Hierarchical Organization of the Semantic Rules for the Images Annotation By Co-Quotation Method

Yassine Ayadi
MIR@CL Laboratory
Multimedia InfoRmation systems and @dvanced Computing Laboratory
University of Sfax, Tunisia
ayadi.yassine@gmail.com

Ikram Amous, Faiez Gargouri
MIR@CL Laboratory
Multimedia InfoRmation systems and @dvanced Computing Laboratory
University of Sfax, Tunisia
ikram.amous@isecs.rnu.tn, faiez.gargouri@gmail.com

***Abstract*— The present paper introduces an approach for image semantic annotation. It discusses work in progress and reports the current state of our approach. This comprises the development of the domain ontology used for annotation, the functionalities for annotating image with an underlying ontology and search features based on these annotations. We describe a method for automatic annotation of images and apply it to and evaluate it on images of inference process.**

Keywords-Automatic Annotation; Semantic; Co-Quotation; Ontology; Inference;

I. INTRODUCTION

During the last decades, a number of digital images have burst with the advent of digital cameras which require effective search methods. However, due to the semantic gap between image visual features and human concepts, most users prefer textual queries. Hence, it is always difficult to find a specific image if we want to show it or share it with another person. In this context, the use of annotation can facilitate the task of images management. Besides, the image annotation establishes the main tool for semantics associated with an image. Moreover, the addition of meta-data to an image enriches its description and allows the construction of more successful consultation tools and visualizations.

Our work objective is to describe the multimedia document contents, facilitate and optimize their search. To do so, we build on the documents annotation by semantics descriptors. With semantics, we imply any information that can be deduced and explicitly specified. We can deduce the Car in the parking without such information being directly mentioned in the document. According to [1], semantics depends on the knowledge level and on the user perception as well as on its objective. Therefore, the semantics of a situation (or of a context) can be differently expressed by diverse users.

Furthermore, labeling the semantic content of images with a set of keywords is a problem known as image annotation. Annotation is used primarily for image database management, especially those using keyword-based search, while not annotated images can be extremely difficult to find in large database.

Once the documents are annotated, they can be used such as [17]. Indeed, there exists much work on the multimedia documents manually annotation, among which we quote: AnnoSearch [2], IMAGINATION [3], IAM@Image CLEFPhoto Annotation [4].

In our study, the idea is to exploit the visual descriptors and topological relationships in image to determine their semantics. Actually, neither tool presents concepts of exactly annotated images.

The continuation of this paper is organized in the following way. First of all, Section 2 presents our proposed annotation approach. Then, Section 3 provides the use of urban ontology which will be used in this work. Afterward, Section 4 illustrates the construction of hierarchy semantic rules where we use the Co-quotation method in the images annotation. Next, Section 5 describes the automatic image annotation. As for Section 6, it presents an enrichment of ontology and process inference. Finally, this paper ends with some concluding remarks and future perspectives.

II. PROPOSED APPROACH

The automatic image annotation is an effective research subject [5]. Its goal is to develop methods that can produce, for a new image, the relevant keywords among an annotation vocabulary. These predictions of keywords can be used to propose the image semantics.

We will describe in this section our annotation approach based on three steps to solve these problems. The proposed approach is illustrated by Figure 1.

- The first step ***Training***: extract the image preliminary characteristics to classify objects (many tools exist permit to automatically associate with some image a characteristics vector).In our study, the idea is to exploit the visual descriptors and topological relationships in image to determine their semantics. Actually, neither tool presenting concepts annotates exactly the images. The existing tools do not combine the object detection and the relation one.

The annotation process, defined in the Training step, is composed of two sub-steps:

I. We start with a set of images, which we call the images of apprenticeship. We proceed in a way that the users select an object of the image manually. The selection of an object makes possible for the user to affect a manual semantics annotation. The tools for image processing determine the low level characteristics defining this object.

II. The second sub-steps consist in building a value matrix of low level descriptors describing the required object. This matrix represents the result of several iterations of the first phase on the basis of image for the same object.

After the apprenticeship phase, the system automatically creates one or several rules describing the object chosen from the first phase.

The main objective of our classification is to associate a unique interpretation from low level descriptors with an image document [6]. The result of the combination of the MPEG-7 descriptors with those of cavities and contour is a well-formed XML file.

