantization; selecting the samples of learning set; supplying with textual metadata, which is used for comparison of adequacy of received results by visual features;*
- *visual characteristics extraction: calculating colour distribution; estimating harmonies' and contrasts' descriptors; establishing local features, based on vector quantization of MPEG-7 descriptors over the tiles of the image;*
- *data delivery: examining colour distribution; multidimensional scaling; knowledge analysis; visualizing extracted feature values, as well as statistical and data mining analysis.*

6.1 Functional Requirements

The first step towards defining a suitable architecture for a CBIR system is to analyze the functional requirements it needs to meet. Our state-of the art review demonstrated that CBIR systems are developed most typically as specialized stand-alone applications or modules and are designed as such. This is a typical approach within an emerging domain but with the growing importance of image retrieval in the modern Web environment what becomes of special importance is how to develop modules for CBIR which could easily be integrated in digital repositories and web portals. This would require analyzing functional requirements for CBIR systems in the context of functional requirements within the current trends in digital archives. In order to address them, we will first present the high-level architecture of modern digital archives.

In 2002 the Consultative Committee for Space Data Systems (CCSDS) prepared a Blue book with technical recommendation that establishes a common framework of terms and concepts which comprise an Open Archival Information System (OAIS) [OAIS, 2002] which had been also adopted as the international standard ISO 14721:2003 Space data and information transfer systems – Open archival information system – Reference model.

This model can be successfully implemented as a common framework with concretizations in application areas for the memory institutions, or the so called GLAM (Galleries, Libraries, Archives, Museums). The functional schema of OAIS (Figure 35) contains six functional entities and related interfaces.

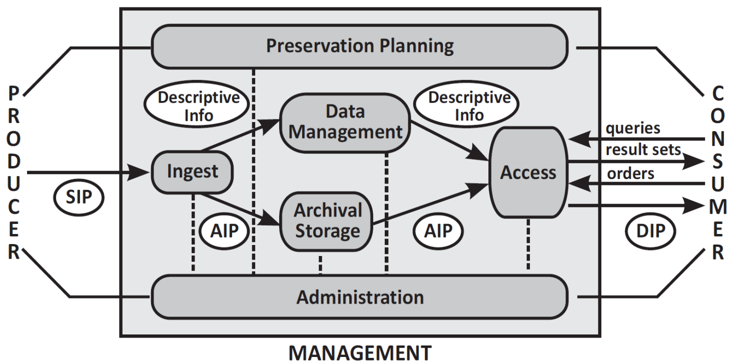


Figure 35. OAIS Functional Entities [OAIS, 2002]

Ingest functions include receiving Submission Information Packages (SIPs), performing quality assurance on SIPs, generating an Archival Information Package (AIP), extracting Descriptive Information from the AIPs for inclusion in the archive database, and coordinating updates to Archival Storage and Data Management. **Archival Storage** provides the services and functions for the storage, maintenance and retrieval of AIPs. **Data Management** provides the services and functions for populating, maintaining, and accessing both Descriptive Information which identifies and documents archive holdings and administrative data used to manage the archive. Data Management functions include administering the archive database functions (maintaining schema and view definitions, and referential integrity), performing database updates (loading new descriptive information or archive administrative data), performing queries on the data management data to generate result sets, and producing reports from these result sets. **Administration** provides the services and functions for the overall operation of the system. **Preservation Planning** provides the services and functions for monitoring the environment of the OAIS and providing recommendations to ensure that the information stored in the OAIS remains accessible to the Designated User Community over the long term, even if the original computing environment becomes obsolete. **Access** provides the services and functions that support Consumers in determining the existence, description, location and availability of information stored in the OAIS, and allowing Consumers to request and receive information products. Access functions include communicating with Consumers to receive requests, applying controls to limit access to specially protected information, coordinating the execution of requests to successful completion, generating responses (Dissemination Information Packages, result sets, reports) and delivering the responses to Consumers.

Within the context of such general digital archive architecture, CBIR-related implementations can be seen as a module which would best fit within the **Data Management** functional entity. However, it would also have influence on Ingest and more specifically on the structure of the SIPs (submission information packages) because the successful implementation of CBIR requires some specific data and metadata. CBIR also enriches the possibilities for delivery and will influence the **Access** functional entity which would accommodate more options for digital content discovery.

This wider context is reflected in the architecture of a CBIR system called APICAS; this system is fine-tuned to the need of IR in the area of digitized art collections. The specialized core part of the system accommodated the necessary specific instrumentarium in terms of algorithms and methods for IR; these are seen as specialized instances of data management tools. At the same time the specific requirements for **Ingest** of specific data necessary for the components of information retrieval and the expanded **Access** possibilities are also highlighted.

6.2 APICAS Architecture

A designated software system "Art Painting Image Colour Aesthetics and Semantics" (APICAS) was developed in order to provide an appropriate environment for realizing examined algorithms and conducting experiments.

In Figure 36, the architecture of the proposed system is shown. The process of extraction of presented in previous chapter visual features is placed in a global frame, where several types of similarity queries, statistical analysis and categorization can be made.

The functional schema of APICAS follows the OAIS standard, escaping blocks, connected with overall administration and preservation planning. The Ingest functions in such experimental system are also very simplified, because the focus is on the extracting of visual metadata and analyzing received features. By this reason concrete realization of archival storage functions is not important.

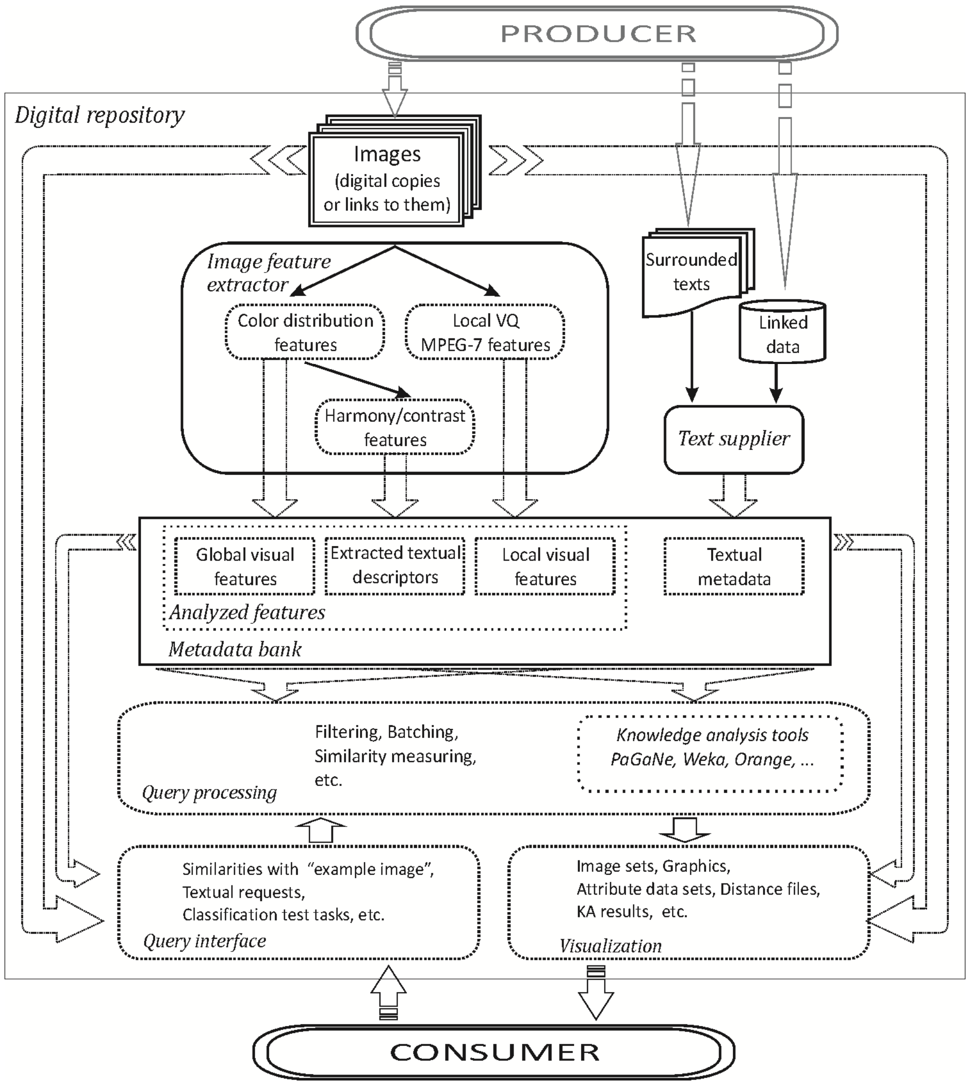


Figure 36. APICAS architecture

Main functions in APICAS are connected with:

- Data entry – establishing connections with image sources as well as supplying controlling textual metadata;
- Feature extraction – such function produces automated metadata for image labelling;

- Query interface – part of user interface functions, connected with receiving of the tasks from the consumer. Here image bank is used in order to select "an example" for searching most similar images to the selected image. The metadata bank is used for constructing "controlled vocabulary", from which user can select desired feature;
- Query processing – analysis of extracted metadata, their potential to answer user's expectations to receive 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 images sets (whole images or patches), attribute data sets, distance files, graphics, knowledge analysis results, etc.

The main goals of APICAS are two-fold:

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

The vividness of the 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 to the society new greater sophistication, complexity, and functionality.

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 [Mitov et al, 2009a, 2009b], which is supplied statistical and attribute analyzing tools as well as a specially designed PGN-classifier, which combines generalization possibilities of Association Rule Classifiers with answer accuracy like K-Nearest Neighbours. The set of resulting association rules for some class value can be examined as a definition of this concept. Obtaining these regularities, highly depends not only by chosen machine learning algorithm for creating these rules, but also on choosing visual attributes which represent the images well.

6.3 APICAS Ground

The system is realized using CodeGear Delphi 2007 for Win32.

As metadata storage space an Arm 32, property of FOI Creative Ltd. was used.

For obtaining the MPEG-7 descriptors APICAS refers to the Multimedia Content Management System MILOS [Amato et al, 2004].

For obtaining the results of multidimensional scaling we used the open component-based data mining and machine learning software suite Orange [Demsar et al, 2004], maintained and developed at the Faculty of Computer and Information Science, University of Ljubljana, Slovenia.

As a clustering algorithm the program "vcluster", which is a part of the CLUTO open source software package [Karypis, 2003], is implemented in the system.

As a knowledge analysis and testing environment we used the data mining analysis environment PaGaNe, realized by our collective in Institute of Mathematics and Informatics [Mitov et al, 2009a&b]. We use the PGN classifier, ArmSquare association rule miner and implemented statistical analyzing tools for checking up our results and extracting regularities for artists' and movements' styles based on the extracted attributes.

As a control environment to the obtained results from the PGN classifier we used Waikato Environment for Knowledge Analysis (Weka) [Witten and Frank, 2005], developed at the University of Waikato, New Zealand.

6.4 APICAS Functionality

The system is built as an environment for carrying out different types of experiments. The starting screen connects several program modules, developed gradually over the years, in a common background.

The main goal of the analysis of colour characteristics distribution is not to use these low-level features themselves, but only as a ground for establishing upper level harmonies/contrast descriptors. Because of this the functions, connected with their analysis, are encapsulated in a separate section.

Functions for calculating different attributes, which become the focus of defining upper-level descriptors (such as harmonies and contrasts, vector quantization of MPEG-7 descriptors for tiles of the images, etc.) lay in the other part of the program.

Obtaining 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:

- Extracting colour distribution using exact quantization of the HSL-colour space [Ivanova et al, 2008];

- Using MPEG-7 Dominant Colour descriptor as a source for determining colour distribution in the image and calculating of harmonies and contrasts features [Ivanova and Stanchev, 2009];
- Using fuzzy calculations for establishing belonging of the colour characteristics into quantized bins [Ivanova et al, 2010/IJAS].

6.4.1 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 (using a corresponded button). The system automatically scans new collections and forms a database. There is a special button which 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 (Figure 37). Access to user defined minimal thresholds, used in definition of different kinds of harmonies and contrast descriptors is also provided here.

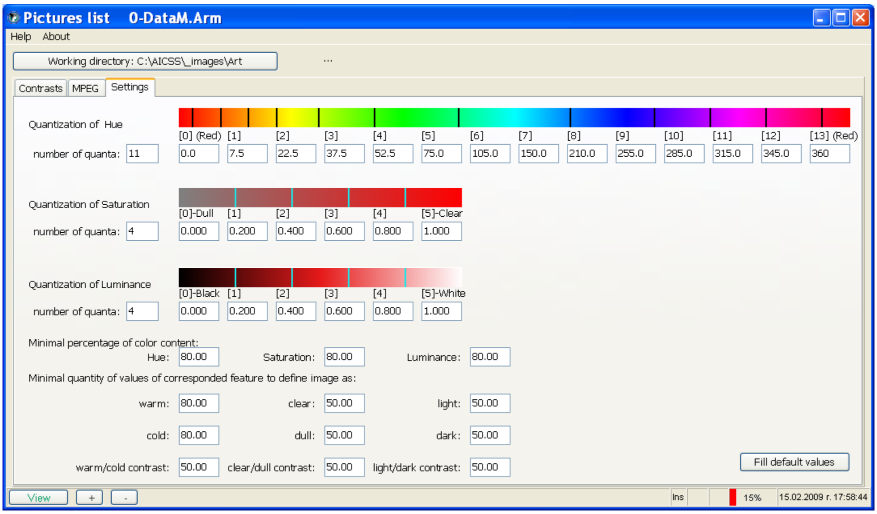


Figure 37. Screen for setting up the quantization parameters and boundaries

➢ *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.

All parameters, connected with choosing and calculating vector quantization features, are shown in Figure 38.

The user can choose which of the MPEG-7 descriptors to be included in the clustering procedure as well as in the process of preparing the datasets for knowledge analysis.

The examined tiles can be given from the whole surface (LR), only from left (L) or right (R) half of the image. On the other side all tiles from a chosen surface can be given in chess-board order starting from the first tile (odd start) or from the second tile (even start). The last criterion is giving all elements, with or without tiles, which are at the boundary.

The numbers of tiling by width and by height as well as the number of resulting clusters are given as parameters.

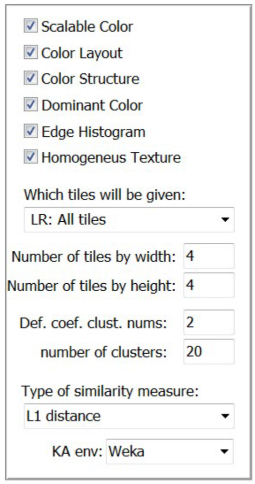


Figure 38. Parameters for Vector Quantization

By default the similarity measure function is based on L_1-metric. The user can choose another metric using the provided button.

A combo box "KA env." allows changing the style of preparing the dataset in a way which is more appropriate for PaGaNe, Weka or Orange.

➢ *Selecting the Samples of Learning Set*

The files that contain samples of learning or examining set can be prepared 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 facilitates the creation of these files (Figure 39) by marking which image to be included in the corresponding set. The activation of the buttons "Write learning set" or "Write examining set" forms the corresponding text file.

The paintings belonging to the learning set or examining set, is extracted from the file selected by the user. The system checks reading samples for availability in the collection and marks them as participants of the learning set. In this way the selected paintings can be fixed and used in different collections.

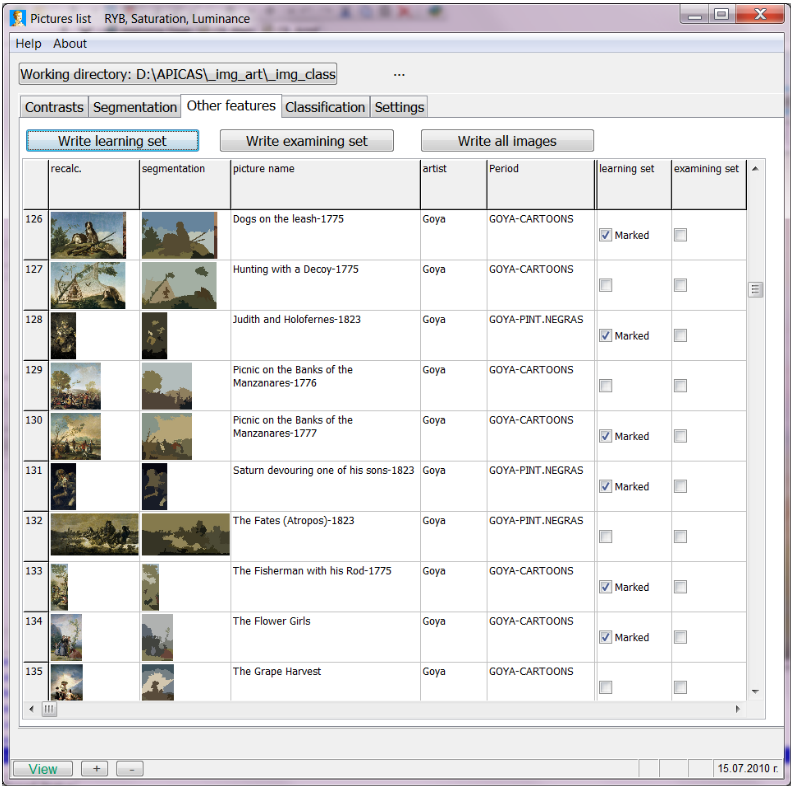


Figure 39. The grid for manual selection of the learning and examining sets

6.4.2 Function for Supplying with Textual Metadata

The process needs textual metadata, which describe different aspects of the image content for the purposes of the used learning algorithms as well as for testing the received results. This information can be received by different ways – filled in manually or automatically extracted from the context. In the Web space this information can be extracted from the Internet page, which contains the examined image. This process taken alone is a separate field for investigation.

For the purposes of the experiments made in this dissertation, we used simple ways for obtaining the metadata needed. The files, used in the collections, contain in their names the information about the artist, the name of the stored painting and sometimes the year when the painting is made. The filenames are used 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, such as dates of birth and death; countries, where mainly they have worked; periods of their creative work. The ontology contains the connections between the described concepts, which allow using the extracted from the filename information to receive all additional information that can be attached to the examined paintings. Other kinds of metadata, such as the theme of the paintings, their genres, used techniques, etc., can be added manually.

The system extracts the names of the picture and the artist and, using this small thesaurus, connects the metadata, extracted by the context, to information for the pictures. The files containing lists of movements, sub-movements and artists found in current collection, are also created or actualized.

This function can be initiated by a designated button. During the primary scanning of the collection the function is activated automatically.

Receiving all these data is capsulated in a separate function, which can be changed with more enhanced instruments in further realization of the system. For instance, some information can be extracted from the context of the image, using the web page where the image is posed. For the experimental case, this apparatus is enough to supply an image with sufficient metadata for conducting categorization and testing tasks.

6.4.3 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 being a part of the process of defining harmonies and contrast descriptors.

The function receives as input:

- The name of the image;
- Which feature or combinations of features have to be calculated. The possible values are: 0: all features; 1: Hue; 2: Saturation; 3: Lightness; 12: Hue&Saturation; 13: Hue&Lightness; 23: Saturation&Lightness. In the case when only part of features are pointed – the corresponded projections are received in zero coordinate for non-selected feature(s);
- Examined area. The system can either take into consideration the whole image (by default) or only a chosen part of it (for instance – rectangle area with left-up corner (33%,33%) and right-bottom corner (66%,66%);
- Examined colour model. The system has possibilities to use HSV or HSL colour model (with exact quantization) or HSL-artist colour model, using fuzzy quantization method.

The function gives each pixel from observed area, converts the colour value from RGB-colour model to colour coordinates in HSV or HSL colour model and as a result of the application of the selected quantization increases the presence of colour with quantized coordinates. In the case of fuzzy quantization, the increase catches neighbour coordinates with corresponding value. Finally the values in the array are normalized.

When the function is used in the case of defining harmonies and contrast features on the base of MPEG-7 descriptors the function uses the Dominant Colour descriptor as a source of colour distribution. The function gives the colour values of Dominant Colour descriptor and applying selected quantization receives coordinates in the array where the percentage of this colour will be placed. The percentage values of Dominant Colour in MPEG-7 standard are coded in 5 bits. The function transforms these values to percentage values ranged from 0 to 100 and puts them in the point with given coordinates. In this way further steps of defining colour harmonies and contrast descriptors are unified.

The output is a three dimensional array

$A = \{A(ih, is, il) \mid ih = -1, \dots, NH - 1;\ is = 0, \dots, NS - 1;\ il = 0, \dots, NL - 1\}$

which contains colour distribution by selected dimensions.

➢ ***Estimating Harmonies' and Contrasts' Descriptors***

Special functions for calculating defined harmonies' and contrast' descriptors are realized in the system.

The function receives as parameters:

- The name of the image;
- The way of receiving colour distribution. Three variants are possible here – HSL-colour model with exact quantization; MPEG-7 Dominant Colour descriptor as a source of colour distribution; and HSL-artist colour model with fuzzy quantization.

For obtaining three dimensional colour distribution array this function refers to a previously discussed function. The further algorithms are already discussed in previous chapter. The resulting features are written in a database.

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

✓ ***Choosing Learning Samples***

Setting up the learning set is executed by choosing the text file that contains learning samples. The function reads text file, check the images for existing in current collection, and writes correct samples in a database.

✓ ***Clustering***

The function passes into several steps:

- Creating tiles from images of the learning set with given parameters (number of tiles by width m and number of tiles by height n);
- Calculating MPEG-7 descriptors for these tiles using the MILOS system;
- For each MPEG-7 descriptor the system builds a dataset, which contains corresponding feature vector for each chosen tile from the learning set of images;
- This dataset is put under clustering procedure "vcluster" with selected number of clusters. As a clustering method "repeated bisections" is used. The similarity between objects is computed using the correlation coefficient. As a clustering criterion function, cluster maximization of sum of square root of sums of similarities of vectors, belonging to given cluster, is used;
- The centroids of each cluster are calculated;
- The value which corresponds to the number of the cluster where the tile belongs to, is assigned to each tile. These results are saved in the database.

✓ ***Finding Most Similar Tiles to the Centroids***

This function 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***

The button "Processing rest of the images" activates the function that scans the database and defines corresponded values for all images which are not in the learning set. 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 is used:
  - *Scalable Colour:* function takes into account the significance of the order of coefficients [Herrmann, 2002]:

$$DY = \{DY_y; y = 1..6\} \quad DY_y = \begin{cases} YDCCoeff & y = 1 \\ YACCoeff_{y-1} & y \neq 1 \end{cases}$$

$$DCb = \{DCb_y; y = 1..3\} \quad DCb_y = \begin{cases} YDCbCoeff & y = 1 \\ YACbCoeff_{y-1} & y \neq 1 \end{cases}$$

$$DCr = \{DCr_y; y = 1..3\} \quad DCr_y = \begin{cases} YDCrCoeff & y = 1 \\ YACrCoeff_{y-1} & y \neq 1 \end{cases}$$

$$dist = \sqrt{\sum_{i=1}^{6} w_{yi}(DY_i - DY_i^{'})^2} + \sqrt{\sum_{i=1}^{3} w_{bi}(DCb_i - DCb_i^{'})^2} + \sqrt{\sum_{i=1}^{3} w_{ri}(DCr_i - DCr_i^{'})^2}$$

where
$$\begin{array}{lllllll} w_{y0} = 3 & w_{y1} = 3 & w_{y2} = 3 & w_{y3} = 1 & w_{y4} = 1 & w_{y5} = 1 & w_{y6} = 1 \\ w_{b0} = 2 & w_{b1} = 2 & w_{b2} = 2 & w_{b3} = 1 & & & \\ w_{r0} = 4 & w_{r1} = 2 & w_{r2} = 2 & w_{r3} = 1 & & & \end{array};$$

– *Edge Histogram:* here is realized one extension in order to capture not only the local edge distribution information but also semi-global and global ones, which is proposed in [Won et al, 2002]:

$$\begin{aligned} Dist(A,B) = & \sum_{i=0}^{79} \left| Local\_A[i] - Local\_B[i] \right| + \\ & 5 \times \sum_{i=0}^{4} \left| Global\_A[i] - Global\_B[i] \right| + \\ & \sum_{i=0}^{64} \left| SemiGlobal\_A[i] - SemiGlobal\_B[i] \right| \end{aligned}$$

where for image A**:** $Local\_A[i]$ represents the reconstructed value of $BinCount[i]$. $Global\_A[]$ represents the normalized bin values for the global edge histograms. $SemiGlobal\_A[]$ represents the normalized histogram bin values for the semi-global edge histograms of the image (Figure 40).

