USING CONTEXT AND FUZZY RELATIONS TO INTERPRET MULTIMEDIA CONTENT

Manolis Wallace, Giorgos Akrivas, Phivos Mylonas, Yannis Avrithis and Stefanos Kollias

Department of Computer Science
School of Electrical and Computer Engineering
National Technical University of Athens, Greece
iavr@image.ntua.gr

ABSTRACT

Object detection techniques are coming closer to the automatic detection and identification of objects in multimedia documents. Still, this is not sufficient for the understanding of multimedia content, mainly because a simple object may be related to multiple topics, few of which are indeed related to a given document. In this paper we determine the thematic categories that are related to a document based on the objects that have been automatically detected in it. Our approach relies on stored knowledge and a fuzzy hierarchical clustering algorithm; this algorithm uses a similarity measure that is based on the notion of context. The context is extracted using fuzzy ontological relations.

1. INTRODUCTION

The advances in computer and data networks along with the success of standardization efforts of MPEG and JPEG boosted the movement of archives towards the conversion of their fragile and manually indexed material to digital, computer accessible data. Database management systems (DBMSs) are being designed that are able to handle access to such types of stored information.

The focus of technological attempts has been on the analysis of digital video, due to its large amounts of spatiotemporal interrelations, which turns it into the most demanding and complex data structure. Current and evolving international standardization activities, such as of the EBU, MPEG-4 [4],[5], MPEG-7 [9],[10], or JPEG-2000 [13] for still images, deal with aspects related to data structures and metadata. The objective is to quickly and efficiently search and retrieve audiovisual material, based on its content.

Current state-of-the-art in content based retrieval is based on a query by example (QbE) approach [6]. In this paradigm, users describe their information needs by providing a sample document that matches their desire. The system extracts syntactic information from the sample document, and then compares this information with the corresponding one from available documents (a somewhat more elegant approach permits users to specify syntactic information directly)[7]. This approach is in general defective, mainly due to the following reasons:

- It is not always possible for the user to locate a sample document in order to initiate the process of multimedia retrieval.
- It is not easy for the system to determine which of the features of the sample document should be considered as matching criteria.
- Numerous relevance feedback loops are usually needed in order for the retrieved documents to start matching the real user information need.

Moreover, both approaches are quite difficult to use when the user is not searching for documents that contain a specific object. Most often, user queries are related to higher level concepts or entities, such as thematic categories, whose mapping to simple objects is not defined uniquely.

An alternative approach to QbE is allowing the user to issue textual queries. These are matched with a textual index that is created either by human experts, or automatically, through multimedia content analysis. Since the former is a painful, expensive and time consuming process, the latter is considered to be a field of great importance. In order to tackle the problem of automatic content analysis, prototype objects, together with their corresponding features, as well as their textual forms, are described and stored in an encyclopedia. Object detection and feature matching algorithms are then applied to all available documents, in order to generate the index [12]. This approach seems to be quite promising, as it is the one that facilitates users the most in forming meaningful queries. Still, the automatic generation of the index is a process that is governed by uncertainty.

In this paper we generate a conceptual thematic categorization, based on such an index, thus extending works such as [12], to a more semantic level. Our approach performs

THIS WORK HAS BEEN PARTIALLY FUNDED BY THE EU IST-1999-20502 FAETHON PROJECT.

a fuzzy hierarchical clustering of the semantic entities, relying on knowledge that is stored in the form of semantic relations. The notion of context has a central role in this process.

The structure of the paper is as follows: In Section 2, we present a novel quasi – taxonomic semantic relation. Based on this relation, after formally defining the problem of thematic categorization in Section 3, in Section 4, we rely on the notion of context in order to detect the thematic categories that are related to a document. In Section 5, we provide examples of the proposed methodology and, in Section 6, we present our concluding remarks, as well as possible extensions to our work.

2. BACKGROUND

The analysis of an automatically generated index, which contains inherent uncertainty, is not possible without the use of knowledge. Thus, a knowledge model is necessary. The knowledge model that is gaining momentum is that of ontologies.

