Semantic Association of Multimedia Document Descriptions through Fuzzy Relational Algebra and Fuzzy Reasoning

G. Akrivas, G. B. Stamou and S. Kollias

Abstract— **According to the emerging MPEG-7 standard, the semantic description of multimedia documents is expressed in terms of semantic entities such as objects, events, concepts, and relations among them. The semantic entities can be used as index terms, in order to support the semantic search process. In this paper, we propose a method that (a) applies fuzzy relational operations (closure, composition) and fuzzy rules to expand a semantic encyclopedia and (b) uses the encyclopedia to associate the semantic entities with the aid of a fuzzy thesaurus. This method is shown to reduce the need for human intervention in creating semantic descriptions of multimedia documents, as well as correct for incompleteness and inconsistency.**

Keywords— **Human – Computer Interaction, Multimedia databases, Information retrieval, Fuzzy relations**

I. INTRODUCTION

Lately, there is a growing interest in the representation, storage and retrieval of multimedia documents containing textual as well as audiovisual information. The need for *content – based retrieval* has resulted in a need for more sophisticated description of multimedia documents. Unlike text documents, whose *content* can be used to describe them, multimedia documents need to have a set of *features* extracted out of the content, as audiovisual data themselves are meaningless to humans. The features are represented via *descriptors*, which are data structures designed to support matching between features. User queries and consequent search refer to specific features of the content. Documents whose features match the user query, in terms of the corresponding descriptors, are retrieved [1].

MPEG-7 [2] is a new standard for the description of multimedia content, designed to assist content-based access to multimedia data [3],[pp 688-695]. The standard defines three kinds of information that comprise the description, that is Creation and Usage Information, Structural Information and Semantic Information [3][ch. 11-12]. The former regards mostly textual information, commonly known as metadata. Structural aspects express a low – level and machine – oriented kind of description, since they describe content in the form of signal segments and their properties. On the other hand, conceptual aspects express a high – level and human – oriented kind of description, since they deal with semantic entities, such as objects, events and relations among them. These differences result in different approaches both for the creation of the description,

The authors are with the Image, Video and Multimedia Laboratory of the National Technical University of Athens, Greece. Contact: http://www.image. ntua.gr/˜gakrivas

with automatic extraction of features complementing manual annotation, and for the query form, with query by example extending traditional term – based queries.

Usage of conceptual description has advantages over structural description, because of its proximity to human understanding of multimedia information. A conceptual description can be either automatic or manual. Currently, automatic analysis algorithms extract specific structural features. These features can, under application – specific assumptions, be mapped to semantic ones. However, the range of entities that can be extracted in this way is limited. A manual description, on the other hand, exploits the analyzing capabilities of a human expert (annotator)[4]. This has its own disadvantages, however, partially because it is tedious and time – consuming and partly because it suffers from subjectivity and partiality; different people will often consider different aspects of the content as important, and exhaustive annotations are considered unrealistic.

Thus, both automatically and manually generated conceptual descriptions tend to be partial and incomplete. Textual information retrieval systems have been using a thesaurus as a means to find terms that are associated to the terms of a document. Since semantic association is usually a matter of degree, thesauri are often fuzzy [3][5][6]. By using the thesaurus, the user query can be *expanded* to contain all associated semantic entities. The expanded query is expected to retrieve more relevant documents, even in the case of incomplete indexing, because of the higher probability that the annotators have included one of the associated entities in their descriptions.

In order to apply the above scheme in multimedia retrieval systems, one needs a mapping between textual terms and semantic entities stored in a knowledge base. Until recently, this was difficult, due to the lack of universally accepted standards for multimedia descriptions. MPEG-7, as well as recent developments towards a universal ontology language, enable different multimedia retrieval systems to disseminate their knowledge. In this paper, we assume that relations among semantic entities have been stored in the knowledge base, in MPEG-7 format. The proposed method employs fuzzy relational operations and fuzzy rules, first to enrich an existing knowledge base (thus reducing the need for manual construction), and second to construct a fuzzy thesaurus based on it. A uniform application of rules enforces the consistency of the data. The system is being used by the FAETHON project (see section V), whose main goal is to semantically unify heterogenous multimedia archives.

The paper is organized as follows. In section II, we present the MPEG-7 semantic entities and relations, and introduce the notions of the semantic encyclopedia and the fuzzy thesaurus. The method for the automatic construction of the thesaurus is also given. In section III, we provide a theoretical framework on algebraic operations on fuzzy relations. In section IV, we describe the fuzzy algebraic and logical rules used to expand the encyclopedia and construct the thesaurus. In section V, we present the integrated system, and its role in a multimedia application. Finally, in section VI we provide conclusions and suggestions for future work.

II. THE KNOWLEDGE BASE AND THE THESAURUS

The MPEG-7 standard refers to the reality in which a description makes sense as a *narrative world*. The *description tools* (data structures) that compose a narrative world are the set of its *semantic entities* S and the set of its *semantic relations* R. S is further partitioned into the sets of objects O, events E, concepts C, places P, times T and states A. Objects and events can be viewed as corresponding to nouns and verbs of natural language, respectively. It is important that there exist in the same set both specific semantic entities (*instances*, e.g. "George", "Spanish soccer team") and classes of objects, (*formal abstractions*, e.g. "human", "soccer team").

