Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 37(4), 2010, Pages 37–49
ISSN: 1223-6934

A Framework for semantic modeling of images

Anca Loredana Ion

Abstract. In this paper we propose a framework for semantic modeling of images. The framework includes components for extraction of low-level image features, for content-based image retrieval and methods for incorporation of semantic knowledge into the retrieval process. The semantic information is represented by semantic association rules, which are used for an interactive annotation of the image data. So, the paper approaches modalities for reducing the semantic gap between the low-level characteristics automatically extracted from the visual content and the high-level concepts. The experiments through the framework were realized on collections of images from nature and medical domain.

2010 Mathematics Subject Classification. Primary 68U10; Secondary 68T10.
Key words and phrases. content-based visual retrieval, image annotation, association rules, image mining.

1. Introduction

The great challenges of the XXI century are the semantic aspects of the visual information. The manual annotation is subjective, great time consumer and leads to incomplete images description. So, new technologies are needed to reduce the costs and inefficiency of manual image annotation.

While the effort in solving the fundamental open problem of robust image understanding continued, we also see people from different fields, 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 content-based image retrieval (CBIR) community [3] [10].

One problem with all current approaches is the reliance on visual similarity for judging the semantic similarity, which may be problematic due to the semantic gap between low-level content and higher-level concepts [11]. Due to the great difficulty in recognizing and classifying the images, the methods that identify the semantic features of images recorded a great success. A lot of researches were developed to discover automatic techniques that could generate the semantic description for multimedia content.

Two of the fundamental directions are:

- Methods based on machine learning; these methods manually annotate the training set of data for generating graphs, statistical models or other indexing techniques for big data collections.

Received September 05, 2010. Revision received October 14, 2010.
This work was supported by the strategic grant POSDRU/89/1.5/S/61968, Project ID 61968 (2009), co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 - 2013.

- Ontology for multimedia data; these methods use the relationships described by the ontology or thesaurus for permitting semantic queries on annotated multimedia data to infer new information.

The remainder of this paper discusses the problem description of the image automatic annotation in Section 2. Section 3 introduces the architecture of our system and describes the interactions between its components. A comparative study on image visual representation is presented in Section 4. The capturing of semantic knowledge and steps required to generate semantic image representations are detailed in Section 5. Finally, in Section 6 we present a summary of the presented approach.

2. Problem description

Retrieval by image content has received great attention in the last decades. Although there exist several CBIR techniques, there are still many unsolved issues in content-based retrieval systems: the difficulty of complete semantic concepts extraction from images, meaning the objects recognition and understanding, known as the problem of semantic concepts extraction [3]; the complexity, ambiguity and subjectivity of human interpretation, known as the problem of semantic concepts interpretation [3].

The problem of semantic concepts extraction appears because a real object can have more appearances, due to physical aspects like light conditions, parameters of the capturing devices like focus, angle, distance, or different expressions, positions, etc. Also, the same real object could have not the same colour, texture or shape. The problem of semantic concepts interpretation is due to a lot of factors, like the cultural differences, education, which affects the user interpretation model. Also, the human perception and judgement are not time invariant.

This study was started from the limitations regarding the researches in multimedia semantic modeling. This paper proposes new approaches for image annotations, like: methods for generation of rules which identify image categories, a method for mapping low-level features to semantic indicators, using the Prolog declarative language, the creation of a representation image vocabulary and syntax, and semantic image classification.

3. Architecture

The principal objective of our annotation system, SIRS, is to provide users an image retrieval system with the capabilities of image classification and assignment of the data to semantic concepts. By analyzing the visual structure of already annotated images, the system provides a semi-automatic annotation which generates descriptions for a new and unlabeled image. Since the system is working semi-automatically, it depends on an expert at the mapping processing step. Figure 1 illustrates the components of our SIRS semantic retrieval system.

Feature Extraction Component. This component mainly provides methods for extracting primitive (visual) characteristics of images. For example, the set of low-level features implemented in SIRS system currently includes color, shape and texture features.