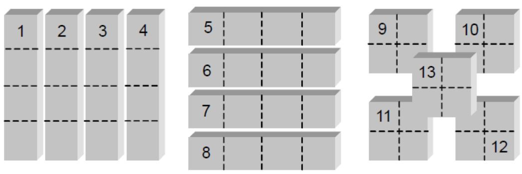


Figure 40. Segments of sub-images for semi-global histograms [Won et al, 2002]

In current experiments we have used the first variant because the clustering procedure, applied on the tiles of the learning set uses L_1-metric.

6.4.4 Functions Connected with Output Information

➢ ***Examining Colour Distribution***

One class of functions, realized in APICAS, is directed to carry out the analysis of distribution of colour characteristics in the images – hue, saturation or luminance, or combination of them (Figure 41). These functions are firstly introduced in [Ivanova et al, 2008], 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.

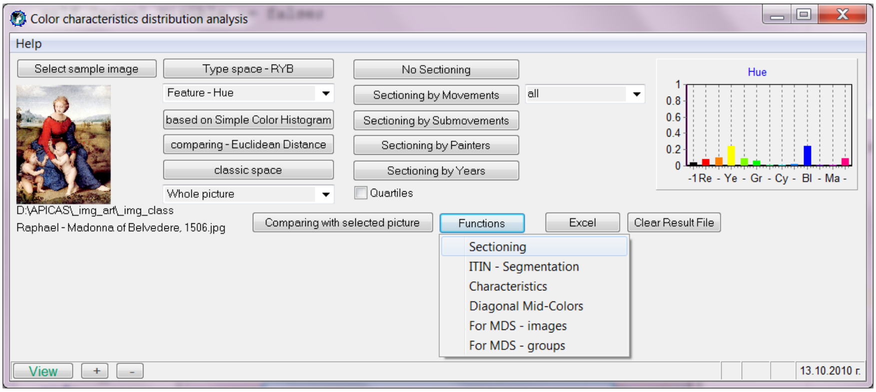


Figure 41. Some of the functions for colour characteristics distribution analysis

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 over:

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

The function selects and/or sections the images in the collection. For Obtaining colour distribution of a given image it refers to already discussed function for calculating colour distribution. If the calculations are already made, the function can overcome calculating the meaning values and only visualizes them using stored information. The result is displayed on the screen and in the same time is recorded as a "csv" file in the text directory of the system.

➢ ***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 42.

Figure 42. Results of calculating of types of harmonies/contrasts for the picture "Annunciation" by Botticelli

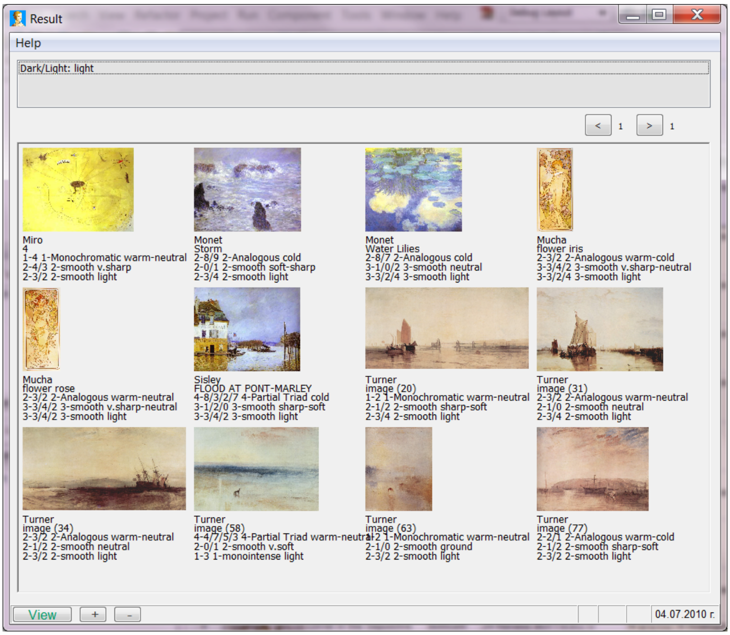


Figure 43. Result of retrieval from the image base with parameter: "Dark/light contrast = Light" (includes "smooth light" and "monointense light")

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

The system allows searching within a collection of images, which has specific combination of the colours, defined by some harmony or contrast. One example is shown in Figure 43.

Another branch of the system allows creating datasets containing extracted attributes or selected part of them labelled with chosen profile such as artist name, movement, scene-type. These datasets can be 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 theme in the images.

➢ *Visualizing the Results of Clustering*

The system allows viewing the results of the clustering procedure showing all tiles, which belong to the selected number of cluster for specific MPEG-7 descriptor. For instance, in Figure 44 part of the tiles, which are in cluster No:2 of Colour Structure Descriptor for 5×5 tiling, is presented.

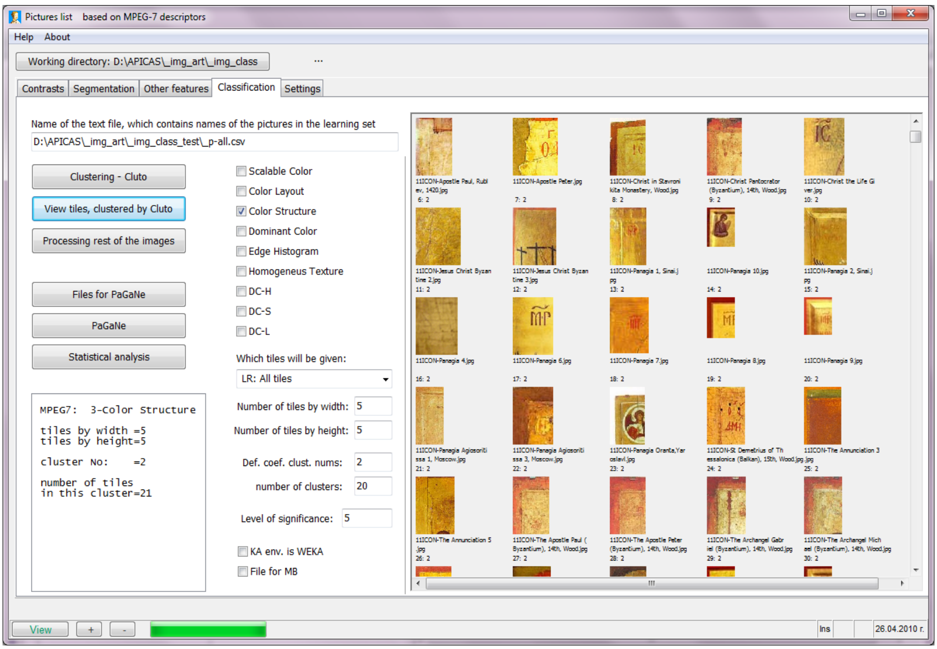


Figure 44. The screenshot of viewing 5×5 tiles, belonging to cluster No:2 of Colour Structure Descriptor

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. The function uses the results of the already discussed function for finding most similar tiles to the centroids.

Figure 45 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.

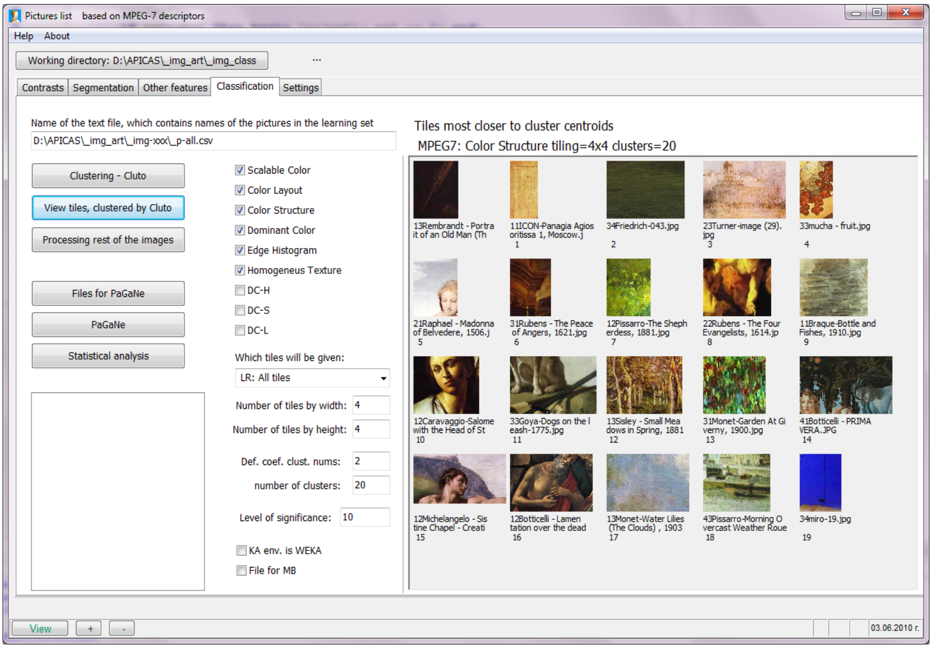


Figure 45. The tiles, most closer to the centroid of Colour Structure Descriptor (tiling 4×4 with 20 clusters)

➢ ***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 file with primary data "tab-delimited format" contains following rows: first row contains names of the attributes, second row point the type of the attribute, the third line contains flag "class" on the position of class-label.

The file containing distance matrix has a special format: the first line in the file has to start with the number of items (the dimension of the matrix), followed by the word "labeled"; the rest of the file is the matrix, where the elements are separated by tabulators. The label is in front of each line. In our case the matrix is given with the lower part.

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 points I and J is calculated as Manhattan distance (L^1-metric) between distances of theirs characteristics: $d(I,J)=d_h(h_i,h_j)+d_s(s_i,s_j)+d_l(l_i,l_j)$. Because of the angular type of hue characteristic, the distance is calculated as $d_h(h_i,h_j)=\min(\left|h_i-h_j\right|, maxh-\left|h_i-h_j\right|)$, where $maxh$ is the number of hue colour bins (Figure 46). For saturation distance and luminance distance the L^1-metric is used.

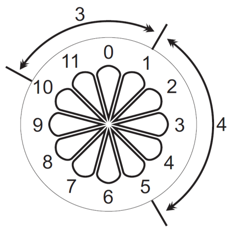


Figure 46. Angular type of distance

- ✓ ***For Harmonies/Contrast Descriptors***

The values of these descriptors are categorical, but can be 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 instance, the distance matrix for hue harmony is shown in Table 1.

Table 1. Distances between values of hue harmony/contrast

Achromatic	0						
Monochromatic	1	0					
Analogous	2	1	0				
Partial Triad	3	3	2	0			
Triad	3	3	2	1	0		
Tetrad	4	4	3	2	2	0	
Complementary	5	5	4	3	3	3	0
	Achromatic	Monochromatic	Analogous	Partial Triad	Triad	Tetrad	Complementary

✓ ***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 included into the system.

The functions allow the user to choose which of the attributes to be included in the dataset: only local parameters, connected with vector quantization of checked MPEG-7 descriptors or together with global ones, which present harmonies and contrast descriptors.

The artist's name, movement, sub-movement, scene-type can be assigned as class labels.

A checkbox "KA environment is Weka" allows to change the format of the output files for PaGaNe, Weka or Orange.

An additional function allows to create a dataset for the 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.

- ***Connection with Knowledge Analysis Tools***

As knowledge analysis tools, we have used several programs, such as PaGaNe, Weka and Orange.

PaGaNe contains several tools that are useful for analysis of attributes and their combinations for forming discriminating semantic and abstraction profiles of the images:

- Tools for analyzing statistical attribute significance;
- Associative rule miner ArmSquare, which allows discovering of frequent mining rules in a global environment as well as sectioned by some criterion;
- Class-association rule PGN classifier, which, by its construction, combines generalization possibilities of rule-based classifiers with answer accuracy like lazy algorithms.

We have used Weka [Witten and Frank, 2005] as a control test-bed for the outcomes of the PGN classifier.

There are buttons available which activate the systems PaGaNe, Weka and Orange.

Conclusion

We have proposed the architecture of an experimental CBIR lab-system, aimed at analyzing different types of visual features, which strives 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).

We have divided the main functions in APICAS into the following groups:

(1) Functions that Serve Data Entry:

- *Choosing the Collection;*
- *Setting up Quantization Parameters and Boundaries;*
- *Setting up Parameters for Vector Quantization;*
- *Selecting the Samples of Learning Set.*

(2) Function for Textual Metadata Supplying.

(3) Functions for Calculating Visual Characteristics (Feature Extraction):

- *Calculating Colour Distribution;*
- *Estimating Harmonies' and Contrasts' Descriptors;*
- *Establishing Local Features, Based on Vector Quantization of MPEG-7 Descriptors over the Tiles of the Image:*
    - *Choosing Learning Samples;*
    - *Clustering;*

- *Finding Most Similar Tiles to the Centroids;*
- *Defining Corresponded Features for the Rest of the Images.*

(4) Functions, Connected with User Interaction:

- *Examining Colour Distribution;*
- *Visualizing Extracted Colour Harmonies and Contrast Feature Values;*
- *Visualizing the Results of Clustering;*
- *Preparing Data for Knowledge Analysis.*

In order to make the knowledge analysis easier we also established connections with PaGaNe, Weka and Orange by special buttons in the system.

All these functions had been implemented in the APICAS system.

7 Experimental Results

Abstract:

We have the following general research goal: showing and proving the added value of the visual, higher-level and local constructs for describing, classifying, retrieving art images. This will be done on one hand by descriptive analysis and predictive analysis and on the other hand by giving applications.

Therefore this chapter has five parts:

- *global framework and description of data sets;*
- *analysis of the visual constructs;*
- *analysis of the higher level harmonies/contrast constructs;*
- *analysis of the local constructs;*
- *applications.*

7.1 Global Framework and Description of the Data Sets

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 represents Iconographical Style from Eastern Medieval Culture (Table 1).

Table 2. 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)

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 digitised 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.

In chapter 5 we defined visual features and higher-level features concerning harmonies and contrast using Itten's theory. Descriptive and predictive analyses will show the added value of the visual and higher-level features. For the predictive analysis we use our own association rule analysis algorithm PGN. PGN is able to retrieve high confidence and low support rules. This property is of special importance when the goal is to learn characteristic patterns in a multi-class classification problem [Mitov et al, 2009a]. For comparison we use representatives of different classification models – OneR and JRip (the Weka implementation of RIPPER) from the group of Decision Rules and J48 (the Weka implementation of C 4.5) as representative of Decision Trees.

The experiments are made using 5-fold cross-validation and applying Chi-merge discretization with 95% significance level for numerical attributes. The confusion matrices of the classifiers for different datasets are given in the Appendix.

The results of the descriptive analysis are compared with domain knowledge. For the predictive analysis we use classical hold-out measures like accuracy for classification and recall and precision for retrieval.

7.2 Analysis of the Visual Features

7.2.1 Descriptive Analysis of the Visual Features

The descriptive analysis of the visual features will be done by quantifying and visualizing uni-dimensional (histograms) and two dimensional projections and by multidimensional scaling. At the pre-processing stage the pixels in the images are converted into the HSL-artist colour model. The quantization of Hue is made to 13 bins, $ih = -1, \ldots, NH - 1$, $NH = 12$, where one value is used for achromatic colours ($ih = -1$) and twelve hues are used for fundamental colours ($ih = 0, \ldots, NH - 1$). For HSL-artist colour model the quantization function is non-linear with respect to taking into account the misplacement of artists' colour wheel and Hue definition in HSL colour space. The quantization intervals are given in Figure 47.

The saturation and lightness in HSL colour model and saturation are linearly quantized into NS-bins ($is = 0, \ldots, NS - 1$), respectively NL-bins ($il = 0, \ldots, NL - 1$). We have used $NS = 10$ and $NL = 10$.

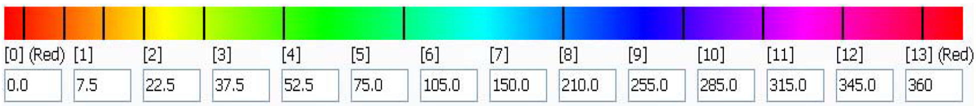


Figure 47. Quantization of Hue

➢ ***Unidimensional Analysis of the Visual Features***

Figure 48 shows the distributions and Table 3 gives percentages of hues in the art painting images for examined movements. The predominate presence of warm colours (red-orange spectrum) in paintings are due to colouring of faces and bodies from one side and using the materials and varnish, which acquired yellowish tinge from other side. Not without importance is the fact that cold hues as blue and green are non-durable under the influence of light. On the other side technical artists practices of some movements did not use the green colour – this colour was replaced by brown. For instance, experiments of Constable at the beginning of XIX century to capture the inevitable light and shade effects of the nature with using more green colours was not accepted from their colleagues [Raychev, 2005].

Table 3. Percentages of hues in the art painting images for examined movements

Hue	Red	R/O	Orange	O/Y	Yellow	Y/G	Green	G/B	Blue	B/P	Purple	P/R
ICON	7.1	18.2	30.5	29.6	4.7	1.3	0.7	0.8	1.2	0.4	0.4	4.7
Renaissance	10.3	26.7	27.8	8.2	2.7	1.0	0.8	2.5	4.4	1.6	1.2	11.5
Baroque	9.5	18.3	26.6	11.0	3.3	0.4	0.1	0.4	0.7	0.4	2.5	23.7
Romanticism	2.3	10.5	25.1	24.6	10.8	4.0	2.4	9.4	5.4	0.5	0.3	3.1
Impressionism	0.9	2.8	8.1	19.9	22.5	8.7	4.6	11.1	16.2	2.2	0.9	2.1
Cubism	3.4	7.2	21.6	24.7	12.1	6.2	4.8	6.4	5.5	1.1	0.8	4.5
Modern art	2.8	9.2	26.9	26.7	10.5	3.7	2.2	3.6	10.7	0.8	0.5	2.2

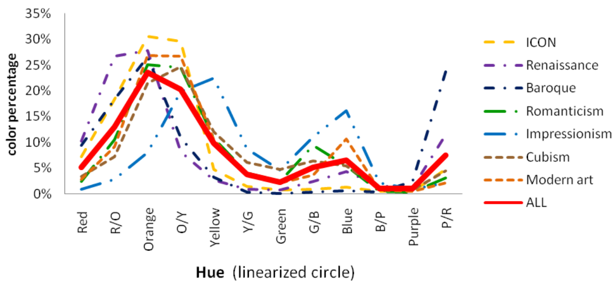


Figure 48. The Hue distribution of all pictures, grouped by movements

Table 4 summarizes percentages and Figure 49 shows the distributions of saturation for each movement. The differences among these movements are obvious. The big difference of the global trend belongs to the Eastern iconographic style, which uses canonical representation of the figures with more schematic lines and pure colours (let's remember that values of saturation near 1 encode the pure colours).

Table 4. Percentages of saturation in the art painting images for examined movements

Saturation	0.0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0
ICON	2.7	4.8	7.3	8.5	10.7	13.1	15.7	13.5	9.7	14.1
Renaissance	12.4	18.4	21.2	18.1	13.0	7.4	4.1	2.2	1.1	2.1
Baroque	4.3	12.1	19.3	15.8	12.6	10.6	8.6	6.3	4.3	5.9
Romanticism	18.1	23.5	23.1	15.6	8.5	3.7	2.6	1.4	1.2	2.3
Impressionism	10.5	16.0	15.8	14.1	13.1	10.2	7.4	5.0	3.3	4.7
Cubism	12.9	19.7	16.9	13.1	10.4	7.7	5.8	4.3	3.2	6.0
Modern art	4.7	9.3	11.6	12.2	11.9	10.8	11.0	9.0	6.3	13.2

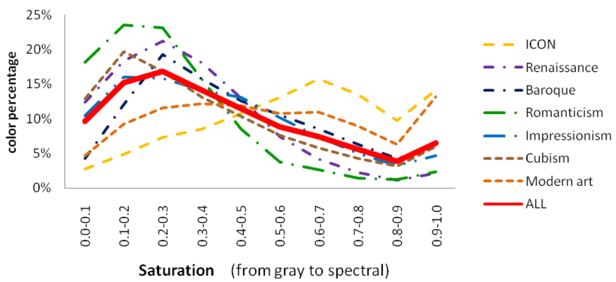


Figure 49. The Saturation distribution of all pictures, grouped by movements

Table 5 gives the percentages and Figure 50 shows the distributions of the luminance within groups determined by movements. Here Baroque highly differs with a big presence of dark colours (values of lightness near 0).

Table 5. Percentages of luminance in the art painting images for examined movements

Luminance	0.0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0
ICON	5.7	11.4	12.6	13.8	15.3	13.6	11.5	9.5	4.8	1.8
Renaissance	11.3	13.6	13.9	12.3	11.3	9.9	9.4	8.8	6.8	2.8
Baroque	39.3	29.4	11.1	6.3	4.3	3.3	2.5	2.0	1.3	0.5
Romanticism	6.6	13.8	10.5	8.6	9.6	9.5	11.1	14.2	13.4	2.7
Impressionism	1.3	5.2	11.0	14.6	16.5	15.8	14.2	11.4	6.9	2.9
Cubism	9.0	11.2	12.0	12.7	13.4	11.5	9.4	8.2	7.0	5.6
Modern art	3.3	7.6	11.0	13.9	13.7	10.9	10.0	11.1	11.1	7.4

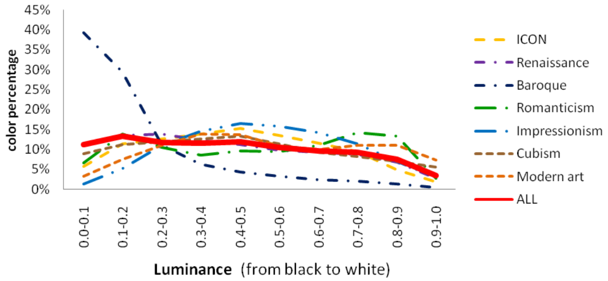


Figure 50. The Luminance distribution of all pictures, grouped by movements

From this unidimensional analysis we learn that distributions of the components are similar when grouped by movements. We do not expect huge accuracy differences between classifiers built for the separate components. Due to its specific luminance distribution we expect the best accuracy results for the Baroque movement. The movement Icons has a specific saturation distribution.

- ***Two Dimensional Projections***

APICAS allows making more complex analysis on combination of projections of the colour space. Some examples are given here. Analysis of the colour distribution on two projections – hue and luminance, has shown a predominance of dark orange colours in art paintings for Baroque. Similar predominance of warm colours, but in light tones is seen in Icons (Figure 51).

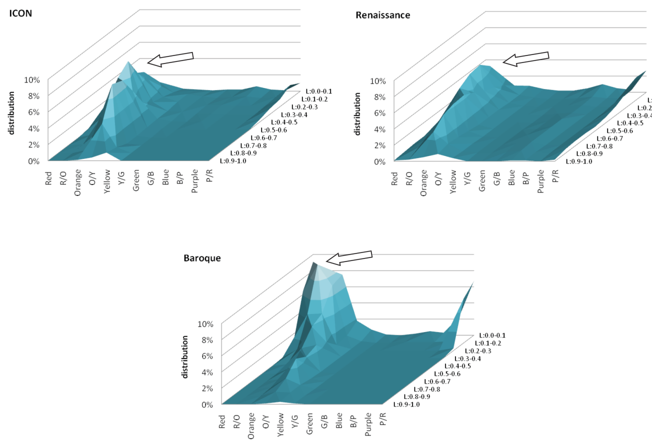


Figure 51. Hue-Luminance Distribution for Icons, Renaissance and Baroque

In more contemporary movements, such as Romanticism, Impressionism, Cubism and Modern art, blue tonality has its strongest presence – lighter in Impressionism and darker in Modern art (Figure 52).

