RULE BASED MODELLING OF IMAGES SEMANTIC CONCEPTS

Stefan Udristoiu, Anca Ion and Dan Mancas
Faculty of Automation, Computers and Electronics, University of Craiova, Bvd.Decebal 107, Craiova, Romania

Keywords: Image Annotation, Semantic Association Rules, Visual Features, Colour, Texture and Shape.

Abstract: In this paper we study the possibilities to discover correlations between visual primitive and high-level characteristics of images, meaning the extraction of semantic concepts. The design and developing of algorithms for image semantic annotation are the main contribution of this paper. The proposed methods are based on developing algorithms that automatically discover semantic rules to identify image categories. A semantic rule is a combination of semantic indicator values that identifies semantic concepts of images. Some models for representing the images and rules are also developed. The annotation methods are not limited to any specific domain and they can be applied in any field of digital imagery.

1 INTRODUCTION

The semantic annotation of visual resources is fundamental for retrieving the quantity of visual digital content, which speedily grows. Such descriptions facilitate the semantic query of multimedia data in terms familiar to the user of a certain domain, permitting the discovery and exploitation of information and knowledge by services, agents and web applications (Hoogs et al., 2003).

Because of visual data quantity and complexity, their annotation is a big time consumer, expensive and very subjective. Even if in the last two decades a big number of techniques were developed, the generation of image semantic annotation remains a significant challenge.

The correlation between the low-level features and high-level concepts is a challenge due to the "semantic gap" (Smeulders et al., 2000).

The fundamental scope of image retrieval is to provide efficient and simple modalities for searching in the image databases (Carneiro et al., 2007). As it is mentioned above, it is difficult to get this target using the traditional retrieval image systems, which do not take into account the semantic aspects of images.

In this paper we describe an automatic method, assisted by user, for generating annotations based on visual features of image regions. The described prototype permits the expert users to generate rules specific to their domain, by submitting to the system significant categorized images from which the system can learn the rules. The semantic rules map the combinations of visual characteristics (colour, texture, shape, position, etc.) to semantic concepts.

The remainder of this paper is structured as follows. Section 2 presents the selection of visual features and the segmentation algorithm. Section 3 presents the mapping between visual features and semantic concepts. Section 4 details the generation of semantic rules, the semantic classification of images, and discusses the experiments. Finally, section 5 summarizes the conclusions of this study.

2 THE IMAGE SEGMENTATION

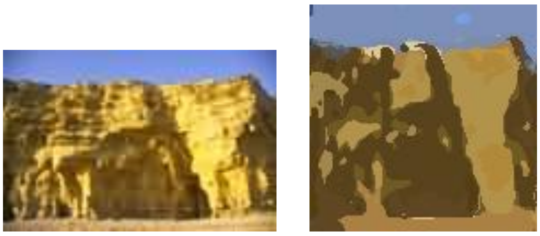
Even if the semantic concepts are not directly related to the visual features (colour, texture, shape, position, dimension, etc.), these attributes capture the information about the semantic meaning (Rasiwasia, et al., 2007).

The ability and efficiency of the colour feature for characterizing the colour perceptual similitude is strongly influenced by the colour space and quantization scheme selection. A set of dominant colours determined for each image provides a compact description, easy to be implemented. The CIE-LUV colour space quantized at 256 colours is used to represent the colour information. Before segmentation, the images are transformed from RGB

Udristoiu S., Ion A. and Mancas D. (2010).
RULE BASED MODELLING OF IMAGES SEMANTIC CONCEPTS.
In *Proceedings of the 2nd International Conference on Agents and Artificial Intelligence - Artificial Intelligence*, pages 540-543

to CIE-LUV colour space and quantized at 256 colours (Smith et al., 1996). The colour regions extraction is realized with K-means clustering algorithm (Berson et al., 1997). This algorithm detects the regions of a single colour. For each colour region, the spatial coherency represents the spatial homogeneity of the region in an image. It is computed for identifying the 8-connected pixels of the same colour in a region.

The figures 1 illustrates the image segmentation process of an image from „cliff" category.



(a) (b)

Figure 1: (a) Original image from category "sunset". (b) Segmented image.

In conformity with the defined characteristics, a region is described by:

-The colour characteristics are represented in the CIE-LUV colour space quantized at 256 colours. A region is represented by a colour index which is, in fact an integer number between 0…255. It is denoted as descriptor F1.