Image Segmentation Component. In order to find out the semantic relations between image categories and 'objects' contained in an image, it should be divided

into objects. For that purpose, an automatic segmentation algorithm based on color-set back projection approach [12] is provided by the segmentation component. The segmentation approach is based on color homogeneity criterion.

Content-Based Visual Retrieval Component. This component manages the visual retrieval process. Beginning with query formulation by image example, the component provides functions for similarity computation between the query object and the image data stored in the database.

Vocabulary Component. This component provides an intermediate level between visual and semantic representation of images. The semantic indicators are the visual elements of the vocabulary and their values are experimentally and manually established.

Semantic Image Annotation Component. This component has the role of deciding categories for new unlabelled images. It includes the following subcomponents:

a.Image Mapping Component. This component has the role of mapping the low-level features of images to semantic indicators.

b.Rules Generation and Classification Component. This component provides methods to generate semantic rules used to recognize the semantic concepts of unlabeled images.

Storage Component. This component manages the physically storage of images and their features.

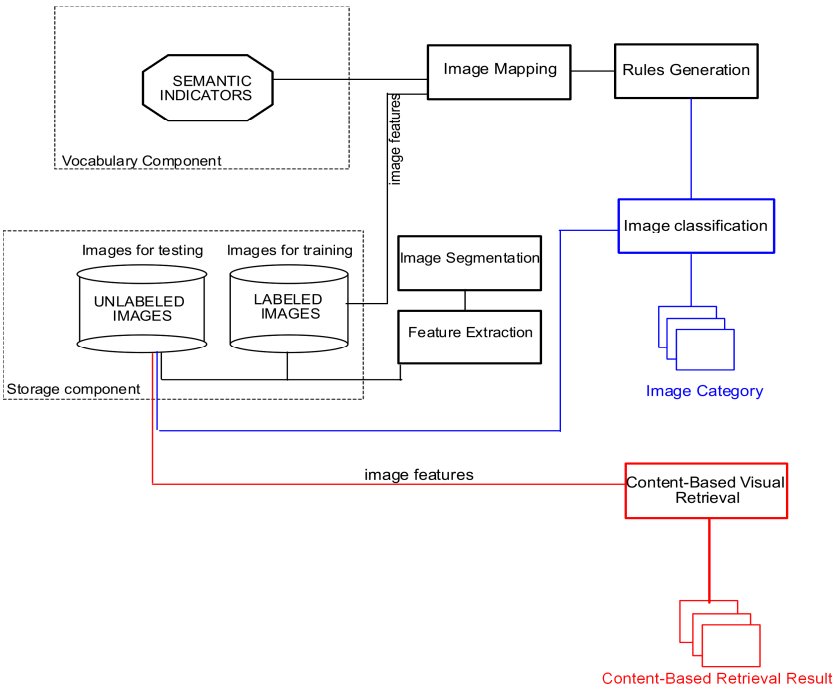

FIGURE 1. Architecture of the SIRS system. Interaction between system components.

4. Content-based image retrieval

Combining knowledge from different areas, like databases, computer vision, artificial intelligence, mathematics, the content-based visual retrieval provides new paradigms and methods for multimedia retrieval.

In our framework, we develop a study of comparing different representations of image visual features, to select the ones with the best retrieval performances and that are not limited to any particular domain. To find a set of performing visual image descriptors for heterogeneous domains (nature, medicine, entertainment, sport, etc.) is a very difficult task, because there is not apriori knowledge to be used.

The motivation of realizing this comparative study is the necessity of finding a good set of visual image descriptors, as a precondition for the accuracy of the retrieval and annotation processing.

The most important characteristics of image databases are: color, texture, and shape. In this section, we present these features, the relationship between them and the results returned by the content-based visual retrieval component of the framework.

4.1. Color feature. Color is a very important feature in many image domains and is the most used feature in the content-based image retrieval systems, because the color characteristic is easy to be detected from images and objects. More, the color is invariant to orientation and scaling and the analysis by color is intuitive. The performance and efficiency of the color feature for characterizing the perceptual similitude is strongly influenced by the selection of the color space and its quantization.