2.1. Ontologies

An ontology is a framework for knowledge representation in which the context determines the intended meaning of each word. A word used in different context may have different meanings. In general, ontologies may be described as follows:

$$O = \{S, \{R_i\}\}, i = 1 \dots n$$
$$R_i : S \times S \to \{0, 1\}, i = 1 \dots n$$

where O is an ontology, S the set of semantic entities it describes and R_i the i-th semantic relation amongst the semantic entities. The formal definition of ontologies also supports an inference layer, but herein we omit it for the sake of simplicity.

Although any type of relation may be contained in an ontology, the two main categories are taxonomic (i.e. ordering) and compatibility (i.e. symmetric) relations. Compatibility relations have traditionally been exploited by information retrieval systems for tasks such as query expansion. They are ideal for the description of similarities of various natures, but fail to assist in the determination of the context of a query or a document; the use of ordering relations is necessary for such tasks [1]. Thus, a main challenge of intelligent information retrieval is the meaningful exploitation of information contained in taxonomic relations of an ontology.

It is well understood that relations among real life entities are always a matter of degree, and are, therefore, best modelled using fuzzy relations. Ontological taxonomies, on the other hand, are crisp in principle. Thus, they fail to fully describe real life concepts, and are limited to α-cuts of the desired relations. This is a very important drawback, that makes such relations insufficient for the services that an intelligent information retrieval system aims to offer.

Tab. 1. The fuzzy semantic relations

Sp	Specialization
Ct	Context
Ins	Instrument
P	Part
Pat	Patient
Loc	Location
Ag	Agent

2.2. The Fuzzy Quasi – Taxonomic Relation

The authors have proposed fuzzy semantic relations that are most suitable for the modelling of real life information [2]. In this section, we present a few commonly encountered semantic relations that can be modelled as fuzzy ordering relations, and propose their combination for the generation of a meaningful, fuzzy, quasi-taxonomic relation. Based on this relation, in the following sections we will explain how the problem of automatic thematic categorization may be tackled.

The *specialization* relation Sp is a fuzzy partial ordering on the set of semantic entities. $Sp(a, b) > 0$ means that the meaning of a “includes ”the meaning of b; the most common form of specialization is sub – classing, i.e. a is a generalization of b. The role of the specialization relation in knowledge – based retrieval is as follows: if a document refers to the meaning of entity b, then it is also related to a, since b is a special case of a. Still, there is no evidence that the opposite also holds; it is obvious that the specialization relation contains important information that can not be modelled in a symmetric relation. The *context* relation Ct is also a fuzzy partial ordering on the set of semantic entities. $Ct(a, b) > 0$ means that b provides the context for a or, in other words, that b is the thematic category that a belongs to. Other relations considered in the following have similar interpretations. Their names and corresponding notations are given in Table 1.

In this work, fuzziness of the aforementioned relations has the following meaning: High values of $Sp(a, b)$, imply that the meaning of b approaches the meaning of a, in the sense that when a document is 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 a document’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 a document implies relation to a. Likewise, the degrees of the other relations can also be interpreted as conditional probabilities or degrees of implied relevance.

The above imply that, for example, $a \neq b \implies Sp(a,b) < 1$ since, if $a \neq b$, then we cannot be sure that both a and b are related to a given document, without first examining the document'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.

A last point to consider is the transitivity of the relations presented above. It is obvious that if b is a specialization of a and c is a specialization of b, then c is a specialization of a. This implies that the specialization relation is transitive. A similar argument can be made for the other relations, as well. Still, the form of transitivity used cannot be $\sup - \min$ transitivity, but one relying on a subidempotent norm. Therefore, we demand that the presented relations are $\sup -t$ transitive, where t is an Archimedean norm.