The spatial relationships constitute the basis of the linguistic descriptions of the spatial configurations. These relations are generally classified a different category [8], [9], [10]: topology, orientation and distance. These can be descended of an explicit declaration on the part of the user or inferred from the existing information. In our proposition, we build on topological relationships. To detect the relationship between two detected objects, we calculate the angle between the including rectangles.

In our context, several object types can be distinguished: car, building, persons, panels, road, etc. These different objects are classified according to classes: means of transport, buildings, the place and objects. In order to do that, we can determine a set of spatial relations that can exist between objects in a picture to know: right, left, behind, in front. [7]

- In the second step, the process describes the Image ***semantic annotation***. This step is the relationships extraction between objects, to build the first level semantic rules (these rules represent a human knowledge). They are stored in a knowledge base. From an archive picture, we loosened a semantic set, represented through the notion of predicate logic. From all the rules, we loosen the various predicates whose semantics is extracted from the image. These predicates can be grouped in elements establishing the rule in relation between elements and result.
- The third step consists in the ***Inference process*** creation to generate a new semantics. The inference is then defined, then, as an action which allows a human or a machine to increase its knowledge. This person or this machine "makes an inference", i.e, it infers a result starting from a set of data. Our process of inference consists in a unit of inference rules, being based on the principle of the front chaining.

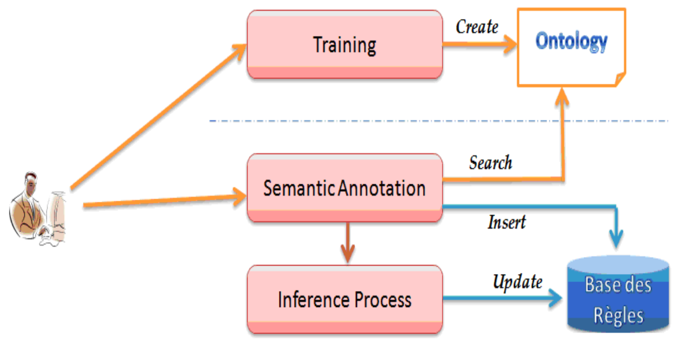

Figure 1. Proposed Demarche

To create a model, we ask to the domain expert to draft a list, as complete as that possible, in natural language, various semantics extracted of the images. He integrates into this list the knowledge environment and the studied context.

III. URBAN ONTOLOGIES

Ontologies are used to formalize the concepts semantics of each domain. We have already used ontology of field formalized in XML (TOWNTOLOGY) which represents the concepts used by the urban image.

Ontology defines a common vocabulary for the researchers who need to share the information in a domain. It includes by the definitions legible by a machine on the basic concepts of this domain and of their relations. [11] [16]

Ontology is a formal explicit description of the concepts in a domain (classes), properties of every concept describing characteristics and attributes (facets).

Ontology as well as all the individual instances of the classes constitutes a knowledge base. The classes constitute the main concepts of several ontologies. They describe the concepts in the domain. A class can have subclasses which represent more specific concepts than the super-class (or superior class). The attributes describe the classes and the instances properties.

Let us illustrate these ideas for our domain. We are interested in the objects component of an image by describing the taxonomy of types represented by the model in Figure 1.

On this model, the Object class is subdivided into subclasses such as "Means", "Place", "Panel", "Buildings" and these in specialized subclasses as "Building", "car" or "Road"

This ontology also contains topological relations such as "Front", "Behind", "to the right", "to the left".

Properties are represented on the model by elementary principles, called Properties. If we assign precise objects in X and Y, these principles will become assertions: ' a car in front of a building '.

IV. HIERARCHY OF SEMANTIC RULES

The manual annotation of semantics remains a problem, because it depends on the user objective and on his knowledge. It thus seems to us thus convenient to find a way of automating the annotation of semantics to improve the

Copyright (c) IARIA, 2010 ISBN: 978-1-61208-104-5

research for the multimedia documents by building on requests.

Several questions, thus, are left to be elucidated:

- How can we extract the semantic contents of the multimedia documents in an automatic way to ensure and facilitate future research for users?
- How can we connect high level knowledge to the low level characteristics of the documents?

The process of images annotation, which we propose, is based primarily on the results provided by Training step. The descriptors, extracted and instantiated automatically, from such a step, allows the construction of our first work scheme.

The result of this step consists in extracting automatically the low level descriptors and deducing the image elements recorded in ontology. Each element can have well determined semantics and can refer to an object (person, car, building…).