The standard defines a rich set of relations among semantic entities. Each relation consists of pairs of semantic entities, and an optional degree of correlation (fuzziness). The semantic relations defined by the standard are shown in Table I. An example of a semantic description based on the principles of MPEG-7 presented above is given in section V.

The knowledge base contains, in the MPEG-7 terminology, the narrative world, to which the specific document semantic descriptions refer. Such a description contains most of the abstract entities and their relations, since they are more likely to appear in multiple descriptions, and some of the frequently encountered non-abstract entities. We will refer to this special kind of description as a *semantic encyclopedia*. Such an encyclopedia is a semantic generalization of the dictionary (the set of index terms) used by textual information retrieval systems. Similarly to the semantic description of a multimedia document, the construction of a semantic encyclopedia is done by a human expert, who is required to define the semantic entities and supply the relations among them.

A *thesaurus* is a small set of semantic relations, whose usefulness is to expand a user's query in order to achieve higher recall. The relations of the thesaurus have two differences, compared to the ones of a semantic description: they are inherently fuzzy, while the MPEG-7 ones are mostly crisp, and they have greater density. In particular, most of the lines of the thesaurus relations are non-empty, i.e. most of the entities are correlated with at least one entity; by contrast, the relations of MPEG-7 have most of their lines empty, i.e. each relation correlates a small subset of the set of entities.

TABLE I
MPEG-7 SEMANTIC RELATIONS AND THEIR PROPERTIES

Relation	Ref	Sym	Tr	c_E	c_I	c_A
accompanier	-1	-1				0.5
agent	-1	-1				0.5
beneficiary	-1	-1				0.5
causer	-1	-1				0.5
component	-1	-1	1		1	1
context	-1	-1	1			0.8
contrast	-1	1				0.8
dependency	1	1	1		0.8	0.8
depiction	-1	-1				0.8
destination	-1	-1				0.6
entailment	-1		1			1
example	-1	-1	1		1	1
experiencer	-1	-1				0.5
goal	-1	-1				0.5
identifier	1	1	1	1		1
instrument	-1	-1				0.5
interchangeable	1	1	1	1		1
interpretation	-1	-1	1			0.7
location	-1	-1				0.6
manner	-1	-1				0.5
opposite	-1	1	-1			0.7
part	-1	-1	1		1	1
path	-1	-1				0.6
patient	-1	-1				0.5
property	-1	-1	1			0.8
quality	-1	-1	1			0.4
representation	-1	-1				0.8
result	-1	-1				0.5
similar	1	1				1
source	-1	-1				0.6
specialization	-1	-1	1		1	1
state	-1	-1				0.5
stimulus	-1	-1				0.5
substance	-1	-1	1			1
summary	-1	-1	1			0.5
symbol	-1	-1	1			0.5
theme	-1	-1				0.5
time	-1	-1				0.6
user	-1	-1				0.6

The proposed thesaurus for multimedia databases is composed of the following relations: equivalence, inclusion, association. These will be denoted with the symbols E, I, A. Since the MPEG-7 relations are mostly crisp, it makes sense to compute the relations of the thesaurus as a weighted union of the MPEG-7 relations:

$$E = \bigcup_i c_{E,i} r_i$$
$$I = \bigcup_i c_{I,i} r_i$$
$$A = \bigcup_i c_{A,i} r_i$$

where $c_{E,i}, c_{I,i}, c_{A,i}$ are the weights of the union, taking values in $[0, 1]$, and r_i are the MPEG-7 relations. The proposed values are shown in Table I, in the last three

columns.

By using the E relation, which is an equivalence relation, the query is expanded to contain terms that are, in some sense, synonymous.

The I relation, which is a partial ordering relation, on the other hand, implies generalization. For example, if the query requests a player, it is natural to retrieve a description containing a goalkeeper, because it is a specialization of a player. On the other hand, if the query requests for a goalkeeper, a description containing any player is not acceptable. This asymmetry is found in many of the relations defined by the MPEG-7 standard.

Finally, the A relation, which is a compatibility relation, implies that two entities are related, however loosely. This relation associates terms that are less strongly related than the other two and is intended to retrieve results that the other two cannot retrieve. Most of the MPEG-7 relations are relevant for this relation.

III. OPERATIONS ON SEMANTIC RELATIONS

A. The theoretical framework

Let us first present in brief the mathematical framework of fuzzy relations. For more information on fuzzy sets theory, we direct the reader to [7].

A *binary fuzzy relation* on a set $S = \{s_1, s_2, \ldots, s_n\}$ is defined as a function $R : S^2 \to [0, 1]$.

The $\sup -t$ composition of two relations P, Q is defined as $[P \overset{t}{\circ} Q](x, y) = \sup_z t(P(x, z), Q(z, y))$ The inverse of a composition is: $(P \circ Q)^{-1} = Q^{-1} \circ P^{-1}$

The identity relation, I, is the identity element of $\sup -t$ composition: $R \overset{t}{\circ} I = I \overset{t}{\circ} R = R, \forall R$.