More formally, the knowledge model presented above may be summarized in the following:

$$O_{\mathcal{F}} = \{S, \{r_i\}\}, i = 1 \ldots n$$
$$r_i = \mathcal{F}(R_i) : S \times S \rightarrow [0,1], i = 1 \ldots n$$

Based on the relations r_i we construct the following semantic relation:

$$T = Tr^t(\bigcup_i r_i^{p_i}), p_i \in \{-1, 1\}, i \in 1 \ldots n$$

where $Tr^t(A)$ is the $\sup -t$ transitive closure of relation A; the transitivity of relation T was not implied by the definition, as the union of transitive relations is not necessarily transitive. In our application we construct the T relation as follows:

$$T = Tr^t(Sp \cup C^{-1} \cup Ins \cup P \cup Pat \cup Loc \cup Ag)$$

Based on the semantics of relations r_i, it is easy to see that T is ideal for the determination of the thematic categories that an entity may be related to, as thematic categories are also semantic entities:

$$TC \subseteq S$$

where $TC = \{tc_i\}, i \in 1 \ldots k$ is the set of thematic categories (for example *ball* and *stadium* may be semantic entities, while *football* and *sports* are both semantic entities and thematic categories). Unfortunately, the example of the T relation has to be omitted for the sake of space.

All the relations used for the generation of T are partial ordering relations. Still, there is no evidence that their union is also antisymmetric. Quite the contrary, T may vary from being a partial ordering to being an equivalence relation. This is an important observation, as true semantic relations also fit in this range (total symmetricity as well as total antisymmetricity often have to be abandoned when modelling real life). Still, the semantics of the used relations, as well as our experiments, indicate that T is "almost" antisymmetric. Therefore, we categorize to it as quasi – ordering or quasi – taxonomic.

3. PROBLEM DEFINITION

Before anything else, let us first present the problem that this work attempts to address, in a more formal manner. The intelligent module presented herein accepts as input the Semantic Index I. This is in fact a fuzzy relation between documents and semantic entities. The semantic index must be normal for each document, i.e.:

$$\forall d \in D \;\; \exists s \in S \text{ such that } I(s,d) = 1$$

Based on this relation, and the knowledge contained in the available semantic relations R_i, the module aims to detect the degree to which a given document $d \in D$ is related to a thematic category $tc \in TC$. We will refer to this degree as $R_{\mathrm{TC}}(tc, d)$. In other words, the module attempts to calculate the relation:

$$R_{\mathrm{TC}} : TC \times D \rightarrow [0,1]$$

In designing an algorithm that is able to calculate this relation, in a meaningful manner, a series of issues need to be tackled:

1. A semantic entity may be related to multiple, unrelated thematic categories.

2. A document may be related to multiple, unrelated thematic categories.

3. The semantic index may have been created in an automated manner. Thus, existence of random, and therefore misleading semantic entities cannot be excluded. For example, entities that correspond to terms that have been used in a metaphorical sense when annotating a documents may be included in the index.

4. Semantic relations are always a matter of degree. Therefore, correlation between a document and a thematic category is also a matter of degree.

In the following, we provide the principles of the proposed approach to the problem of thematic categorization. According to issue 1, it is necessary for the algorithm to be able to determine which thematic categories are indeed related to a given document. In order for this task to be performed in a meaningful manner, the common meaning of the remaining entities that index the given document needs to be considered as well.

On the other hand, when a document is related to more than one, unrelated thematic categories, as issue 2 points out, we should not expect all the terms that index it to be related to one another, or to each one of the thematic categories in question. Quite the contrary, we should expect most entities to be related to just one of these thematic categories. Therefore, a clustering of semantic entities, based on the their common meaning, needs to be applied.

In this process, entities that are misleading (eg. entities that resulted from the use of terms in a metaphorical sense) will probably not be found similar with other entities that index a document. Therefore, the cardinality of the clusters may be used to tackle issue 3.

Finally, issue 4 is easily solved by allowing the algorithm to be fuzzy. In the following, we proceed with the presentation of an algorithm which complies with the above principles.