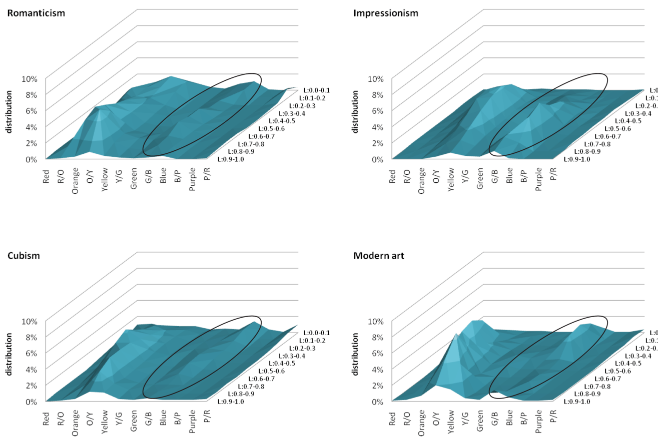
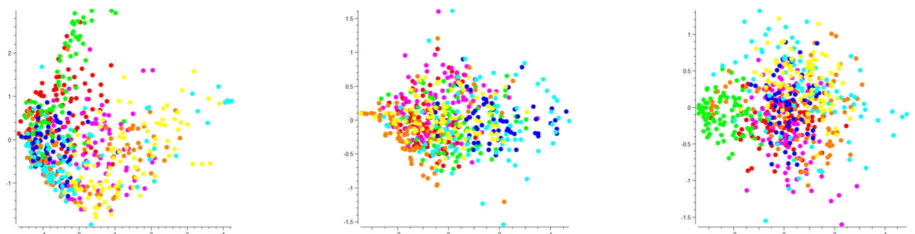


Figure 52. Hue-Luminance Distribution for Romanticism, Impressionism, Cubism and Modern art

➢ ***Multidimensional Scaling***

The results of multidimensional scaling for the distributions of hue, saturation and luminance are shown on Figure 53. The tendency of grouping by hue values of different movements, the relative separating of saturation values for Icons and Modern Art or distinctness of Baroque in luminance values can also be seen here.



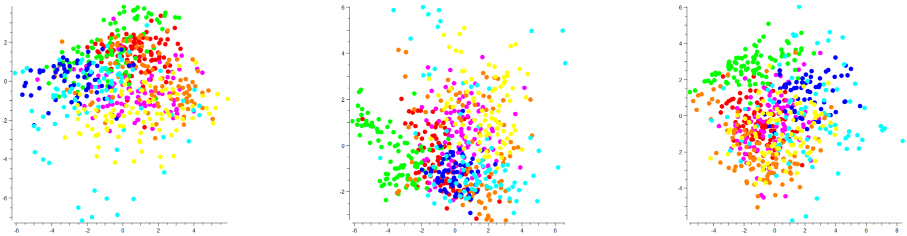
a) Hue (avg.stress 30.835) *b) Saturation (avg.stress 0.058)* *c) Luminance (avg.stress 0.062)*

- 12-Icon
- 14-Renaissance
- 17-Baroque
- 18-Romanticism
- 19-Impressionism
- 20-Cubism
- 20-ModernArt

Figure 53. Results of multidimensional scaling for colour distribution characteristics

As we can see the average stress for Hue is very big, which means that 2-dimensions are insufficient to represent the dataset. Stress values for Saturation and Luminance are agreeable.

Decreasing dimensionality using multidimensional scaling loses possibilities to make analysis of exact presence of combination of factors, but allow showing the tendencies of grouping or merging of images, belonging to different movements (Figure 54).



a) Hue-Saturation (avg.stress 0.164) *b) Hue-Luminance (avg.stress 0.168)* *c) Saturation-Luminance (avg.stress 0.127)*

● 12-Icon ● 14-Renaissance ● 17-Baroque ● 18-Romanticism
● 19-Impressionism ● 20-Cubism ● 20-ModernArt

Figure 54. Results of multidimensional scaling for colour distribution characteristics and their combinations

These results provide the evidence that using these colour characteristics as a basis for defining higher level features is reasonable. Firstly we will analyse whether these colour characteristics can classify the different movements and artists.

7.2.2 Predictive Analysis of the Visual Features

The visual features can be used to classify/retrieve movements and artists styles. We made three-fold cross validation using the datasets that contains hue values, saturation values, luminance values separately and all three together. We analyzed the results of OneR, JRip, J48, and PGN, comparing average accuracies and confusion matrices.

Table 6. Accuracy – movements

Database	OneR	JRip	J48	PGN
Hue	27.83	34.00	39.00	42.83
saturation	34.83	33.00	35.33	36.50
luminance	30.67	35.00	38.50	45.83
HSL	33.50	49.00	47.00	63.17

Table 7. Accuracy – artists' names

Database	OneR	JRip	J48	PGN
hue	15.33	24.67	24.17	29.83
saturation	18.83	17.17	22.5	24.67
luminance	17.83	26.33	26.17	32.00
HSL	18.83	36.67	37.17	49.33

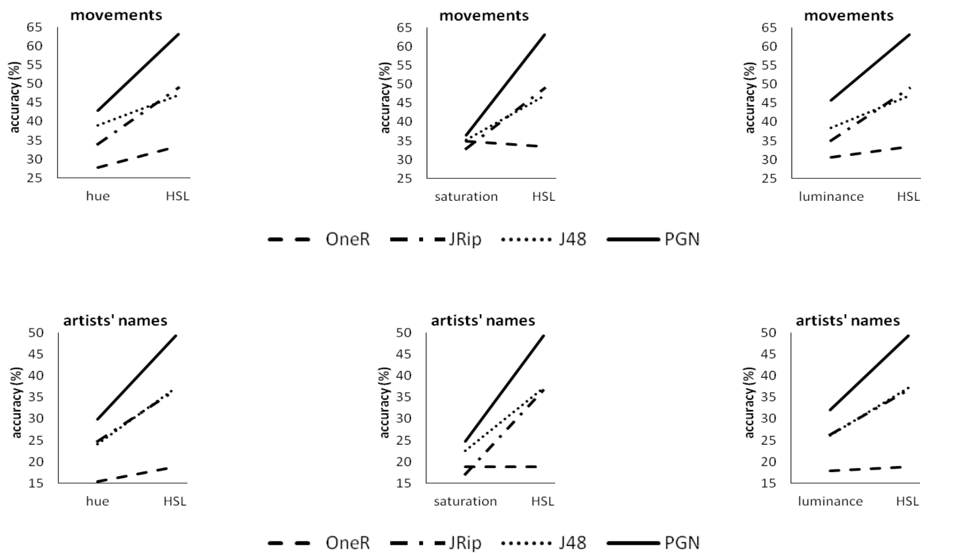


Figure 55. The accuracies of different classifiers by hue, saturation, luminance separately and all three together

Table 6 and Table 7 and Figure 55 show the accuracies by different classifiers by distribution of hue, saturation, luminance separately and all three together.

As expected the accuracies obtained by the classifiers based on one colour component are similar and we have an increase in accuracy by combining the components. The table shows however a curiosity. As we can see, examining all attributes together does not increase the accuracy of the OneR classifier for movements. In the three fold-cases for "HSL" dataset OneR choose "v0" attribute as most appropriate, but not "s7" or "s8" as in the case of "Saturation" dataset, it leads to decreasing of overall accuracy in HSL dataset than in simpler one "Saturation" dataset.

As we can see PGN shows the best accuracies from examined models for all datasets. Additionally PGN shows the best possibilities to explore specific combinations of attribute values; it achieves the biggest increase of accuracy by examining all three characteristics together.

Figure 56 and Figure 57 show the confusion matrices for movements and for artists' names respectively. In the visualization of confusion matrices, the darker a square is, the bigger is the percentage of images following into corresponded square.

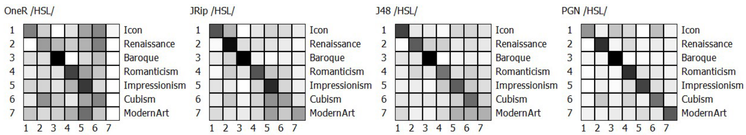


Figure 56. Confusion matrices for HSL features, movements as class labels

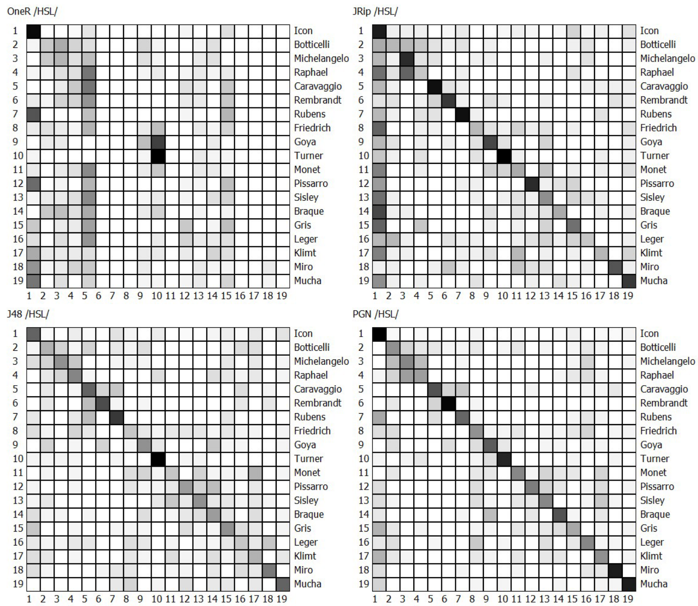


Figure 57. Confusion matrices for HSL features, artists' names as class labels

Analyzing the movements results three patterns immediately get attention. First, the Baroque movement is the easiest to predict, OneR fails to predict Modern Art, PGN is the only classifier with a smooth consistent black/gray downwards diagonal. The first pattern repeats patterns seen in the descriptive analysis. It seems that Modern Art pictures cannot be characterized with one visual attribute. The characteristic PGN rules can better discriminate than J48 rules especially between the movements Romanticism, Impressionism, Cubism and Modern Art. Let's mention again the specifics of the PGN against other classifiers. All other classifiers take into account in one on other manner the support, controversially to PGN, which focuses primarily on the confidence of the association rules and only in a later stage on the support of the rules.

Analysing the artist results the three mentioned patterns are confirmed and two new ones are seen: the presence of vertical lines (dark or light) and the presence of "movement" squares.

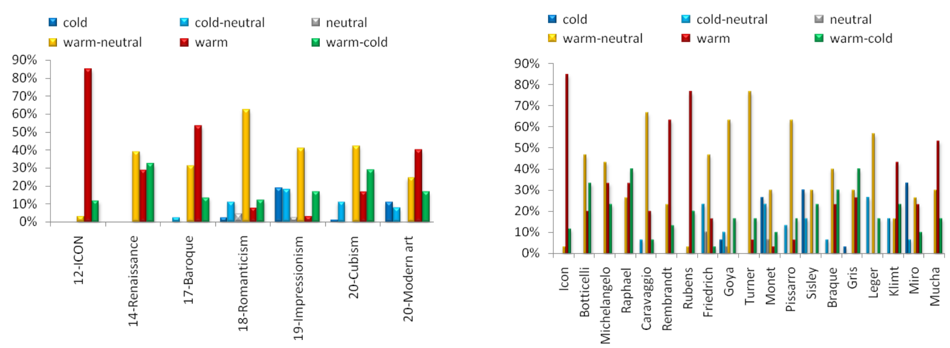
It is clear that based on visual characteristics OneR is not able to classify the different artist paintings. JRip predicts almost 25% of the paintings as Icon (the vertical line in the JRip confusion matrix). The datasets that we use here are specific because all artists are represented with equal numbers of paintings, and all selected movements contain also fixed number of artists, i.e. the distributions are equal. The exception is Icons, which are twice more than each artist and two-thirds than the movements. Because of this, we can see for the precision of Icons the tendencies of losing percentages for movements and enforcing ones for artists for OneR, JRip and J48 – here and in consequent analyses.

The grey squares show some common tendencies of recognizing or misplacing the class labels. For instance, it is interesting that the Renaissance painters Botticelli, Michelangelo and Raphael are not recognized correctly but are misclassified mainly within their own group. Icons, Michelangelo, Caravaggio, Rembrandt, Rubens, Turner, Pissarro, Miro and Mucha are easier to classify.

7.3 Analysis of the Harmonies/Contrast Descriptors

7.3.1 Descriptive Analysis of the Harmonies/Contrast Descriptors

In chapter 5 we defined higher-level features concerning harmonies and contrast using Itten's theory. Some examples of distribution of defined harmonies and contrast descriptors by movements or artists styles are presented and explained by domain knowledge.

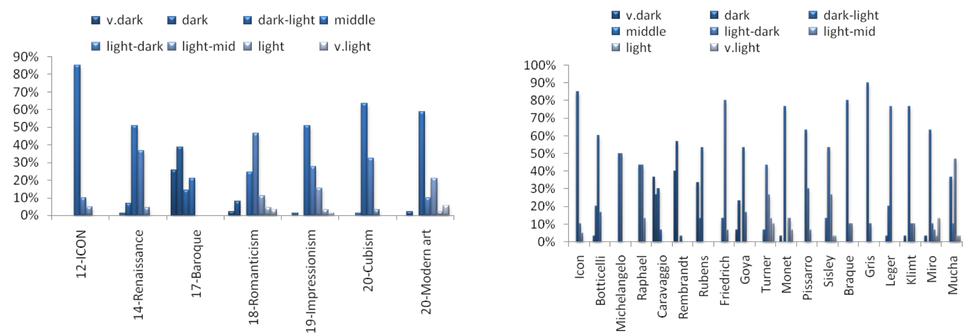


a) movements *b) artists' names*

Figure 58. Distribution of paintings, based on cold/warm contrast.

Figure 58 shows the distribution of images, based on cold/warm contrast. The high predominance of warm paintings in Icon style can be explained with the Orthodox tradition for using gold paints as well as red colour, which is main symbol of sacrificing and martyrdom. The big presence of dark warm colours is specific for the Baroque. Presenting the nature in paintings is typical for the Romanticism, which leads to forcing the presence of cold (green and blue) tones. This tendency increases in the Impressionism. Intensive study of nature led the Impressionists to an entirely new colour rendition. Study of sunlight, which alters the local tones of natural objects, and study of light in the atmospheric world of landscape, provided the Impressionist painters with new essential patterns [Itten, 1961].

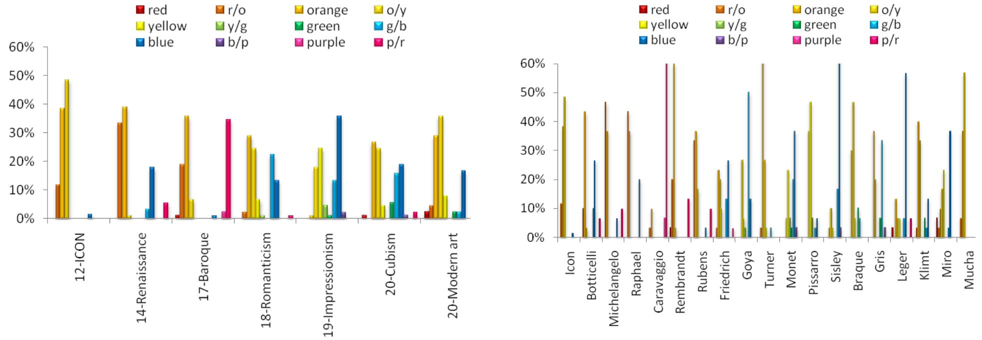
Figure 59 shows the distribution of lightness in paintings from different movements and artists. The big presence of dark colours and dark-light contrast is typical for Baroque. This is connected with using the techniques of oil-paints, which gives very deep dark effects in the paintings from one side and with typical using of light-dark contrast in this movement. This fact is connected not only with searching of maximal expression with applying this tool in the paintings, but also with the practice of this epoch to paint in the candle lights in studios [Raychev, 2005].



a) movements *b) artists' names*

Figure 59. Distribution of paintings, based on light/dark contrast.

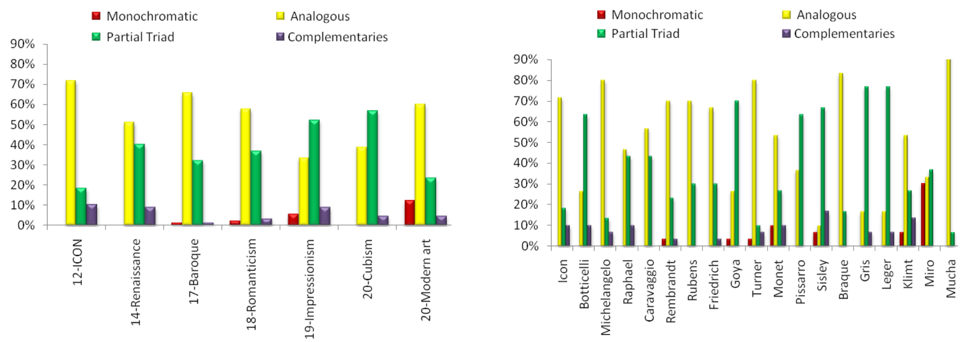
Figure 60 shows the distribution of images, based on the first dominant hue. As we have observed in our work [Ivanova et al, 2008] the colours around orange are frequently dominant colours in the paintings in classic art. More modern movements tend to use different colours as dominant.



a) movements *b) artists' names*

Figure 60. Distribution of paintings, based on first dominant hue

Figure 61 shows the distribution of hue contrasts. As we can see partial triads are used in multiple cases of natural paintings, for instance Pissarro and Sisley. Instead of high abstractionism of Cubism such colour combinations are used often in the works of Gris and Leger. The triads exist in paintings with scene presentation from authors, which techniques are based mainly on hue contrasts, such as Botticelli and Goya. Monochromaticity and analogous harmonies are presented in artworks of painters, where other key expressions are used, for instance light-dark contrast in Baroque artists, gradient expressions in Braque style, Miro's abstract paintings, etc. [Koenig, 2010].



a) movements *b) artists' names*

Figure 61. Percentage of different hue contrasts in the paintings

Generally we can say that there are some tendencies for using some colour combinations in different movements or artists.

7.3.2 Predictive Analysis of the Harmonies/Contrast Features

The harmonies/contrast features try to extract very global colour combinations constructs. As we have already mentioned, there are numerous reasons that influence the choice of colour combinations – the thematic of the painting, fashion style, philosophical visions of the painter, its current emotional condition, etc. Because of this we do not expect that such features can be used for exact classification of movements or artists. We put these descriptors into the classification task in order to see whether there are some tendencies.

Table 8 and Figure 62 show the accuracies of different classifiers, based on harmonies/contrast descriptors with movements and artists' names as class labels. Taking into account the facts that the features are too global and the numbers of class labels are great, we receive acceptable results. Here the best classification model is J48 following by PGN. Considering the movements results we see that OneR now fails to predict Icons but is able to predict Modern Art. Impressionism is classified well due to the frequent use of partial triads in natural paintings.

Table 8. Accuracy of different classifiers – based of harmonies/contrast descriptors

Database	**OneR**	**Jrip**	**J48**	**PGN**
movements	27.83	36.67	45.00	41.67
artists' names	15.50	21.33	30.17	29.83

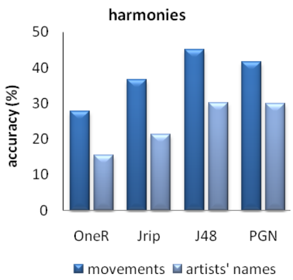


Figure 62. Accuracy of different classifiers – based on harmonies/contrast descriptors

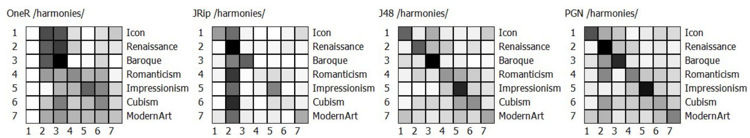


Figure 63. Confusion matrices for harmonies/contrast features, movements as class labels

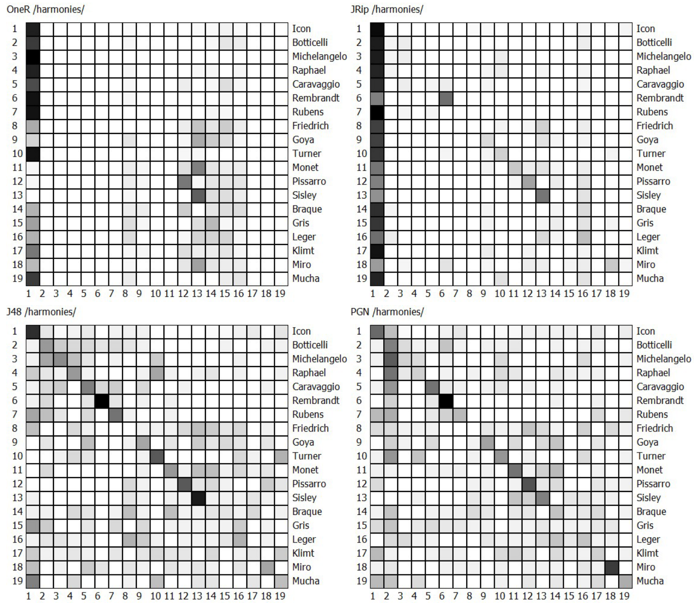


Figure 64. Confusion matrices for harmonies/contrast features, artists' names as class labels

The more detailed analysis on confusion matrices, presented in Figure 63 and Figure 64, shows also that in this case rule-based classifiers OneR and JRip do not produce good classification models, creating rules that overuse Renaissance in the case of movements, as well as Icons in the case of artists' names. In general we can say that the harmonies/contrast features are too global and not the best choice for classifying artist's paintings. The paintings of Rembrandt and Sisley are exceptions of this general rule.

7.4 Analysis of the Local Features

The first step is the development of a system that optimal uses the MPEG7 local features, extracted by proposed method (chapter 5), in respect of: the type of descriptors; position of tiles; number of clusters. MPEG-7 descriptors are complex descriptors. If we use these descriptors as described in chapter 5, we will obtain a vector with more than 300 attributes. Local features can capture more detailed information that can be useful for characterizing the artist's styles and movements, but it also introduces redundancy. This redundancy causes computational problems and can degenerate the results of the classifiers. The research challenge is to become local sensitive without significant loss of accuracy in the classification and retrieval tasks.

7.4.1 Choice of Evaluation Function for Significance of Attributes

We have processed the datasets under the procedures of attribute selection in order to receive the order of significance of attributes for prediction. We have implemented the Chi-square evaluation method. As datasets we have used 3×3, 4×4, 5×5, 6×6 and 7×7 tiling and different numbers of clusters – 20, 40 and 60. 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.

7.4.2 Selection of MPEG7 Descriptors

Figure 65 shows the distribution of significance of MPEG-7 descriptors for class prediction. As it is shown, the *Colour Structure* (*CS*) descriptor is the most informative for our datasets.

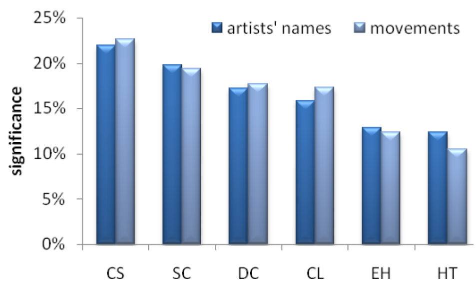


Figure 65. Average distribution of significance by MPEG-7 descriptors over datasets, using different tiling and clustering, using Chi-square evaluation method

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

- The artist's palettes, which are captured in colour 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.

7.4.3 Optimize Spatial Granularity

We have made the analysis of the significance of the 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 66 shows the distribution 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 [Arnheim, 1974]. We intend to use this fact in further investigation with analyzing the tiles only of the right half of the image.

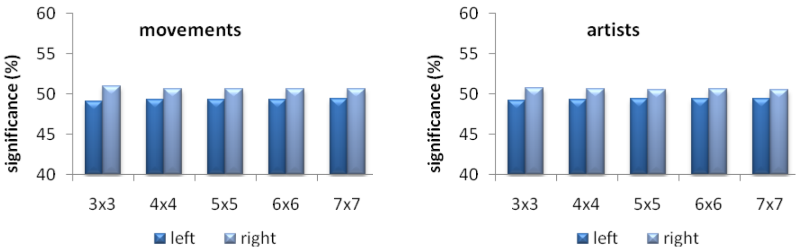


Figure 66. Distribution of significance
of left side and right side of the images with different tiling.

Figure 67 shows the distribution 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.

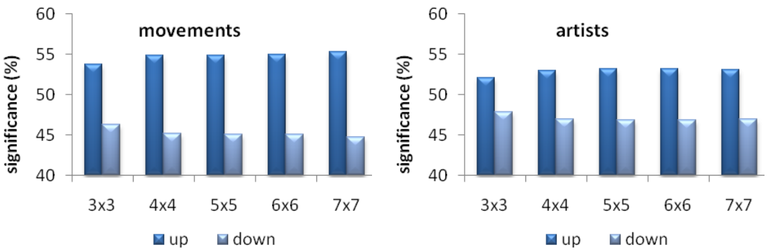
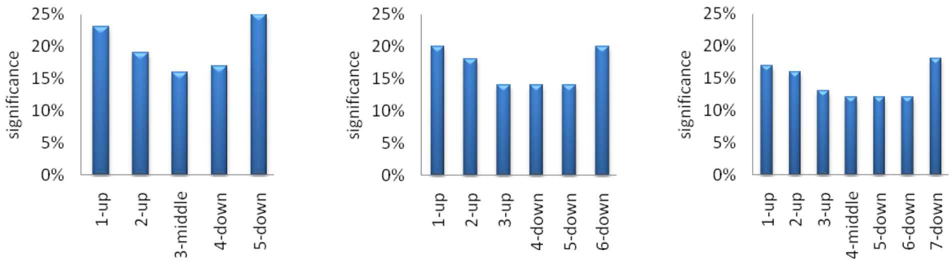


Figure 67. Distribution of significance
of upper and lower zone of the images with different tiling.

