

Using ontologies and fuzzy relations in multimedia personalization

Phivos Mylonas¹ and Manolis Wallace²

¹ *Image, Video and Multimedia Laboratory
School of Electrical Engineering
National Technical University of Athens
fmylonas@image.ntua.gr*

² *Department of Computer Science
University of Indianapolis, Athens Campus
wallace@uindy.gr*

Abstract

In this paper we extend on previous work in order to address the problem of multimedia personalization at a semantic level. In different previous works we have developed algorithms to address computationally efficient handling of large but sparse fuzzy relations, and theory to address knowledge representation, thematic categorization and user modeling. In this work we take two further steps: i) we integrate ontologies in our original knowledge modeling approach and ii) we explain how these diverse algorithms and methodologies can be combined in order to approach a greater goal, that of semantic multimedia personalization.

1. Introduction

For years information retrieval was considered to be synonymous to text retrieval [8]. The Text REtrieval Conference (TREC) [11] was traditionally considered as the most established conference in the field of information retrieval and a system winning a TREC competition was automatically considered as thoroughly successful.

At the same time, people working in multimedia related fields, such as image processing, video processing, video synthesis and so on, having managed to solve simpler problems were starting to acknowledge the closing of the semantic gap as the common ultimate goal of their individual works.

Nowadays, with the semantic gap still being as wide as always, text and multimedia communities are both realizing that the amount of accessible, processable and searchable data currently available brings two more parameters in play; those of personalization and semantics. With multimedia retrieval building upon and borrowing from text retrieval in many aspects, most current information retrieval problems can generally be classified under the generic umbrella of semantic multimedia personalized retrieval.

Ontologies [10], a formal way to model knowledge, provide a solid basis upon which semantic representation and semantic operations could be developed. On the other hand, their size alone can make their handling, let alone their practical utilization problematic.

In this paper we integrate ontologies and various independent previous works in order to provide a framework upon which semantic multimedia personalization can be based. In [1] we had described the fuzzy properties a knowledge representation scheme should have in order to be informative enough for semantic operations. In section 2 we integrate this approach with modern ontological modeling. In [14] we had proven that the processing time of large sparse fuzzy binary relations could be reduced from days to milliseconds. In section 3 we briefly review the approach and confirm that it is also applicable for the chosen ontological representation. In [12] we had described a way to automatically identify the topics a multimedia document was related to. In section 4 we explain how this can still be achieved when using the

ontological knowledge representation of section 2. Finally, in section 5 we explain how the above can be used in order to extract the user profile structures described in [13]. In section 6 we list our concluding remarks.

2. Ontological knowledge representation

Among all possible knowledge representation formalisms, ontologies [4], [6], present a number of advantages. In the context of this work, ontologies are suitable for expressing multimedia content semantics in a formal machine-processable representation that allows manual or automatic analysis and further processing of the extracted semantic descriptions. As an ontology is a formal specification of a shared understanding of a domain, this formal specification is usually carried out using a subclass hierarchy with relationships among the classes, where one can define complex class descriptions (e.g. in Description Logics (DLs) [3] or Web Ontology Language (OWL) [17]). Amongst all possible ways to describe ontologies one can be formalized as follows:

$$O = \{C, \{R_{ab}\}\}, \text{ where } R_{ab} : C \times C \rightarrow \{0,1\} \quad (1)$$

In equation (1) O is an ontology, a and b are two concepts $a, b \in C$ belonging to the set C of concepts described by the ontology and R_{ab} is the semantic relation amongst these concepts. The proposed knowledge model is based on a set of concepts and relations between them, which form the basic elements towards semantic interpretation of the present research effort. Although almost any type of relation may be included to construct such knowledge representation, the two categories commonly used are taxonomic (i.e. ordering) and compatibility (i.e. symmetric) relations. However, as extensively discussed in [1], compatibility relations fail to assist in the determination of the context and the use of ordering relations is necessary for such tasks. Thus, a first main challenge is the meaningful exploitation of information contained in taxonomic relations for the task of context exploitation within semantic image segmentation and object labeling.