The properties of reflectivity ($I \subseteq R$), symmetricity ($R = R^{-1}$) and $\sup -t$ transitivity ($R \overset{t}{\circ} R \subseteq R$) are often encountered. A transitive closure of a relation is the smallest transitive relation that contains the original relation [7]. The transitive closure of a relation is given by the formula $Tr(R) = \bigcup_{n=1}^{\infty} R^{(n)}$, where $R^{(n)} = R \overset{t}{\circ} R^{(n-1)}$ and $R^{(1)} = R$

Similar operations can be applied in order to construct a reflective ($Ref(R) = R \cup I$), antireflective ($Aref(R) = R \cap \bar{I}$) and symmetric ($Sym(R) = R \cup R^{-1}$) relation. On the other hand, an antisymmetric closure requires one of the associations among two semantic entities to be eliminated (set to zero). In this work, the smaller association (the one with the lowest value) is eliminated.

The properties of equivalence (reflective, symmetric and transitive), compatibility (reflective and symmetric) and ordering (antisymmetric and transitive) are often found in fuzzy relations. These properties solely arise from the meaning of relations, and do not depend on specific entities. However, data supplied during the annotation procedure, are not always consistent with them. For example, the annotator might have stated that a goalkeeper is a player, and that a player is an athlete, but not that a goalkeeper is an athlete. A transitive closure will correct this incon-

sistency. Table I shows the properties possessed by the MPEG-7 relations in columns 2-5. A value of 1 means that a relation possesses a property (reflectivity, symmetricity, transitivity), while a value of -1 implies the opposite property (antireflectivity, antisymmetricity, antitransitivity).

A part of a specialization relation, before the total ordering closure is shown in table II. It can be seen that footballPlayer-ag is a specialization of athlete-ag, and that footballPlayer-ag has a number of specializations. In table III, it can be seen that after the total ordering closure, the athlete-ag has the same specializations.

B. The equivalence closure of a relation

In this subsection, we show that the relation $Eq(R) = Tr(Sym(Ref(R)))$ is the equivalence closure of R and give some of its properties.

Theorem 1: 1. The composition of a reflective relation with a reflective relation is a reflective relation

2. The transitive closure of a symmetric relation is a symmetric relation

3. The transitive closure of a reflective relation is a reflective relation

4. $Sym(Ref(A)) = Ref(Sym(A))$

5. $Ref(Tr(A)) = Tr(Ref(A))$

Proof:

1. Since A and B are reflective, $A(x, x) = B(x, x) = 1$. $\left(A \overset{i}{\circ} B\right)(x, x) = \sup_z i[A(x, z), B(z, x)] =$
$i[A(x, x), B(x, x)]$ (monotonicity of t-norms)
$= i(1, 1) = 1$ (boundary condition of t-norms)

2. As already mentioned, the transitive closure of A is defined as $Tr(A) = \bigcup_{n=1}^{\infty} A^{(n)}$. Therefore, it suffices to prove that $A^{(n)}$ is symmetric, for all n. The proof of symmetricity is inductive:

- First, we prove that $A \overset{i}{\circ} A$ is symmetric:
Since A is symmetric, $A(x, y) = A(y, x)$.
$\left(A \overset{i}{\circ} A\right)(x, y) = \sup_z i[A(x, z), A(z, y)] =$
$\sup_z i[A(y, z), A(z, x)]$ (commutativity of t-norms)
$= \left(A \overset{i}{\circ} A\right)(y, x)$

- Next, we prove that if $A^{(n)}$ is symmetric, then $A^{(n+1)}$ is symmetric.
$A^{(n)} = \left(A^{(n)}\right)^{-1} \implies A \overset{i}{\circ} A^{(n-1)} = \left(A \overset{i}{\circ} A^{(n-1)}\right)^{-1}$ (definition of $A^{(n)}$)
$= \left(A^{(n-1)}\right)^{-1} \overset{i}{\circ} A^{-1}$ (inversion of composition)
$= A^{(n-1)} \overset{i}{\circ} A$ (symmetricity of $A^{(n-1)}$ and A)
$A^{(n+1)} = A \overset{i}{\circ} A^{(n)} = A \overset{i}{\circ} A \overset{i}{\circ} A^{(n-1)} = A \overset{i}{\circ} A^{(n-1)} \overset{i}{\circ} A$ (associativity of composition and the above property)
$\left(A^{(n+1)}\right)^{-1} = \left(A \overset{i}{\circ} A^{(n)}\right)^{-1} = \left(A^{(n)}\right)^{-1} \overset{i}{\circ} A^{-1} = \left(A^{(n)}\right) \overset{i}{\circ} A = A \overset{i}{\circ} A^{(n-1)} \overset{i}{\circ} A = A^{(n+1)}$ (by the above property)

3. Because $Tr(A) = \bigcup_{n=1}^{\infty} A^{(n)}$ and prop. 1 above, $Tr(A)$ is a union of reflective relations, hence a reflective relation.