The developed comparative study of methods for color representation includes the following descriptors, some of them from MPEG-7 standard. The first representation is the colour histogram represented in HSV color space quantized at 166 colors [12]. The HSV color space offers an intuitive representation of color and approximates the way in which humans visualize and manipulate color. The transformation from RGB to HSV color space is nonlinear, but irreversible. The quantization of HSV color space provides 18 colors, 3 saturations, 3 values, and 4 grey levels, meaning 18x3x3+4=166 distinct colors in HSV color space [12].

The second descriptors is the MPEG-7 - color structure descriptor (CSD) represented in HMMD color space quantized at 128 colors [7]. In HMMD color space, the nuance (H) has the same definition as the "hue" channel of the HSV color space, and the max (M) and min(m) channels are the maximum, respectively minimum between the values of R, G, B channels. The diff(D) component is defined as the difference between the maximum(M) and minimum(m) values. The CSD descriptor represents the local color structure from an image, counting the number of appearances of a color into a block of dimension 8x8, which scans the image. The histogram represents the number of block in which appears each of the 128 quantized colors.

The third descriptor is the dominant color descriptor represented in CIE-LUV color space [7]. The method for extracting the dominant color is based on K-means algorithm [2]. The descriptor includes the dominant colors, their percentage, and their spatial coherency.

The average normalized modified retrieval rate (ANMRR) [7] is computed for testing the efficiency and performance of the color descriptors. To compare the developed methods, the following conditions were set:

1. The database of about 1200 images is created, including images from nature [16] and medical collections [15]: clouds, flowers, sunset, sunrise, leaves, trees, duodenal ulcer, gastric ulcer, gastric cancer, esophagitis, rectocolitis, etc.

TABLE 1. The average normalized modified rate for color feature

Category	HSV166	HMMD128	Dominant Color
Flower	0.2588	0.259	0.2713
Leaves	0.2620	0.262	0.265
Clouds	0.2060	0.2073	0.2236
Sunset	0.1941	0.1947	0.197
Sunrise	0.1940	0.1943	0.197
Cliff	0.192	0.192	0.195
Mountains	0.2	0.26	0.27
Desert	0.19	0.194	0.195
Duodenal ulcer	0.2267	0.2267	0.29
Gastric ulcer	0.22	0.21	0.3
Gastric cancer	0.19	0.20	0.3

2. The three color descriptors are computed for each image from the database.
3. The relevant images were established for each experiment.
4. For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 1.

By analysing the results of the content-based retrieval using the color feature, the best results were recorded using the color histogram in the HSV color space and the color histogram in HMMD color space.

4.2. Texture feature. The texture is another important characteristic taken into consideration for classifying and recognizing the material surfaces. In the current framework, two methods are developed: Gabor filter and co-occurrence matrices. For the first method, we interpret the hue and saturation channels like polar coordinates to allow the direct use of the HSV color space for Fourier transform [8]. This technique is used for the extraction of Gabor characteristics for color texture.

The color space HSV is a non-linear transformation of the RGB color space. The H, S, V components closely correspond to the human color perceptions. Starting from the representation of the HSV color space, we may represent the color in complex. The affix of any point from the cone base can be computed as: z_M = S (cosH + i sinH). So, the saturation is interpreted as the magnitude and the hue as the phase of the complex value b; the value channel is not included [8]. The advantages of this representation of complex color are the simplicity due to the fact that the color is now a scalar and not a vector, and the combination between channels is done before filtering. So, the color can be represented in complex as in equation (1):

$$b(x,y) = S(x,y) \cdot e^{iH(x,y)} \tag{1}$$

The computation of Gabor characteristics for the image represented in the HS-complex space is similar with the one for the unichrome Gabor characteristics, because the combination of color channels is done before filtering, as in equation (2):

$$C_{f,\phi} = (\sum_{x,y}(FFT^{-1}P(u,v) \cdot M_{f,\phi}(u,v)))^2 \tag{2}$$

The Gabor characteristics vector is created using the value $C_{f,\phi}$ computed for 3 scales and 4 orientations, as in equation (3):

TABLE 2. The average normalized modified rate for texture feature.