In addition, for a knowledge model to be highly descriptive, it must contain a large number of distinct and diverse relations among concepts. A major side effect of this approach is the fact that available information will then be scattered among them, making each one of them inadequate to describe a context in a meaningful way. Consequently, relations need to be combined to provide a view of the knowledge that suffices for context definition and estimation. In this

work we utilize ten types of relations, whose semantics are defined in the MPEG-7 standard [9] and are presented in the following Table 1:

<i>Name</i>	<i>Example</i>
SpecializationOf <i>Sp</i>	“wrc” is a specialization of “rally”
PartOf <i>P</i>	“leg” is a part of “human body”
PropertyOf <i>Pr</i>	“democracy” is property of “United Kingdom”
ExampleOf <i>E</i>	“George” is example of “human”
inContextOf <i>Ct</i>	“Gone With the Wind” is in context of “American Civil War”
InterpretationOf <i>In</i>	“War Crimes” and “Strategy” are interpretations of “Hiroshima Bomb”
MemberOf <i>M</i>	“player” is member of “team”
AgentOf <i>Ag</i>	“player” is agent of “kick”
ResultOf <i>Res</i>	“goal” is result of “kick”
LocationOf <i>Loc</i>	“Paris” is the location of “game”

Table 1. Categories of MPEG-7 semantic relations

A last point to consider when designing such a knowledge model is the fact that real-life data often differ from research data. Real-life information is in principal governed by uncertainty and fuzziness, thus its modeling is based on fuzzy relations. For the problem at hand, the above commonly encountered crisp relations can be modeled as fuzzy ordering relations and can be combined for the generation of a meaningful fuzzy taxonomic relation. Consequently, to tackle such complex types of relations we propose a “fuzzification” of the previous ontology definition, as follows:

$$O_F = \{C, \{r_{ab}\}\}, r_{ab} = F(R_{ab}) : C \times C \rightarrow [0,1] \quad (2)$$

In equation (2) O_F defines a “fuzzified” ontology, C is again the set of all possible concepts it describes and r_{ab} denotes a fuzzy relation amongst the two concepts a and b . More specifically, given a universe U a crisp set C is described by a membership function $\mu_C : U \rightarrow \{0,1\}$, whereas a fuzzy set F on C is described by a membership function $\mu_F : C \rightarrow [0,1]$. We may describe the fuzzy set F using the sum notation [7]:

$$F = \sum_i c_i / w_i = \{c_1 / w_1, c_2 / w_2, \dots, c_n / w_n\}, \text{ where } i \in N_n,$$

$n = |C|$ is the cardinality of C , $w_i = \mu_F(c_i)$ or, more simply, $w_i = F(c_i)$ and $c_i \in C$. According to [5] a fuzzy relation on C is a function $r_{ab} : C \times C \rightarrow [0,1]$ and its inverse relation is defined as $r_{ab}^{-1} = r_{ba}$. Based on the relations r_{ab} and, for the purpose of current work, we construct the following relation T with the use of the above set of fuzzy taxonomic relations:

$$T = Tr^t \left(Sp \cup P^{-1} \cup Pr^{-1} \cup E \cup Ct \cup In \cup M \cup Ag \cup Res \cup Loc \right) \quad (3)$$

In the aforementioned relations, fuzziness has the following meaning: High values of $Sp(a;b)$, for instance, imply that the meaning of b approaches the meaning of a , in the sense that when an image is semantically related to b , then it is most probably related to a as well. On the other hand, as $Sp(a;b)$ decreases, the meaning of b becomes “narrower” than the meaning of a , in the sense that an image’s relation to b will not imply a relation to a as well with a high probability, or to a high degree. Summarizing, the value of $Sp(a;b)$ indicates the degree to which the stored knowledge shows that an occurrence of b in an image implies relation to a . Likewise, the degrees of the other two relations can also be interpreted as conditional probabilities or degrees of implied relevance. The latter imply that, for example, $a \neq b \Rightarrow Sp(a;b) < 1$ since, if $a \neq b$, then we cannot be sure that both a and b are related to a given image, without first examining the image’s context; at this point it is important to remind the reader that a and b are not terms but concepts, which means that $a \neq b$ indicates/ensures a difference in a conceptual level.