4. $Ref\left(Sym\left(A\right)\right) = A \cup A^{-1} \cup I$

$Sym\left(Ref\left(A\right)\right) = A \cup I \cup \left(A \cup I\right)^{-1} = A \cup I \cup A^{-1} \cup I = A \cup A^{-1} \cup I = Ref\left(Sym\left(A\right)\right)$

5. $Ref\left(Tr\left(A\right)\right) = \bigcup_{n=1}^{\infty} A^{(n)} \cup I$

$Tr\left(Ref\left(A\right)\right) = \bigcup_{n=1}^{\infty} \left(A^{(n)} \cup I\right) = \bigcup_{n=1}^{\infty} A^{(n)} \cup I = Ref\left(Tr\left(A\right)\right)$ ∎

Thus $Eq(A)$ is indeed the equivalence closure of A. Moreover, $\sup-t$ composition preserves the reflectivity property. Before continuing, it is interesting to note some properties that do *not* hold:

• The composition of a symmetric relation with an equivalence relation is generally *not* a symmetric relation

• The composition of a transitive relation with an equivalence relation is generally *not* a transitive relation

• The composition of an equivalence relation with an equivalence relation is generally *not* an equivalence relation

In our application, relations are updated based on other relations through operations such as composition. Since composition does not preserve the properties, with the exception of reflectivity, it is necessary to apply a closure. Moreover, as we will see in section V, operations and closures are applied repeatedly, until the relations no longer change. Theorem 2.8 shows that, for equivalence relations, it is sufficient to perform composition and closure just once.

Theorem 2: The following properties are satisfied by the equivalence closure of the sup-t composition:

1. If A is reflective, then $A \subseteq A \overset{i}{\circ} A$

2. If A is reflective, then $\forall B, B \subseteq A \overset{i}{\circ} B$ and $B \subseteq B \overset{i}{\circ} A$

3. If B is reflective, then $\forall A, Eq\left(A \overset{i}{\circ} B\right) \subseteq Eq\left(A \overset{i}{\circ} B\right) \overset{i}{\circ} B$ and $Eq\left(B \overset{i}{\circ} A\right) \subseteq Eq\left(B \overset{i}{\circ} A\right) \overset{i}{\circ} B$

4. $A \subseteq Eq\left(A\right)$

5. If A is reflective, then $\forall B, B \subseteq Eq\left(A \overset{i}{\circ} B\right)$ and $B \subseteq Eq\left(B \overset{i}{\circ} A\right)$

6. If A is reflective and transitive, then $A = A \overset{i}{\circ} A$

7. If A is reflective, then $\forall B, Eq\left(A \overset{i}{\circ} B\right) \supseteq Eq\left(A \overset{i}{\circ} B\right) \overset{i}{\circ} B$ and $Eq\left(B \overset{i}{\circ} A\right) \supseteq Eq\left(B \overset{i}{\circ} A\right) \overset{i}{\circ} B$

8. If A and B are reflective, then:

$Eq\left(A \overset{i}{\circ} B\right) = Eq\left(A \overset{i}{\circ} B\right) \overset{i}{\circ} B$ and

$Eq\left(B \overset{i}{\circ} A\right) = Eq\left(B \overset{i}{\circ} A\right) \overset{i}{\circ} B$

Proof:

1. $\left(A \overset{i}{\circ} A\right)(x,y) = \sup_z i\left[A\left(x,z\right), A\left(z,y\right)\right]$

for $z = x$: $\left(A \overset{i}{\circ} A\right) = i\left[A\left(x,x\right), A\left(x,y\right)\right] = i\left[1, A\left(x,y\right)\right] = A\left(x,y\right)$. Therefore, $A(x,y) \leq \left(A \overset{i}{\circ} A\right)(x,y)$.

2. $I \subseteq A \implies I \overset{i}{\circ} B \subseteq A \overset{i}{\circ} B$ (because composition is monotonically increasing)

$\implies B \subseteq A \overset{i}{\circ} B$

$I \subseteq A \implies B \overset{i}{\circ} I = B \subseteq B \overset{i}{\circ} A$

3. Directly, by 2.

4. $A \subseteq Ref\left(A\right) \subseteq Sym\left(Ref\left(A\right)\right) \subseteq Tr\left(Sym\left(Ref\left(A\right)\right)\right)$

5. $B \subseteq A \overset{i}{\circ} B$ (by prop. 2 above)

$\subseteq Eq\left(A \overset{i}{\circ} B\right)$ (by prop. 4 above)

Likewise, $B \subseteq B \overset{i}{\circ} A \subseteq Eq\left(B \overset{i}{\circ} A\right)$

6. By prop. 1 and the definition of transitivity, i.e. $A \overset{i}{\circ} A \subseteq A$.

7. $B \subseteq Eq\left(A \overset{i}{\circ} B\right)$ (by prop. 5)

$\implies Eq\left(A \overset{i}{\circ} B\right) \overset{i}{\circ} B \subseteq Eq\left(A \overset{i}{\circ} B\right) \overset{i}{\circ} Eq\left(A \overset{i}{\circ} B\right)$ (because composition is monotonically increasing)

$= Eq\left(A \overset{i}{\circ} B\right)$ (by prop. 6)

8. By props. 3 and 7. ∎

Application of rules in our system tends to increase the percentage of non-zero elements of the relations, by introducing numerous too small elements. Therefore, elements of a relation which are smaller than a threshold a, for example $a = 0.05$, are considered irrelevant and are eliminated (set to zero). The following theorem shows the preservation of relation properties after elimination of small elements.