4. THE ALGORITHM FOR THEMATIC CATEGORIZATION

The proposed approach may be decomposed into the following steps:

- Perform a fuzzy clustering of semantic entities, using their common meaning as clustering criterion in order to determine the count of distinct topics that a document is related to.
- Find the thematic categories that are related to each cluster.
- Aggregate the findings for each cluster in order to acquire an overall result for the whole document.

Each of the above steps uses the taxonomy relation, in addition to the index. In the following, after discussing the notion of "common meaning", we elaborate on each of these steps.

4.1. The notion of context

In general, the term *context* refers to whatever is common among a set of elements. In this work, where the elements are semantic entities and documents, the term context may refer to the common meaning of a set of entities, or to the overall topic of a document, respectively.

A document is represented only by its mapping to semantic entities, via the semantic index I. Therefore, the context of a document is again defined via the semantic entities that are related to it. The fact that relation T described in subsection 2.2 is (almost) an ordering relation allows us to use it in order to define, extract and use the context of a document, or a set of semantic entities in general. Relying on the semantics of the T relation, we define the *context* $K(s)$ of a semantic entity $s \in S$ as the set of its descendants in relation T:

$$K(s) = T_{\leq}(s)$$

This set also includes the semantic entity in question.

Assuming that a set of entities $S' \subset S$ is crisp, i.e. all considered entities belong to the set with degree one, the context of the group, which is again a set of semantic entities, can be defined simply as the set of their common descendants.

$$K(S') = \bigcap_i K(s_i), s_i \in S'$$

Obviously, as more entities are considered, the context becomes narrower, i.e. it contains less entities and to smaller degrees. When the definition of context is extended to the case of fuzzy sets of semantic entities, this inequality must still hold. Moreover, we demand that the following are satisfied as well:

- $S'(s) = 0 \implies K(S') = K(S' - \{s\})$, i.e. no narrowing of context.
- $S'(s) = 1 \implies K(S') \subseteq K(s)$, i.e. full narrowing of context.
- $K(S')$ decreases monotonically with respect to $S'(s)$.

Taking these into consideration, we demand that, when S' is fuzzy, the "considered" context $\mathcal{K}(s)$ of s, i.e. the entity's context when taking its degree of participation to the set into account, becomes low when the degrees of taxonomy are low and the degree of participation $S'(s)$ is high. Therefore:

$$cp(\mathcal{K}(s)) \doteq cp(K(s)) \cap (S'(s) \cdot S)$$

where cp is an involutive fuzzy complement, and $\cap$ and $\cup$ correspond to a t-norm and a t-conorm which are dual, with respect to cp. By applying de Morgan's law, we obtain:

$$\mathcal{K}(s) \doteq K(s) \cup cp(S'(s)) \tag{1}$$

Then the set's context is easily calculated as follows:

$$K(S') = \bigcap_i \mathcal{K}(s_i), s_i \in S'$$

Considering the semantics of the T relation and the process of context determination, it is easy to realize that when the entities in a set are highly related to a common meaning, the context will have high degrees of membership for the entities that represent this common meaning. Therefore, the height of the context $h(K(S'))$ may be used as a measure of the semantic correlation of entities in set S'. We will refer to this measure as *intensity* of the context.

4.2. Hierarchical clustering

Before actually extracting thematic category information from the set of semantic entities $I(d)$ that are related to a document d via the semantic index I, in order to support the possibility of existence of multiple distinct topics in a single document, the support of the document's description, i.e. the set

$$I(d) = \{s \in S : I(s, d) > 0\}$$

of the entities that are related to it needs to be clustered to groups, according to the topics they are related to.

The general structure of agglomerative clustering algorithms, adjusted for the needs of the problem at hand, is as follows [11]:

1. When considering document d, turn each semantic entity $s \in I(d)$ into a singleton, i.e. into a cluster of its own.
2. For each pair of clusters c_1, c_2 calculate a compatibility indicator $CI(c_1, c_2)$. The CI is also referred to as cluster similarity, or dissimilarity, measure.
3. Merge the pair of clusters that have the best CI. Depending on whether this is a similarity or a dissimilarity measure, the best indicator could be the maximum or the minimum operator, respectively.
4. Continue at step 2, until the termination criterion is satisfied. The termination criterion most commonly used is the definition of a threshold for the value of the best compatibility indicator.