Category	Co-occurrence Matrix	Gabor Filter
Sunrise	0.232	0.239
Leaves	0.3037	0.3047
clouds	0.2883	0.4
Sunset	0.23	0.35
Cliff	0.19	0.2
Mountains	0.19	0.19
Desert	0.15	0.20
Fire	0.22	0.22
See	0.31	0.31
Duodenal ulcer	0.236	0.259
Gastric ulcer	0.24	0.24
Gastric cancer	0.231	0.235

$$V = (C_{0,0}, C_{0,1}, ...C_{2,3}) \tag{3}$$

The similitude between the texture characteristics of the query image Q and the target image T is defined by the metric defined in equation (4):

$$D^2(Q,T) = \sum_f \sum_\phi d_{f\phi}(Q,T) \tag{4}$$

where $d_{f\phi}(Q,T) = (V^Q - V^T)^2$.

The co-occurrence matrix is another method studied and implemented in this software system. The co-occurrence matrix is based on the repeated occurrence of some configurations of pixels intensity in the texture. These configurations vary with rapidity for thin texture and slower for the roughly texture. The occurrence of these intensity configurations can be described by the co-occurrence matrices.

In the case of color images, one matrix was computed for each of the three channels (R, G, B). For an image f(x, y), the co-occurrence matrix $h_{d\phi}(i,j)$ is defined so that each entry (i, j) is equal to the number of times for that $f(x_1, y_1) = i$ and $f(x_2, y_2) = j$, where $(x_2, y_2) = (x_1, y_1) + (dcos\phi, dsin\phi)$. This leads to three quadratic matrices of dimension equal to the number of the color levels presented in an image (256 in our case) for distance d=1 and orientation $\phi = 0$.

The classification of texture is based on the characteristics extracted from the co-occurrence matrices [13]: energy, entropy, maximum probability, contrast, cluster shade, cluster prominence, correlation.

The average normalized modified retrieval rate (ANMRR) [7] is computed for testing the efficiency and performance of the texture descriptors. To compare the developed methods, the following conditions were set:

1. The database of about 1200 images is created, including images from nature [16] and medical collections [15]: clouds, flowers, sunset, sunrise, leaves, trees, ulcer, gastric cancer, etc.

2. The two texture descriptors are computed for each image from the database.

3. The relevant images were established for each experiment.

4. For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 2.

By analysing the results of the content-based retrieval using the texture feature, the best results are obtained by the co-occurrence matrices.

4.3. Shape feature. The region methods for shape detection are the most promising descriptors for image retrieval based on shape [4]. The region-based methods use the information along the contour and from inside shape, so the edge detection is not necessary.

In this framework, three shape descriptors were developed: the geometric moment(GMD), the Zernike moment and the eccentricity. The technique based on geometric moment invariants for shape representation is extremely used in object recognition [9]. The geometric moment invariants are derived from shape moments that are invariant to 2D shape geometric transformation. The central moments of (p+q) order of a bi-dimensional shape represented by f(x, y) function are given by equation (5):

$$\mu_{pq} = \sum_{p} \sum_{q} (x - \overline{x})^p (y - \overline{y})^q f(x, y) \tag{5}$$

where $\overline{x} = \mu_{10}/m$, $\overline{y} = \mu_{01}/m$, and m is is the weight of the region shape. A vector consisting of seven moments, invariant to translation, rotation and scaling, are used to index each shape in the database. The distance between two feature vectors is determined by the city block distance. The values of computed moment invariants are usually small, and the values of moment invariants of higher order are closer to zero in some cases. The advantage of GMD is because it offers a compact representation of shape and the computation is low.