Theorem 3: For all $a \in [0, 1)$, if the elements of a $\sup - \min$ equivalence relation, which are less than or equal to a, are eliminated, the result is a $\sup - \min$ equivalence relation. Moreover, no Archimedian t-norm preserves equivalence under such an operation.

Proof: Proving preservation for the properties of reflectivity and symmetricity is trivial. We only prove the preservation of transitivity. Let A be a $\sup - \min$ transitive relation, i.e. $\sup_z \min\left[A\left(x,z\right), A\left(z,y\right)\right] \leq A\left(x,y\right), \forall x, y \in X$. Therefore, all elements of $S = \{\min\left[A\left(x,z\right), A\left(z,y\right)\right] : z \in X\}$ are less than or equal to $A\left(x,y\right)$. If we eliminate all elements of A, which are less than or equal to a, then:

• If $a < A\left(x,y\right)$, then some of the elements of S may be eliminated but the property $\sup(S) \leq A\left(x,y\right)$ will still hold, hence the transitivity property is preserved.

• If $a \geq A\left(x,y\right)$, then $A\left(x,y\right)$ will be eliminated, therefore we must prove that all elements of S are eliminated too. Since $\forall z \in S.z \leq A\left(x,y\right)$, then at least one of $\{A\left(x,z\right), A\left(z,y\right)\}$ is less than or equal to $A\left(x,y\right)$; therefore, it is eliminated, and the minimum is also eliminated.

The reason that no Archimedian t-norm preserves transitivity under elimination of small elements is the property of subidempotency ($t(x,x) < x$) that Archimedian t-norms satisfy. Because of subidempotency, there may be two elements $\{A\left(x,z\right), A\left(z,y\right)\}$, which are both greater than a (and therefore not eliminated), while $t(A\left(x,z\right), A\left(z,y\right)) \leq a$. Thus, even though $A\left(x,y\right)$ is eliminated, it is possible that $\sup_z t\left[A\left(x,z\right), A\left(z,y\right)\right] > 0$. Transitivity is therefore lost. ∎

A consequence of the above theorem is that, although many forms of transitivity can be considered, each corresponding to a different t-norm, any archimedian transitivity requires a transitive closure after elimination of small elements.

IV. FUZZY SEMANTIC RULES

Rules are laws that concern the semantic relations and originate from their meaning. The rules are classified into the following categories:

Relation – specific rules : These are rules that involve specific relations, and are a direct consequence of their meaning. These rules contain *symbolic* semantic entities, meaning that they apply to all semantic entities.

Semantic entity – specific rules : These are rules that involve specific semantic entities. These rules are useful because they provide a means to specify the meaning and significance of specific semantic entities, for example to specify what instances of abstract entities have in common.

The general form of a semantic rule is: $R(A) = F(R_1(A), R_2(A), \ldots, R_m(A))$, where A is a set of (symbolic or literal) semantic entities and $R(A) = \{(a,b) \backslash a, b \in A\}$. The above definition implies that the new value of a semantic relation is a function of the values of all relations.

In our system, we use two special classes of rules, algebraic rules and fuzzy logical rules. The former class associates relations through relational operations, such as composition, intersection and union and is presented in subsection IV-A. On the other hand, logical rules connect the elements of the fuzzy relations with fuzzy logical functions such as conjunction, disjunction and negation; these cannot be evaluated through a single algebraic operations. An algorithm that evaluates these rules is presented in subsection IV-B.

A. Algebraic rules

In this subsection, the following operations are studied on semantic relations:

$$R' = R \cup (R \circ A)$$
$$R' = R \cup (A \circ R)$$

where R' denotes relation R after the application of the rule.

The first operation applies the following rule: for all a, b, c, if $R(a,b)$ and $A(b,c)$, then the relation R is updated by having the element $R(a,c)$ inserted to it. The second operation applies the following rule: for all a, b, c, if $R(a,b)$ and $A(c,a)$, then $R(c,b)$. R' denotes the relation R after the application of the rule. We will use the term *compositional rules* for this kind of rules.

The E relation of the thesaurus can be used to expand the relations of the encyclopedia through composition, in the following manner: for all a, b, c, if $R(a,b)$ and b is equivalent to c, then $R(a,c)$. The operation that applies the rule is: $R' = R \circ E$. We note that because E is reflective, Theorem 2.2 implies that a union is not needed. Likewise,

composition from the left side. Obviously, since both rules need to be applied, the union of the two compositions must be used: $R' = (E \circ R) \cup (R \circ E)$.