Similar analysis in respect to the vertical position of the tiles is shown on Figure 68. 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).

This fact can also be explained with differences in the composition in different styles [Arnheim, 1974]. 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' palette and brushwork.



(a) $n = 5$ *(b)* $n = 6$ *(c)* $n = 7$

Figure 68. Distribution of significance of the tiles by position of height, $j \in 1...n$ *(up to down)*

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.

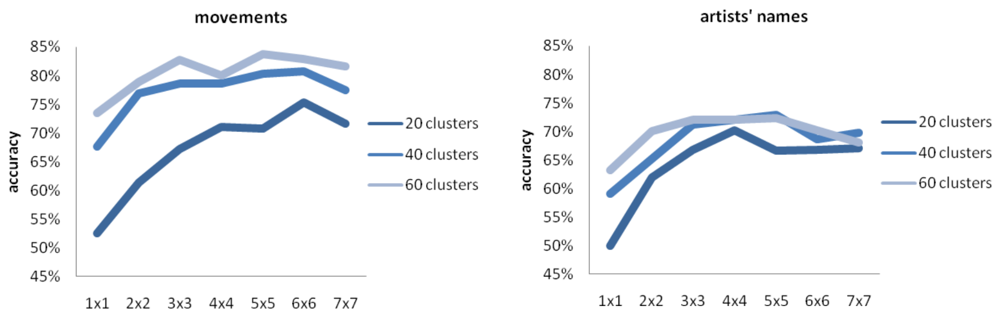


Figure 69. Classification accuracy for datasets using 20, 40, 60 clusters and different numbers of tiling

The results displayed in Figure 69, 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.

7.4.4 Predictive Analysis of the VQ-MPEG Descriptors

Table 9 and Figure 70 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 9. Accuracy 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

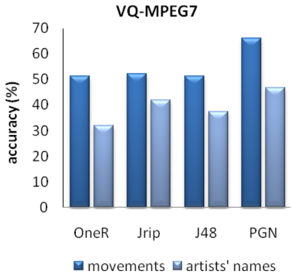


Figure 70. Accuracy of different classifiers using VQ-MPEG7 descriptors

Figure 71 and Figure 72 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 analysing 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 favour 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.

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

The grey squares in Figure 72 show some common tendencies of recognizing or misplacing the class labels. Again, it is interesting to see that the Renaissance painters Botticelli, Michelangelo and Raphael are not recognized correctly but are misclassified mainly within their own group.

Figure 73 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.

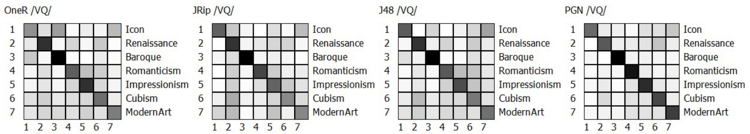


Figure 71. Confusion matrices for VQ-MPEG7 descriptors, movements as class labels

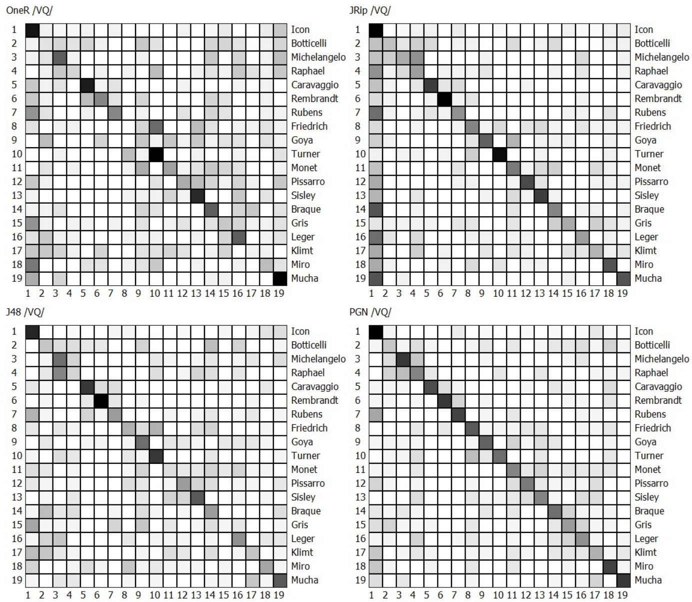


Figure 72. Confusion matrices for VQ-MPEG7 descriptors, artists' names as class labels

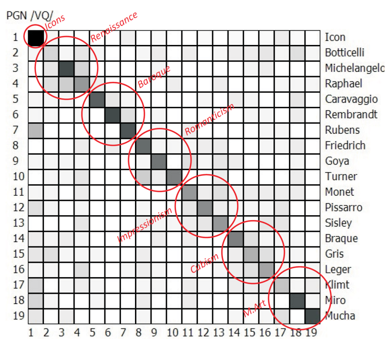


Figure 73. Visualisation of confusion matrix of PGN
for VQ-MPEG7 descriptors, artists' names as class label with grouping by movements.

On the base of the confusion matrix for VQ-MPEG7 descriptors, given in the Appendix, Recall, Precision and F-measure are calculated as follows:

– Recall is the number of correct answers over the number of the images, belonging to the corresponded class (diagonal value over the sum by row);
– Precision is the number of correct answers over the total number of the images assigned to this class by classifier (diagonal value over the sum by column);
– F-measure is calculated as a harmonic mean of precision and recall: $F = 2 * \dfrac{precision * recall}{precision + recall}$

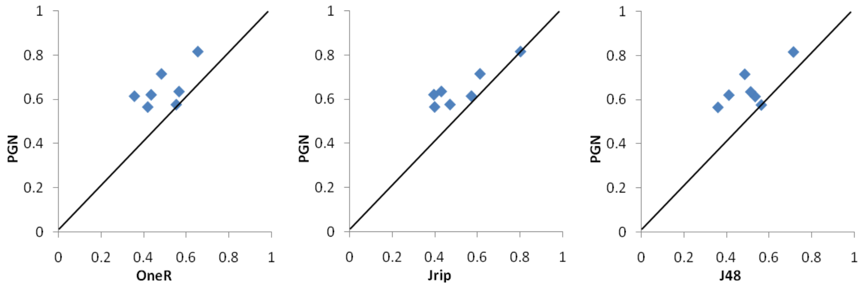


Figure 74. Scatterplots of F-measure of PGN against other classifiers for VQ-MPEG7 dataset – movements

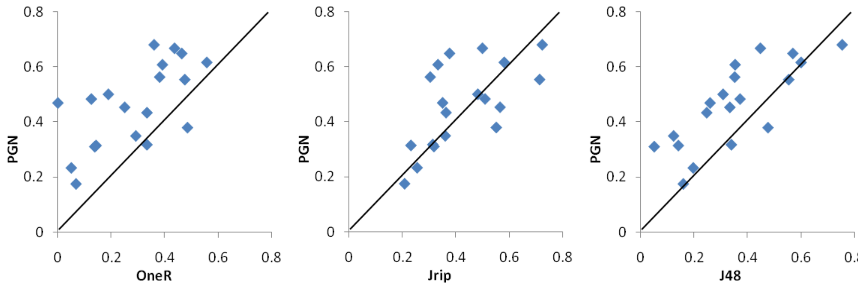


Figure 75. Scatterplots of F-measure of PGN against other classifiers for VQ-MPEG7 dataset – artists' names

The scatterplot of F-measures of PGN against other classifiers (Figure 74 and Figure 75) shows that PGN has not only the best accuracy, but has also better local behaviour 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.

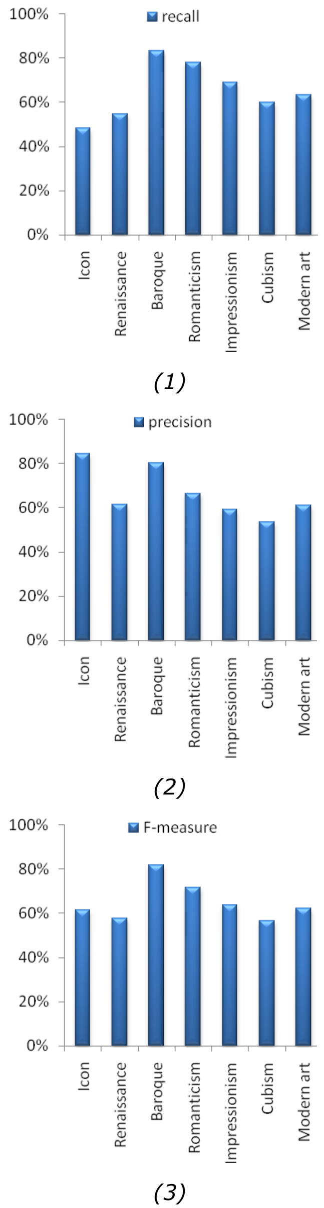


Figure 76. (1) Recall, (2) Precision and (3) F-measure for the examining set, classes "movements"

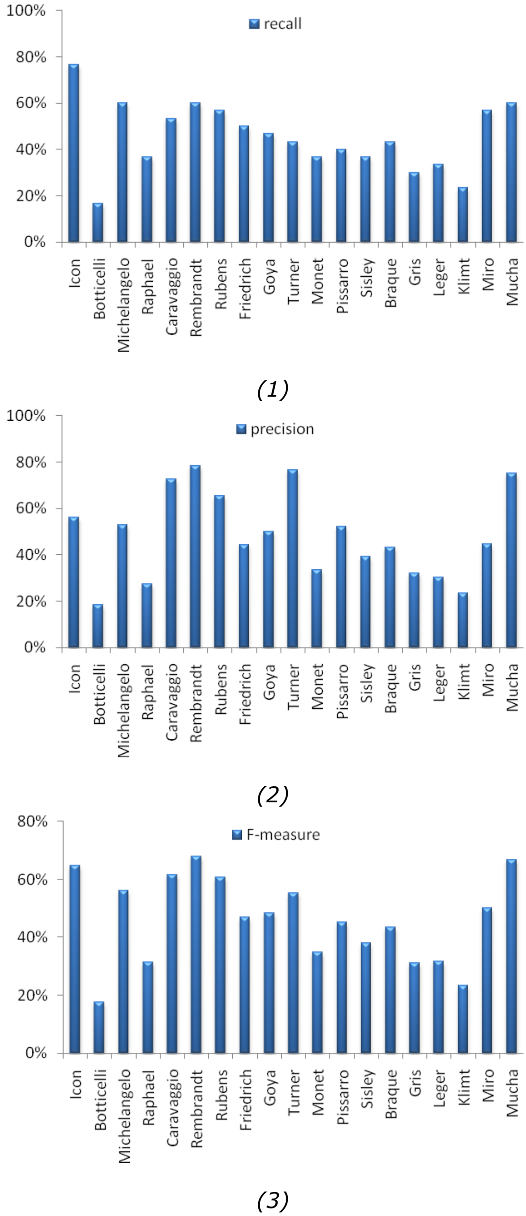


Figure 77. (1) Recall, (2) Precision and (3) F-measure for the examining set, classes "artists' names"

The results presented in Figure 76 show that the measures for some movements (for instance, Baroque) are very good. The nature of a part of paintings from Romanticism, presenting landscapes [Clark, 1969], 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 77 shows 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.

7.5 Applications

In the previous experiments we analysed the paintings of movements and the paintings of different artists. The tools developed in APICAS can also be used to profile or contrast artist or to analyze different periods of a particular artist.

7.5.1 Example of the Variety of Specifics

Itten argued that each artist has his own private conception of colour harmony [Itten, 1961]. This fact does not deny existing of objective rules for constructing colour harmonies and contrasts, but each person has individual subjective opinion, or so cold subjective colour. Among painters, Itten perceived three different attitudes towards problems of colour.

- First there are the "epigones", having no colouration of their own but composing after the manner of their teachers or other exemplars;
- The second group is that of the "originals", those who paint as they themselves are. They compose according to their subjective timbre. Though the theme changes, the chromatic expression of their paintings remains the same;
- The third group is that of the "universalists", artists who compose from inclusive, objective considerations. Each of their compositions, according to the subject to be developed, has a different colour treatment. That there should be but few painters in this group is understandable, for their subjective timbre must comprehend the entire colour circle, and this happens rarely.

In [Ivanova et al, 2008] we have presented two histograms for Hue distribution of the paintings of Titian and Morisot, which are the artists with minimum and maximum standard deviation for the vectors based on the Hue characteristic. The quantization of colour hue was based on HSL colour model. In order to unify the exposition here we have remade these graphics in HSL-artist colour model.

Figure 78 presents the distribution of hue in Titian's paintings and Figure 79 presents hue distribution of the paintings of Morisot. The idea is to verify the common belief that if some characteristic is relatively stable (measure of stability is reciprocal to the numeric value of the standard deviation) for some authors, for others it completely does not matter. In the case, Titian uses almost the same colour relationships in his paintings, while in the pictures of Morisot there are significantly different relationships between the colours.

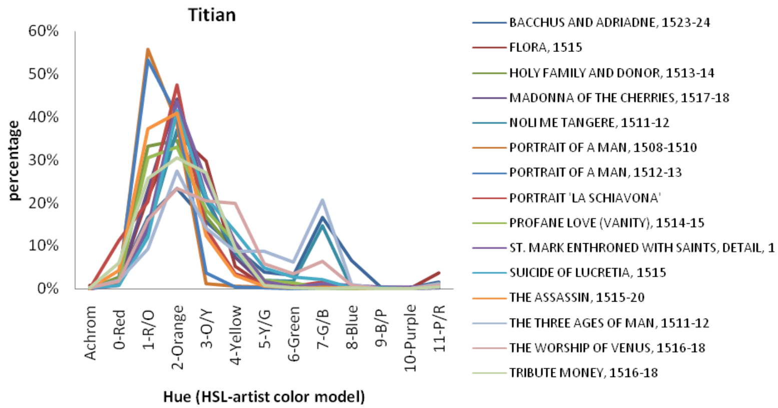


Figure 78. The Hue Distributions of paintings of Titian (minimal standard deviation=0.1654)

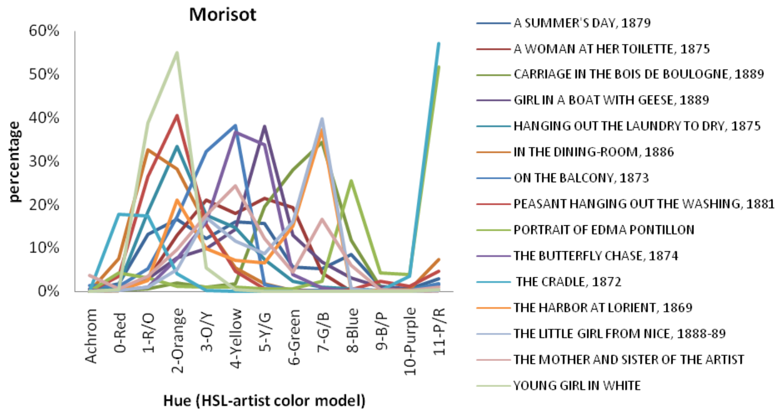


Figure 79. The Hue Distributions of paintings of Morisot (maximal standard deviation=0.4061)

7.5.2 Analysis of Different Periods of Goya's Paintings

During his life one artist can pass different stages of professional growth and usually the creative work of every artist is divided into periods. A typical example is Picasso, who changed his style completely several times.

Here we choose to make an analysis of the Goya's paintings. Francisco José de Goya y Lucientes (30 March 1746 – 16 April 1828) was a Spanish Romantic painter and printmaker regarded both as the last of the Old Masters and as the first of the moderns. During his creative work he also has quite different periods. We will examine three of them.

Between 1775 and 1791 Goya designed tapestry cartoons, used to decorate the walls of El Escorial and the Palacio Real del Pardo, the newly built residences of the Spanish monarchs near Madrid. He found the format to be limiting though, because being inherently matte tapestry was unable to capture complex colour shift or texture, and was unsuited to the impasto and glazing techniques he was applying to his painted works.

From 1783 Goya was appointed the court painter to the Spanish Crown, and through his works were both a commentator on and chronicler of his era. He made numerous portraits of Charles IV, Fernando VII, royal family pictures and many other nobles.

In later life Goya bought a house, called *Quinta del Sordo* ("Deaf Man's House"), and during the period 1819–1823 painted many unusual paintings on canvas and on the walls, including references to witchcraft and war. This set of paintings, received the name The Black Paintings (Pinturas Negras). They portray intense, haunting themes, reflective of the artist's fear of insanity, and his outlook on humanity.

We have obtained from the open access virtual gallery of Prado museum, which has the riches collection of Goya's paintings, digitised artworks that present these three main branches of Goya's creative work.

We have analyzed:

- 60 Cartoons;
- 15 Pinturas Negras (there are no more such paintings);
- 30 Royal Family Portraits.

Results of Multidimensional scaling for Goya's paintings, divided into three groups – Cartoons (blue), Pinturas Negras (red), Royal Family Portraits (green) is presented in Figure 80. It is seen that Luminance brings the biggest separation between cartoons and two others. In parallel, difference between Pinturas Negras and Royal Family Portraits is more distinctive in Hue. As we can see from the Figure 80 there are tendencies of division between three groups – maybe most distinctive by the Luminance, after that by the Hue. These results give us assurance to continue analysis with higher level features, which are based on visual ones.

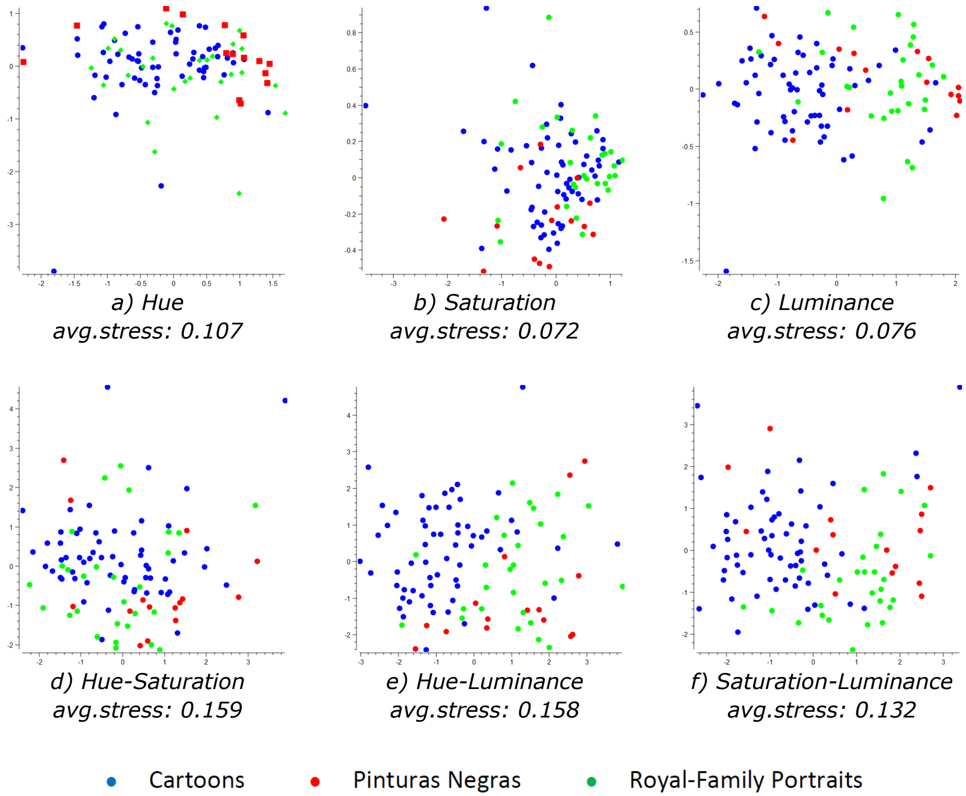


Figure 80. MDS Results for three Goya's paintings periods

We calculated harmonies/contrast descriptors using HSL-artists colour wheel and fuzzy quantization of colour characteristics.

Table 10. Accuracy, received from PGN for different amount of the learning set for dataset with harmonies/contrast attributes

Database	Number of instances in Learning Set	Number of patterns	Accuracy (%)
1:4	20.6	16.6	68.93
1:3	25.8	18.8	69.90
1:2	34.3	22.0	71.84
1:1	51.5	27.0	75.73
2:1	68.7	34.3	77.67
3:1	77.3	36.3	78.40
4:1	82.4	38.4	79.80

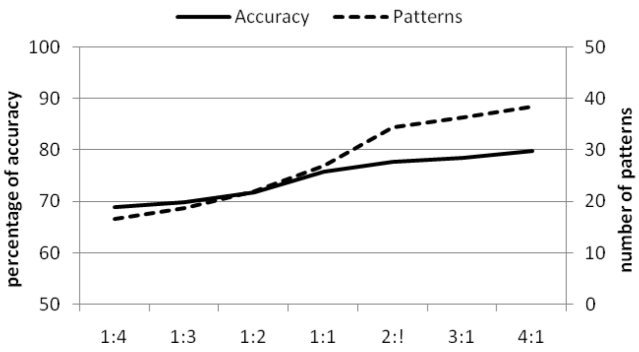


Figure 81. Goya's paintings – average number of patterns and accuracy

Different numbers of folds allow to test different numbers of examples in the learning set as well as to average the results.

As we can see, the dataset that describes Goya's paintings using only harmonies/contrast attributes defines relatively well different kinds of paintings and creates low number of pattern in the recognition set. Although the harmonies/contrast attributes were considered as too global for discriminating within movements and artists, this application shows the added value of these attributes for discriminating between different periods of an artist.

Conclusion

We collected an experimental dataset that includes 600 paintings of 18 representative artists from different movements of West-European fine arts (Renaissance, Baroque, Romanticism, Impressionism, Cubism, and Modern Art) and Eastern Medieval Icons.

Using the possibilities of APICAS for extracting examined features and data mining tools from PaGaNe, Orange, and Weka we made several kinds of descriptive and predictive analysis:

- *The first part of experiments is focused on the analysis of the colour distribution characteristics. From one side, we find some specifics of given movements, which distinct them from the others. From other side, the examination of the features for all paintings revealed some common trends for art images. We have used later these results as a normalization factor in the process of defining the colour harmonies and contrast descriptors;*
- *The second part of experiments is focused on results of automatic annotation of the images with harmonies' and contrasts' descriptors. These high level features can be used not only in the processes of categorization where cultural influences and specific techniques are*

revealed, but also by extracting the images with given characteristics, which are closely connected with the emotional responses evoked by an image. These features refer to corresponded elements of the abstract space of image content;

- *The third group of experiments is focused on examination of the the significance of local features, extracted by the proposed method, in respect of: the type of underlying MPEG-7 descriptors; position of tiles; number of clusters. The goal was to establish additional decreasing of examined features in order to accelerate the process without significant losing of accuracy in the classification tasks.*
- *Special group of experiments are focused on the evaluation of the classification accuracy by different types of used classifiers. The goal was to confirm the proposition that PGN classifier has good representation for such kind of features compared with the other well known classifiers. This give the opportunity to use the rules, produced by PGN classifier as a profile of given class, such as movement, artist, artist period, genre, scene type, etc.;*
- *As example of this proposition, the last group of experiments is focused to show the results of classification by PGN on the base of painting periods in Goya's creative works.*

8 Conclusions and Future Work

The main goals of this work were:

- to provide a detailed analysis of the colour theories, especially on existing interconnections in successful colour combinations;
- to formalize them in order to implement automated extraction from digitized artworks.

The extracted features were successfully used for:

- similarity search with selected image by one or more of the extracted features;
- search of images that satisfied user queries featuring contrasts' characteristics;
- investigation on the possibilities to integrate such characteristics within specialized resource discovery (searching for distinctive feature of movements, artists or artists' periods).