$P(a;b)$ means that b is a part of a , e.g. a could be a *boat* and b could be a *sail*. So it is expected that the role of $P(a;b)$ is the opposite of that of $Sp(a;b)$, i.e. $Sp(a;b)$ means that the meaning of a “includes” the meaning of b . Finally, $Pr(a;b)$ means that b is a property of a . Based on these roles and semantic interpretations of all semantic relations, it is easy to see that

(3) combines them in a straightforward and meaningful way, utilizing inverse functionality where it is semantically appropriate, i.e. where the meaning of one relation is semantically contradictory to the meaning of the rest on the same set of concepts. Finally, the transitive closure Tr^t is required in order for T to be taxonomic, as the union of transitive relations is not necessarily transitive, as discussed in [2].

Representation of our concept-centric contextual knowledge model follows the Resource Description Framework (RDF) standard proposed in the context of the Semantic Web [15]. RDF is the framework in which Semantic Web metadata statements are expressed and usually represented as graphs. Relation T can be visualized as a graph, in which every node represents a concept and each edge between two nodes constitutes a contextual relation between the respective concepts. Additionally each edge has an associated

degree of confidence, which represents the fuzziness within the context model. Representing the graph in RDF is a straight forward task, since RDF structure itself is based on a similar graph model.

Reification [16] was used in order to achieve the desired expressiveness and obtain the enhanced functionality introduced by fuzziness. Representing the degree of confidence associated with each relation is carried out using reification, i.e. making a statement about the statement, which contains the degree information. Representing fuzziness with reified statements is a novel but acceptable way, since the reified statement should not be asserted automatically. For instance, having a statement such as: “*Sky PartOf BeachScene*” and a degree of confidence of 0.75 for this statement, does obviously not entail, that *sky* is always a part of a *beach scene*.

A clarifying example of an instance of the part relation P derived from the tennis sports domain is provided in the following Figure 1 and Figure 2, both in terms of RDF/XML syntax and graph visualization. As already discussed, $P(x;y) > 0$ means that y is part of x . Following the RDF terminology, the subject *ball* has *partOf* as its predicate and *tennis* forms the object. Additionally, the proposed reification process introduces a statement about the former statement on the *partOf* resource, by stating that 0.90 is the degree of confidence to this relation.

```
<rdf:Description rdf:about="#s1">
  <rdf:subject rdf:resource="#&dom;ball"/>
  <rdf:predicate rdf:resource="#&dom;partOf"/>
  <rdf:object rdf:resource="#&dom;tennis"/>
  <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
  <context:partOf rdf:datatype="http://www.w3.org/2001/XMLSchema#float">
    0.90</context:partOf>
</rdf:Description>
```

Figure 1. Fuzzy relation representation: RDF reification.

A visualization of a graph fragment of the contextualized *tennis* ontology is presented in Figure 2; it represents relation *inContextOf* (Ct).

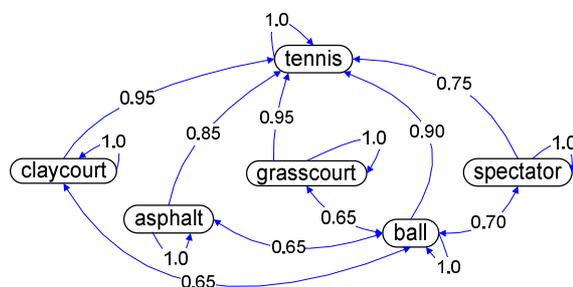


Figure 2. A fragment of the tennis ontology.