Other examples of compositional rules are:
- $\text{example}' = \text{example} \cup (\text{specialization} \circ \text{example})$
- $\text{similar}' = \text{similar} \cup (\text{specialization}^{-1} \circ \text{specialization})$
- $\text{example}'(\text{soccerPlayer}, :) = \text{example}(\text{soccerPlayer}, :) \cup (\text{example}(soccerTeam, :) \circ \text{member})$

where $\text{example}(a, :)$ denotes the a-line of the example relation.

A small part of the example relation (the actual one has 29 elements), before the expansion, is shown in table IV. A number of expansions, for example that Maldini-ag is, apart from defender-ag, an athlete-ag, can be seen in table V.

B. Evaluation of fuzzy logical rules

In order to support symbolic variables for semantic rules, the set of fuzzy logical rules contains a separate universe of symbolic semantic entities, the symbolic universe. Each symbolic entity has the `isSymbolic` attribute set true and a pointer to its current value (the current semantic entity).

Each rule consists of an `antecedent` (the *if* part of the rule) and a `consequence` (the *then* part of the rule). The antecedent is composed of a proposition, i.e. a fuzzy function of the values of the semantic relations, while the consequence consists of symbolic relations, containing elements that are a function of the truth value of the proposition. The symbolic relations contain pointers to the ordinary relations. In order to apply the rule, the antecedent is *evaluated*, and the value of the proposition sets the new value(s) of the result. A proposition is defined as either the fuzzy function of one or two propositions, or as the value of a relation.

The algorithm for the evaluation of a rule is the following:
1. Set all pointers of symbolic variables to null; the symbolic variables are *open*.
2. Evaluate the antecedent, based on its contents (which may include evaluating other propositions). Each time an open symbolic variable is encountered, a loop begins, where the symbolic variable takes all possible values; the `isSymbolic` attribute is set false and the variable is *closed*.
3. For each value of the closed variable, re-evaluate the rule.
4. Once the antecedent is evaluated, the elements of the consequence of the rules are set, based on the truth value of the antecedent.

Finally, a fuzzy logical rule is shown below:

```
<rule>
    <proposition operation="min">
    <proposition r="example" e1="shoot-ev" e2="a"/>
    <proposition operation="min">
        <proposition r="example" e1="player-ag" e2="b"/>
        <proposition operation="min">
        <proposition r="agent" e1="a" e2="b"/>
        <proposition operation="min">
            <proposition r="example" e1="goal-ev" e2="c"/>
            <proposition r="result" e1="a" e2="c"/>
        </proposition>
        </proposition>
    </proposition>
    </proposition>
```

TABLE II
The specialization relation, before closure

source	target
athlete-ag	footballPlayer-ag
footballPlayer-ag	defender-ag
footballPlayer-ag	forward-ag
footballPlayer-ag	goalkeeper-ag
footballPlayer-ag	teamMate-ag
footballPlayer-ag	opponent-ag
footballPlayer-ag	midfielder-ag

```
    <result>
    <semanticRelation id="result1" source="symbolicSet1" target=
        "symbolicSet1" semRelationRef="causer">
        <element source="c" target="b"/>
    </semanticRelation>
    </result>
</rule>
```

The antecedent of the rule consists of a proposition, which must be evaluated. If it is true, then a semantic relation element is created. The proposition of the example is the conjunction of two propositions (hence the "min" operation). The first proposition asserts that e1 has example e2. Similarly, the second proposition is a conjunction of a number of propositions, which can be either assertions of the form $R(a,b)$, or operations on propositions.

The result of the rule is a semantic relation element in the relation "causer". The c,b entities are references to the entities encountered in the antecedent.

Thus, the meaning of this rule is: if a is a player and if a shoots and if the shoot results in a goal, then a is the causer of the goal. In our example, Maldini-ag makes the shoot Shoot1-ev, and Shoot1-ev results into Goal1-ev. Therefore, Maldini-ag is the causer of Goal1-ev. Thus, this expansion would be helpful if the query was about goals resulted by Maldini. If there was uncertainty for one of the propositions, for example that Maldini-ag is a player to a degree of 0.9, then the result of the rule would also have a fuzzy degree, 0.9 in this example.

V. The integrated system and its role in semantic unification of heterogenous multimedia archives

In this section, we describe a system that receives as input a semantic encyclopedia and produces an improved (in terms of consistency and completeness) version of the description and the fuzzy thesaurus. We note that, although the relations of the semantic description and the fuzzy thesaurus serve different purposes, they are treated similarly by this procedure, since rules can be applied to either kind of relations. Moreover, the thesaurus is created (not augmented), by the rules of section II.

The procedure is outlined as follows:
1. Closure of relations: the relations are closed, based on the property information stored in them.
2. Application of fuzzy rules: all rules, including the ones that construct the thesaurus, are applied.