In order to achieve these goals, several tasks were addressed:

1. An overview of the area of resource discovery in the case of art images had been made. Different kinds of existing gaps between extraction of low-level features by the computer and user expectation for satisfying higher objective (semantic) and subjective (aesthetic and emotional) expectations. Analysis of existing taxonomies of image content as extracted by the viewer of an image was made.
2. An overview of content-based image retrieval was made addressing: process description; different methods of feature extraction; various types of similarity measures used in the process of image retrieval; as well as techniques for improving the quality of image retrieval.
3. A brief review of existing colour theories from different points of view (physiological, psychological, social) was made. A succinct historical overview of attempts to find colour interconnections and mutual influences of colours was presented. A colour model, which combines the benefits of three others was constructed.

4. An overview of already existing art image analysis systems and a study of the connection of the reviewed systems with the taxonomy of art image content had been made.
5. Visual low-level features, which represent colour distribution in art images, were proposed to be used as a ground for constructing higher-level concepts.
6. The classification of harmonies and contrasts from the point of view of three main characteristics of the colour – hue, saturation and luminance, was made. A formal description of defined harmonies and contrasts was established.
7. A method for extracting local features based on tiling the image and applying vector quantization of MPEG-7 descriptors over the tiles, has been described and implemented.
8. Architecture for experimental CBIR-system is presented. A program system APICAS ("Art Painting Image Colour Aesthetics and Semantics") was developed in order to supply an appropriate environment for realizing proposed algorithms and for conducting experiments.
9. An experimental base including 600 paintings of 18 representative artists from Renaissance, Baroque, Romanticism, Impressionism, Cubism, and Modern Art and one group of images from Eastern Iconographical Style was chosen.
10. Experiments were made to evaluate the features' added value. The descriptive analysis shows the common trends and specifics of examined features projections. The predictive analysis with classification, especially tree classifiers and associative classifiers, shows the benefits of using such features within the recognition process of artists' styles, movements or groups of movements.

Main contributions can be summarized as:

1. A new **methodological approach** to integration of colours' characteristics for CBIR within digital art repositories.
2. Development of a **colour model** appropriate for contrast characteristics extraction as a combination of three other models; the model is innovative because it is easy for human comprehension while also allows for efficient conversion from RGB.
3. A **formal description** of harmonies and contrasts from the point of view of three main characteristics of the colour – hue, saturation and luminance.
4. An **architecture and implementation** of experimental CBIR-system. The implementation includes a mixture of newly developed and existing open source components.

5. Thorough **experiments** on the use of the system for different tasks (similarity search, user queries, predictive analysis).

The near plans for further research are focused on:

- Analyzing the possibilities of using SIFT-descriptors [Lowe, 1999] as a ground for defining upper-layer concepts;
- Focusing on the processes of throwing out redundant attributes in order to achieve more clear and faster results;
- Applying already extracted as well as new developed attributes and corresponding methods in the field of analysis Eastern Iconographical painting schools (especially Bulgarian tradition) and themes within the icons.

As main conclusion:

We proposed, developed and analysed a variety of different types of visual and upper level features, which strive to narrow the semantic and abstraction gap between low-level automatic visual extraction and high-level human expression. We are convinced that the vividness of proposed features will open the door for indexing and searching in paintings repositories, according to such characteristics of their content and can be used as a step in the transition from Web 2.0 to Web 3.0.

Appendix

The results which had been discussed in the experimental chapter are presented here. The tables contain summaries of the confusion matrices received from the 5-fold cross-validation of each classifier across the examined dataset.

Table 11. Confusion matrices of different classifiers using low level HSL descriptors – class "movements"

OneR /HSL/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	18	7	0	4	13	17	1	60
Renaissance	0	21	15	11	16	27	0	90
Baroque	7	7	54	0	8	14	0	90
Romanticism	4	8	6	40	20	12	0	90
Impressionism	9	9	6	16	42	8	0	90
Cubism	5	23	12	5	19	26	0	90
Modern art	12	8	10	17	31	12	0	90
	55	83	103	93	149	116	1	600
JRip /HSL/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	28	12	0	4	7	6	3	60
Renaissance	1	60	8	8	7	5	1	90
Baroque	5	11	63	3	1	3	4	90
Romanticism	1	12	7	42	14	11	3	90
Impressionism	7	5	0	11	53	10	4	90
Cubism	5	13	3	11	27	24	7	90
Modern art	9	6	5	9	24	13	24	90
	56	119	86	88	133	72	46	600
J48 /HSL/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	33	4	3	1	7	7	5	60

Renaissance	1	43	12	13	6	9	6	90
Baroque	7	7	67	0	3	4	2	90
Romanticism	5	11	7	37	9	13	8	90
Impressionism	4	5	2	14	42	14	9	90
Cubism	7	6	7	6	13	39	12	90
Modern art	10	4	3	10	21	21	21	90
	67	80	101	81	101	107	63	600
PGN /HSL/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	22	3	12	8	0	13	2	60
Renaissance	0	63	3	11	3	6	4	90
Baroque	3	3	78	1	1	2	2	90
Romanticism	1	5	2	63	9	7	3	90
Impressionism	0	5	1	10	58	10	6	90
Cubism	2	16	5	13	5	44	5	90
Modern art	4	4	6	6	11	8	51	90
	32	99	107	112	87	90	73	600

Table 12. Confusion matrices of different classifiers using low level Hue descriptors – class "movements"

OneR /Hue/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	7	15	24	2	4	1	7	60
Renaissance	7	52	18	5	2	3	3	90
Baroque	12	33	19	2	5	8	11	90
Romanticism	4	14	16	4	27	8	17	90
Impressionism	1	2	5	3	52	5	22	90
Cubism	4	7	16	2	29	10	22	90
Modern art	3	8	12	3	32	9	23	90
	38	131	110	21	151	44	105	600
JRip /Hue/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	11	4	8	16	11	3	7	60
Renaissance	0	35	14	12	17	2	10	90
Baroque	1	19	39	8	12	1	10	90
Romanticism	1	5	5	20	33	4	22	90
Impressionism	0	3	7	13	53	7	7	90
Cubism	2	5	8	17	32	11	15	90
Modern art	3	2	8	11	28	3	35	90
	18	73	89	97	186	31	106	600
J48 /Hue/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row

Icon	19	11	9	9	0	7	5	60
Renaissance	5	54	17	4	3	6	1	90
Baroque	5	29	37	8	0	5	6	90
Romanticism	5	15	10	19	17	16	8	90
Impressionism	5	1	0	10	45	18	11	90
Cubism	7	9	8	10	16	27	13	90
Modern art	3	9	8	12	13	12	33	90
	49	128	89	72	94	91	77	600
PGN /Hue/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	14	5	12	9	2	11	7	60
Renaissance	4	52	19	3	2	9	1	90
Baroque	10	14	51	3	2	3	7	90
Romanticism	6	10	4	16	24	14	16	90
Impressionism	2	2	0	8	55	18	5	90
Cubism	2	9	4	14	13	35	13	90
Modern art	4	7	13	6	14	12	34	90
	42	99	103	59	112	102	83	600

Table 13. Confusion matrices of different classifiers using low level Saturation descriptors – class "movements"

OneR /Sat/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	43	1	5	0	2	8	1	60
Renaissance	0	41	19	14	12	4	0	90
Baroque	19	13	41	0	9	8	0	90
Romanticism	5	13	4	63	2	3	0	90
Impressionism	16	21	23	10	11	9	0	90
Cubism	11	24	21	9	15	10	0	90
Modern art	30	18	18	10	7	7	0	90
	124	131	131	106	58	49	1	600
JRip /Sat/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	18	0	3	0	5	28	6	60
Renaissance	1	26	2	10	23	25	3	90
Baroque	6	8	29	1	16	28	2	90
Romanticism	0	9	5	52	8	14	2	90
Impressionism	8	11	3	9	25	33	1	90
Cubism	2	7	5	8	22	43	3	90
Modern art	11	6	13	8	13	34	5	90
	46	67	60	88	112	205	22	600

J48 /Sat/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	29	3	12	0	1	6	9	60
Renaissance	1	39	4	9	12	15	10	90
Baroque	17	7	29	4	11	15	7	90
Romanticism	2	16	3	55	4	7	3	90
Impressionism	7	19	13	10	20	15	6	90
Cubism	6	22	12	8	8	26	8	90
Modern art	22	15	15	10	4	10	14	90
	84	121	88	96	60	94	57	600
PGN /Sat/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	20	0	12	1	4	6	17	60
Renaissance	0	34	8	13	16	13	6	90
Baroque	8	4	38	6	12	15	7	90
Romanticism	2	12	4	47	5	13	7	90
Impressionism	10	14	10	9	24	15	8	90
Cubism	4	14	15	10	12	27	8	90
Modern art	16	10	10	8	11	6	29	90
	60	88	97	94	84	95	82	600

Table 14. Confusion matrices of different classifiers using low level Luminance descriptors – class "movements"

OneR /Lum/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	0	21	1	7	14	17	0	60
Renaissance	0	21	13	9	10	36	1	90
Baroque	0	11	62	0	3	12	2	90
Romanticism	0	15	11	20	25	17	2	90
Impressionism	0	3	1	25	46	5	10	90
Cubism	0	32	5	4	18	25	6	90
Modern art	0	14	1	22	32	11	10	90
	0	117	94	87	148	123	31	600
JRip /Lum/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	1	9	10	2	24	4	10	60
Renaissance	0	23	20	5	21	8	13	90
Baroque	0	3	80	2	4	0	1	90
Romanticism	0	5	26	26	20	2	11	90
Impressionism	1	6	19	6	38	3	17	90
Cubism	1	9	26	4	24	7	19	90
Modern art	1	3	22	4	22	3	35	90

	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
	4	58	203	49	153	27	106	600
J48 /Lum/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	10	16	1	3	11	14	5	60
Renaissance	8	30	15	9	10	14	4	90
Baroque	2	7	71	2	2	3	3	90
Romanticism	11	10	12	24	16	12	5	90
Impressionism	8	8	1	9	32	12	20	90
Cubism	7	26	2	8	6	34	7	90
Modern art	4	10	4	11	19	12	30	90
	50	107	106	66	96	101	74	600
PGN /Lum/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	6	18	2	7	13	11	3	60
Renaissance	8	38	10	13	5	15	1	90
Baroque	0	7	77	2	1	1	2	90
Romanticism	2	7	8	44	13	11	5	90
Impressionism	9	8	2	9	30	16	16	90
Cubism	4	21	5	13	8	29	10	90
Modern art	1	7	2	7	14	8	51	90
	30	106	106	95	84	91	88	600

Table 15. Confusion matrices of different classifiers using Harmonies/Contrasts descriptors – class "movements"

OneR /harmonies/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	0	23	22	4	0	4	7	60
Renaissance	0	31	37	10	1	6	5	90
Baroque	0	32	49	0	1	8	0	90
Romanticism	0	19	17	22	14	16	2	90
Impressionism	0	1	14	14	30	23	8	90
Cubism	0	11	25	10	13	23	8	90
Modern art	0	16	20	14	13	15	12	90
	0	133	184	74	72	95	42	600
JRip /harmonies/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	20	29	2	0	0	2	7	60
Renaissance	1	76	2	2	2	5	2	90
Baroque	4	38	47	0	1	0	0	90
Romanticism	3	59	1	7	11	7	2	90
Impressionism	0	45	1	3	37	1	3	90
Cubism	1	63	4	0	10	8	4	90

Modern art	7	44	5	3	5	1	25	90
	36	354	62	15	66	24	43	600
J48 /harmonies/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	27	3	14	3	2	2	9	60
Renaissance	3	42	18	14	3	8	2	90
Baroque	9	7	66	3	4	0	1	90
Romanticism	3	12	7	27	22	13	6	90
Impressionism	1	1	1	13	56	11	7	90
Cubism	2	10	11	8	24	29	6	90
Modern art	16	10	6	10	15	10	23	90
	61	85	123	78	126	73	54	600
PGN /harmonies/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	25	12	8	1	3	6	5	60
Renaissance	2	55	13	10	3	5	2	90
Baroque	6	22	46	5	2	5	4	90
Romanticism	4	26	2	27	12	11	8	90
Impressionism	4	3	4	12	51	10	6	90
Cubism	3	21	10	11	15	19	11	90
Modern art	11	16	5	10	7	14	27	90
	55	155	88	76	93	70	63	600

Table 16. Confusion matrices of different classifiers using VQ-MPEG7 descriptors – class "movements"

OneR /VQ/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	19	6	17	1	3	2	12	60
Renaissance	6	54	2	11	2	14	1	90
Baroque	14	2	67	1	0	5	1	90
Romanticism	3	9	2	42	17	13	4	90
Impressionism	3	7	2	11	54	10	3	90
Cubism	8	13	6	9	11	38	5	90
Modern art	7	11	10	8	12	8	34	90
	60	102	106	83	99	90	60	600
JRip /VQ/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	29	10	2	0	6	1	12	60
Renaissance	0	56	3	6	11	12	2	90
Baroque	2	9	70	2	4	2	1	90
Romanticism	0	13	3	51	14	4	5	90
Impressionism	1	13	1	9	43	15	8	90

Cubism	4	23	4	5	12	32	10	90
Modern art	12	19	1	4	17	4	33	90
	48	143	84	77	107	70	71	600

J48 /VQ/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	28	0	7	0	2	7	16	60
Renaissance	2	53	4	11	2	13	5	90
Baroque	8	3	66	1	1	8	3	90
Romanticism	1	6	2	43	17	15	6	90
Impressionism	2	6	1	12	46	15	8	90
Cubism	5	18	5	10	10	35	7	90
Modern art	7	12	6	10	10	8	37	90
	53	98	91	87	88	101	82	600

PGN /VQ/	Icon	Renaissance	Baroque	Romanticism	Impressionism	Cubism	Modern art	Sum by row
Icon	29	4	3	1	6	6	11	60
Renaissance	0	49	5	6	8	17	5	90
Baroque	2	0	75	4	4	3	2	90
Romanticism	1	4	2	70	6	3	4	90
Impressionism	2	6	1	8	62	7	4	90
Cubism	1	11	3	8	8	54	5	90
Modern art	2	4	3	8	8	8	57	90
	37	78	92	105	102	98	88	600

Table 17. Confusion matrices of different classifiers using low level HSL descriptors – class "artists' names"

OneR /HSL/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	46	0	0	1	5	0	0	0	0	0	0	2	0	0	6	0	0	0	0	60
Botticelli	1	5	8	4	3	0	0	0	4	1	0	0	1	3	0	0	0	0	0	30
Michelangelo	0	5	7	3	6	0	0	0	3	1	0	0	2	3	0	0	0	0	0	30
Raphael	1	1	3	3	13	0	0	0	1	2	0	1	2	1	2	0	0	0	0	30
Caravaggio	0	0	2	5	13	0	0	0	0	0	0	3	2	0	5	0	0	0	0	30
Rembrandt	2	1	4	4	10	0	0	0	0	0	0	2	2	0	5	0	0	0	0	30
Rubens	16	0	0	0	6	0	0	0	0	0	0	1	0	0	7	0	0	0	0	30
Friedrich	2	3	2	1	7	0	0	0	2	6	0	0	2	1	4	0	0	0	0	30
Goya	0	2	2	0	0	0	0	0	6	18	0	0	0	2	0	0	0	0	0	30
Turner	1	0	0	0	0	0	0	0	1	24	0	0	2	1	1	0	0	0	0	30
Monet	2	0	2	3	11	0	0	0	2	5	0	0	2	2	1	0	0	0	0	30
Pissarro	14	0	1	0	7	0	0	0	0	0	0	1	1	0	6	0	0	0	0	30
Sisley	1	2	2	3	11	0	0	0	1	2	0	1	0	3	4	0	0	0	0	30
Braque	0	5	6	3	8	0	0	0	2	4	0	0	1	0	1	0	0	0	0	30
Gris	5	0	2	2	9	0	0	0	0	0	0	5	0	1	6	0	0	0	0	30
Leger	3	0	3	1	10	0	0	0	0	2	0	4	2	2	3	0	0	0	0	30
Klimt	8	2	2	2	3	0	0	0	2	2	0	2	0	4	3	0	0	0	0	30
Miro	10	1	2	4	5	0	0	0	2	0	0	0	2	1	1	0	0	2	0	30
Mucha	13	0	3	0	4	0	0	0	1	1	0	1	2	0	4	0	0	1	0	30
	125	27	51	39	131	0	0	0	27	68	0	23	23	24	59	0	0	3	0	600

OneR /HSL/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	32	2	2	2	0	1	3	0	0	1	2	1	0	1	6	1	2	0	4	60
Botticelli	6	5	5	4	2	1	2	0	1	0	0	0	0	0	1	2	0	0	1	30
Michelangelo	5	1	15	2	2	0	0	1	1	0	0	0	1	0	0	2	0	0	0	30
Raphael	10	2	11	4	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	30
Caravaggio	5	1	0	0	17	2	1	0	0	0	0	0	0	0	0	1	1	1	1	30
Rembrandt	2	2	0	1	3	14	1	0	2	1	0	0	1	0	0	0	0	2	1	30
Rubens	5	1	0	2	2	0	17	1	0	0	0	0	0	0	0	0	2	0	0	30
Friedrich	11	0	0	1	1	0	0	5	2	1	3	1	2	0	1	0	0	0	2	30
Goya	5	2	0	1	0	1	0	2	13	2	0	0	3	0	0	0	0	0	1	30
Turner	4	1	0	2	0	0	0	2	0	18	1	0	0	0	0	0	1	0	1	30
Monet	9	0	0	1	0	0	0	0	3	3	6	0	4	0	1	0	1	0	2	30
Pissarro	7	0	0	0	0	0	1	1	0	0	0	15	1	2	3	0	0	0	0	30
Sisley	10	1	0	0	0	0	0	1	0	1	2	1	8	0	3	2	1	0	0	30
Braque	13	1	1	0	0	0	0	0	2	0	2	1	2	6	0	0	1	0	1	30
Gris	11	0	0	4	0	0	0	0	1	0	1	0	3	0	10	0	0	0	0	30
Leger	5	5	1	0	0	2	1	1	0	0	1	2	3	1	3	4	0	0	1	30
Klimt	9	1	1	0	0	0	1	1	2	0	5	0	1	0	1	0	5	0	3	30
Miro	1	1	0	0	1	4	1	0	0	0	4	2	2	0	0	1	0	12	1	30
Mucha	8	1	0	1	0	0	0	2	0	0	0	1	0	2	1	0	0	0	14	30
	158	27	36	25	28	25	29	17	27	27	27	24	31	12	32	13	14	15	33	600

OneR /HSL/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	28	1	0	3	1	0	5	2	0	3	0	1	0	2	3	1	4	1	5	60
Botticelli	1	7	4	1	4	0	2	1	3	0	0	0	0	0	1	3	3	0	0	30
Michelangelo	3	4	10	5	0	0	1	1	1	0	0	0	0	0	2	0	2	0	1	30
Raphael	0	2	4	11	1	0	0	0	0	3	0	0	0	3	2	1	2	0	1	30
Caravaggio	0	2	0	1	14	4	5	0	0	0	0	0	0	0	0	3	0	1	0	30
Rembrandt	0	2	0	0	6	16	1	0	1	0	0	0	0	0	0	2	0	2	0	30
Rubens	2	0	0	2	6	0	18	0	0	0	0	0	0	0	2	0	0	0	0	30
Friedrich	3	1	0	4	2	0	0	5	2	1	2	2	2	1	2	1	1	0	1	30
Goya	0	4	0	1	0	4	0	1	10	2	0	0	0	5	1	1	0	0	1	30
Turner	1	1	1	1	0	0	1	0	0	23	0	1	1	0	0	0	0	0	0	30
Monet	0	1	0	1	0	0	0	3	4	1	5	0	4	1	1	1	7	0	1	30
Pissarro	1	0	0	1	0	0	0	2	1	1	2	9	4	6	1	0	2	0	0	30
Sisley	1	2	0	0	0	0	0	0	0	2	5	2	10	2	2	1	2	1	0	30
Braque	3	1	1	0	1	0	0	0	3	0	1	2	3	9	2	0	3	0	1	30
Gris	5	0	0	1	0	0	1	1	1	0	0	3	1	1	10	3	2	0	1	30
Leger	2	1	1	2	2	0	0	3	2	0	1	0	1	1	2	5	2	5	0	30
Klimt	3	2	0	1	0	0	0	2	1	0	3	1	1	3	1	3	7	0	2	30
Miro	3	1	0	0	0	1	2	2	0	0	1	1	0	2	0	3	2	12	0	30
Mucha	0	0	0	0	0	0	2	4	0	1	0	3	0	2	2	0	2	0	14	30
	56	32	21	35	37	25	38	27	29	37	20	25	27	38	34	28	41	22	28	600

OneR /HSL/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	47	1	0	1	0	0	0	2	0	0	1	0	0	1	3	0	1	1	2	60
Botticelli	0	10	4	2	2	1	1	2	0	1	2	0	1	0	1	2	1	0	0	30
Michelangelo	0	6	11	6	0	0	1	0	0	0	0	0	0	0	1	4	0	0	1	30
Raphael	1	4	9	9	0	0	0	0	0	0	0	0	0	1	2	3	1	0	0	30
Caravaggio	0	0	0	0	16	4	5	0	0	0	0	1	0	0	0	1	1	2	0	30
Rembrandt	0	2	0	0	4	23	0	1	0	0	0	0	0	0	0	0	0	0	0	30
Rubens	8	0	0	0	3	0	14	1	0	0	0	0	0	0	1	2	1	0	0	30
Friedrich	3	3	0	0	0	0	0	10	1	1	2	0	1	1	2	4	0	1	1	30

Goya	0	3	0	0	0	3	0	3	15	2	0	0	1	1	0	1	1	0	0	30
Turner	1	0	1	2	0	0	0	1	3	20	0	0	0	0	1	0	0	1	0	30
Monet	0	2	0	0	0	0	0	2	2	0	11	1	4	2	2	0	4	0	0	30
Pissarro	3	0	0	0	0	0	0	3	0	0	1	12	4	2	0	1	3	0	1	30
Sisley	3	1	0	0	0	0	0	3	1	0	2	2	11	0	0	1	5	1	0	30
Braque	2	0	0	0	0	0	1	0	6	0	0	2	0	16	1	0	2	0	0	30
Gris	7	2	0	1	0	0	0	1	1	0	2	1	2	3	8	0	1	0	1	30
Leger	2	3	0	0	1	0	0	5	0	1	0	0	3	0	2	11	1	1	0	30
Klimt	7	2	1	0	0	0	0	2	1	0	2	0	2	2	0	0	10	0	1	30
Miro	3	0	0	0	0	0	1	0	0	0	1	0	0	0	0	2	2	21	0	30
Mucha	4	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	1	21	30
	91	39	26	21	26	31	23	37	30	26	24	19	29	30	24	32	35	29	28	600

Table 18. Confusion matrices of different classifiers using low level Hue descriptors – class "artists' names"

OneR /Hue/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	33	0	14	0	7	0	0	0	1	0	0	1	1	0	0	3	0	0	0	60
Botticelli	13	0	9	0	4	0	0	0	0	0	0	0	1	0	0	1	0	0	2	30
Michelangelo	11	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30
Raphael	10	0	19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30
Caravaggio	18	0	2	0	2	0	0	0	1	0	0	0	0	0	2	2	0	0	3	30
Rembrandt	13	0	7	0	4	0	0	0	2	0	0	0	1	0	0	0	0	0	3	30
Rubens	15	0	12	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	30
Friedrich	7	0	2	0	0	0	0	0	2	0	0	11	3	0	0	3	0	0	2	30
Goya	3	0	0	0	0	0	0	0	1	0	0	9	6	0	0	5	0	0	6	30
Turner	14	0	9	0	3	0	0	0	0	0	0	1	1	0	0	2	0	0	0	30
Monet	3	0	0	0	1	0	0	0	4	0	0	10	5	0	0	3	0	0	4	30
Pissarro	0	0	0	0	1	0	0	0	0	0	0	23	1	0	0	1	0	0	4	30
Sisley	1	0	0	0	2	0	0	0	4	0	0	7	4	0	0	3	0	0	9	30
Braque	5	0	0	0	0	0	0	0	2	0	0	14	2	0	0	2	0	0	5	30
Gris	5	0	3	0	4	0	0	0	2	0	0	3	3	0	1	3	0	0	6	30
Leger	12	0	0	0	2	0	0	0	1	0	0	0	4	0	2	6	0	0	3	30
Klimt	9	0	1	0	2	0	0	0	2	0	0	5	2	0	0	3	0	0	6	30
Miro	2	0	4	0	1	0	0	0	0	0	0	13	5	0	1	1	0	0	3	30
Mucha	8	0	3	0	2	0	0	0	4	0	0	0	4	0	0	6	0	0	3	30
	182	0	104	0	37	0	0	0	26	0	0	97	43	0	7	45	0	0	59	600