The two key points in hierarchical clustering are the identification of the clusters to merge at each step, i.e. the definition of a meaningful measure for CI, and the identification of the optimal terminating step, i.e. the definition of a meaningful termination criterion.

When clustering semantic entities, the ideal similarity measure is one that quantifies their semantic correlation. In subsection 4.1 we have defined such a measure; it is the height of their common context. Therefore, the merging of clusters will be based on this measure.

The process of merging should terminate when the entities are clustered into sets that correspond to distinct topics. We may identify such sets by the fact that their common contexts will have low, if not zero, intensity. Therefore, the termination criterion shall be a threshold on the intensity of the common meaning, i.e. a threshold on the selected compatibility measure.

4.3. Fuzzy clustering

Hierarchical clustering methods are more flexible than their partitioning counterparts, in that they do not need the number of clusters as an input. Still, they are less robust in other ways:

- They only create crisp clusterings, i.e. they do not support degrees of membership in their output.
- They only create partitions, i.e. they do not allow for overlapping among the detected clusters.

Both of the above are great disadvantages for the problem at hand, as they are not compatible with the task's semantics: in real life, a semantic entity may be related to a topic to a degree other than 1 or 0, and may also be related to more than one distinct topics.

In order to overcome such problems, we describe in the following a method for fuzzyfication of the partitioning. In this way the clusters' cardinalities will be corrected, so that they may be used in subsection 4.4 for the meaningful extraction of thematic categories.

Each cluster c is described by the crisp set of semantic entities S_c that belong to it. Using those, we may create a fuzzy classifier, i.e. a function C_c that will measure the degree of correlation of a semantic entity s with the cluster c.

$$C_c : S \to [0, 1]$$

Obviously, a semantic entity should be considered correlated with c, if it is related to the common meaning of the semantic entities in S_c. Therefore, the quantity

$$Cor_1(c, s) = h(K(S_c \cup \{s\}))$$

where $h(\cdot)$ symbolizes the height of a fuzzy set, is a meaningful measure of correlation. Of course, not all clusters are equally compact; we may measure cluster compactness using the similarity among the entities it contains, i.e. using the intensity of the cluster's context. Therefore, the aforementioned correlation measure needs to be adjusted, to the characteristics of the cluster in question:

$$Cor_2(c, s) = \frac{Cor_1(c, s)}{h(K(c))}$$

It is easy to see that this measure obviously has the following properties:

- $Cor_2(c, s) = 1$ if the semantics of s imply it should belong to c. For example $Cor_2(c, s) = 1, \forall s \in S_c$
- $Cor_2(c, s) = 0$ if the semantics of s imply it should not belong to c.
- $Cor_2(c, s) \in (0, 1)$ if s is neither totally related, nor totally unrelated to c.

These are the properties that we wish for the cluster's fuzzy classifier, so:

$$C_c(s) \doteq Cor_2(c, s)$$

Using such classifiers, we may expand the detected crisp partitions, as to include more semantic entities, as follows: partition c is replaced by cluster

$$c' = \sum_{s \in I(d)} s / C_c(s)$$

Obviously $c' \supseteq c$.

4.4. Extraction of thematic categories

Thematic categories are semantic entities that have been selected as having a special meaning for the system; more formally:

$$TC \subset S$$

This simplifies the process of automatic thematic categorization: We have already explained that the context of a set of semantic entities is a fuzzy set of semantic entities; this contains the entities that describe the common meaning of the original set. The thematic categories that are contained in the context of a cluster of semantic entities are obviously thematic categories that are related to the whole document. Based on this concept, in the following we present a method for automatic thematic categorization of documents.