TABLE III
The specialization relation, after closure

source	target
athlete-ag	defender-ag
athlete-ag	teamMate-ag
athlete-ag	goalkeeper-ag
athlete-ag	midfielder-ag
athlete-ag	forward-ag
athlete-ag	opponent-ag
athlete-ag	footballPlayer-ag
footballPlayer-ag	teamMate-ag
footballPlayer-ag	forward-ag
footballPlayer-ag	midfielder-ag
footballPlayer-ag	goalkeeper-ag
footballPlayer-ag	opponent-ag
footballPlayer-ag	defender-ag

TABLE IV
A part of the example relation, before expansion

source	target
defender-ag	Maldini-ag
defender-ag	Aldair-ag
forward-ag	Romario-ag
forward-ag	Massaro-ag
goal-ev	Goal1-ev
goalkeeper-ag	Pourliotopoulos-ag
midfielder-ag	Berti-ag
shoot-ev	Shoot1-ev

TABLE V
A part of the example relation, after expansion

value	source	target
0.7	athlete-ag	Massaro-ag
0.7	athlete-ag	Berti-ag
0.7	athlete-ag	Aldair-ag
0.7	athlete-ag	Maldini-ag
0.7	athlete-ag	Romario-ag
1	defender-ag	Aldair-ag
1	defender-ag	Maldini-ag
0.7	footballPlayer-ag	Pourliotopoulos-ag
0.7	footballPlayer-ag	Massaro-ag
0.7	footballPlayer-ag	Romario-ag
0.7	footballPlayer-ag	Aldair-ag
0.7	footballPlayer-ag	Maldini-ag
1	goal-ev	Goal1-ev
0.7	kick-ev	Shoot1-ev
1	shoot-ev	Shoot1-ev

3. Encyclopedia expansion with the E relation of the thesaurus: the procedure defined in subsection IV-A is applied.

4. Elimination of relation elements less than a certain threshold.

5. Repetition of the procedure until convergence.

Step 5 is needed because output of the system is the set of relations, which is also its input; therefore, steps 1 - 4 must be reapplied as many times as necessary, i.e. until the relations no longer change. In order to ensure that the loop eventually terminates, we apply the restriction that the result of each semantic rule is used in disjunction (an OR operation) with the current membership grade. Hence, membership grades are increasing. On the other hand, membership grades are upper-bounded by 1. A sequence of numbers that is both increasing and upper-bounded converges, hence the procedure converges. Additionally, we must also ensure that convergence occurs after a limited number of steps. A way to achieve this is to stop the process when the relations do not change more than a threshold. Obviously, the threshold must be greater than, or equal to the value of a used for elimination of small values.

The system for encyclopedia expansion and thesaurus construction is available for download at:
`http://www.image.ece.ntua.gr/~gakrivas/FuzzyThesaurus/` .

XML is used for modelling knowledge. The data type definitions of the encyclopedia, in XML Schema, as well as samples of an encyclopedia are also available in the same address. A brief description of the Schema is given below.

The encyclopedia consists of a semantic `universe`, semantic `relations`, semantic `rules` and, finally, the `thesaurus`. The universe is composed of `sets` of semantic `entities`, and each set can contain subsets of semantic entities. A semantic relation contains a pointer to the source and the target set, a sequence of elements and a set of relation properties. Each relation element contains a pointer to the source semantic entity, a pointer to the target semantic entity, and an optional real value, which is the degree of correlation between the two entities (the default value is 1). The thesaurus relations are contained in the `thesaurus` element of the encyclopedia.

The sample encyclopedia contains semantic entities in the thematic category of football. Altogether, the encyclopedia contains 109 semantic entities. The semantic relation elements are 236, before the expansion, and 413, after the expansion (the reflective and symmetric components are not counted). Moreover, the constructed thesaurus contains 4 elements for the E relation, 253 elements for the I relation and 392 elements for the A relation.

Our system constitutes a module of the FAETHON project, `http://www.image.ece.ntua.gr/faethon`. The main goal of FAETHON is to provide homogenized access to multiple multimedia archives. The user's query is translated into proprietary archive queries. FAETHON solves the problem of query translation by using a semantic encyclopedia like the one described in section II. Thus, the terms of the query are mapped into semantic entities, and the consistency and completeness of the automatically expanded encyclopedia ensures that the query translation is also complete and consistent. The thesaurus is used in order to expand the user's query, in the usual manner of information retrieval systems.

VI. Conclusions and future work

The main contribution of this paper is the automation of the thesaurus construction, as well as a complete and consistent automatic expansion of expert – supplied knowledge. This is achieved via fuzzy rules, which must also be supplied by the expert. It is interesting to note that the part of the same rules can also be used to expand semantic descriptions of individual multimedia documents, because of the MPEG-7 language used.

A possible extension to this work is exploitation of the fuzzy thesaurus for expanding a semantic query. This is not a trivial task, because of the need to direct the expansion towards the context of the query, as well as towards the preferences and interests of the individual user.

Another question is the adequacy of MPEG-7, a standard for describing multimedia documents, for describing multimedia knowledge bases. Since MPEG-7 is based on XML, the necessary clearly defined semantics of ontological languages, such as OIL, is missing. For example, inferences such as subsumption, which are standard in OIL are not directly achieved by our system. Although many attempts have been made, an ontological language for describing multimedia "concepts" is still missing.