OneR /Hue/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	39	1	0	4	0	1	1	0	0	4	0	0	0	0	2	1	2	0	5	60
Botticelli	17	3	4	1	1	0	1	0	1	1	0	0	0	0	0	0	1	0	0	30
Michelangelo	13	1	3	5	1	4	0	0	0	2	0	0	0	0	0	0	1	0	0	30
Raphael	17	0	3	5	0	2	1	0	0	1	0	0	0	0	0	0	0	1	0	30
Caravaggio	6	1	1	1	17	1	1	0	0	0	0	0	0	0	0	1	0	1	0	30
Rembrandt	15	0	2	1	3	5	0	0	0	1	0	0	0	0	0	0	1	1	1	30
Rubens	19	1	0	0	0	1	5	0	0	1	0	0	0	1	0	0	0	1	1	30
Friedrich	17	1	0	0	0	0	0	0	2	1	2	1	2	0	0	1	0	0	3	30
Goya	18	0	0	0	0	0	0	0	6	1	0	0	2	1	1	0	0	0	1	30
Turner	16	0	2	1	0	2	0	0	0	8	0	0	0	1	0	0	0	0	0	30
Monet	12	0	0	0	0	0	0	0	6	0	6	3	1	0	0	0	0	1	1	30
Pissarro	11	0	0	0	0	0	0	1	1	0	0	10	0	2	0	0	1	2	2	30
Sisley	16	0	0	0	0	0	0	1	1	0	0	2	6	1	0	2	0	1	0	30
Braque	17	0	0	0	0	0	0	1	0	1	1	0	0	3	0	1	1	1	4	30
Gris	18	2	0	2	0	0	1	0	1	1	0	0	0	0	4	1	0	0	0	30

Leger	15	3	0	0	0	0	0	0	2	0	0	2	4	1	0	3	0	0	0	30
Klimt	14	1	1	0	0	0	0	0	2	3	1	0	2	1	0	0	4	0	1	30
Miro	13	1	2	0	0	0	0	0	1	1	2	2	0	0	0	0	0	8	0	30
Mucha	15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	13	30
	308	15	18	20	22	16	10	3	23	26	12	20	17	12	7	10	12	17	32	600

OneR /Hue/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	19	4	1	1	3	3	5	2	2	3	1	1	1	1	5	0	1	0	7	60
Botticelli	4	5	4	1	4	1	1	1	1	1	0	0	0	0	3	3	1	0	0	30
Michelangelo	1	3	7	6	1	1	1	3	0	2	0	0	0	0	3	1	0	0	1	30
Raphael	1	0	8	6	2	3	3	1	0	0	0	0	0	1	3	0	1	0	1	30
Caravaggio	3	1	1	0	17	1	3	0	0	0	0	0	0	0	0	2	1	1	0	30
Rembrandt	2	1	6	0	4	5	1	1	0	2	0	0	0	1	3	2	0	0	2	30
Rubens	6	2	0	2	1	3	6	0	1	1	0	1	0	2	2	0	2	0	1	30
Friedrich	1	2	1	0	1	0	2	2	3	0	0	2	3	5	1	2	1	1	3	30
Goya	1	0	0	0	0	0	0	1	3	0	0	4	4	8	2	3	0	0	4	30
Turner	3	1	2	0	1	3	0	2	1	11	0	0	0	1	2	0	1	1	1	30
Monet	1	0	0	0	0	0	1	4	2	0	5	2	1	1	3	3	3	4	0	30
Pissarro	0	0	0	0	0	0	0	4	3	0	1	9	1	8	1	0	1	2	0	30
Sisley	2	1	0	0	0	0	0	4	4	0	2	1	6	4	1	0	4	1	0	30
Braque	3	1	0	0	0	0	0	3	3	1	1	1	0	10	2	1	0	2	2	30
Gris	2	1	0	1	1	1	1	2	4	1	0	1	1	1	6	4	1	1	1	30
Leger	4	4	0	1	1	0	1	0	3	1	2	0	2	1	4	4	1	1	0	30
Klimt	6	1	1	0	1	0	0	3	2	1	1	2	2	2	3	2	1	1	1	30
Miro	1	0	3	0	1	0	0	1	2	1	6	3	2	2	1	0	3	4	0	30
Mucha	3	1	0	0	0	0	0	0	2	0	0	0	0	2	2	1	0	0	19	30
	63	28	34	18	38	21	25	34	36	25	19	27	23	50	47	28	22	19	43	600

OneR /Hue/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	24	2	2	3	2	5	4	2	0	2	1	0	1	4	2	3	1	0	2	60
Botticelli	4	3	4	5	3	0	0	0	2	1	0	0	0	0	4	2	2	0	0	30
Michelangelo	1	2	3	6	1	7	2	3	0	0	0	0	0	0	0	1	3	1	0	30
Raphael	7	3	8	5	0	1	0	2	1	0	0	0	0	0	1	1	0	0	1	30
Caravaggio	2	1	1	0	16	2	2	0	0	0	0	0	0	0	1	3	0	2	0	30
Rembrandt	5	0	8	1	2	6	1	3	0	3	0	0	0	0	0	0	0	0	1	30
Rubens	8	0	1	3	1	2	7	2	0	2	0	0	0	0	3	0	0	0	1	30
Friedrich	3	1	2	0	0	0	1	0	3	0	2	1	2	0	3	5	1	2	4	30
Goya	3	0	0	0	0	0	1	2	10	1	1	2	3	2	2	2	0	0	1	30
Turner	7	0	2	0	0	2	1	1	0	14	0	1	0	0	2	0	0	0	0	30
Monet	1	1	0	0	0	0	0	0	4	0	9	2	3	1	0	2	4	3	0	30
Pissarro	0	0	0	0	0	0	0	3	4	0	3	15	2	2	0	0	0	0	1	30
Sisley	1	1	0	1	0	0	0	0	4	0	1	2	11	1	1	3	0	3	1	30
Braque	1	0	0	0	0	0	2	1	5	0	0	1	2	8	2	1	0	1	6	30
Gris	3	0	1	3	0	0	2	0	5	1	0	0	3	1	8	0	0	1	2	30
Leger	1	3	1	1	1	0	0	0	2	1	4	1	5	1	2	6	0	1	0	30
Klimt	6	0	1	1	0	1	1	1	2	0	4	0	2	2	1	2	5	0	1	30
Miro	2	0	0	1	2	0	0	2	1	2	2	2	1	2	0	1	0	11	1	30
Mucha	2	0	0	0	0	1	3	0	1	0	0	1	0	1	1	0	2	0	18	30
	81	17	34	30	28	27	27	22	44	27	27	28	35	25	33	32	18	25	40	600

Table 19. Confusion matrices of different classifiers using low level Saturation descriptors – class "artists' names"

OneR /Sat/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	46	0	0	1	5	0	0	0	0	0	0	2	0	0	6	0	0	0	0	60
Botticelli	1	5	8	4	3	0	0	0	4	1	0	0	1	3	0	0	0	0	0	30
Michelangelo	0	5	7	3	6	0	0	0	3	1	0	0	2	3	0	0	0	0	0	30
Raphael	1	1	3	3	13	0	0	0	1	2	0	1	2	1	2	0	0	0	0	30
Caravaggio	0	0	2	5	13	0	0	0	0	0	0	3	2	0	5	0	0	0	0	30
Rembrandt	2	1	4	4	10	0	0	0	0	0	0	2	2	0	5	0	0	0	0	30
Rubens	16	0	0	0	6	0	0	0	0	0	0	1	0	0	7	0	0	0	0	30
Friedrich	2	3	2	1	7	0	0	0	2	6	0	0	2	1	4	0	0	0	0	30
Goya	0	2	2	0	0	0	0	0	6	18	0	0	0	2	0	0	0	0	0	30
Turner	1	0	0	0	0	0	0	0	1	24	0	0	2	1	1	0	0	0	0	30
Monet	2	0	2	3	11	0	0	0	2	5	0	0	2	2	1	0	0	0	0	30
Pissarro	14	0	1	0	7	0	0	0	0	0	0	1	1	0	6	0	0	0	0	30
Sisley	1	2	2	3	11	0	0	0	1	2	0	1	0	3	4	0	0	0	0	30
Braque	0	5	6	3	8	0	0	0	2	4	0	0	1	0	1	0	0	0	0	30
Gris	5	0	2	2	9	0	0	0	0	0	0	5	0	1	6	0	0	0	0	30
Leger	3	0	3	1	10	0	0	0	0	2	0	4	2	2	3	0	0	0	0	30
Klimt	8	2	2	2	3	0	0	0	2	2	0	2	0	4	3	0	0	0	0	30
Miro	10	1	2	4	5	0	0	0	2	0	0	0	2	1	1	0	0	2	0	30
Mucha	13	0	3	0	4	0	0	0	1	1	0	1	2	0	4	0	0	1	0	30
	125	27	51	39	131	0	0	0	27	68	0	23	23	24	59	0	0	3	0	600

OneR /Sat/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	47	0	1	0	1	0	0	0	0	0	0	4	0	0	1	0	0	4	2	60
Botticelli	24	1	1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	30
Michelangelo	22	1	3	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	1	30
Raphael	19	0	1	0	2	0	1	0	0	2	1	1	0	0	0	3	0	0	0	30
Caravaggio	19	0	0	0	7	2	0	0	0	0	1	0	0	0	0	0	0	0	1	30
Rembrandt	21	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	3	30
Rubens	20	0	0	1	2	0	1	0	0	0	0	1	0	0	0	0	3	0	2	30
Friedrich	19	2	0	0	0	0	0	0	0	3	0	1	0	0	0	3	0	2	0	30
Goya	17	0	0	0	0	0	0	1	6	5	0	0	0	1	0	0	0	0	0	30
Turner	10	0	0	0	0	0	0	0	3	16	0	0	0	0	0	0	1	0	0	30
Monet	21	0	0	0	1	0	0	1	0	3	3	0	0	0	0	1	0	0	0	30
Pissarro	19	0	0	0	1	0	1	0	0	0	0	5	1	0	0	2	0	0	1	30
Sisley	22	0	1	0	0	0	1	0	0	0	3	0	0	0	0	1	0	2	0	30
Braque	20	1	1	0	2	1	0	0	2	0	1	0	0	1	0	0	1	0	0	30
Gris	26	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	30
Leger	21	0	0	1	1	1	0	0	0	0	0	1	0	1	0	1	0	2	1	30
Klimt	20	2	1	0	1	0	0	0	1	1	0	0	1	2	1	0	0	0	0	30
Miro	18	0	0	0	0	2	0	0	0	0	0	1	0	0	0	1	0	7	1	30
Mucha	21	0	0	0	1	1	1	0	0	2	0	1	1	0	0	0	0	0	2	30
	406	8	9	2	24	10	6	2	13	32	9	16	3	5	2	17	5	17	14	600

OneR /Sat/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	31	0	1	1	0	2	7	0	0	0	2	3	0	0	1	1	3	4	4	60
Botticelli	1	7	4	2	0	0	1	0	5	0	3	0	3	2	1	1	0	0	0	30
Michelangelo	1	2	6	0	0	0	0	3	4	1	0	0	3	4	0	1	3	1	1	30
Raphael	0	3	1	3	5	0	0	3	2	1	4	0	0	2	1	3	1	0	1	30

Caravaggio	0	0	2	1	12	7	1	0	0	0	1	0	0	3	0	0	2	0	1	30
Rembrandt	2	2	1	1	5	7	1	0	0	0	1	2	0	2	2	2	1	1	0	30
Rubens	9	0	1	0	1	0	5	0	0	0	0	2	4	0	3	0	3	0	2	30
Friedrich	1	4	1	2	1	0	1	3	3	2	2	2	0	1	2	3	0	2	0	30
Goya	0	4	1	0	0	0	0	4	7	3	3	0	2	5	1	0	0	0	0	30
Turner	1	0	0	1	1	0	0	1	6	13	2	0	1	1	1	2	0	0	0	30
Monet	1	6	1	2	0	1	1	2	2	2	4	0	2	1	1	3	1	0	0	30
Pissarro	9	0	1	0	2	0	2	2	0	0	0	5	2	1	2	2	2	0	0	30
Sisley	3	3	4	0	1	1	1	2	1	0	8	1	3	0	2	0	0	0	0	30
Braque	0	5	3	1	3	0	0	2	6	0	2	1	1	4	1	0	0	1	0	30
Gris	3	2	0	1	2	0	2	0	0	0	2	2	2	1	9	1	2	0	1	30
Leger	1	1	1	1	3	1	1	1	0	0	5	1	1	2	2	4	1	3	1	30
Klimt	3	5	0	0	1	0	0	0	3	1	1	2	0	4	3	0	3	2	2	30
Miro	7	4	2	1	1	0	1	2	0	0	0	0	0	2	1	2	0	6	1	30
Mucha	10	1	1	0	5	1	1	0	1	1	0	1	1	0	1	0	1	2	3	30
	83	49	31	17	43	20	25	25	40	24	40	22	25	35	34	25	23	22	17	600

OneR /Sat/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	28	0	2	1	0	2	3	0	0	0	1	7	1	0	2	0	2	5	6	60
Botticelli	0	7	2	3	1	0	1	2	2	0	1	2	2	3	2	2	0	0	0	30
Michelangelo	0	3	6	1	2	3	0	0	3	1	1	0	4	4	0	0	2	0	0	30
Raphael	2	2	2	2	2	3	0	0	0	1	3	1	2	2	3	3	2	0	0	30
Caravaggio	1	0	2	0	14	3	0	0	0	0	3	0	1	0	1	3	0	0	2	30
Rembrandt	1	1	3	1	6	6	1	1	0	0	0	2	0	1	2	2	1	0	2	30
Rubens	10	0	0	1	1	0	5	0	0	0	1	2	1	0	3	1	3	0	2	30
Friedrich	0	3	3	1	0	4	1	3	0	1	2	2	2	0	1	4	0	2	1	30
Goya	0	2	3	0	0	0	0	2	6	4	3	0	2	6	0	0	2	0	0	30
Turner	0	0	0	0	0	1	0	1	4	19	1	0	1	0	1	1	0	0	1	30
Monet	0	2	2	2	3	2	0	2	2	2	3	0	5	1	0	2	0	0	2	30
Pissarro	8	1	0	0	1	0	2	0	0	0	0	8	2	0	3	1	3	0	1	30
Sisley	1	3	5	2	1	1	0	0	1	0	5	2	2	0	1	3	3	0	0	30
Braque	1	2	3	2	1	2	0	1	4	0	4	0	0	5	2	1	2	0	0	30
Gris	5	1	0	4	2	0	1	2	0	1	2	3	1	0	6	1	0	0	1	30
Leger	1	4	1	1	0	3	0	1	1	0	3	1	1	1	1	8	0	3	0	30
Klimt	5	4	1	0	1	0	0	1	2	2	2	2	1	1	3	0	3	0	2	30
Miro	8	0	4	1	1	1	0	0	0	0	0	0	0	1	0	1	1	10	2	30
Mucha	6	0	1	0	4	2	1	0	0	2	0	4	1	0	1	0	0	1	7	30
	77	35	40	22	40	33	15	16	25	33	35	36	29	25	32	33	24	21	29	600

Table 20. Confusion matrices of different classifiers using low level Luminance descriptors – class "artists' names"

OneR /Lum/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	20	0	0	0	0	0	3	0	0	8	0	0	1	13	6	1	0	0	8	60
Botticelli	10	0	0	0	4	3	3	2	0	0	0	0	0	2	3	1	0	0	2	30
Michelangelo	19	0	0	0	0	0	2	0	0	0	0	0	0	6	1	1	0	0	1	30
Raphael	9	0	0	0	0	1	1	0	0	4	0	0	0	6	4	2	0	0	3	30
Caravaggio	0	0	0	0	16	6	1	3	0	0	0	0	0	0	1	1	0	2	0	30
Rembrandt	0	0	0	0	12	3	2	0	0	0	0	0	0	0	3	0	0	10	0	30
Rubens	4	0	0	0	5	2	4	3	0	0	0	0	0	1	7	3	0	0	1	30
Friedrich	5	0	0	0	1	0	1	3	0	3	0	0	2	6	3	2	0	0	4	30
Goya	8	0	0	1	2	1	1	3	0	1	0	0	0	4	3	1	0	3	2	30
Turner	2	0	0	0	0	0	3	1	0	17	0	0	4	2	0	1	0	0	0	30
Monet	8	0	0	0	3	0	0	0	0	4	0	0	0	10	0	1	0	0	4	30

Pissarro	4	0	0	0	0	0	0	0	0	7	0	0	2	11	0	0	0	0	6	30
Sisley	7	0	0	0	0	0	2	0	0	4	0	0	2	9	0	1	0	0	5	30
Braque	9	0	0	0	0	0	0	0	0	1	0	0	1	15	0	0	0	0	4	30
Gris	14	0	0	0	1	0	0	2	0	0	0	0	0	1	10	2	0	0	0	30
Leger	8	0	0	0	4	0	3	3	0	0	0	0	0	0	8	2	0	0	2	30
Klimt	7	0	0	0	1	1	2	0	0	4	0	0	1	8	2	1	0	0	3	30
Miro	5	0	0	0	1	0	1	1	0	8	0	0	0	4	1	2	0	6	1	30
Mucha	3	0	0	0	0	0	0	2	0	8	0	0	0	8	0	0	0	0	9	30
	142	0	0	1	50	17	29	23	0	69	0	0	13	106	52	22	0	21	55	600

OneR /Lum/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	39	1	2	1	0	0	1	2	2	1	1	0	1	4	4	0	0	0	1	60
Botticelli	13	2	1	0	3	1	7	1	0	0	0	0	0	0	1	1	0	0	0	30
Michelangelo	21	1	2	3	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	30
Raphael	22	0	2	1	1	0	0	0	0	1	0	0	0	1	1	1	0	0	0	30
Caravaggio	2	1	0	0	16	5	3	0	1	0	0	0	0	0	0	1	0	1	0	30
Rembrandt	0	0	0	0	5	21	0	1	1	0	0	0	0	0	0	0	0	2	0	30
Rubens	8	1	0	0	5	0	16	0	0	0	0	0	0	0	0	0	0	0	0	30
Friedrich	12	0	0	0	0	0	1	5	1	2	0	0	2	0	1	3	0	1	2	30
Goya	16	0	0	0	4	3	0	0	2	2	0	0	2	1	0	0	0	0	0	30
Turner	10	0	0	1	0	0	0	1	0	12	0	1	0	0	0	0	0	2	3	30
Monet	16	0	2	0	0	0	2	0	1	0	1	1	1	3	1	0	1	0	1	30
Pissarro	23	0	0	1	0	0	0	0	0	2	1	0	2	0	0	0	0	0	1	30
Sisley	13	0	0	1	0	0	0	1	1	1	0	1	7	0	0	0	0	0	5	30
Braque	21	0	1	0	0	0	0	0	1	0	0	1	1	3	0	0	0	1	1	30
Gris	22	0	1	1	0	0	0	2	0	0	1	0	1	0	1	1	0	0	0	30
Leger	19	0	0	0	1	0	1	2	1	0	0	0	0	0	0	4	0	2	0	30
Klimt	15	0	2	0	1	0	0	0	2	1	1	0	1	1	1	1	3	0	1	30
Miro	2	0	0	0	1	2	1	2	3	0	0	1	0	0	0	2	0	16	0	30
Mucha	16	0	0	1	0	0	0	0	0	1	1	0	3	1	0	0	0	0	7	30
	290	6	13	10	37	32	33	17	16	23	6	5	21	14	11	15	4	25	22	600

OneR /Lum/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	10	5	7	1	1	0	0	4	3	0	9	3	3	5	6	0	2	0	1	60
Botticelli	2	7	2	0	3	0	5	2	0	0	3	0	0	0	2	1	2	0	1	30
Michelangelo	7	2	5	0	0	0	0	2	0	0	0	1	0	5	4	2	2	0	0	30
Raphael	4	2	3	3	0	0	1	1	0	2	1	2	0	4	3	1	1	0	2	30
Caravaggio	0	2	0	1	14	6	3	0	0	0	0	0	0	0	0	1	2	1	0	30
Rembrandt	0	1	0	0	6	20	1	0	0	0	0	0	0	0	0	0	0	2	0	30
Rubens	1	2	0	0	6	0	14	1	1	0	1	0	0	0	1	1	2	0	0	30
Friedrich	4	0	0	4	0	0	1	3	2	2	5	0	1	1	1	4	1	0	1	30
Goya	3	3	1	0	1	3	0	1	5	1	0	1	1	3	2	3	1	1	0	30
Turner	1	0	0	2	0	0	1	4	1	9	4	1	1	0	1	0	0	0	5	30
Monet	5	2	1	0	0	0	0	0	0	2	4	3	3	5	1	0	3	0	1	30
Pissarro	4	0	0	1	0	0	0	0	0	3	4	5	2	3	0	0	4	0	4	30
Sisley	4	1	1	0	0	0	0	1	1	1	4	3	4	5	0	1	1	0	3	30
Braque	2	1	3	0	0	0	0	0	2	0	5	2	1	9	2	0	1	0	2	30
Gris	5	5	1	0	0	0	2	3	3	0	0	0	0	1	7	2	1	0	0	30
Leger	1	3	2	2	1	0	0	5	1	0	0	0	1	0	0	9	1	2	2	30
Klimt	3	2	1	0	1	0	1	2	3	0	2	4	2	1	4	0	2	0	2	30
Miro	1	1	0	0	1	1	0	2	2	1	0	1	0	1	1	1	1	15	1	30
Mucha	0	0	2	1	0	0	0	0	0	3	2	8	0	2	0	0	0	0	12	30
	57	39	29	15	34	30	29	31	24	24	44	34	19	45	35	26	27	21	37	600

OneR /Lum/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	10	1	7	7	0	0	1	3	3	2	4	3	2	5	6	0	5	0	1	60
Botticelli	2	3	4	0	1	1	8	1	2	0	1	0	0	2	3	2	0	0	0	30
Michelangelo	8	0	6	3	0	0	1	0	0	0	0	0	0	4	5	1	1	0	1	30
Raphael	9	0	4	6	0	0	1	0	0	1	1	2	0	0	4	0	0	0	2	30
Caravaggio	0	1	0	0	18	3	4	0	1	0	1	0	0	0	0	1	0	1	0	30
Rembrandt	0	0	0	0	10	17	1	1	1	0	0	0	0	0	0	0	0	0	0	30
Rubens	2	3	1	0	5	0	16	1	0	0	0	0	0	0	0	0	1	1	0	30
Friedrich	3	0	0	0	0	0	1	9	3	3	0	0	1	0	3	3	1	0	3	30
Goya	5	1	2	2	2	2	1	0	6	1	1	0	1	2	1	0	2	0	1	30
Turner	1	0	0	1	0	0	0	2	0	20	0	2	1	0	0	0	2	0	1	30
Monet	4	1	1	1	0	0	2	1	1	0	4	1	3	5	2	0	0	2	2	30
Pissarro	6	0	0	0	0	0	0	3	0	1	2	3	4	1	0	0	2	1	7	30
Sisley	3	1	1	0	0	0	0	2	1	3	1	2	7	1	0	2	0	0	6	30
Braque	10	0	0	1	0	0	0	1	2	0	3	2	1	6	1	0	1	1	1	30
Gris	4	1	1	1	0	0	2	1	1	0	2	1	4	0	10	2	0	0	0	30
Leger	3	2	1	2	2	0	0	1	2	1	1	1	0	0	2	10	0	1	1	30
Klimt	4	1	0	1	1	0	1	0	1	0	1	0	3	3	2	1	8	1	2	30
Miro	0	1	0	0	0	2	0	1	1	1	1	0	0	0	0	1	0	21	1	30
Mucha	2	0	0	2	0	0	0	0	0	2	0	5	2	2	0	0	2	1	12	30
	76	16	28	27	39	25	39	27	25	35	23	22	29	31	39	23	25	30	41	600

Table 21. Confusion matrices of different classifiers using Harmonies/Contrasts descriptors – class "artists' names"