-The spatial coherency represents the region descriptor, which measures the spatial compactness of the pixels of same colour. It is denoted as descriptor F2.

-A seven-dimension vector (maximum probability, energy, entropy, contrast, cluster shade, cluster prominence, correlation) represents the texture characteristic. It is denoted as descriptor F3.

-The region dimension descriptor represents the number of pixels from region. It is denoted as descriptor F4.

-The spatial information is represented by the centroid coordinates of the region and by minimum bounding rectangle. It is denoted as descriptor F5.

-A two-dimensional vector (eccentricity and compactness) represents the shape feature. It is denoted as descriptor F6.

3 MAPPING VISUAL FEATURES TO SEMANTIC INDICATORS

The visual vocabulary developed for image annotation is based on the concepts of semantic indicators, whereas the syntax captures the basis models of human perception about patterns and semantic categories.

In this study, the representation language is simple, because the syntax and vocabulary are elementary. The language words are limited to the name of semantic indicators. Being visual elements, the semantic indicators are, by example, the colour (colour-light-red), spatial coherency (spatial coherency-weak, spatial coherency-medium, spatial coherency-strong), texture (energy-small, energy-medium, energy-big, etc.), dimension (dimension-small, dimension-medium, dimension-big, etc.), position (vertical-upper, vertical-center, vertical-bottom, horizontal-upper, etc.), shape (eccentricity-small, compactness-small, etc.).

The syntax is represented by a 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 values domains of visual characteristics were manually experimented on images of WxH dimension.

A value of colour semantic indicator is associated to each region colour in the CIE-LUV colour space quantized at 256 colours. The colour correspondence between the mathematical and semantic indicator values is determined based on experiments effectuated on a training image database.

A hierarchy of values, which are mapped to semantic indicator values, is also determined for the other features: colour, texture, shape, dimension, spatial coherency, and position.

At the end of the mapping process, a figure is represented in Prolog by means of the terms *figure(ListofRegions)*, where *ListofRegions* is a list of image regions.

The term *region(ListofDescriptors)* is used for region representation, where the argument is a list of terms used to specify the semantic indicators. The term used to specify the semantic indicators is of the form:

descriptor(DescriptorName, DescriptorValue).

The model representation of an image from *cliff* category can be observed in the bellow example:

figure([
region([descriptor(colour,dark-brown),
descriptor(horizontal-position, center),
descriptor(vertical-position,center),
descriptor(dimension,big),descriptor(shape-eccentricity, small),
descriptor(texture-probability, medium),
descriptor(texture-inversedifference, medium),
descriptor(texture-entropy,small),
descriptor(texture-energy,big),
descriptor(texture-contrast,big),
descriptor(texture-correlation, big)]), …

4 IMAGE SEMANTIC RULES

The process of the automated generation of semantic rules and image annotation is the following:

1. The learning phase: rules generation

A semantic rule is of the form:

"semantic indicators -> category"

he stages of the learning process are:

-relevant images for a semantic concept are used for learning it.

-each image is automatically processed and segmented and the primitive visual features are computed, as it is described in section 2.

-for each image, the primitive visual features are mapped to semantic indicators, as it is described in section 3.

-the rule generation algorithms are applied to produce rules, which will identify each semantic category from database.

2. The image testing/annotation phase has as scope the automatic annotation of images.

-each new image is processed and segmented in regions,

-for each new image the low-level characteristics are mapped to semantic indicators,

-the classification algorithm is applied for identifying the image category/semantic concept.

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.

4.1 The Description of the Algorithm for Rule Generation

The algorithm for semantic rules generation is based on A-priori algorithm of finding the frequent itemsets (Berson et al., 1997; Frawley et al., 1991).

The choice of the itemsets and transactions is a domain dependent problem. In the case of market analysis, the itemsets are products, and the transactions are itemsets brought together.

The scope of image association rules is to find semantic relationships between image objects. For using association rules that discover the semantic information from images, the modelling of images in the terms of itemsets and transactions is necessary:

-the set of images within the same category represents the transactions set,

-the itemsets are the colours of image regions,

-the frequent itemsets represent the itemsets with support bigger or equal than the minimum support (min_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.

