Using Context and a Genetic Algorithm for Knowledge-Assisted Image Analysis

Stamatia Dasiopoulou, George Th. Papadopoulos, Phivos Mylonas, Yannis Avrtihis, and Ioannis Kompatsiaris

Abstract— In this poster, we present an approach to contextualized semantic image annotation as an optimization problem. Ontologies are used to capture general and contextual knowledge of the domain considered, and a genetic algorithm is applied to realize the final annotation. Experiments with images from the beach vacation domain demonstrate the performance of the proposed approach and illustrate the added value of utilizing contextual information.

Index Terms— knowledge-assisted image analysis, ontologies, context modelling, genetic algorithm.

I. INTRODUCTION

Given the continuously increasing information flow, providing tools and methodologies for (semi-) automatically extracting content descriptions at a conceptual level is a key factor for enabling users to effectively access the information needed. Incorporating knowledge has been acknowledged as probably the only viable solution to overcome the limitations resulting from attempts to imitate the way users assess similarity by focusing on the definition of suitable descriptors that could be automatically extracted and of appropriate metrics. Going though the relevant literature, one can distinguish two main streams in the reported knowledge-driven approaches with respect to the adopted knowledge acquisition and processing strategy: implicit, realized by machine learning methods, and explicit, realized by model-based approaches. The former provide relatively powerful methods for discovering complex and hidden relationships between image data and corresponding conceptual descriptions. Model-based analysis approaches on the other hand make use of explicitly defined prior knowledge, attempting to provide a coherent domain model to support symbolic inference; as a result issues are raised with respect to the entailed complexity and the completeness of knowledge acquisition and construction.

In order to benefit from the advantages of each category and overcome their individual limitations, we propose an approach coupling explicit prior knowledge and a genetic algorithm for domain-specific semantic image analysis. Ontologies, being the leading edge technology for knowledge sharing and reuse and providing well-defined inferences, have been selected for representation. The employed knowledge considers both

S. Dasiopoulou and G.Th. Papadopoulos are with the Information Processing Laboratory, Electrical and Computer Engineering Department, Aristotle University of Thessaloniki, Greeece, and the Informatics and Telematics Institute, Centre for Research and Technology Hellas (email: dasiop@iti.gr; papad@iti.gr). I. Kompatsiaris is with Informatics and Telematics Institute, Centre for Research and Technology Hellas (email: ikom@iti.gr). P. Mylonas and Y. Avrithis are with the Image, Video and Multimedia Laboratory, School of Electrical and Computer Engineering, National Technical University Athens, Greece (email: fmylonas@image.ntua.gr; iavr@image.ntua.gr).
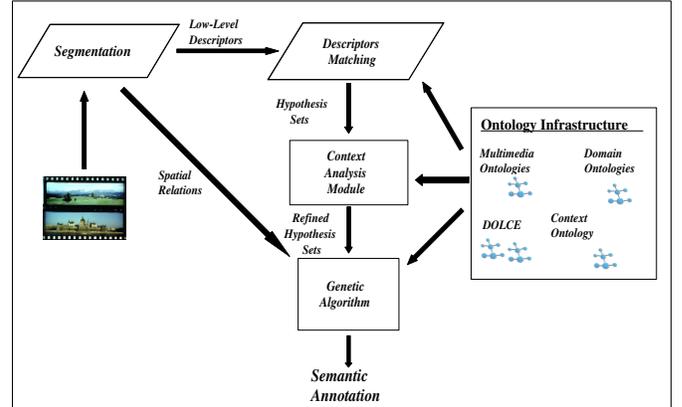
Fig. 1. Overall architecture

high- and low-level information in order to provide the means for driving the semantic analysis and annotation. High-level knowledge includes the domain concepts of interest, their relations and contextual knowledge in terms of fuzzy ontological relations. Low-level knowledge on the other hand, consists of the low-level visual and structural descriptions required for the actual analysis process.

II. OVERALL ARCHITECTURE

The overall architecture of the proposed framework is illustrated in Fig. 1. To represent the required knowledge components, the ontology infrastructure introduced in [2] has been employed and appropriately extended to provide support for contextual knowledge modelling. Analysis starts by segmentation, and subsequently low-level descriptors and spatial relations are extracted for the generated image segments. An extension of the Recursive Shortest Spanning Tree (RSST) algorithm has been used for segmenting the image [1], while descriptors extraction is based on the guidelines given by the MPEG-7 eXperimentation Model (XM). The image preprocessing stage completes with the extraction of spatial relations between adjacent image segments, where fuzzy definitions have been employed [4]. Once the low-level descriptors are available, an initial set of hypotheses is generated for each image segment based on the distance between each segment extracted descriptors and the domain concepts prototypical descriptors that are included in the knowledge base. Thereby, a set of plausible annotations with corresponding degrees of confidence are produced for each segment, which are refined in the sequel based on the provided fuzzy contextual knowledge. The refined hypotheses sets along with the segments spatial

TABLE I

FUZZY SEMANTIC RELATIONS.

