AN ARCHITECTURE FOR MULTIMEDIA ANALYSIS AND RETRIEVAL BASED ON FUZZY DESCRIPTION LOGICS

Nikolaos Simou and Thanos Athanasiadis and Stefanos Kollias

Image Video and Multimedia Systems Laboratory,
School of Electrical and Computer Engineering,
National Technical University of Athens

ABSTRACT

Effective management and exploitation of multimedia documents requires extraction of the underlying semantics. In this paper we propose a methodology for semantic indexing and retrieval of images, based on techniques of image segmentation and classification combined with fuzzy reasoning. In the proposed knowledge-assisted analysis architecture a segmentation algorithm firstly generates a set of over-segmented regions. After that a region classification process is employed to assign semantic labels using a confidence degree and simultaneously merge regions based on their semantic similarity. This information comprises the assertional component of a fuzzy knowledge base which is used for the refinement of mistakenly classified regions and also for the extraction of rich implicit knowledge used for global image classification. This semantic metadata of images is stored in a knowledge repository by the fuzzy reasoning engine, also permitting image retrieval and ranking.

1. DESCRIPTION OF WORK

During the last decade a dramatically large increase of the digital multimedia content has occurred. The main reason that led to this change was the broad availability and use of digital devices, not only by professional or very experienced users, but almost by everyone. As a consequence new research interests have been emerged having as primary aim, among others, the efficient management of multimedia content providing mechanisms for indexing and retrieval. The first attempts to meet this target were based on metadata generation for the multimedia documents [1]. Different types of metadata were used which can be separated, based on their context, to low level like MPEG-7 descriptors, and higher level like thematic categorization [2].

Most approaches related to semantic-based analysis and indexing are grounded on the implementation of semantic concept detectors, evaluation of which indicates that a good level of maturity has been reached [3]. At the same time, image segmentation and object recognition have been used simultaneously, instead of sequentially, aiming improvement on both domains [4, 5]. This extracted information by multimedia analysis algorithms requires effective management capable of inferring complicated concepts. In this context, various attempts have been examined using taxonomies and ontologies. Ontologies are based on Description Logics (DLs) [6] that are a family of knowledge representation languages. However, despite the rich expressiveness of DLs, they lack the ability to deal with vague and uncertain information which is very common in multimedia content. This was the reason that a variety of DLs capable of handling imprecise information, like probabilistic and fuzzy [7] have been proposed.

This research was supported by the European Commission under contract FP6-027026 K-SPACE.

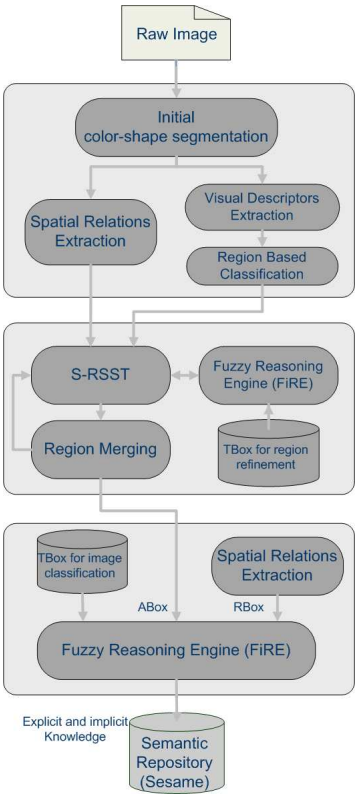

Fig. 1. Overview of the proposed architecture.

In this paper we present a knowledge assisted image analysis and automatic semantic annotation methodology consisting of several novel and state-of-the-art techniques. The architecture of our proposal is shown in Fig. 1. As can be seen, we initially segment an image based on color and shape criteria. It is well known that with such criteria image segmentation algorithms fail to extract semantically meaningful objects. For that reason we introduce a novel semantic region growing methodology which incorporates object detection simultaneously with region merging, providing better input for region-based classification. To further improve this operation we use fuzzy reasoning engine FiRE together with a terminology (TBox) defined in order to improve region-based classification by incorporating spatial relations with neighborhood information. The

previous step provides a list of concepts (together with degrees of confidence for each one) that have been linked to the image. Our purpose is to use this information in order to extract additional, implicit knowledge and also infer abstract concepts on a global image basis. Towards this aim, FiRE is employed using this time a different terminology especially for this purpose. The final results are stored in an online semantic repository, in a strictly structured format, allowing query mechanisms for semantic retrieval and ranking.

1.1. Knowledge-assisted analysis

Knowledge-assisted analysis, in the context of this work, deals with the very important and difficult task of the instantiation of a fuzzy knowledge base by processing a multimedia document. The optimum operation of this task is crucial for our architecture since mistaken instantiation of the fuzzy knowledge base would lead to mistaken inferred knowledge. In order to achieve this objective a semantic variation of RSST segmentation is employed together with fuzzy reasoning engine FiRE.

Traditional RSST is a bottom-up segmentation algorithm that begins from the pixel level and iteratively merges similar neighbor regions until certain termination criteria are satisfied. We used a modified version of RSST, called Semantic RSST (S-RSST) with novel (dis)similarity and termination criteria. The criterion for ordering the edges is not a distance metric based on low-level features, but the semantic similarity measure [5].