Zernike moments are another method used with success for shape detection. Teague [14] proposed the utilization of orthogonal moments for image reconstruction and introduced the Zernike moments that permit the independent construction of invariant moments of arbitrary order. The complex Zernike moments are derived from Zernike polynomials:

$$V_{nm}(x, y) = V_{nm}(\rho cos\theta, \rho sin\theta) = R_{nm}(\rho) exp(j\omega\theta) \tag{6}$$

where R_{nm} is expressed as in equation (7):

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2} - s)!(\frac{n-|m|}{2} - s)!} \rho^{n-2s} \tag{7}$$

where ρ is the radius from (x, y) to the centroid of the shape, θ is the angle between the radius ρ and the axis x, n and m are integers such that: $n - |m| = par$, $|m| \leq n$. The Zernike polynomials are a complete set of orthogonal functions with complex values over the entire unity circle, i.e. $x_2 + y_2 = 1$. The Zernike complex moments of m order with n repetitions are defined as equation (8):

$$A_{nm} = \frac{n+1}{\pi} \sum_{x} \sum_{y} f(x, y) V^*_{nm}(x, y) \tag{8}$$

where $x^2 + y^2 \leq 1$ and * means the complex conjugate.

The precision of the shape representation depends on the number of moments. We took into consideration the first 36 moments of maximum order 10 to equilibrate the efficiency and precision. The similitude between two shapes indexed by Zernike descriptors is determined by the city block distance between the two corresponding vectors of Zernike moments.

TABLE 3. The average normalized modified rate for shape feature.

Category	GM	Zernike	Eccentricity
Sunrise	0.1930	0.198	0.172
Leaves	0.1897	0.199	0.174
Clouds	0.17	0.19	0.17
Sunset	0.1930	0.198	0.172
Cliff	0.1942	0.1943	0.1941
Mountains	0.23	0.35	0.19
Desert	0.21	0.26	0.19
Fire	0.198	0.198	0.196
See	0.22	0.226	0.21
Duodenal ulcer	0.23	0.28	0.198
Gastric ulcer	0.23	0.27	0.198
Gastric cancer	0.21	0.27	0.2

Eccentricity descriptor is another global descriptor taken into account, that characterizes the shape at a general level [4]:

$$eccentricity = \frac{I_{min}}{I_{max}} = \frac{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2}}{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \tag{9}$$

where μ_{pq} is the central moment of order (p, q).

The average normalized modified retrieval rate (ANMRR) [7] is computed for testing the efficiency and performance of the shape descriptors. To compare the developed methods, the following conditions were set:

1. The database of about 1200 images is created, including images from nature [16] and medical collections [15]: clouds, flowers, sunset, sunrise, leaves, trees, ulcer, gastric cancer, etc.
2. The three shape descriptors are computed for each image from the database.
3. The relevant images were established for each experiment.
4. For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 3.

By analysing the results of the content-based retrieval from Table 3, we observe that all three methods have proximate results.

5. Capturing semantic knowledge

The image annotation process means the correlation between keywords and images for capturing their semantic contents. The annotation process should assign the image data to one or more predefined categories. For that, the proposed method uses a vocabulary-based knowledge and semantic rules, which serve as a source of semantic types and their relations.

Accordingly, the image annotation process includes the following steps:

- The automatic segmentation of images and the indexing of resulted colour regions.
- The mapping of visual features of images to semantic indicators.
- The definition of a knowledge database, using the declarative language, Prolog, which facilitates the mapping process.

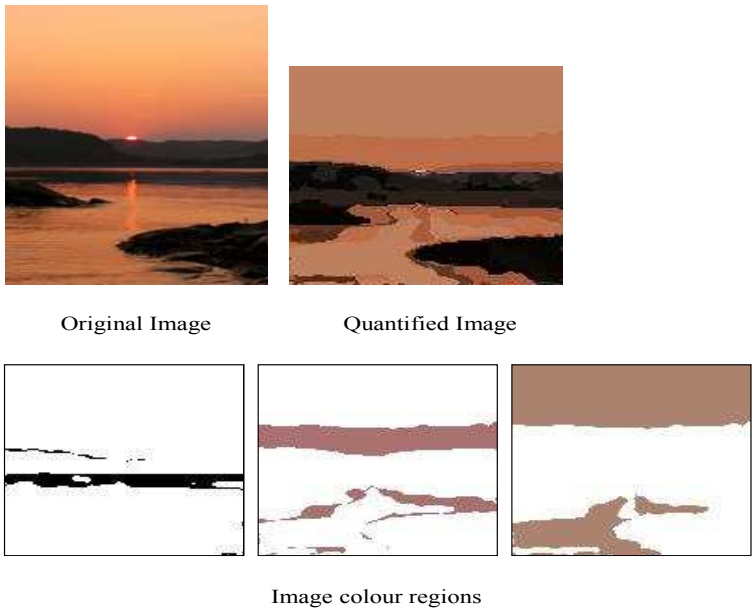

FIGURE 2. Colour segmented regions of the "sunset" image

- The automatic discovery of semantic inference rules for discovery the semantic concepts from images.
- The representation of the semantic rules, using the declarative language, Prolog, to easier infers them to any domain.