References

[1] "Introduction to mpeg-7," ISO/IEC JTC1/SC29/WG11 N4032, Mar. 2001, singapure.

[2] Special issue on MPEG-7, June 2001. [Online]. Available: http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=20050

[3] *The MPEG Home Page*, Iso/iec jtc1/sc29 wg11 ed. http://mpeg.telecomitalialab.com/.

[4] G. Votsis, A. Drosopoulos, G. Akrivas, V. Tzouvaras, and Y. Xirouhakis, "An mpeg-7 compliant integrated system for video archiving, characterization and retrieval," in *IASTED International Conference on Signal and Image Processing (SIP2000)*, Las Vegas, Nevada, USA, Nov. 2000. [Online]. Available: http://www.image.ece.ntua.gr/%7Egakrivas/ggtp.pdf

[5] G. B. D.H. Kraft and G. Passi, "Fuzzy set techniques in information retrieval," in *Fuzzy Sets in Approximate Reasoning and Information Systems*, ser. THE HANDBOOKS OF FUZZY SETS, H. P. James C. Bezdek, Didier Dubois, Ed. Kluwer Academic Publishers, July 1999, vol. 5, ch. 8.

[6] S. Miyamoto, *Fuzzy sets in Information Retrieval and Cluster Analysis*. Kluwer Academic Publishers, 1990.

[7] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic; Theory and Applications*. Upper Saddle River, N. Y.: Prentice Hall, 1995.

[8] J. H. F. Martinez-Trinidad and J. H. Ruiz-Shulcloper, "Fuzzy clustering of semantic spaces." [Online]. Available: citeseer.nj.nec.com/493600.html

[9] G. Akrivas and G. Stamou, "Fuzzy semantic association of audiovisual document descriptions," in *Proc. of Int. Workshop on Very Low Bitrate Video Coding (VLBV)*, Athens, Greece, Oct. 2001. [Online]. Available: http://www.image.ece.ntua.gr/ %7Egakrivas/vlbv01%20version%208.pdf

[10] S. K. N. Tsapatsoulis, Y. Avrithis, "Facial image indexing in multimedia databases," *Pattern Analysis & Applications*, vol. 4, no. 2-3, pp. 93 – 107, June 2001, springer-Verlag London Ltd, ISSN: 1433-7541. [Online]. Available: http://www.springerlink.com/app/home/content.asp?wasp= 2fa16nwtwgckn223r%ecn&referrer=contribution&format= 2&page=1&pagecount=15

[11] Y. Avrithis, Y. Xirouhakis, and S. Kollias, "Affine-invariant curve normalization for object shape representation, classification, and retrieval," *Machine Vision and Applications*, vol. 13, no. 2, pp. 80 – 94, Nov. 2001, springer-Verlag Heidelberg , ISSN: 0932-8092. [Online]. Available: http://www.springerlink. com/app/home/content.asp?wasp=h82pypwtthdmgc6uh% qrl&referrer=contribution&format=2&page=1&pagecount=15

[12] M. Wallace and G. Stamou, "Towards a context aware mining of user interests for consumption of multimedia documents," in *Proceedings of the IEEE International Conference on Multimedia and Expo.* IEEE, Aug. 2002.

[13] G. Akrivas, M. Wallace, G. Andreou, G. Stamou, and S. Kollias, "Context sensitive semantic query expansion," in *IEEE International Conference on Artificial Intelligence Systems, AIS-02*, Russia, Sept. 2002. [Online]. Available: http: //www.image.ece.ntua.gr/%7Egakrivas/AIS-cameraReady.pdf
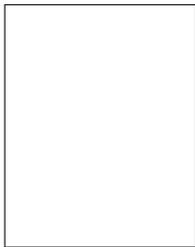
Giorgos Akrivas received his diploma from the department of Electrical and Computer Engineering of National Technical University of Athens, Greece in 1999. He is currently working toward the Ph.D. degree in intelligent video retrieval in the Image, Video and Multimedia Laboratory of NTUA. His research interests include multimedia content description, computer vision, intelligent retrieval and fuzzy logic.
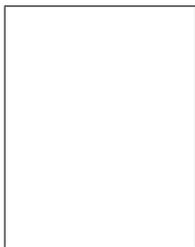
Giorgos B. Stamou obtained the Diploma in Electrical and Computer Engineering from NTUA in 1994 and the PhD in Computational Intelligence from NTUA in 1998. He is currently a researcher of the Institute for Communication and Computer Systems and NTUA. His research interests include fuzzy sets, neural networks, hybrid intelligent systems, computer vision, semantic video analysis and human-computer interaction. He has published over 60 papers in international journals, books and conferences.

Stefanos Kollias obtained his Diploma from NTUA in 1979, his M.Sc. in Communication Engineering in 1980 from UMIST in England and his Ph.D in Signal Processing from the Computer Science Division of NTUA. He is with the Electrical Engineering Department of NTUA since 1986 where he serves now as a Professor. Since 1990 he is Director of the Image, Video and Multimedia Systems Laboratory of NTUA. He has published more than 180 papers in the above fields, 80 of which in international journals.