OneR /harmonies/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	53	0	0	0	0	0	0	1	0	0	0	0	1	1	3	1	0	0	0	60
Botticelli	23	0	0	0	0	0	0	1	0	0	0	0	0	1	3	2	0	0	0	30
Michelangelo	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30
Raphael	26	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	30
Caravaggio	21	0	0	0	0	0	0	1	0	0	0	2	0	1	4	1	0	0	0	30
Rembrandt	28	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	30
Rubens	28	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	30
Friedrich	10	0	0	0	0	0	0	0	0	0	0	2	7	3	6	2	0	0	0	30
Goya	5	0	0	0	0	0	0	3	0	0	0	1	10	6	4	1	0	0	0	30
Turner	28	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	30
Monet	1	0	0	0	0	0	0	2	1	0	0	4	15	3	2	2	0	0	0	30
Pissarro	1	0	0	0	0	0	0	2	1	0	0	16	1	3	3	3	0	0	0	30
Sisley	0	0	0	0	0	0	0	1	0	0	0	0	19	4	4	2	0	0	0	30
Braque	9	0	0	0	0	0	0	2	2	0	0	6	0	3	4	4	0	0	0	30
Gris	11	0	0	0	0	0	0	4	1	0	0	0	4	8	1	1	0	0	0	30
Leger	9	0	0	0	0	0	0	4	0	0	0	1	5	5	5	1	0	0	0	30
Klimt	16	0	0	0	0	0	0	1	1	0	0	4	2	2	3	1	0	0	0	30
Miro	8	0	0	0	0	0	0	1	0	0	0	3	11	3	2	2	0	0	0	30
Mucha	23	0	0	0	0	0	0	2	0	0	0	0	0	0	2	3	0	0	0	30
	330	0	0	0	0	0	0	27	6	0	0	40	77	45	48	27	0	0	0	600

OneR /harmonies/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	54	0	0	0	0	0	0	1	1	0	0	0	0	0	0	2	0	0	2	60
Botticelli	25	1	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	30

Michelangelo	25	1	2	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	30
Raphael	24	1	0	0	0	0	0	0	1	2	0	0	1	0	0	1	0	0	0	30
Caravaggio	23	0	2	0	1	2	0	0	1	0	0	0	0	0	0	1	0	0	0	30
Rembrandt	14	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	30
Rubens	28	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30
Friedrich	21	0	0	0	0	0	0	0	0	1	0	1	5	0	0	2	0	0	0	30
Goya	21	0	0	0	0	0	0	0	4	0	0	0	3	1	0	1	0	0	0	30
Turner	22	0	0	0	0	1	0	0	1	5	0	0	0	0	0	0	0	1	0	30
Monet	15	0	0	0	0	0	0	0	1	0	6	2	3	1	0	2	0	0	0	30
Pissarro	14	0	0	0	0	0	0	0	0	0	1	10	0	2	0	3	0	0	0	30
Sisley	12	0	0	0	0	0	0	0	0	0	0	0	15	0	0	3	0	0	0	30
Braque	23	1	0	0	0	0	0	0	1	0	0	0	0	1	0	4	0	0	0	30
Gris	22	0	0	0	0	0	0	0	1	0	0	0	1	2	0	3	0	1	0	30
Leger	16	2	0	0	0	0	0	0	0	0	0	0	4	1	0	7	0	0	0	30
Klimt	26	0	0	0	0	0	0	0	2	0	0	1	0	0	0	1	0	0	0	30
Miro	11	0	1	0	0	3	0	0	1	0	0	2	2	0	1	2	0	6	1	30
Mucha	24	0	0	0	0	0	0	0	0	3	0	0	0	0	0	3	0	0	0	30
	420	6	8	0	2	22	0	1	14	13	7	16	34	8	1	37	0	8	3	600
OneR /harmonies/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	33	4	1	2	3	0	2	2	2	0	0	0	0	1	1	2	3	0	4	60
Botticelli	1	8	4	3	3	2	3	2	1	0	0	0	0	0	1	0	0	1	1	30
Michelangelo	1	7	9	5	2	0	1	0	0	4	0	0	0	0	0	1	0	0	0	30
Raphael	4	2	1	8	1	0	0	0	1	7	1	0	1	2	0	1	0	0	1	30
Caravaggio	1	3	4	0	10	3	4	0	3	0	0	1	0	0	0	0	0	1	0	30
Rembrandt	1	3	0	0	2	20	2	0	1	0	1	0	0	0	0	0	0	0	0	30
Rubens	7	5	2	0	3	0	11	1	0	0	1	0	0	0	0	0	0	0	0	30
Friedrich	5	0	0	3	1	0	0	0	1	1	1	3	5	4	2	3	0	1	0	30
Goya	0	2	0	0	5	0	0	1	7	0	2	0	4	2	3	2	0	2	0	30
Turner	0	0	1	3	2	0	0	0	1	13	0	0	0	0	0	3	1	0	6	30
Monet	0	3	0	0	0	0	0	0	1	0	8	1	5	5	1	3	1	2	0	30
Pissarro	0	0	0	0	0	0	0	3	1	0	1	13	2	3	1	0	0	4	2	30
Sisley	1	0	0	0	0	0	0	2	1	0	1	1	18	1	1	1	2	1	0	30
Braque	1	1	0	3	2	0	1	5	1	0	5	1	1	2	2	1	3	0	1	30
Gris	8	4	0	0	2	0	3	0	3	0	1	1	1	1	1	4	0	1	0	30
Leger	1	2	2	1	1	0	0	6	4	1	0	1	1	1	2	6	0	1	0	30
Klimt	4	1	1	2	1	0	2	0	2	3	2	2	0	3	0	1	2	0	4	30
Miro	5	0	1	0	4	2	0	2	1	0	2	1	2	1	0	2	0	7	0	30
Mucha	10	0	0	3	0	0	0	1	0	5	0	0	2	2	0	0	2	0	5	30
	83	45	26	33	42	27	29	25	31	34	26	25	42	28	15	30	14	21	24	600
OneR /harmonies/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	26	11	1	0	0	2	4	0	0	1	1	2	2	2	2	0	3	0	3	60
Botticelli	1	11	3	1	1	5	1	0	1	1	2	0	0	1	0	1	1	0	0	30
Michelangelo	1	14	4	3	0	1	0	0	0	3	0	0	0	0	0	3	0	0	1	30
Raphael	0	12	2	4	0	1	0	1	0	3	2	0	0	2	1	1	1	0	0	30
Caravaggio	1	9	0	0	12	2	1	0	1	2	0	1	0	1	0	0	0	0	0	30
Rembrandt	0	4	0	0	2	22	0	0	0	0	0	0	0	1	0	0	1	0	0	30
Rubens	6	7	0	0	2	4	6	0	0	0	0	0	0	0	0	0	3	0	2	30
Friedrich	3	3	1	1	1	0	0	0	2	1	1	5	3	1	0	4	0	3	1	30
Goya	0	4	1	0	3	0	0	1	8	0	3	0	3	5	0	0	2	0	0	30
Turner	1	9	0	5	0	0	0	1	0	9	0	1	0	0	0	3	0	0	1	30
Monet	1	1	0	0	1	1	1	0	1	0	12	1	4	6	0	0	1	0	0	30
Pissarro	3	0	0	0	0	0	0	1	1	0	1	15	2	3	0	0	2	1	1	30
Sisley	0	0	0	1	0	0	0	2	1	0	5	4	11	0	2	2	1	1	0	30
Braque	1	5	0	0	1	1	1	3	2	0	2	2	0	5	1	2	3	0	1	30
Gris	3	5	0	2	3	2	2	1	1	0	1	1	2	1	1	1	2	2	0	30

Leger	0	3	3	1	1	0	0	0	2	1	1	1	3	5	2	5	1	1	0	30
Klimt	6	1	0	0	1	2	1	0	0	3	4	2	0	3	1	1	4	0	1	30
Miro	1	2	0	0	0	1	1	1	0	1	0	1	1	1	0	2	1	17	0	30
Mucha	6	5	0	3	0	0	0	0	0	1	0	2	0	2	0	1	3	0	7	30
	60	106	15	21	28	44	18	11	20	26	35	38	31	39	10	26	29	25	18	600

Table 22. Confusion matrices of different classifiers using VQ-MPEG7 descriptors – class "artists' names"

OneR /VQ/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	35	2	1	1	2	0	2	0	0	0	0	2	0	1	3	0	0	3	8	60
Botticelli	1	2	3	2	3	1	0	0	4	2	0	0	0	3	3	2	0	1	3	30
Michelangelo	1	2	12	1	0	0	0	0	1	0	0	0	0	4	0	3	0	1	5	30
Raphael	0	1	4	3	0	1	0	1	0	5	0	0	1	2	1	4	2	1	4	30
Caravaggio	4	1	2	0	17	1	2	0	0	0	0	1	0	0	0	1	0	1	0	30
Rembrandt	4	1	0	0	5	9	2	0	3	0	0	0	0	3	2	0	0	1	0	30
Rubens	8	2	0	0	1	0	9	0	1	0	0	2	0	3	2	0	0	1	1	30
Friedrich	2	0	1	1	0	0	0	0	3	11	0	0	5	2	0	1	1	2	1	30
Goya	0	5	0	0	0	3	0	0	4	1	4	1	4	1	2	2	0	2	1	30
Turner	0	0	1	0	0	0	0	5	1	19	0	0	1	0	0	1	0	0	2	30
Monet	0	2	0	1	0	0	0	1	6	0	7	1	2	4	3	2	0	0	1	30
Pissarro	4	1	0	0	0	0	0	0	0	2	0	6	5	3	1	3	1	0	4	30
Sisley	1	0	0	0	0	0	0	1	1	1	3	1	16	0	0	4	0	2	0	30
Braque	1	1	2	0	0	0	0	0	3	2	0	0	0	12	0	3	4	0	2	30
Gris	8	1	0	1	1	0	1	0	2	0	1	2	0	2	4	3	1	1	2	30
Leger	2	4	1	1	0	0	0	1	1	1	0	0	1	0	3	12	0	1	2	30
Klimt	4	4	3	1	0	3	0	0	2	2	3	2	0	1	3	0	1	1	0	30
Miro	10	0	0	0	2	2	0	1	2	4	0	0	1	0	0	1	0	5	2	30
Mucha	6	0	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	19	30
	91	29	33	12	31	20	16	10	34	50	18	18	36	42	28	42	10	23	57	600
OneR /VQ/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	42	0	1	2	2	0	3	0	0	1	0	0	0	0	2	1	1	2	3	60
Botticelli	5	5	2	4	3	0	1	1	0	0	3	0	1	3	0	0	0	1	1	30
Michelangelo	5	4	7	9	0	0	0	0	1	0	0	0	0	0	0	3	0	0	1	30
Raphael	9	1	3	8	0	1	1	1	2	0	0	0	0	1	0	2	0	0	1	30
Caravaggio	5	1	0	2	16	2	3	0	0	0	0	1	0	0	0	0	0	0	0	30
Rembrandt	3	1	0	0	1	21	1	2	0	0	0	0	0	0	0	1	0	0	0	30
Rubens	12	1	0	1	2	1	9	0	0	0	0	0	0	0	1	1	0	2	0	30
Friedrich	3	0	0	1	0	0	1	10	1	3	0	2	3	1	0	2	1	1	1	30
Goya	4	1	0	0	0	0	0	2	13	0	6	0	1	1	1	0	0	1	0	30
Turner	1	0	1	1	0	1	0	4	0	20	0	0	1	0	0	0	0	1	0	30
Monet	5	0	0	0	0	0	2	0	0	0	11	1	3	4	0	0	3	1	0	30
Pissarro	7	0	0	0	0	0	1	1	0	0	1	15	2	0	0	0	1	1	1	30
Sisley	6	0	0	2	0	0	0	0	1	0	3	1	16	1	0	0	0	0	0	30
Braque	14	0	0	1	0	0	0	0	0	0	2	0	0	10	0	1	1	0	1	30
Gris	2	1	1	1	1	0	2	2	0	0	2	0	0	4	7	0	4	2	1	30
Leger	12	2	0	3	0	0	0	1	1	0	1	1	0	0	0	8	0	1	0	30
Klimt	7	0	0	4	0	1	0	0	2	0	2	1	0	0	2	2	6	1	2	30
Miro	7	1	0	0	0	1	0	3	0	2	0	1	1	0	0	0	0	14	0	30
Mucha	14	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	14	30
	163	18	16	39	25	28	24	27	21	26	31	23	28	25	14	21	17	28	26	600

OneR /VQ/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	39	0	0	1	0	0	2	2	0	0	0	1	0	0	2	0	2	5	6	60
Botticelli	0	5	3	3	2	2	0	0	4	0	0	0	0	6	0	4	0	0	1	30
Michelangelo	0	1	13	4	0	0	0	0	0	2	1	0	1	2	0	1	2	0	3	30
Raphael	0	2	11	4	0	0	0	0	0	3	0	1	4	3	0	2	0	0	0	30
Caravaggio	2	1	0	0	18	3	3	0	0	0	0	0	0	0	1	0	0	2	0	30
Rembrandt	0	1	0	0	2	23	2	0	1	0	0	0	0	0	0	0	1	0	0	30
Rubens	7	1	0	0	4	1	9	0	0	0	1	0	1	2	0	1	2	0	1	30
Friedrich	2	0	2	1	0	0	0	7	3	7	0	0	3	4	0	0	0	1	0	30
Goya	0	2	0	0	1	1	0	1	13	0	2	2	1	0	2	2	2	1	0	30
Turner	1	0	3	2	0	0	0	2	1	18	0	0	0	2	0	0	0	0	1	30
Monet	3	2	0	1	0	0	0	1	4	0	3	2	2	4	2	4	1	1	0	30
Pissarro	2	1	0	1	0	0	0	3	2	0	1	9	4	3	0	1	1	0	2	30
Sisley	0	1	0	2	0	0	0	0	2	0	4	4	15	0	0	1	0	1	0	30
Braque	1	6	2	3	0	0	1	0	3	1	0	1	0	9	0	0	0	0	3	30
Gris	8	1	1	1	0	0	4	0	5	0	2	0	1	3	1	2	0	0	1	30
Leger	1	4	5	0	0	0	0	2	0	1	1	1	0	1	0	10	0	2	2	30
Klimt	6	5	2	1	0	0	0	0	1	0	3	2	0	3	0	1	5	0	1	30
Miro	4	0	0	1	3	1	0	5	1	2	1	1	1	0	1	0	0	8	1	30
Mucha	1	0	2	2	0	0	0	1	0	1	0	0	0	1	1	0	5	1	15	30
	77	33	44	27	30	31	21	24	40	35	19	24	33	43	10	29	21	22	37	600

OneR /VQ/	Icon	Botticelli	Michelangelo	Raphael	Caravaggio	Rembrandt	Rubens	Friedrich	Goya	Turner	Monet	Pissarro	Sisley	Braque	Gris	Leger	Klimt	Miro	Mucha	
Icon	46	2	0	0	0	0	3	0	1	0	0	1	0	1	0	1	4	1	0	60
Botticelli	1	5	1	3	1	0	0	1	2	0	2	0	1	2	2	3	1	4	1	30
Michelangelo	0	3	18	5	0	0	0	0	1	0	0	0	0	0	1	2	0	0	0	30
Raphael	0	4	6	11	2	0	0	0	0	2	0	0	2	0	0	0	1	1	1	30
Caravaggio	1	1	1	0	16	3	1	0	0	0	0	1	0	2	1	0	0	3	0	30
Rembrandt	1	1	0	1	0	18	4	0	1	0	0	0	0	0	1	1	0	2	0	30
Rubens	8	0	0	0	0	0	17	0	1	0	2	0	0	0	0	1	1	0	0	30
Friedrich	0	0	0	2	0	2	0	15	1	1	1	2	2	0	0	1	1	2	0	30
Goya	1	1	1	1	0	0	0	1	14	0	4	1	3	2	0	0	0	1	0	30
Turner	0	1	1	3	0	0	0	6	2	13	0	0	2	0	0	0	0	2	0	30
Monet	1	1	0	2	0	0	0	0	0	0	11	2	4	1	2	2	2	2	0	30
Pissarro	4	1	0	1	0	0	0	1	0	0	3	12	2	1	1	1	3	0	0	30
Sisley	0	0	0	3	0	0	0	2	1	1	2	3	11	1	0	3	3	0	0	30
Braque	2	1	2	1	0	0	0	0	2	0	2	0	0	13	4	1	1	1	0	30
Gris	3	4	1	1	1	0	0	2	0	0	1	0	0	2	9	4	1	1	0	30
Leger	1	1	0	4	0	0	1	2	1	0	1	0	1	2	4	10	1	0	1	30
Klimt	5	0	1	2	1	0	0	1	1	0	2	1	0	2	2	2	7	1	2	30
Miro	5	1	0	0	1	0	0	2	0	0	2	0	0	0	1	0	0	17	1	30
Mucha	3	0	2	0	0	0	0	1	0	0	0	0	0	1	0	1	4	0	18	30
	82	27	34	40	22	23	26	34	28	17	33	23	28	30	28	33	30	38	24	600

References

[Agarwal, 2009] Agarwal, A.: Web 3.0 concepts explained in plain English. 30.05.2009. http://www.labnol.org/internet/web-3-concepts-explained/8908/

[Agrawal et al, 1993] Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. Proc. of the ACM SIGMOD Int. Conf. on Management of Data, Washington, DC, 1993, pp. 207-216.

[Agrawal, 2009] Agrawal, R.: Narrowing Down the Semantic Gap Between Content and Context Using Multimodal Keywords. PhD Thesis, Wayne State University, 2009, 135 p.

[Albers, 1963] Albers, J. Interaction of Colors. New Haven: Yale University Press, 1963, Rev. ed., 1975.

[Amato et al, 2004] 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: LNCS, Springer Berlin/Heidelberg, vol. 3232, 2004, pp. 14-25.

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

[Batko et al, 2009] Batko, M., Dohnal, V., Novak, D., Sedmidubsky, J.: MUFIN: A multi-feature indexing network. Second Int. Workshop on Similarity Search and Applications, 2009, pp. 158-159.

[Bayardo, 1997] Bayardo, R.: Brute-force mining of high-confidence classification rules. Third Int. Conf. on Knowledge Discovery and Data Mining (KDD'97), 1997, pp. 123-126.

[Bellman, 1961] Bellman, R.: Adaptive Control Processes: a Guided Tour. Princeton Univ. Press, 1961.

[Berezhnoy et al, 2005] Berezhnoy, I., Postma, E., van den Herik, H.: Authentic: computerized brushstroke analysis. Proc. IEEE Int. Conf. on Multimedia and Expo (ICME'05), 2005, pp. 1586-1588.

[Berezhnoy et al, 2007] Berezhnoy, I., Postma, E., van den Herik, H.: Computer analysis of Van Gogh's complementary colours. Pattern Recogn. Lett. 28, 6, 2007, pp. 703-709.

[Birren, 1981] Birren, F.: History of Colour in Painting. New York: Reinhold, 1981, 304 p.

[Boriah et al, 2008] Boriah, S., Chandola, V., Kumar, V.: Similarity Measures for Categorical Data: A Comparative Evaluation, In Proceedings of 2008 SIAM Data Mining Conference, 2008, Atlanta, pp. 243-254.

[Broek et al, 2006] van den Broek, E., Kok, T., Schouten, T., Hoenkamp, E.: Multimedia for Art ReTrieval (M4ART). Proc. of SPIE - Multimedia Content Analysis, Management, and Retrieval, 6073, 2006, pp. 1-12.

[Burford et al, 2003] Burford, B., Briggs, P., Eakins, J.: A taxonomy of the image: on the classification of content for image retrieval. Visual Communication, 2(2), 2003, pp. 123-161.

[Burgi and Monbaron, 2008] Burgi, P., Monbaron, P.: Content-based image retrieval integrated into Fedora. Third Int. Conf. on Open Repositories, Southampton, UK, 2008.

[Carson et al, 2002] Carson, C., Belongie, S., Greenspan, H., Malik J.: Blobworld: image segmentation using expectation-maximization and its application to image querying. IEEE Trans. on Pattern Analysis and Machine Intelligence, 24/8, 2002, pp. 1026-1038.

[Castelli and Bergman, 2002] Castelli, V., Bergman, L. (eds.): Image Databases: Search and Retrieval of Digital Imagery, John Wiley & Sons, 2002.

[Cha et al, 1999] Cha, S.-H., Shin, Y.-C., Srihari, S.: Algorithm for the Edit Distance between Angular Type Histograms. Technical report, St.Univ. of New York at Buffalo, 1999, 9 pp.

[Chavez et al, 2001] Chavez, E., Navarro, G., Baeza-Yates, R., Marroquin, J.: Searching in metric spaces. ACM Computing Surveys, 33(3), 2001, pp. 273-321.

[Chen et al, 2001] Chen, Y., Zhou, X., Huang T.: One-class SVM for learning in image retrieval. Proc. IEEE Int. Conf. on Image Processing, vol. 1, 2001, pp. 34-37.

[Chen et al, 2005a] Chen, C.-C., Wactlar, H., Wang, J., Kiernan, K.: Digital imagery for significant cultural and historical materials – An emerging research field bridging people, culture, and technologies. Int. J. Digital Libraries, 5, 4, 2005, pp. 275–286.

[Chen et al, 2005b] Chen, Y., Wang, J., Krovetz, R. CLUE: Cluster-based retrieval of images by unsupervised learning, IEEE Trans. on Image Processing, 14/8, 2005, pp. 1187-1201.

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

[Cohen, 1995] Cohen, W.: Fast effective rule induction. In Proc. of the 12th Int. Conf. on Machine Learning, Lake Taho, California, Morgan Kauffman, 1995.

[Colman, 2006] Colman A.: A Dictionary of Psychology, 2nd ed., Oxford University Press, Oxford, 2006.

[Colombo et al, 1999] Colombo, C., Del Bimbo, A., Pala, P.: Semantics in visual information retrieval. IEEE Trans. on Multimedia 6, 3, 1999, pp. 38-53.

[Comaniciu and Meer, 1999] Comaniciu, D., Meer, P.: Mean shift analysis and applications. Seventh Int. Conf. on Computer Vision, Kerkyra, Greece, 1999, pp. 1197-1203.

[Corridoni et al, 1999] Corridoni, J. M., del Bimbo, A., Pala, P.: Image retrieval by color semantics. Multimedia Systems, 7, 3, 1999, pp. 175-183.

[Croft, 1995] Croft, W.: What Do People Want from Information Retrieval? (The Top 10 Research Issues for Companies that Use and Sell IR Systems). Center for Intelligent Information Retrieval Computer Science Department, University of Massachusetts, Amherst, 1995.

[Crucianu et al, 2004] Crucianu, M., Ferecatu, M., Boujemaa, N.: Relevance feedback for image retrieval: a short survey. State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access and Interaction Including Data Models and Languages (DELOS2 Report), 2004, 20 p.

[Dasiapoulou et al, 2007] Dasiapoulou, S., Spyrou, E., Kompatsiaris, Y., Avrithis, Y., Stintzis, M.: Semantic processing of color images. Color Image Processing: Methods and Applications, ch. 11, CRC Press, Boca Raton, USA, 2007, pp. 259-284.

[Datta et al, 2006] Datta, R., Joshi, D., Li, J., Wang, J.: Studying aesthetics in photographic images using a computational approach. Proc. of the European Conf. on Computer Vision, Graz, Austria, LNCS, vol. 3953, Part III, 2006, pp. 288-301.

[Datta et al, 2008] Datta, R., Joshi, D., Li, J., and Wang, J. Z. 2008. Image retrieval: ideas, influences, and trends of the new age. ACM Computing Surveys, 40, 2, Article 5, 2008, 60 p.

[Datta, 2009] Datta, R.: Semantics and Aesthetic Inference for Image Search: Statistical Learning Approaches. PhD thesis, the Pennsylvania State University, 2009, 264 p.

[Daubechies, 1988] Daubechies, I.: Orthonormal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics. 41/7, 909-996, 1988.

[DCMI, 2009] Interoperability for Dublin Core Metadata http://dublincore.org/documents/interoperability-levels/

[de Berg et al, 2000] de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf O.: Computational Geometry: Algorithms and Applications (2nd revised ed.), Springer-Verlag, 2000, 367 p.

[Dehmer et al, 2006] Dehmer, M., Emmert-Streib, F., Wolkenhauer, O.: Perspectives of graph mining techniques. Rostocker Informatik Berichte, 30/2, 2006, pp. 47-57.

[Del Bimbo and Pala, 1997] Del Bimbo, A., Pala, P.: Visual querying by color perceptive regions. Fourth DELOS Workshop Image Indexing and Retrieval, San Miniato, 1997, pp. 2-6.

[Demartines and Herault, 1997] Demartines, P., Herault, J.: Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets. IEEE Trans. Neural Networks, 8(1), 1997, pp. 148-154.

[Demsar et al, 2004] 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.

[Deng and Manjunath, 2001] Deng, Y., Manjunath, B.: Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Analysis and Machine Intelligence, 23(8), 2001, pp. 800-810.