First of all, the process of fuzzy hierarchical clustering has been based on the crisp set $I(d)$, thus ignoring fuzziness in the semantic index. In order to incorporate this information in the clusters of semantic entities considered for the process of thematic categorization, we adjust the degrees of membership for them as follows:

$$c''(s) = t(c'(s), I(s, d)) \tag{2}$$

where t is a fuzzy norm. The semantic nature of this operation demands that t is an Archimedean norm. From each one of those clusters, we may extract the corresponding thematic categories. In the following we shall refer to a random fuzzy cluster c'' and its corresponding fuzzy set of thematic categories $R_{\mathrm{TC}}(c'')$.

Obviously, thematic categories that are not contained in the context of c'' cannot be selected as being related to it. Therefore

$$R_{\mathrm{TC}}(c'') \subseteq R^1_{\mathrm{TC}}(c'') \doteq w(K(c'') \cap TC) \tag{3}$$

where w is a weak *modifier*. Modifiers, which are also met in the literature as *linguistic hedges* [8], are used (in this work) to adjust mathematically computed values so as to match their semantically anticipated counterparts. Specifically, our experiments indicate, for example, that a value of 0.7 for the expression $K(c'') \cap TC$ corresponds to a great degree of relevance, and should, therefore, be adjusted accordingly.

In the case that the semantic entities that index document d are all clustered in a unique cluster c'', then $R_{\mathrm{TC}}(d) = R^1_{\mathrm{TC}}(c'')$ is a meaningful approach. On the other hand, when more than one clusters are detected, then cluster cardinalities have to be considered as well.

Clusters of extremely low cardinality probably only contain misleading entities, and therefore need to be ignored in the estimation of $R_{\mathrm{TC}}(d)$. On the contrary, clusters of high cardinality almost certainly correspond to the distinct topics d is related to, and need to be considered in the estimation of $R_{\mathrm{TC}}(d)$. The notion of "high cardinality" is modelled with the use of a "big" fuzzy number L. $L(a)$ is the truth value of the preposition "the cardinality of a is high".

The set of thematic categories that correspond to a document is computed from the remaining clusters, after adjusting membership degrees according to scalar cardinalities, as follows:

$$R_{\mathrm{TC}}(d) \doteq \underset{c'' \in G}{u} (R_{\mathrm{TC}}(c'')) \tag{4}$$

$$R_{\mathrm{TC}}(c'') = R^1_{\mathrm{TC}}(c'') \cdot L(|c''|) \tag{5}$$

where u is a fuzzy co-norm, G is the set of fuzzy clusters that have been detected in $I(d)$ and have had their membership degrees adjusted according to equation 2, and $|b|$ is the scalar cardinality of set b.

It is easy to see that $R_{\mathrm{TC}}(d, tc)$ will be high if a cluster c'', whose context contains tc, is detected in $I(d)$, and additionally, the cardinality of c is high (i.e. the cluster is most probably not comprised of misleading entities) and the degree of membership of tc in the context of c'' is high.

5. EXAMPLES

In this paper we tackle the problem of automatic thematic categorization with tools borrowed from the field of fuzzy set theory. The best performance has been observed, by trial and error work, when using the operators mentioned below:

- In subsection 2.2, the t-norm used for the transitive closure of relation T is Yager's t-norm with parameter 3 [8].
- In equation 1, the co-norm u used is the bounded sum, while the complement cp of choice is the standard complement: $cp(a) = 1 - a$
- In equation 2, the t-norm used is the product.
- In equation 3, the modifier used is $w(a) = \sqrt{a}$

Tab. 2. Semantic Entity names

S. Entity	Mnem.		S. Entity	Mnem.
arts	**art**		uniform	unf
tank	tnk		lawn	lwn
missile	msl		goal	gol
scene	scn		shoot	sht
war	**war**		tier	tir
cinema	**cnm**		river	riv
performer	prf		speak	spk
sitting person	spr		F16	f16
explosion	exp		football player	fpl
missile launch	lms		goalkeeper	glk
screen	scr		**theater**	**thr**
football	**fbl**		fighter-plane	far
curtain	crn		seat	sit