In the framework of this work, three domain ontologies were developed for representing knowledge components under the proposed approach, namely the beach, tennis and motorsports domains. These application domains were selected based on their popularity and amount of multimedia content available, whereas the proposed knowledge infrastructure was manually constructed by domain experts.

3. Transitive Closure

In the above knowledge representation structure the weak part is the combination of different relations towards the generation of an practically exploitable knowledge view. The careful reader may have noticed that a transitive closure operation is required, which typically has a complexity of $O(n^4)$. With the amount of concepts typically involved in an adequate ontological representation of a domain, the practical application of the known transitive closure algorithm is impossible. Given the sparse nature of ontological relations, it is possible to construct a computationally efficient algorithm customized for the needs of transitive closure of ontological knowledge representations [14].

The heart of this algorithm is in the incremental update. When a single element is added to an already transitive relation the property of transitivity is damaged, but only locally. Thus, a computationally cheap operation can be utilized, instead of a complete transitive closure operation, in order to restore the property of transitivity. Following this observation, the proposed approach starts from an empty relation and continues by adding one by one the elements in the (sparse) ontological relation, restoring transitivity after each step. The result, for the typical ontological relation, is a complexity of $O(n \log^4)$ compared to a complexity of $O(n^4)$ for the standard approach.

4. Topic Extraction

Although most multimedia processing groups are working towards the identification of the objects contained in a multimedia document, the greater ultimate goal is the estimation of the higher level concepts related to each document; identifying the individual objects is most often the means that the goal.

Of course, even if the participating objects are readily available, directly through analysis algorithms or indirectly through annotation, the estimation of the

underlying topics can be a difficult task. The main problems involved in the process are that

1. A concept may be related to multiple, unrelated topics.
2. A document may be related to multiple, unrelated thematic categories.
3. The list of concepts contained in a multimedia document may have been created in an automated manner. Thus, existence of random and therefore misleading concepts cannot be excluded. For example, concepts that correspond to terms that have been used in a metaphorical sense when annotating a document may be included in the list.
4. Semantic ontological relations are a matter of degree. Therefore, correlation between a document and a topic is also a matter of degree.

In order to detect topics in a multimedia document with the highest possible confidence an agglomerative clustering approach may be used. Clearly, the definition of iter-cluster distances is the fine issue that determines the success of the approach. Our proposal is to use the intensity of the common context, as defined in [2]. This indicates the degree to which the ontological knowledge indicates that the concepts contained in two clusters are indeed semantically related.

The agglomerative clustering approach helps us overcome most of the problems faced in the process. For example, by ignoring clusters of low cardinality we remove concepts that are misleading and should not be considered in the estimation of related topics.

5. User Modeling

Detecting dominant topics in a multimedia document is in essence a process dual to the process of detecting dominant topics in a wide list of documents in a user's usage history. As a result, a similar agglomerative clustering approach can be utilized in order to extract user preferences based on usage history, when using a content based filtering.

Again, the fine issue one needs to tackle is that of defining the inter-cluster distance between clusters in a meaningful way. Given the topic estimation methodology presented in the previous section, it is easy to develop a vector based document similarity metric, with each document represented in a vector space much like in vector space models of information retrieval, the main difference being that the vocabulary used to form the vector only contains higher level concepts, i.e. topics.

As explained in [13], extracted preferences (topics accompanied by a clear overall positive indication) should be stored independently from each other, so that when utilized in the consequent process of multimedia personalization inter-preference noise is avoided. On the other hand, all dislikes (topics accompanied by a clear overall negative indication) may be stored as one, as the discrimination between them has no practical value.

6. Conclusion

This paper comes after a list of previous works on various theoretical and practical topics and before another similar list. Its purpose is to identify independent previous works that may be combined in order to come closer to the ultimate goal of multimedia personalization, indicate the way these works can be combined and identify main directions for future and further research.