The combination of these elements with space relations [14] creates the first level of semantics rules describing the image content. This first level is created manually.

An image can refer to several semantic rules. By using this context, makes it possible to us to gather the semantic rules by topic.

The purpose of the Co-quotation method [12], used in bibliometry since 1973, is to create starting from scientific articles of the same field of research and by using their bibliographical references, of the relations between these articles.

This method rests on the assumption that two bibliographical references of unspecified dates, frequently quoted together, have a parity set of themes. In the same way that for the table of the couplings, the matrix of Co-quotations is built by each line is the studied set quotations i, each column is the quotations set j and the elements set x_{ij} of the matrix corresponds to the number of documents which quoted documents i and j in same time.

The use of the bonds in order to annotate a resource also applies to the images.

The use of the Co-quotation method in the images annotation can help us to use the annotations of close references by topic in order to annotate new images.

The following figure (Figure 2) is an extract of the quotation graph: the image ***I*** cite semantic rules ***RG$_i$***, ***RG$_{i+1}$***, ***RG$_{i+2}$*** and ***RG$_n$***. In this case, these rules Co-are quoted at least once by image ***I***.

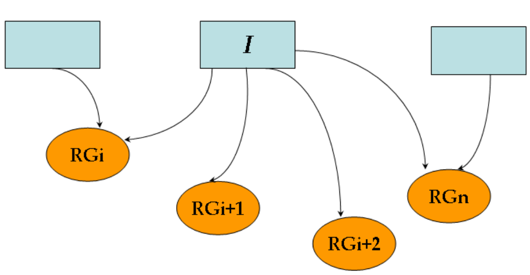

Figure 2. Extract of the Quotation Graph

The method of Co-quotation [12] is used to calculate the resemblance between the semantic rules and not between the images.

The following figure (Figure 3) is an example of graph of Co-quotation. Value 2 between rules ***RG$_i$*** and ***RG$_{i+2}$*** indicates that these two rules are quoted together by two images.

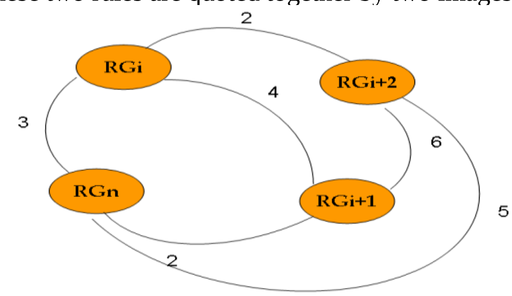

Figure 3. Example of Graph of Co-Quotation

The Co-quotation matrix is a representation of the Co-quotation graph; it corresponds to a square matrix.

C_{i+2n}: (i+2 Line, n Column) the Co-quotation frequency of ***RG$_{i+2}$*** and ***RG$_n$***, which is equal to 5, because they are quoted together by five Images.

	RGi	RGi+1	RGi+2	RGn
RGi	0	4	2	3
RGi+1	4	0	6	2
RGi+2	2	6	0	5
RGn	3	2	5	0

Figure 4. The Co-quotation matrix

The result of the distance function [13]

$$S_{i,j} = 1/C(i,j)2 \quad (1)$$

Being in the interval [0,1]. Two semantic rules are quoted units, the more the distance S (I, J) will be close to zero.

As soon as the semantic rules of a document image are quoted, we build the graph of distance and the matrix of distance MC.

Matrix MC of the example is the following one:

	RGi	RGi+1	RGi+2	RGn
RGi	0	1/C(i,j)2	1/C(i,j)2	1/C(i,j)2
RGi+1	1/C(i,j)2	0	1/C(i,j)2	1/C(i,j)2
RGi+2	1/C(i,j)2	1/C(i,j)2	0	1/C(i,j)2
RGn	1/C(i,j)2	1/C(i,j)2	1/C(i,j)2	0

Figure 5. The Co-quotation Distance

Copyright (c) IARIA, 2010 ISBN: 978-1-61208-104-5

In order to represent association (Rules/Images), we use the concepts lattice. The concepts lattices are used in the search for information to refine or generalize the request user. [15]

The Galois lattice or concepts lattice is a mathematical structure making it possible to represent the not disjoined classes subjacent with an objects set [18].

Context: a context is a triplet K = (O, A, ζ) where O is an objects or individuals set, A is an attributes or properties set and ζ is a binary relation between O and A.