1.2. Fuzzy reasoning and querying

Description Logic (DL) f-$\mathcal{SHIN}$ is a fuzzy extension of DL $\mathcal{SHIN}$ [8] and it similarly consists of an alphabet of distinct concepts names ($\mathbf{C}$), role names ($\mathbf{R}$) and individual names ($\mathbf{I}$). Using DLs the construction of new concepts and roles is possible. For that purpose DLs include a set of constructors to construct concept and role descriptions. These constructors specify the name of the DL language [6] and in the case of f-$\mathcal{SHIN}$ these are $\mathcal{ALC}$ constructors (i.e. negation $\neg$, conjunction $\sqcap$, disjunction $\sqcup$, full existential quantification $\exists$ and value restriction $\forall$) extended by transitive roles ($\mathcal{S}$), roles hierarchy ($\mathcal{H}$), inverse roles ($\mathcal{I}$), and number restrictions ($\mathcal{N} \leq, \geq$).

FiRE[1] is a Java based fuzzy reasoning engine currently supporting f-$\mathcal{SHIN}$ that can be used either as an API by another application or by using its graphical user interface. FiRE was enhanced by the functionalities of the RDF-Store Sesame (Sesame 2 beta 6). In the proposed architecture the RDF Store is used as a back end for storing and querying RDF triples in a sufficient and convenient way, while the reasoner is the front end that the user can use in order to store and query a fuzzy knowledge base. By that way, a user is able to access data from a repository, apply any of the available reasoning services on this data and then store back in the repository the implicit knowledge extracted from them. Since in our case we extend classical assertions to fuzzy assertions, new methods of querying such fuzzy information are possible. More precisely, in [9] authors extend ordinary conjunctive queries to a family of significantly more expressive query languages, which are borrowed from the fields of fuzzy information retrieval [10]. These queries were implemented by the use of query language SPARQL[11] that is supported by Sesame implementing in that way semantic indexing and retrieval of images.

[1]FiRE can be found at `http://www.image.ece.ntua.gr/~nsimou/FiRE/` together with installation instructions and examples

2. CONCLUSIONS AND FUTURE WORK

In this short paper we briefly presented an architecture for combining image analysis techniques with expressive fuzzy description logics. Semantic image segmentation comprises the assertional component of a fuzzy knowledge base that is used for refining region-based classification and also for the extraction of rich implicit knowledge on a global image basis. This is achieved with fuzzy reasoning that incorporates region classification and spatial information. Additionally, through FiRE API we are able to access and query an online fuzzy knowledge-base, using Sesame repository, allowing the user to perform semantic retrieval and ranking of images based on semantic metadata extracted either directly from the content or implicitly through reasoning.

Future work can be pursued towards two directions: i) Enlarge the semantic vocabulary, by adding more concept detectors (and more robust), enriching in that way the available knowledge and ii) extend the proposed architecture to video sequences, permitting semantic browsing of scenes.

3. REFERENCES

[1] J. Smith, *Video Indexing and Retrieval using MPEG-7*, CRC Press, 2003.

[2] Ph. Mylonas, Th. Athanasiadis, M. Wallace, Y. Avrithis, and S. Kollias, "Semantic representation of multimedia content: Knowledge representation and semantic indexing," *Multimedia Tools and Applications*, vol. 39, no. 3, pp. 293–327, 2008.

[3] C. Snoek, B. Huurnink, L. Hollink, M. de Rijke, G. Schreiber, and M. Worring, "Adding semantics to detectors for video retrieval," *Transactions on Multimedia*, vol. 9, no. 5, pp. 975986, August 2007.

[4] E. Borenstein and E. Sharon, "Combining top-down and bottom-up segmentation," in *Proceedings of 8th Conference on Computer Vision and Pattern Recognition Workshop*, 2004.

[5] Th. Athanasiadis, Ph. Mylonas, Y. Avrithis, and S. Kollias, "Semantic image segmentation and object labeling," *Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 3, pp. 298–312, March 2007.

[6] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, *The Description Logic Handbook: Theory, implementation and applications*, Cambridge University Press, 2003.

[7] G. Stoilos, G. Stamou, V. Tzouvaras, J. Pan, and I. Horrocks, "Reasoning with very expressive fuzzy description logics," *Journal of Artificial Intelligence Research*, vol. 30, no. 5, pp. 273–320, 2007.

[8] I. Horrocks, U. Sattler, and S. Tobies, "Reasoning with Individuals for the Description Logic $\mathcal{SHIQ}$," in *CADE-2000*, David MacAllester, Ed. 2000, number 1831 in LNAI, pp. 482–496, Springer-Verlag.

[9] J. Pan, G. Stamou, G. Stoilos, and E. Thomas, "Expressive querying over fuzzy DL-Lite ontologies," in *Proceedings of the International Workshop on Description Logics*, 2007.

[10] V. Cross, "Fuzzy information retrieval," *Journal of Intelligent Information Systems*, vol. 3, pp. 29–56, 1994.

[11] E. Prud'hommeaux and A. Seaborne, "SPARQL query language for RDF," 2006, W3C Working Draft, http://www.w3.org/TR/rdf-sparql-query/.