5.1. Rule generation and image annotation. The selection of the visual feature set and the image segmentation algorithm are the definitive stage for the semantic annotation process of images. After we performed a large set of experiments we inferred the importance of semantic concepts in establishing the similitude between images. Even if the semantic concepts are not directly related to the visual features (colour, texture, shape, position, dimension, etc.), these attributes capture the information about the semantic meaning.

Using the results of experiments realized in Section 4, the HSV colour space quantized at 166 colours is used for representing the colour features, the co-occurrence matrix is used for representing the texture features, and the eccentricity is used for representing the shape features.

Before segmentation, the images are transformed from RGB to HSV colour space and quantized to 166 colours. The colour regions extraction is realized with the colour set back projection algorithm [12]. This algorithm detects the regions of a single colour [5]. The results of the segmentation algorithm on the "sunset" image includes four regions with colors black, red and two browns and can be observed in figure 2:

Each region is described by the following visual characteristics:

-The colour characteristics are represented in the HSV colour space quantized at 166 colours. A region is represented by a colour index, which is in fact an integer number between 0..165.

-The spatial coherency represents the region descriptor which measures the spatial compactness of the pixels of same colour.

-A seven-dimension vector (maximum probability, energy, entropy, contrast, cluster shade, cluster prominence, correlation) represents the texture characteristic.

-The region dimension descriptor represents the number of pixels from region.

-The spatial information is represented by the centroid coordinates of the region and by minimum bounded rectangle.

-An eccentricity represents the shape feature.

For mapping the visual features of images to semantic interpretation, a vocabulary and a syntax are used [6]. The vocabulary words are limited to the name of semantic indicators, which are visual elements, namely: the colour semantic indicator with values like light-red, dark-green, medium yellow, a.s.o, the spatial coherency semantic indicator with values like weak, medium, strong, the texture semantic indicator with values like energy-small, energy-medium, energy-big, probability-medium, a.s.o, the dimension semantic indicator with values like small, medium, big, a.s.o., the position semantic indicator with values like vertical-center, horizontal- upper, a.s.o, the shape semantic indicator small, big, medium. The syntax is represented by the model, which describes the images in terms of semantic indicators values. The values of each semantic descriptor are mapped to a value domain, which corresponds to the mathematical descriptor.

The representation with semantic indicators of the "sunset" image from Figure 2 is:

```
image([
    region([
        descriptor(color, dark-black),
        descriptor(horizontal, center),
        descriptor(vertical, center),
        descriptor(dimension, small),
        descriptor(eccentricity, big),
        descriptor(texture-probability, big),
        descriptor(texture-difference, big),
        descriptor(texture-entropy, medium),
        descriptor(texture-energy, big),
        descriptor(texture-contrast,big),
        descriptor(texture-correlation,big)]),
    region([
        descriptor(color,medium-red),
        descriptor(horizontal, center),
        descriptor(vertical, center),
        descriptor(dimension, medium),
        descriptor(eccentricity, small),
        descriptor(texture-probability, medium),
        descriptor(texture-difference, medium),
        descriptor(texture-entropy, big),
        ..,
    region([
        descriptor(color, medium-brown),
        descriptor(horizontal, left),
        descriptor(vertical, bottom),
```

```
    descriptor(dimension, medium),
    ..,
region([
    descriptor(color, medium-brown),
    descriptor(horizontal, center),
    descriptor(vertical, bottom),
    descriptor(dimension, big),
    ..
    ])])
```

The algorithms for semantic rules generation are based on A-priori algorithm of finding the frequent itemsets [1]. The scope of image association rules is to find semantic relationships between image objects.