Abbreviation	Name
P	PartOf
Sp	SpecializationOf
Pr	PropertyOf
Ct	ContextOf
Loc	Location
Ins	InstrumentOf
Pat	PatientOf

relations are eventually passed to the genetic algorithm, which based on the provided domain knowledge decides the optimal semantic interpretation.

III. CONTEXT MODELLING AND ANALYSIS

A fuzzified ontology, defined on top of the domain one, has been built to represent the modelled contextual knowledge. To provide a sufficiently descriptive context model, we selected a set of diverse relations from the ones included in the MPEG-7 specification (Table **??**) and extended their definition to support uncertainty. More specifically, the defined fuzzified contextual domain knowledge consists of the set

$$O_F = \left\{ C, \left\{ r_{c_i,c_j} \right\} \right\}, i,j = 1..n, i \neq j, \quad (1)$$

where O_F forms a domain-specific "fuzzified" ontology, C is the set of all possible concepts it describes, $r_{c_i,c_j} = F(R_{c_i,c_j}) : C \times C \to [0,1], R_{c_i,c_j} : C \times C \to \{0,1\}, i,j \in N$ denotes a fuzzy ontological relation amongst two concepts c_i, c_j and R_{c_i,c_j} is a crisp semantic relation amongst the two concepts. For the representation the RDF language has been selected and reification was used in order to achieve the desired expressiveness. The refinement of the initial hypotheses' degrees is performed based on the readjustment algorithm presented in [3], where the notion of overall *context relevance* of a concept to the *root element* is employed to tackle cases where a concept is related to multiple ones.

IV. GENETIC ALGORITHM

Under the proposed approach, each chromosome represents a possible solution. To determine the degree to which each interpretation is plausible, the employed fitness function has been defined as follows:

$$f(C) = \lambda \times FS_{norm} + (1-\lambda) \times SC_{norm}, \quad (2)$$

where C denotes a particular Chromosome, FS_{norm} refers to the degree of low-level descriptors similarity, and SC_{norm} stands for the degree of consistency with respect to the provided spatial domain knowledge. The variable λ is introduced to adjust the degree to which visual similarity and spatial consistency should affect the final outcome. SC_{norm} and FS_{norm} are computed as follows:

$$FS_{norm} = \frac{\sum_{i=0}^{N-1} I_M(g_i) - I_{min}}{I_{max} - I_{min}}, \quad (3)$$

where I_{min} is the sum of the minimum degrees of confidence assigned of each region hypotheses set and I_{max} the sum of the maximum degrees of confidence values respectively.
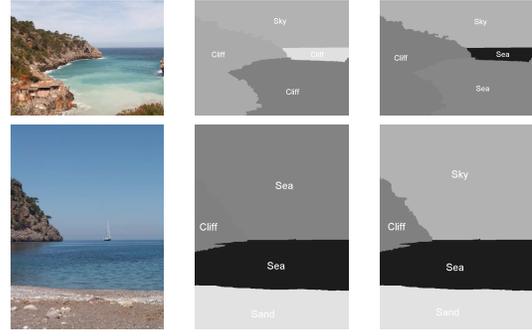

Fig. 2. Exemplar results for the beach domain, depicting the original image, and the GA application without and with context utilization.

$$SC_{norm} = \frac{SC+1}{2} \ and \ SC = \frac{\sum_{r=1}^{M} I_{s_r}(g_i,g_j)}{M}, \quad (4)$$

where M denotes the number of relations in the constraints that had to be examined.

V. EXPERIMENTAL RESULTS AND CONCLUSIONS

The proposed semantic image analysis framework was tested on the beach vacations domain on the following concepts: *Sea*, *Sky*, *Sand*, *Plant*, *Cliff* and *Person*. The employed descriptors are Scalable Color, Homogeneous Texture, Edge Histogram and Region Shape, while the implemented fuzzy spatial relations include the eight diagonal relations, i.e., *left*, *above*, *above-left* etc. Indicative results are given in Fig. 2, illustrating the added value of utilizing contextual information and the performance of the genetic algorithm.

REFERENCES

[1] T. Adamek, N. O'Connor, N. Murphy, *Region-based Segmentation of Images Using Syntactic Visual Features*. Workshop on Image Analysis for Multimedia Interactive Services, (WIAMIS), Montreux, Switzerland, 2005.
[2] S. Bloehdorn, K. Petridis, C. Saathoff, N. Simou, V. Tzouvaras, Y. Avrithis, I. Kompatsiaris, S. Staab, M. G. Strintzis, *Semantic Annotation of Images and Videos for Multimedia Analysis*. 2nd European Semantic Web Conference (ESWC), Heraklion, Greece, May 2005.
[3] Ph. Mylonas, Th. Athanasiadis and Y. Avrithis, *Improving image analysis using a contextual approach*. Internationl Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Seoul, April 2006.
[4] P. Panagi, S. Dasiopoulou, G.Th. Papadopoulos, I. Kompatsiaris and M.G. Strintzis, *A Genetic Algorithm Approach to Ontology-Driven Semantic Image Analysis*. IEE International Cofnerence on Visual Information Engineering (VIE), Bangalore, India, Sept 2006.