Image Annotation Based on Semantic Rules

A.L. Ion

Software Engineering Department, Faculty of Automation, Computers and Electronics, Craiova, Romania
anca.ion@software.ucv.ro

Abstract. For developing image navigation systems, we need tools to realize the semantic relationship between user and database. In this paper we develop algorithms that automatically generate semantic rules that identify image categories and introduce the cognitive dimension in the retrieval process. The semantic rules are represented in Prolog and can be shared and modified depending on the updates in the respective domain.

1 Introduction

Content-based image retrieval (CBIR) is described as any technology that helps to organize digital picture archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR [6, 8].

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 CBIR community [5].

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 [1] between low-level content and higher-level concepts [6].

The popular online photo-sharing Flickr [3], which hosts hundreds of millions of pictures with diverse content, the video sharing and distribution forum YouTube have brought in a new revolution for multimedia usage. In [6] it is supposed that image retrieval will enjoy a success story in the coming years. Although great progress has been made in image retrieval and browsing systems since the early, none of the existing methods captures enough of the semantic-related information to be used as a navigation tool in a general content image database. Our inability to capture “image semantics” comes from the incompatibility of the information we are able to compute directly from image data, and our subjective interpretation of the same data [4].
The proposed study is based on:

a) the understanding of semantic categories taking into account the human visual perception,
b) the study of human judgment regarding the similitude between images for extracting the significant and discriminate attributes of semantic categories,
c) the design and implementation of algorithms for extracting the image characteristics, similitude metrics, semantic vocabulary, that offers important semantic elements in the retrieval and categorization, semantic rules generation, image classification.

Our study is done on nature categories, where subjects as water, sky/clouds, mountains, snow, rain, sun rise are considered to be important cues. Also the colour composition and colour features played an important role in comparing nature images. Within these categories, spatial organization, spatial frequency, or dimension do influence similarity judgments. Exceptions from these ideas are some categories as flowers, fruits, exotic animals, which contain strong hues (dark, medium red, yellow, blue, green, pink, etc) that are not in the description of nature. This study is an extension of the previous work [7] and started from the limitations regarding the researches in multimedia semantic modelling. In this study, we propose 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.

2 The Image Segmentation

The selection of the visual feature set and the image segmentation algorithm is the definitive stage for the semantic annotation process of images. By doing a large set of experiments, we deduce 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.

Before to be segmented, the images are transformed from RGB to HSV colour space and quantized to 166 colours. The extraction of single colour regions is realized by applying the modified colour set back projection algorithm [10].

The implementation of this algorithm is described in pseudo-code:

**Algorithm.** Segmentation of an image in regions having a single colour

**Input:** image I, colour set C

**Output:** the set of single colour regions, R and the set of spatial coherencies of each region M.

**Method:**

1. InitStack(S)
2. Visited = ∅