A context K = (O, A, ζ) can be represented in the table form, where a line corresponds to an object with its attributes.

Lattice: a lattice is an ordered Set in which two unspecified elements have an upper limit and a lower limit. A complete lattice is a lattice for which any element has an upper limit and a lower limit.

Galois correspondence: Is the context K = (O, A, ζ), f an application P (O) in P (A) and g an application P (A) in P (O), f and g defined:

- f: P(O)→ P(A) f (Oi) = {a ∈ A| (o,a) ∈ A, ∀o ∈ Oi} intention ;
- g: P(A)→P(O) g(Ai) = {o ∈ O| (o,a) ∈ A, ∀a ∈ Ai} extension ;

The couple (f, g) is called the Galois correspondence on K.

Formal concept: Are Oi ⊆ O and Ai ⊆ A, (Oi, Ai) is a concept if:

- Oi is the extension of Ai;
- Ai is the intention of Oi;

Oi = g (Ai) and Ai = f (Oi)

Galois lattices: Are f: O → A and g : A → O two functions defined on the lattices (O,≤O) and (A, ,≤A), such as (f, g) is a Galois correspondence.

Either G= {(o, a), where o is an element of O and where a is an element of A, such as o = g (a) and a = f (o)}. That is to say ≤ the relation of order defined by: (o1, a1) ≤ (o2,a2) if a1 ≤A a2. (G,≤) is a Galois lattice.

Example:

The following table represents the correspondence between six images answers of the five rules {RG1, RG2, RG3, RG4, RG5}

The Rules are:

- RG1, Car → Transport Means
- RG2, Taxi → Car
- RG3, Car in front of Car → Parking
- RG4, Taxi in front of Taxi → Parking
- RG5, Transport Means in front of Transport Means → Parking

TABLE I. IMAGE/ RULES ASSOCIATION

	RG1	RG2	RG3	RG4	RG5
I1	X	X			X
I2			X	X	X
I3	X			X	X
I4	X				X
I5				X	X
I6			X		X

The example of Galois correspondence is:

- O1 = {I3, I4} → f(O1) = {RG1, RG5}
- A1 = {RG1, RG5} → g(A1) = {I1, I3, I4}

In this example we have the couple ({RG5}, {I1, I3, I4})

In this example, we have the couple ({RG5}, {I1, I3, I4}) which means that the result of the request with rule RG5 gives for answer the images I1, I3, I4.

We illustrate the result of the Bordat algorithm [14] on the example of the preceding table.

C = (Ø, {I1, I2, I3, I4, I5})

δC = max {fC(I1), fC(I2), fC(I3), fC(I4), fC(I5), fC(I6)}
= max {{RG1, RG2, RG5}, {RG3, RG4, RG5}, {RG1, RG4, RG5}, {RG1, RG5}, {RG4, RG5}, {RG3, RG5}}
= max {{RG1, RG2, RG5}, {RG3, RG4, RG5}, {RG1, RG4, RG5}}

In this case the direct successors of C are:

C1 = ({I1}, {RG1, RG2, RG5})
C2 = ({I2}, {RG3, RG4, RG5})
C3 = ({I3}, {RG1, RG4, RG5}}

In the same way, one calculates the direct successors of C1:

δC1 = max {fC1(I1), fC1(I2), fC1(I3), fC1(I4), fC1(I5), fC1(I6)}
= max {{RG5}, {RG1, RG5}, {RG1, RG5}, {RG5}, {RG5}}

The C1 Successors are:

C4 = ({I1, I3, I4}, {RG1, RG5})

The continuation of the direct successor's calculation is made same manner. The result of the example is the following (Figure 6)

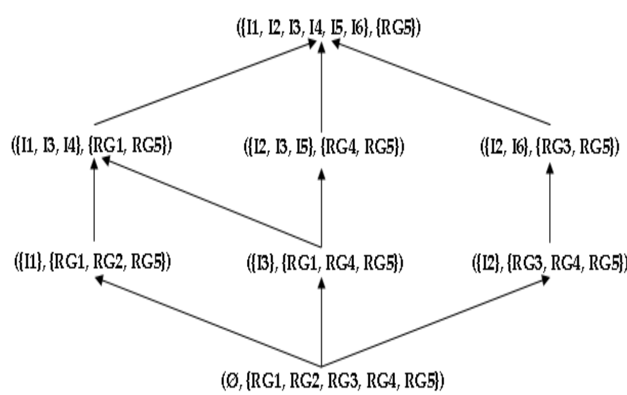

Figure 6. Galois Lattice

The lattice structure is used in order to extract the hierarchical relations between the semantic rules. We build the hierarchy of the semantic rules in order to keep only one occurrence of the rules. We leave the rules set of the more high level and one eliminates the occurrences from each element in the lower levels.