In our system, the learning of semantic rules is continuously made, because when a categorized image is added in the learning database, the system continues the process of rules generation. The methods used in this study bring important improvements related to the detailed descriptions of images, which are necessary for defining relationships between objects/regions, classes of visual characteristic, objects/regions and classes of visual characteristics. Two methods for semantic rule generation are developed in this study, because we considered that the colour could be the most discriminated features for some image categories/keywords, while for the others, all the features could be necessary to describe them. For the first variant of the algorithm, the images and regions are modeled in the terms of itemsets and transactions. The image set with the same category represents the transactions and the itemsets are the colours of image regions. The frequent itemsets represent the itemsets with support bigger or equal than the minimum support. A subset of frequent itemsets is also frequent. The itemsets of cardinality between 1 and k are iteratively found (k-length itemsets). The frequent itemsets are used for rule generation.

A semantic rule generated by this method has the following form:

C_1 (union of semantic indicators of regions with color C_1) and ... and C_n (union of semantic indicators of regions with color C_n) $\rightarrow$ *category*

In the second variant (Algorithm 3.4), the rule generation algorithm takes into account all region features, not only the colour, as in the first variant. This algorithm is based on "region patterns" and necessitates some computations, being necessary a pre-processing phase for determining the visual similitude between the image regions from the same category.

In the pre-processing phase, the region patterns, which appear in the images, are determined. So, each image region Reg_{ij} is compared with other image regions from the same categories. If the region Reg_{ij} matches other region Reg_{km}, having in common the features on the positions $n_1, n_2, , n_c$, then the generated region pattern is $SR_j(-,-,-,n_1,n_2,,n_c,-,-)$, and the other features are ignored.

A semantic rule generated by this method has the following form:

$SR_1(-,-,n_1^{'},n_2^{'},..n_c^{'},-,-)$ and,.., and $SR_n(-,-,n_1,n_2,..n_c,..) \rightarrow$ *category*

The generated rules identify the image category or category keywords. Each semantic rule is associated by support that represents the percent of transactions that contain both the precedent and the head and the body of the rule and confidence that represents the ratio between the number of transactions that contain both head and the body of the rule and the number of transactions that contain only the head of the rule.

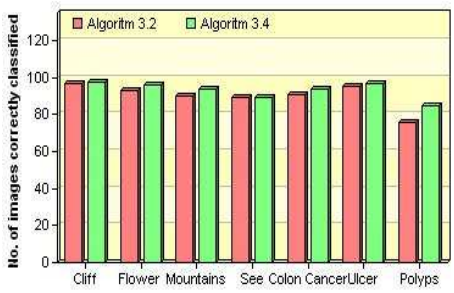

FIGURE 3. Category vs. percent of images correctly classified by the system using Algorithm 3.2 and Algorithm 3.4

Before classification, the new unlabeled image is automatically processed, the semantic representation of the image being generated, using the Prolog [6]. A semantic rule matches an image, if all the characteristics, which appear in the body of the rule, also appear in the image characteristics. The semantic rules that match the image are sorted based on confidence, in descending order. The semantic rules with the confidence bigger than the minimum confidence are the ones that classified the image.

For testing the efficiency and performance of rules generation (Algorithm 3.2 and Algorithm 3.4), for each image category, the percent of images correctly classified by the two algorithms is computed. The minimum support was established to 40% and the minimum confidence to 70%.

In the experiments realized through this study, two databases are used for the learning and testing process, so the inference of the learned semantic rules on other images than the ones used in the learning process is much more difficult.

The database used for learning contains 400 images from different nature categories and medical domains and is used to learn the correlations between images and semantic concepts. The database used in the learning process is categorized into 60 semantic concepts and diagnoses.The database used for testing contains 800 images from different nature categories and medical domains. The keywords are manually added to each category, for describing the images used for learning process. The descriptions of these images are made from simple concepts like "flower, mushrooms" to complex ones "mountains, cliff, lake". In average, 3.5 keywords were used to describe each category. The process of manual annotation of images used for learning semantic rules took about 7 hours.