Going further from our previous works, we have focused on using an ontological modeling for knowledge, thus allowing our approach to utilize knowledge bases developed by other researchers and making it easier for it to be integrated with other modern approaches. We have also provided a computationally efficient solution to the handling of the considered ontological relations and outlined the ways in which topic estimation and user modeling could be approached.

Clearly, this is only the beginning. The semantic gap has yet to be bridged, and a definite answer to the personalization problem has yet to be given. As a main field for further research we certainly identify the semi-automated generation of the required ontological knowledge. Where other approaches focus on crisp ontological relations our approach is based on a fuzzified version of this knowledge; manually specifying these fuzzy degrees in both time consuming and subjective, so an automated or semi-automated methodology to do this is certainly desired.

Other than that, much more is to be done in the field of experimental evaluation of methodologies, for all or us working in the field of personalization, as ground truth is practically impossible to generate.

References

- [1] G. Akrivas, G. Stamou and S. Kollias, "Semantic Association of Multimedia Document Descriptions through Fuzzy Relational Algebra and Fuzzy Reasoning", *IEEE Trans. on Systems, Man, and Cybernetics*, part A, Volume 34 (2), March 2004
- [2] G. Akrivas, M. Wallace, G. Andreou, G. Stamou and S. Kollias, "Context – Sensitive Semantic Query Expansion", In *Proceedings of the IEEE International Conference on Artificial Intelligence Systems (ICAIS)*, Divnomorskoe, Russia, September 2002
- [3] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, "The Description Logic Hand-book: Theory, Implementation and Application", Cambridge University Press, 2002
- [4] T.R. Gruber, "A Translation Approach to Portable Ontology Specification", *Knowledge Acquisition 5: 199-220*, 1993
- [5] G. Klir and B. Yuan, "Fuzzy Sets and Fuzzy Logic, Theory and Applications", New Jersey, Prentice Hall, 1995
- [6] A. Maedche, B. Motik, N. Silva and R. Volz, "MAFRA - An Ontology MAPPING FRAMework in the Context of the SemanticWeb", In *Proceedings of the Workshop on Ontology Transformation at ECAI2002*, Lyon, France, July 2002
- [7] S. Miyamoto, "Fuzzy Sets in Information Retrieval and Cluster Analysis", Kluwer Academic Publishers, Dordrecht / Boston / London, 1990
- [8] G. Salton, M.J. McGill, "Introduction to modern information retrieval", McGraw-Hill, 1983.
- [9] T. Sikora, "The MPEG-7 Visual standard for content description - an overview", *Special Issue on MPEG-7, IEEE Trans. on Circuits and Systems for Video Technology*, 11/6:696-702, June 2001
- [10] S. Staab and R. Studer, "Handbook on ontologies, international handbooks on information systems", Heidelberg: Springer-Verlag, 2004
- [11] Text REtrieval Conference, <http://trec.nist.gov/>
- [12] M. Wallace, G. Akrivas, Ph. Mylonas, Y. Avrithis, S. Kollias, "Using context and fuzzy relations to interpret multimedia content", In *Proceedings of the 3rd International Workshop on Content-Based Multimedia Indexing (CBMI)*, IRISA, Rennes, France, September 2003
- [13] M. Wallace, Y. Avrithis, G. Stamou, S. Kollias, "Knowledge-based Multimedia Content Indexing and Retrieval", in Stamou G., Kollias S. (Editors) *Multimedia Content and Semantic Web: Methods, Standards and Tools*, Wiley, 2005
- [14] M. Wallace, Y. Avrithis, S. Kollias, "Computationally efficient sup-t transitive closure for sparse fuzzy binary relations", *Fuzzy Sets and Systems* 157(3), pp. 341-372, 2006
- [15] W3C, RDF, <http://www.w3.org/RDF/>
- [16] W3C, RDF Reification, http://www.w3.org/TR/rdf-schema/#ch_reificationvocab
- [17] W3C Recommendation, OWL Web Ontology Language Reference, <http://www.w3.org/TR/owl-ref/>, 10 February 2004