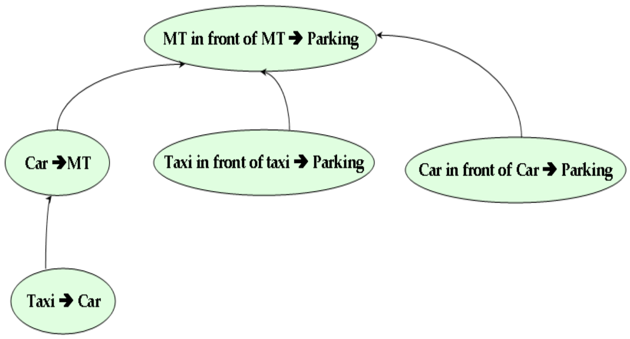

Figure 7. Semantics Rules Lattice

Figure 7, present the result of the semantic rules hierarchical.

V. Automatic Image Annotation

The presentation of the imported annotations is made by defining a multi-criterion choice to select annotations to be used in the following phase.

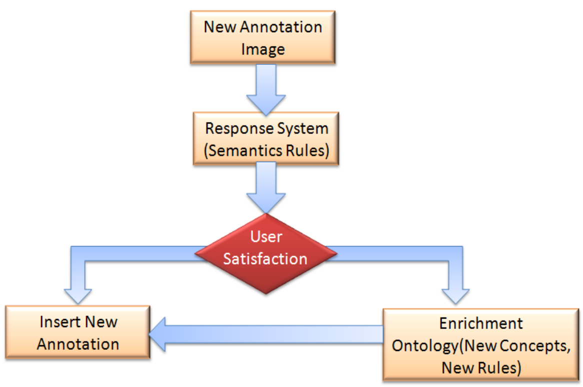

Figure 8. Automatic Image Annotation

Each Rule semantic has a rate (utilization Ratio of the rule for forthcoming annotations). This rate is a value included/understood an interval [0, 1].

The rate is calculated by the following formula:

$$\text{Fact-salt} / \text{Fact-Aff} \qquad (2)$$

Such as:

- ***Fact-salt***: It is the number of times that the rule at summer chosen as a solution for the annotation of a new image
- ***Fact-Aff***: It is the number of times that the rule at summer suggested as a solution for the annotation of a new image.

```
Algorithm: Automatic Annotation
Data
   C ← {C1, C2,…,Cn] Concepts Set of New Image
   (With Ontology)
   ReG ← {RG1, RG2, …, RGn} Rules Set
   List : List of Rules Semantic
   U Type Rules Semantic
   A Type Rule
Begin
//Extract all the Rules for each concepts couples
List ← Extract_Rules (Ci, Cj);
U = List Beginning
Repeat
// extract all the low level rules than the rule U
A ← Extract_rules_lowlevel (U);
//Add the new rules at the list
End Automatic Annotation
```

The goal in this section is to import and order the semantic rules quoted by a document image I.

We retain the following criteria to order the annotations as a whole ***ReG***; The rate of selection of the imported rules. If the rule appears in several images, the largest rate of selection will be considered (maximum).

VI. Ontology Enrichment and Inference Process

Ontological engineering consists of the search for general, reusable, shareable and durable concepts to build a model of knowledge able to help people solve problems [MIZ 04].

Because our step of annotation is based on ontology, we also dealt with the problem of the enrichment of ontology. This enrichment will also be used to refine or enrich the automatic annotation by the documents multi-media as Image type.

We should note that ontology is a set of concepts connected by the relation of Specialization/ Generalization and other like synonymy.

The principle of enrichment in our step does not include/understand the suppression and the transformation of concepts, but earlier to add principle again is that of the semantic rules (semantic rules is a set of concepts connected to each other by relations which provide one or more semantics which can be the same concepts of ontology).

We were interested within the framework of our work in used ontology TOWNTOLOGY.

This ontology is described by the concepts of the urban field and the relations which connect them. Figure 9 illustrates an extract of this ontology:

- A node represents a concept, represented by a circle in the figure (for example the C1 concept).
- Concepts are connected by directed arcs defining the relation of Specialization/ Generalization, here the C2 concept is a specialization of the C1 concept.