For testing the efficiency and performance of rules generation (Algorithm 3.2 and Algorithm 3.4), for each image category, the percent of images correctly classified by the two algorithms is computed, as can be vizualise in Figure 3.

From the experiments, we deduce that Algorithm 3.4 records better results, because it selects the images characterized with greater apparition probability and it offers greater generality.

6. Conclusion

The main motivation for the developed work is the fact that users are highly interested in querying images at conceptual and semantic level, not only in terms of low-level features. The need of enhancement of the retrieval performance and

the importance of 'semantic meaning' makes a detailed image annotation necessary. Presently, most of the image database systems utilize manual annotation, where users assign some descriptive keywords to images. Although this process eliminates the uncertainty of fully automatic annotation, it requires a high effort in exchange. In summary, since it is very difficult to automatically construct semantic knowledge from the extracted low-level features and map them on human perception, methods which combine both approaches are of great interest.

References

[1] R. Agrawal, T. Imielinski and A. Swami, Mining Association Rules between Sets of Items in Large Databases, *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* (1993), Washington, DC, 207–216.

[2] A. Berson and S.J. Smith, *Data Warehousing, Data Mining, and OLAP*, McGraw-Hill, New York, 1997.

[3] G. Carneiro, A. Chan, P. Moreno and N. Vasconcelos, Supervised learning of semantic classes for image annotation and retrieval, *IEEE Pattern Anal. Machine Intelligence* **29** (2007), no. 3, 394–410.

[4] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision*, Addition-Wesley, 1992.

[5] A.L. Ion, Methods for Knowledge Discovery in Images, *Information Technology and Control* **38** (2009), no.1, 43–49.

[6] A.L. Ion, Image Annotation Based on Semantic Rules, *Human-Computer Systems Interaction: Backgrounds and Applications* **60** (2009), 83–94.

[7] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan and A. Yamada, Color and Texture Descriptors, *IEEE Trans. Circuits and Systems for Video Technology* **11** (2001), no. 6, 703–715.

[8] C. Palm, D. Keysers, T. Lehmann and K. Spitzer, Gabor Filtering of Complex Hue/Saturation Images for Color Texture Classification, *JCIS 2000, Joint Conference on Information Sciences - International Conference on Computer Vision, Pattern Recognition, and Image Processing*, Atlantic City, NJ, USA, **2** (2000), 45–49.

[9] R.J. Prokop and A.P. Reeves, A Survey of Moment-based Techniques for Unoccluded Object Representation and Recognition, *Graphical Models and Image Processing* **54** (1992), 438–460.

[10] N. Rasiwasia, P.J. Moreno and N. Vasconcelos, Bridging the Gap: Query by Semantic Example, *IEEE Transactions On Multimedia* **9** (2007), no. 5, 923–938.

[11] A. Smeulders, M.Worring, S. Santini, A. Gupta and R. Jain, Content-based image retrieval: The end of the early years, *IEEE Trans. Pattern Anal. Machine Intelligence* **22** (2000), no. 12, 1349–1380.

[12] J.R. Smith and S.-F. Chang, VisualSEEk: a fully automated content-based image query system, *The Fourth ACM International Multimedia Conference and Exhibition*, Boston, MA, USA, 1996.

[13] H.Tamura, S. Mori and T. Yamawaki, Texture Features Corresponding to Visual Perception, *IEEE Transactions on Systems, Man, and Cybernetics* **8** (1978), no. 6, 1264–1274.

[14] M.R. Teague, Image Analysis Via the General theory of Moments, *Journal of Optical Society of America* **70** (1980), no. 8, 920–930.

[15] The Gatrolab Image Library, `http://www.gastrolab.net/`.

[16] ImagesFree, Free Images Collection, `http://imagesfree.net/`.

(Anca Loredana Ion) Department of Software Engineering, Faculty of Automation, Computers and Electronics, Bvd. Decebal, Nr. 107, 200440, Craiova, Dolj, ROMANIA
E-mail address: `anca.ion@software.ucv.ro`