In our method, the Apriori algorithm is used for discovering the semantic association rules between primitive characteristics extracted from images and categories/semantic concepts, which images belong to. The semantic association rules have the body composed by conjunctions of semantic indicators, while the head is the category/semantic concept. A semantic rule describes the most frequent characteristics for each category, based on Apriori rule generation algorithm.

The rules are represented in Prolog as facts of the form:

rule(Category, Score, ListofRegionPatterns).

The patterns from ListofRegionPatterns are terms of the form:

regionPattern(ListofPatternDescriptors).

The patterns from the descriptors list specify the set of possible values for a certain descriptor name. The form of this term is:

descriptorPattern(descriptorName,ValueList).

4.2 Semantic Image Classification

The classifier represents the set of semantic rules used to predict the category of images from the test database. Being given a new image, the classification process searches in the rules set for finding its most appropriate category. Images are processed and are represented by means of semantic indicators as Prolog facts. The semantic rules are applied to the set of image facts, using the Prolog inference engine.

A semantic rule matches an image if all characteristics which appear in the body of the rule also appear in the image characteristics.

4.3 Experiments

In the experiments realized through this study, two databases are used for testing the learning process. The database used for learning contains 200 images from different nature categories and is used to learn the correlations between images and semantic concepts. The database used in the learning process is categorized into 50 semantic concepts. The system learns each concept by submitting appreciatively 20 images per category. The testing database contains 500 unclassified images

The performance metrics, precision and average normalized modified retrieval rate (ANMRR), are computed to evaluate the efficiency and accuracy of the rules generation and annotation methods (Manjunath et al., 2001). These parameters are computed as average for each image category as in Table 1:

Table 1: The precision and ANMRR computed for each image category.

Category	Precision	ANMRR
Fire	0.77	0.39
Iceberg	0.71	0.34
Tree	0.65	0.45
Sunset	0.89	0.14
Cliff	0.93	0.11
Desert	0.89	0.11
Red Rose	0.75	0.20
Elephant	0.65	0.43
Mountain	0.85	0.16
See	0.91	0.09
Flower	0.77	0.31

As it can be observed from the experiments, the results are strongly influenced by the complexity of each image category. Actually, the results of experiments are very promising, because they show a small average normalized modified retrieval rate and a good precision for the majority of the database categories, making the system more reliable.

5 CONCLUSIONS

In this study we propose methods for semantic image annotation based on visual content.

By comparison to other image annotation methods, our proposed and developed methods have some advantages:

- The entire process is automated, and a great number of semantic concepts can be defined.
- These methods can be easily extended to any domain, because the visual features, semantic indicators remain unchanged, and the semantic rules are generated based on the set of example labelled images used for learning semantic concepts.
- The spatial information is taken into account and it offers rich semantic information about the relationships of the image colour regions (left, right, center, bottom, and upper).
- The Prolog logic programming used to model images and semantic rules facilitates the interaction with them in a easier way.

The proposed methods have the limitation that they can't learn every semantic concept, due to the fact that the segmentation algorithm is not capable to segment images in real objects. Improvements can be brought using a segmentation method with greater semantic accuracy.

REFERENCES

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

Carneiro, G., Chan, A., Moreno, P., and Vasconcelos, N., 2007. Supervised learning of semantic classes for image annotation and retrieval. In *IEEE Pattern Analysis Machine Intelligence*, vol. 29(3), pp. 394–410.

Hoogs, A., Rittscher, J., Stein, G., and Schmiederer, J., 2003. Video content annotation using visual analysis and a large semantic knowledge base. In Proceedings *of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 327 – 334.

Frawley, W. J., G. Piatetsky-Shapiro, and C. J. Matheus, 1991. *Knowledge Discovery in Databases, chapter Knowledge Discovery in Databases: An Overview*. MIT Press.

Manjunath, B. S, Salembier, P., and Sikora, T., 2001. *Introduction to MPEG-7: Multimedia Content Description Standard*. Wiley, New York.

Rasiwasia, N., Moreno, P. J., Vasconcelos, N., 2007. Bridging the Gap: Query by Semantic Example. In *IEEE Transactions On Multimedia*, vol. 9(5), pp. 923-938.

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

Smeulders, A. W., M. Worring, S. Santini, A. Gupta, and R. Jain, 2000. Content-Based Image Retrieval at the End of the Early Years. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22(12), pp. 1349–1380.