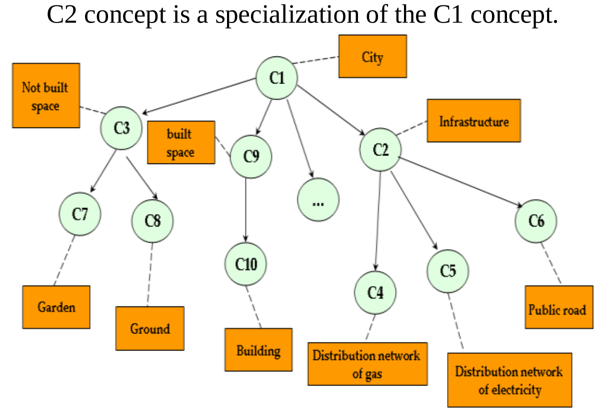

Figure 9. Example TOWNTOLOGY Ontology

Copyright (c) IARIA, 2010 ISBN: 978-1-61208-104-5

The idea of our approach is based on the integration of the semantic rules within ontology. This stage of enrichment breaks up into two phases, the first stage is the manual integration of the rules within ontology; the second phase is an enrichment starting from the phase of interrogation.

- ***Manual Enrichment:*** Consist in adding semantic rules within descriptive XML file of ontology in order to build a first level of knowledge.
- ***Automatic Enrichment:*** is based on the exploitation of new annotations to enrich ontology.

Let us illustrate automatic enrichment for an example. A user ***U*** annotates a new image according to the following stages:

1. Initially, U seized the new image I, it obtains as an answer RG1, RG2… RG5 semantic rules.
2. The user can not be satisfied and it adds new concepts or semantic new rules
3. The U user finishes his annotation when it satisfies the result.

The increase in the information level which one can lay out on a system is essential for the improvement of this system control and the processes which are integrated there (automation, maintenance…).

Two primary sources give access to this information. The first one starts from human expert knowledge which gives rather qualitative information on what the studied system is, while the second is the data acquisition directly on the system, giving rather quantitative information.

The inference is an action which allows a human or a machine to increase its knowledge. This person or this machine "makes an inference", i.e, it infers a result starting from a set of data.

Our inference process is composed of a set of inference rules, being based on the principle of the front chaining. In what follows a formal specification of all the inference rules as well as an application of these rules on the collected image basis are presented.

A. *The inference algorithm:*

```
Start
OPEN ← Semantic Rule
Repeat
     Ü ← Beginning OPEN
     Repeat
        To observe inference rules
        To add in end of OPEN semantic new Rules
     Until I = End Rules
     To add U to Close
Until OPNE = Ø
End
```

B. *Inference Rules*

Rule 1:

```
BE: Elements Base
BR: Relations Base
BGR: Rules Base
∀ {OB1, OB2, S1, S2} ∈ BE
∀ R ∈ BR
∃ RG1 a Rule, RG2 a Rule /
     RG1 = OB1, R, OB2 → S1
     RG2 = OB1 → S2
     ∃ RG3 a new Rule /
     RG3 = S2, R, OB2 → S1
```

Rule 2:

```
BE: Elements Base
BR: Relations Base
BGR: Rules Base
∀ {OB1, OB2, S1, S2} ∈ BE
∃ RG1 a Rule, RG2 a Rule /
     RG1 = OB1 → S1
     RG2 = S1 → S2
     ∃ RG3 a new Rule /
     RG3 = OB1 → S2
```

VII. CONCLUSION & PERSPECTIVE

A system of research for image adapted to the needs of the users is capable of extracting the image semantics. However, the ditch between the low levels attributes and the semantic knowledge is the main obstacle in the construction of reliable semantics for the image research.

In this paper we proposed an approach which allows the discovery of semantic information from the low level image. Our approach is interested in the semantic description of the objects of a given image. We presented our vision for semantic annotation and inference to support the discovery of general image.

Our system is work in progress, and we are actively experimenting with implementation alternatives. As continuation of this work, the semantic representing the semantic and role relationships between the concepts will be constructed from our current sentence level semantic trees.