Tab. 3. The thematic categorization fuzzy relation

s_1	s_2	$TC(s_1, s_2)$	s_1	s_2	$TC(s_1, s_2)$
war	unf	0.90	war	exp	0.60
war	far	0.80	fbl	gol	0.80
war	tnk	0.80	fbl	sit	0.60
war	msl	0.80	cnm	sit	0.60
thr	scn	0.90	fbl	sht	0.90
thr	prf	0.90	fbl	tir	0.80
thr	spr	0.80	fbl	fpl	0.90
war	lms	0.70	fbl	lwn	0.90
cnm	scr	0.90	cnm	spr	0.80
fbl	spr	0.60	thr	sit	0.60
thr	crn	0.70	art	thr	0.80
far	f16	1.00	art	cnm	0.80
fpl	glk	1.00			

- In equation 4, the standard co-norm (max) is used.

Moreover,

- In Section 4.2, the threshold used for the termination criterion of the clustering algorithm is 0.3.
- In equation 5, large fuzzy number L is defined as the triangular fuzzy number $(1.3, 3, \infty)$.

In the following, we demonstrate the efficiency of our approach when using the aforementioned operators and thresholds, by presenting some numerical examples.

The semantic entities included in the exemplar system are shown in Table 2, with thematic categories shown in boldface. The thematic categorization relation TC is shown in Table 3. Zero elements of the relation, as well as elements that are implied by reflexivity are omitted, for the sake of simplicity.

In the following, we apply our algorithm on a set of manually indexed documents; this emulates automatic semantic indexing, while taking into account weaknesses of multimedia analysis techniques we have discussed in Section 1. A portion of the semantic index is shown in Table 4. The results of the algorithm are shown in Table 5.

Document d_1 contains a shot from a theater hall. The play is war-related. We can see that objects and events are detected with a limited degree of certainty. Furthermore, detected entities are not always directly related to the overall topic of the document (for example a "tank" may appear in a shot from a theater, as a part of the play, but this is not a piece of information that can aid in the process of thematic categorization). The algorithm ignores "tank" and "speak".

Document d_2 contains a shot from a cinema hall. The film is again war-related. Although some entities are common between d_1 and d_2 (and they are related to both "theater" and "cinema"), the algorithm correctly detects that in this case the overall topic is different. This is accomplished by considering that "screen" alters the context and thus the overall meaning.

As a last example, let us present document d_3, which is a sequence of shots from a news broadcast. Due to the diversity of stories presented in it, the semantic entities that are detected and included in the index are quite unrelated to each other:
$d_3 =$ spr/0.9 + unf/0.8 + lwn/0.5 + gol/0.9 + tir/0.7 + spk/0.9 + glk/0.8 + sht/0.5 + prf/0.7 + sit/0.9 + crn/0.7 + scn/0.8 + tnk/0.9 + msl/0.8 + exp/0.9 + riv/1

After the consideration of the fuzziness of the index, the following five fuzzy clusters of entities are created:
$c_1 =$ spk/0.9
$c_2 =$ riv/1.0
$c_3 =$ spr/0.9 + prf/0.7 + sit/0.77 + crn/0.7 + scn/0.8
$c_4 =$ spr/0.9 + lwn/0.5 + gol/0.9 + tir/0.7 + glk/0.8 + sht/0.5 + sit/0.9
$c_5 =$ unf/0.8 + tnk/0.9 + msl/0.8 + exp/0.9

First of all, we can observe that the algorithm successfully identifies the existence of more than one distinct topics in the document. Furthermore, entities such as "seat" and "sitting-person" are assigned to more than one clusters, as they are related to more than one of the contexts that are detected in the document.In the following steps of the algorithm, the first two clusters are ignored, due to their small scalar cardinality.