[Deza and Deza, 2009] Deza, M.-M., Deza, E.: Encyclopedia of Distances. Springer, XIV, 2009, 590 p.

[Dobreva and Chowdhury, 2010] Dobreva, M., Chowdhury, S.: A User-Centric Evaluation of the Europeana Digital Library. In: ICADL 2010, LNCS 6102, 2010, pp. 148-157.

[Eakins and Graham, 1999] Eakins, J., Graham, M.: Content-based Image Retrieval. University of Northumbria at Newcastle. Report: 39, JSC Technology Application Programme, 1999, 65 p.

[Eidenberger, 2003] Eidenberger, H.: Distance measures for MPEG-7-based retrieval. Fifth ACM SIGMM Int. Workshop on Multimedia Information Retrieval, 2003, pp. 130-137.

[Eiseman, 2006] Eiseman, L.: Color: Messages and Meanings. A PANTONE Color Resource, Hand Books Press, 2006, 146 p.

[Enser et al, 2006] Enser, P., Sandom, Ch., Lewis, P., Hare, J.: The reality of the semantic gap in image retrieval. Tutorial, 1st International Conference on Semantic and Digital Media Technologies, Athens, Greece, 2006. http://eprints.ecs.soton.ac.uk/13272/

[Esposito et al, 2002] Esposito, F., Malebra, D., Tamma, V., Bock H.: Classical resemblance measures, Analysis of Symbolic Data, Springer, 2002, pp. 139-152.

[Estrada, 2005] Estrada, F.: Advances in Computational Image Segmentation and Perceptual Grouping. PhD Thesis, Graduate Department of Computer Science, University of Toronto, 2005, 209 p.

[Fellbaum, 1998] Fellbaum, C. (Ed.): WorldNet: An Electronic Lexical Database. MIT Press, Massachusetts, USA, 1998.

[Flickner et al, 1995] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by image and video content: the QBIC system. Computer, September, 1995, pp. 23-32.

[Freeman, 1975] Freeman, J.: The modelling of spatial relations. Computer Graphics and Image Processing, 4/2, 1975, pp. 156-171.

[Gaede and Günther, 1998] Gaede, V., Günther, O.: Multidimensional access methods. ACM Computing Surveys, 30/2, 1998, pp. 170-231.

[Gage, 1993] Gage, J.:Colour and Culture: Practice and Meaning from Antiquity to Abstraction, Thames and Hudson, London, 1993.

[George, 2008] George, C., 2008. User-Centred Library Websites. Usability Evaluation Methods. Chandos publishing, 231 pp.

[Gevers, 2001] Gevers, T.: Color in image search engines. Principles of Visual Information Retrieval, ed. M. Lew, 2001, 34 p.

[Gheyas and Smith, 2010] Gheyas, I., Smith, L.: Feature subset selection in large dimensionality domains. Elsevier, Pattern Recognition, 43 (1), 2010, pp. 5-13.

[Goldhammer, 1981] Goldhammer K.: "Farbe, liturgisch" in Reallexikon zur Deutschen Kunstgeschichte. Published by the Zentralinstitut für Kunstgeschichte München, vol. VII, Munich, 1981.

[Gong et al, 1996] Gong, Y., Chuan, C., Xiaoyi, G.: Image indexing and retrieval using color histograms. Multimedia Tools and Applications, vol. 2, 1996, pp. 133-156.

[Gradmann, 2010] Gradmann, S., 2010. Knowledge = Information in Context: on the Importance of Semantic Contextualisation in Europeana. Europeana White Paper 1, 19 pp.

[Grosky et al, 2008] Grosky, W., Agrawal, R., Fotouchi, F.: Mind the gaps - finding the appropriate dimensional representation for semantic retrieval of multimedia assets. In Semantic Multimedia and Ontologies, Springer London, 2008, pp. 229-252.

[Han and Kamber, 2006] Han, J., Kamber, M.: Data Mining: Concepts and Techniques, Second ed., Morgan Kaufmann Publishers, 2006, 772 p.

[Hering, 1964] Hering, E. Outlines of a Theory of the Light Sense. Harvard University Press, Cambridge, Mass, 1964, 317 p.

[Herrmann, 2002] Herrmann, S.: MPEG-7 Reference Software. Munich University of Technology, 2002. http://www.lis.e-technik.tu-muenchen.de/research/bv/topics/mmdb/e_mpeg7.html

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

[Holtzschue, 2006] Holtzschue, L.: Understanding Colors, John Wiley & Sons, 2006, 196 p.

[Huber, 1985] Huber, P.: Projection pursuit. The Annals of Statistics, 13/2, 1985, pp. 435-475.

[Hung et al, 2007] Hung, Sh.-H., Chen, P.-H., Hong, J.-Sh., Cruz-Lara, S.: Context-based image retrieval: A case study in background image access for multimedia presentations. IADIS International Conference WWW/Internet 2007, Vila Real, Portugal, INRIA-00192463, ver.1, 2007, 5 p., http://hal.inria.fr/inria-00192463/en/

[Hurtut, 2010] Hurtut, T. 2D Artistic Images Analysis, a Content-based Survey. http://hal.archives-ouvertes.fr/hal-00459401_v1/

[Hurvich and Jameson, 1957] Hurvich, L.; Jameson, D.: An opponent-process theory of color vision. Psychological Review, 64 (6/I), 1957, pp. 384-404.

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

[ISO/IEC 15938-3] 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

[ISO/IEC JTC 1/SC 29/WG 11] WG11: The Moving Picture Experts Group "Coding of Moving Pictures and Audio" http://www.itscj.ipsj.or.jp/sc29/29w12911.htm

[Itten, 1961] Itten, J.: The Art of Color: the Subjective Experience and Objective Rationale of Color, Reinhold Publishing Corporation of New York, 1961.

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

[Ivanova et al, 2008] Ivanova, K., Stanchev, P., Dimitrov, B.: Analysis of the distributions of color characteristics in art painting images. Serdica Journal of Computing, 2/2, Sofia, 2008, pp. 111-136.

[Ivanova et al, 2010/ICDEM] 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. Second Int. Conf. on Data Engineering and Management ICDEM, India, 2010, LNCS 6411, pp. 146-153.

[Ivanova et al, 2010/IJAS] 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, 2010, pp. 474-484.

[Ivanova et al, 2010/MCIS] Ivanova, K., Stanchev, P., Vanhoof, K., Ein-Dor, Ph.: Semantic and abstraction content of art images. Proc. of Fifth Mediterranean Conf. on Information Systems, Tel Aviv, Israel, 12-14.09.2010, AIS Electronic Library, paper 42, http://aisel.aisnet.org/mcis2010/42.

[Ivanova et al-TR, 2010] Ivanova, K., Stanchev, P., Dobreva, S., Markov, K., Mitov, I.: Analysis of the Existing Practices for Features Extraction from Image Content. Tech. Report on Contract D002-308/2008, Institute of Mathematics and Informatics, Sofia, Bulgaria, 2010, 90 p.

[Jaimes and Chang, 2002] Jaimes, A., Chang, S.-F.: Concepts and techniques for indexing visual semantics. Image Databases: Search and Retrieval of Digital Imagery. John Wiley & Sons, 2002, pp. 497-565.

[Janson, 2004] Janson, A.: History of Art. Prentice Hall, 6th edition, 2004, 1032 p.

[Jörgensen, 2001] Jörgensen, C.: Introduction and overview. Journal of the American Society for Information Science and Technology, 52(11), 2001, pp. 906-910.

[Karypis, 2003] Karypis, G.: CLUTO: A Clustering Toolkit Release 2.1.1. Tech. Report: #02-017, University of Minnesota, Department of Computer Science, Minneapolis, 2003, 66 p.

[Kato, 1992] Kato, T.: Database architecture for content-based image retrieval. Proc. of the SPIE – The International Society for Optical Engineering, San Jose, CA, USA, vol. 1662, 1992, pp. 112-113.

[Keren, 2002] Keren, D.: Painter identification using local features and Naïve Bayes. Proc. of the 16th Int. Conf. on Pattern Recognition, Quebec, Canada, vol. 2, 2002, pp. 474-477.

[Kim and Lee, 2003] Kim, H., Lee, J.: HMMD color space and method for quantizing color using HMMD space and color spreading. United States Patent 6633407, 2003, http://www.freepatentsonline.com/6633407.html

[Koenig, 2010] Koenig, B.: Color Workbook, Prentice Hall, Third edition, 2010, 242 p.

[Koskela et al, 2002] Koskela, M., Laaksonen, J., Oja, E.: MPEG-7 descriptors in content-based image retrieval with PicSOM system. Recent Advances in Visual Information Systems, LNCS, Springer Berlin/Heidelberg, vol. 2314, 2002, pp. 69-92.

[Kuhn, 1955] Kuhn, H.: The Hungarian method for the assignment problem. Naval Research Logistics Quarterly, vol. 2, 1955, pp. 83-97.

[Li and Wang, 2004] Li, J., Wang, J.: Studying digital imagery of ancient paintings by mixtures of stochastic models. IEEE Trans. on Image Processing, 13/3, 2004, pp. 340-353.

[Li and Wang, 2008] Li, J., Wang J.: Real-time computerized annotation of pictures. IEEE Trans. on Pattern Analysis and Machine Intelligence, 30/6, 2008, pp. 985-1002.

[Lombardi et al, 2005] Lombardi, T., Sung-Hyuk, C., Tappert, Ch.: A lightweight image retrieval system for paintings. Proc. of SPIE – Volume Storage and Retrieval Methods and Applications for Multimedia, vol. 5682, 2005, pp.236-246.

[Lowe, 1999] Lowe, D.: Object recognition from local scale-invariant features. Proc. of the Int. Conf. on Computer Vision. 2, 1999, pp. 1150–1157.

[Lux and Chatzichristofis, 2008] Lux, M., Chatzichristofis, S.: Lire: Lucene Image REtrieval – an extensible Java CBIR library. Proc. of the 16th ACM Int. Conf. on Multimedia, Vancouver, Canada, 2008, pp. 1085-1088.

[Lux, 2009] Lux. M.: Caliph & Emir: MPEG-7 photo annotation and retrieval. Proc. of the 17th ACM Int. Conf. on Multimedia, Beijing, China, 2009, pp. 925-926.

[Luxburg, 2006] Luxburg, U.: A Tutorial on Spectral Clustering. Tech. Report No. TR-149, Max Planck Institute for Biological Cybernetics, 2006, 32 p.

[Maimon and Rokach, 2005] Maimon, O., Rokach, L.: Decomposition Methodology for Knowledge Discovery and Data Mining. Vol. 61 of Series in Machine Perception and Artificial Intelligence. World Scientific Press, 2005.

[Maitre et al, 2001] Maitre, H. Schmitt, F. Lahanier, C.: 15 years of image processing and the fine arts. Proc. of Int. Conf. on Image Processing, vol. 1, 2001, pp. 557-561.

[Marchenko et al, 2006] Marchenko, Y., Chua, T., Jain, R.: Semi-supervised Annotation of Brushwork in Painting Domain Using Serial Combinations of Multiple Experts. Tech. Report, NUS, Singapore, 2006.

[Marchenko et al, 2007] Marchenko, Y., Chua, T., Jain R.: Ontology-based annotation of paintings using transductive inference framework. Advances in Multimedia Modeling (MMM 2007), Springer, LNCS 4351, Part I, pp. 13-23.

[Markov et al, 2008] Markov, K., Ivanova, K., Mitov, I., Karastanev, S.: Advance of the access methods. Int. Journal Information Technologies and Knowledge, 2/2, 2008, pp. 123-135.

[Mattison, 2004] Mattison, D.: Looking for good art. Searcher – The Magazine for Database Professionals, vol. 12, 2004, part I – number 8, pp. 12-35; part II – number 9, pp. 8-19; part III – number 10, pp. 21-32.

[Mitov et al, 2009a] 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. In World Scientific Proceedings Series on Computer Engineering and Information Science, No: 2, pp. 279-286.

[Mitov et al, 2009b] 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. Int. Book Series Information Science & Computing – Book No: 14. New Trends in Intelligent Technologies, Sofia, 2009, pp. 31-39.

[Moënne-Loccoz, 2005] Moënne-Loccoz, N.: High-Dimensional Access Methods for Efficient Similarity Queries. Tech. Report No: 0505, University of Geneva, Computer Vision and Multimedia Laboratory, 2005, 27 p.

[MPEG-7:4062, 2001] MPEG-7, Visual experimentation model (xm) version 10.0. ISO/IEC/ JTC1/SC29/WG11, Doc. N4062, 2001.

[Munsell, 1969] Munsell, A.: A Grammar of Colors. New York, Van Nostrand Reinhold Co., 1969.

[Nasrabadi and King, 1988] Nasrabadi, N., King, R.: Image coding using vector quantization: a review. IEEE Trans. on Communications, 36/8, 1988, pp. 957-971.

[Natsev et al, 2007] Natsev, A., Tešić, J., Xie, L., Yan, R., Smith, J.: IBM multimedia search and retrieval system. Proc. of the 6th ACM Int. Conf. on Image and Video Retrieval (CIVR '07), Amsterdam, The Netherlands, 2007, pp. 645-645.

[Nieddu and Rizzi, 2003] Nieddu, L., Rizzi, A.: Proximity Measures in Symbolic Data Analysis. Statistica, vol.63, n. 2, 2003, pp. 195-212.

[OAIS, 2002] Reference Model for an Open Archival Information System (OAIS): Blue book. Consultative Committee for Space Data Systems, January 2002, 148 p.

[O'Connel et al, 2009] O'Connel, M., Airey, R., Craze, R.: The Ultimate Illustrated Encyclopedia of Signs, Symbols & Dream Interpretation. Hermes House, 2007, 512 p.

[Ooi et al, 1993] Ooi, B., Sacks-Davis, R., Han, J.: Indexing in Spatial Databases. Tech. Report. 1993, 71 p.

[Papageorgiou and Protopapas, 2003] Papageorgiou, H., Protopapas, A.: CIMWOS: A multimedia, multimodal and multilingual indexing and retrieval system. Digital Media Processing for Multimedia Interactive Services, World Scientific Publishing Co., 2003, pp. 563-568.

[Pavlov et al, 2010] Pavlov, R., Paneva-Marinova, D., Goynov, M., Pavlova-Draganova, L.: Services for Content Creation and Presentation in an Iconographical Digital Library. Serdica J. of Computing, 4/2, 2010, pp.279-292.

[Pavlova-Draganova et al, 2010] Pavlova-Draganova, L., Paneva-Marinova, D., Pavlov, R., Goynov, M.: On the Wider Accessibility of the Valuable Phenomena of the Orthodox Iconography through a Digital Library. Third Int. Euro-Mediterranean Conf. EuroMed, Lemesos, Cyprus, 2010, Archaeolingua, pp. 173-178.

[Perner, 2003] Perner, P.: Data Mining on Multimedia Data. Springer-Verlag NY, 2003, 141 p.

[Pickett et al, 2000] Pickett, J. (ed.): The American Heritage Dictionary of the English Language. Houghton Mifflin Co., 2000.

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

[Raychev, 2005] Raychev, R.: The Colors in the Art. Sofia. Lik Publ., 2005, 170 p. (in Bulgarian)

[Rubner et al, 1998] Rubner, Y., Tomasi, C., Guibas, L.: A metric for distributions with applications to image databases. In IEEE Int. Conf. on Computer Vision, 1998, pp. 59-66.

[Sablatnig et al, 1998] Sablatnig, R., Kammerer, P., Zolda, E.: Hierarchical classification of paintings using face- and brush stroke models. Proc. of the 14th IEEE Int. Conf. on Pattern Recognition, vol. 1, 1998, pp. 172-174.

[Saraceno et al, 1999] Saraceno, C., Reiter, M., Kammerer, P., Zolda, E., Kropatsch, W.: Pictorial portrait indexing using view-based eigen-eyes. Visual Information and Information Systems, LNCS, vol. 1614, 1999, pp. 649-656.

[Saul and Roweis, 2000] Saul, L., Roweis, S.: An Introduction to Locally Linear Embedding. Tech. Report, AT&T Labs and Gatsby Computational Neuroscience Unit, 2000, 13 p.

[Schallauer et al, 2006] Schallauer, P., Bailer, W., Thallinger, G.: A description infrastructure for audiovisual media processing systems based on MPEG-7. Journal of Universal Knowledge Management, 1/1, 2006, pp. 26-35.

[Shi and Malik, 2000] Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 22/8, 2000, pp. 888-905.

[Smeulders et al, 2000] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content based image retrieval at the end of the early years. IEEE Trans. on Pattern Analysis and Machine Intelligence, 22/12, 2000, pp. 1349-1380.

[Smith, 1978] Smith, A.: Color gamut transform pairs. Computer Graphics. 12/3, 1978, pp. 12-19.

[Smith, 2002] Smith, L.: A Tutorial on Principal Components Analysis, 2002. http://csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.

[Spillmann, 1992] Spillmann, W.: Color systems. Color Consulting, H. Linton, New York, 1992, pp. 169-183.

[Stanchev et al, 2006] Stanchev, P., Green Jr., D., Dimitrov, B.: Some issues in the art image database systems. Journal of Digital Information Management, 4/4, 2006, pp. 227-232.

[Striker and Dimai, 1997] Stricker, M., Dimai, A.: Spectral covariance and fuzzy regions for image indexing, Machine Vision and Applications, No:10, 1997, pp. 66-73.

[Tremeau et al, 2008] Tremeau, A., Tominaga, S., Plataniotis, K.: Color in image and video processing: most recent trends and future research directions. Journal Image Video Processing, vol. 1, 2008, pp. 1-26.

[Tu and Zhu, 2002] Tu, Z., Zhu, S.: Image segmentation by data-driven Markov chain Monte Carlo. IEEE Trans. Pattern Analysis and Machine Intelligence, 24/5, 2002, pp. 657-673.

[Veltkamp and Tanase, 2000] Veltkamp, R., Tanase, M.: Content-Based Image Retrieval Systems: A Survey. Technical Report UU-CS-2000-34, 2000, http://www.aa-lab.cs.uu.nl/cbirsurvey/cbir-survey/

[VizIR, 2003] VizIR: a framework for visual information retrieval. Institute of Software Technology and Interactive Systems. 2003. http://cbvr.ims.tuwien.ac.at/

[Vullo et al, 2010] Vullo, G., Innocenti, P., Ross, S.: Interoperability for digital repositories: towards a policy and quality framework. Fifth Int. Conf. on Open Repositories (OR2010), Madrid, Spain, 2010

[W3C, 2011] W3C: CSS Color Module Level 3. W3C Recommendation 07 June 2011. http://www.w3.org/TR/2011/REC-css3-color-20110607/, http://www.w3.org/TR/css3-color/#hsl-color

[Walch and Hope, 1995] Walch, M., Hope, A.: Living Colors: The Definitive Guide to Color Palettes Through the Ages. Chronicle Books, 1995, 164 p.

[Wang et al, 2001] Wang, J., Li, J., Wiederhold, G.: SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE Trans. on Pattern Analysis and Machine Intelligence, 23/9, 2001, pp. 947-963.

[Wang et al, 2002] Wang, J., Li, J., Chen, Ch. Interdisciplinary research to advance digital imagery indexing and retrieval technologies for Asian art and cultural heritages. Proc. of the 4th ACM Int. Workshop on Multimedia Information Retrieval, at the Int. Conf. on Multimedia, Juan Les Pins, France, 2002, 6 p.

[Wang et al, 2003] Wang, S., Chia, L., Deepu, R.: Efficient image retrieval using MPEG-7 descriptors. Int. Conf. on Image Processing, 2003, pp. 509-512.

[Wang et al, 2006] Wang, J., Boujemaa, N., del Bimbo, A., Geman, D., Hauptmann, A., Tesic, J. Diversity in multimedia information retrieval research. Proc. of the ACM SIGMM Int. Workshop on Multimedia Information Retrieval (MIR) at the Int. Conf. on Multimedia, 2006, pp. 5-12.

[Ward et al, 2005] Ward, A., Graham, M., Riley, K., Sheen, N.: Enhancing a historical digital art collection: evaluation of content-based image retrieval on Collage. Digital Art History: a Subject in Transition, Computers and the History of Art, vol. 1, 2005, pp. 101-112.

[Wei-ning et al, 2006] Wei-ning, W., Ying-lin, Y., Sheng-ming, J.: Image retrieval by emotional semantics: a study of emotional space and feature extraction. IEEE Int. Conf. on Systems, Man and Cybernetics, vol. 4, 2006, pp. 3534-3539.

[Witten and Frank, 2005] Witten, I., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition, Morgan Kaufmann, San Francisco, 2005, 525 p.

[Won et al, 2002] Won, Ch., Park, D., Park, S.: Efficient use of MPEG-7 edge histogram descriptor. ETRI Journal, 24/1, 2002, pp. 23-30.

[Yang et al, 2005] Yang, N., Dong, M., Fotouhi, F.: Semantic feedback for interactive image retrieval. Proc. of the 12th ACM Int. Conf. on Multimedia, 2005, pp. 415-418.

[Yang et al, 2008] Yang, N., Chang, W., Kuo, C., Li, T.: A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval, Journal of Visual Communication and Image Representation, 19/2, 2008, pp. 92-105.

[Zaiane and Antonie, 2005] Zaiane, O., Antonie, M.-L.: On pruning and tuning rules for associative classifiers. In Proc. of Int. Conf. on Knowledge-Based Intelligence Information & Engineering Systems, LNCS, Vol. 3683, 2005, pp.966-973.

[Zhou and Huang, 2003] Zhou, X., Huang, T.: Relevance feedback in image retrieval: a comprehensive review. Multimedia Systems, 8/6, 2003, pp. 536-544.

Vita – Krassimira Ivanova

Krassimira Minkova Ivanova was born on August 16, 1965 in Troyan, Bulgaria. After graduating from the secondary school – profile Mathematics, in her home town, from 1983 till 1988 she studied Mathematics at Sofia University "St. Kliment Ohridsky", where she obtained her Master of Science degree in Computer Science.

Since 1989 she has been working at the Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences where she is currently senior research assistant.

Her research interests are in Data Mining, Image Retrieval, Databases, Information Systems, and User-system Interaction applied in areas such as: Digitization of Cultural Heritage, Analysis of Economical Processes, Disaster Risk Management, and Education.

She obtained additional professional qualifications in Methodology of Teaching Mathematics (1992, Sofia University) and Accounting (1998, Private Economical School "BulProgress").

Till the end of 2010 she authored or co-authored 77 scientific publications in journals and peer-reviewed conferences.

Since 1990 she has been a principal co-developer of the Complex FOI, applied in numerous Bulgarian companies, which provided an excellent framework to investigate workflow, software, psychological, and social aspects of business informatics and human-computer interaction. She also served as a principal co-developer of the data mining environment system PaGaNe, which provides a convenient workplace for examining different tools supporting the process of knowledge discovery.

Since 2009 she is carrying out a doctoral research project at the Hasselt University, which outcomes are presented in this dissertation.

Publications

Overviews, ideas, realization and experiments, included in this phd thesis are presented and discussed in several scientific forums:

- Ivanova K., Stanchev, P., Dimitrov, B.: "Analysis of the distributions of color characteristics in art painting images", *Serdica Journal of Computing*, Vol.2, Num.2, Sofia, 2008, pp. 111-136.
- Ivanova K., Stanchev, P.: "Color harmonies and contrasts search in art image collections", *First International Conference on Advances in Multimedia MMEDIA 2009*, 20-25.07.2009, Colmar, France, pp. 180-187. ***Best paper award at MMEDIA 2009***.
- Ivanova, K., Stanchev, P., Vanhoof, K.: "Automatic tagging of art images with color harmonies and contrasts characteristics in art image collections", *International Journal on Advances in Softwar*e, Vol.3, No.3&4, 2010, pp. 474-484.
- 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", *Second International Conference on Data Engineering and Management ICDEM 2010*, 29-30.07.2010, India, LNCS, Vol. 6411, pp. 146-153.
- Ivanova, K., Stanchev, P., Vanhoof, K., Ein-Dor, Ph.: Semantic and abstraction content of art images. *Proc. of Fifth Mediterranean Conf. on Information Systems*, Tel Aviv, Israel, 12-14.09.2010, AIS Electronic Library, paper 42, http://aisel.aisnet.org/mcis2010/42

www.uhasselt.be

Universiteit Hasselt | Campus Diepenbeek
Agoralaan | Gebouw D | BE-3590 Diepenbeek | België
Tel.: +32(0)11 26 81 11