In this paper, we have described an interface for image annotation based on user-formulated semantic inference rules. The aim of this study was to determine the characteristics that suit the semantic inferencing and Rules. One of the significant findings was that knowledge of multimedia and image analysis terms is both a prerequisite and impediment to obtaining good results. We still found, however, that the results of applying rules defined by domain experts were significantly less than those defined by the authors.

REFERENCES

[1] A. Boucher and T. Le. "Comment extraire la sémantique d'une image ?". In 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, SETIT 2005, Tunisia. pp. 295-306. March 2005.

[2] X. Wang, L. Zhang, F. Jing, and W. Ma. "AnnoSearch : Image Auto-Annotation by search". In. Computer Vision and Pattern Recognition, 2006. IEEE Computer Society Conference on (2006), pp. 1483 – 1490. June 2006.

[3] A. Walter and Gabor Nagypal. "The Combinaison of Techniques for Automatic Semntic Image Annotation Generation in the IMAGINATION Application". In. ESWC 2008/ 5th European Semantic Web Conference. Tenerife, Sapin. 01- 05 June 2008. pp. 879-883. June 2008.

[4] J. Hare and P. Lewis. "IMA@Image CLEFPhoto Annotation 2009 : Naïve application of a linear-algebric semantic space Corfu". In: CLEF 2009 Workshop, 30 September - 2 October 2009, Corfu, Greece. p. 66-71. October 2009.

[5] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid. "Apprentissage de distance pour l'annotation d'images par plus proches voisins". in: Reconnaissance des Formes et Intelligence Artificielle. CAEN. Inria-00439309, version 1. January 2010.

[6] Y. Ayadi, I. Amous, A. Jedidi, and F. Gargouri. "Towards a Semi-Automatic Description of the Multimedia Image Documents". In. The International Business Information Management Conference (8th IBIMA) on June 20, 21, and 22, 2007 in Dublin, Ireland. pp. 203-209. June 2007.

[7] Y. Ayadi, I. Amous, A. Jedidi, and F. Gargouri. "Automatic Annotation Process with Objects and Relationships Detection of Images". 5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, SETIT 2009, Tunisia. pp. 56-61. March 2009.

[8] O. Bedel, S. Ferré, and O. Ridoux. "Handling Spatial Relations in Logical Concept Analysis To Explore Geographical Data". Int. Conf. Formal Concept Analysis (2008) pp. 241—257.

[9] A. Dia Miron, J. Gensel, M. Villanova-Oliver, and H. MARTIN. "Towards the Geo-spatial Querying of the Semantic Web with ONTOAST". 7th International Symposium on Web and Wireless GIS (W2GIS 2007), Cardiff, UK, 28-29 November 2007. pp. 121-136. November 2007.

[10] T. Leonard. "How language structures space". in H. Pick & L. Acredolo, éd., Spatial orientation: theory, research and application, New York, Plenum Press. 1983.

[11] A. Keita, R. Laurini, and C. Roussey. "Towards an Ontology for Urban Planning: The Towntology Project". In Proceedings of the 24th UDMS Symposium, Chioggia, October 27-29, 2004, pp. 12.-24.

[12] E. Garfield. "Co-Citation analysis of the scientific literature: Henry small on mapping the collective mind of science". Essays of an information scientist: Of Nobel Class, Women in Science, Citation Classics and Other Essays, 15 (19), 1993. pp. 293-303.

[13] L. Abrouk. "Annotation de documents par le contexte de citation basée sur une ontologie". Thèse de Doctorat. Soutenue le 27 novembre 2006. Université de Montpelier II.

[14] E. Mephu Nguifo and P. Njiwoua. 'Treillis de concepts et classification supervisée". Technique et Science Informatiques, 24(4) 2005 pp. 449–488

[15] N. Messai. "Treillis de galois et ontologies de domaine pour la classification et la recherche de sources de données génomiques". Master. University Henri Poincare. Juin 2004.

[16] A. Keita, C. Roussey, and R. Laurini. "Un outil d'aide à la construction d'ontologies pré-consensuelles : le projet Towntology". In actes du 24ème congrès de Informatique des Organisations et Systèmes d'Information et de Décision (INFORSID), Tunis 31 Mai-4 Juin 2006, pp. 911-926.

[17] A. Jedidi, I. Amous, and F. Sèdes. "Documents semi-structurés et métadonnées contribution à la réingénierie de collections de documents". Dans : RSTI ISI bases de données semi-structurées, Hermes - Lavoisier, Vol. 8, N. 5-6,. Decembre 2003.pp. 153-172.