6. CONCLUSIONS

In this paper, we started by presenting a fuzzy, quasi – ordering, semantic relation defined on the set of semantic entities. Continuing, we identified the main obstacles that have to be faced in the process of automatic detection of thematic categories that are related to a semantically indexed document,

Tab. 4. The semantic index

s	$d_1(s)$		s	$d_2(s)$
prf	0,9		spr	0.9
spr	0.9		spk	0.8
spk	0.6		sit	0.9
sit	0.7		scr	1.00
crn	0.8		tnk	0.4
scn	0.9			
tnk	0.7			

Tab. 5. The result of thematic categorization

	d_1	d_2	d_3
arts	0.84	0.73	0.85
cinema		0.74	0.86
theater	0.89		0.33
football			0.77
war			0.77

and explained how this can be achieved, using the notion of context; our approach relies on fuzzy hierarchical clustering of the fuzzy index.

The method presented in this paper has been developed and tested in the experimental prototype of the FAETHON multimedia information retrieval system [3]. FAETHON possesses an experimental semantic encyclopedia, as described in Section 2. It contains definitions for numerous semantic entities, about 20% of which are thematic categories, as well as definitions for various semantic relations. FAETHON also possesses a fuzzy semantic index for numerous documents from a/v archives.

Thematic categorization is exploited in numerous ways. As most important we may mention definition and extraction of user preferences at a semantic level, providing of efficient content browsing services to users, timely estimation of the content of relevance feedback based on thematic categorization of documents and automatic suggestion of documents that are related to the document a user is currently viewing.

A major area of future research for this work is the selection of optimal fuzzy operators for most meaningful semantic output. Our findings so far indicate that this selection is not independent from the knowledge itself. In other words, different semantic encyclopedias may perform best for different choices of operators. Thus, the connection between encyclopedia content and operator selection is also an interesting area for research.

7. REFERENCES

[1] G. Akrivas, M. Wallace, G. Andreou, G. Stamou and S. Kollias, ”Context - Sensitive Semantic Query Expansion”, Proceedings of ICAIS, Divnomorskoe, Russia, September 2002.

[2] G. Akrivas G. and G. Stamou, Fuzzy Semantic Association of Audiovisual Document Descriptions, Proceedings of VLBV, Athens, Greece, Oct. 2001.

[3] Avrithis Y. and Stamou G., FAETHON: Unified Intelligent Access to Heterogenous Audiovisual Content, Proceedings of VLBV, Athens, Greece, Oct. 2001.

[4] Battista, S., Casalino, F., Lande C.: MPEG-4: A Multimedia Standard for the Third Millenium, Part 1. IEEE Multimedia 6 (4), 74-83, 1999.

[5] Battista, S., Casalino, F., Lande, C.: MPEG-4: A Multimedia Standard for the Third Millenium, Part 2. IEEE Multimedia 7 (1), 76-84, 2000.

[6] Ciocca G. and Schettini R., A relevance feedback mechanism for content – based image retrieval, Information Processing and Management, Vol. 35, 605-632, 1999.

[7] Di Sciascio, E., Donini, F.M., Mongiello, M., A Logic for SVG Documents Query and Retrieval, Proceedings of the Workshop on Multimedia Semantics at SOFSEM, Milovy, Czech Republic, November 2002.

[8] G. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic, Theory and Applications, New Jersey, Prentice Hall, 1995.

[9] Nack, F., Lindsay, A.: Everything You Wanted to Know About MPEG-7: Part 1. IEEE Multimedia 6 (3), 65-77, 1999.

[10] Nack, F., Lindsay, A.: Everything You Wanted to Know About MPEG-7: Part 2. IEEE Multimedia 6 (4), 64-73, 1999.

[11] S. Theodoridis and K. Koutroumbas, Pattern Recognition, Academic Press, 1998.

[12] Tsechpenakis G., Akrivas G., Andreou G., Stamou G. and Kollias S., Knowledge - Assisted Video Analysis and Object Detection, EUNITE, Albufeira, Portugal, September 2002.

[13] ISO/IEC JTC1/SC29/WG1 N1646R: JPEG 2000 Part I Final Committee Draft Version 1.0 (2000).