ELECTRONIC **WORKSHOPS IN COMPUTING**

Series edited by Professor C.J. van Rijsbergen

Ian Ruthven (Ed)

Miro'95

Proceedings of the Final Workshop on Multimedia Information Retrieval
(Miro '95)
Glasgow, Scotland
18-20 September 1995

Paper:

A Multifacet Formal Image Model for Information Retrieval

Mourad Mechkour

Published in collaboration with the
British Computer Society

BCS

Springer

A Multifacet Formal Image Model for Information Retrieval.

Mourad Mechkour
CLIPS-IMAG
BP 53, 38041 Grenoble cedex 9
France
Mourad.Mechkour@imag.fr
Tel : (33) 76 63 59 76, Fax : (33) 76 44 66 75.

April 23, 1996

Abstract

This paper presents an extended model for image representation and retrieval ($EMIR^2$). This model combines different interpretations of an image to build a complete description of it, each interpretation being represented by a particular view. The set of views considered in $EMIR^2$ include the physical view and the logical view, which is an aggregation of four main views: the structural view, the spatial view, the perceptive view, and the symbolic view. A description of the model concepts is given using a mathematical notation, yieldingthe framework $EMIR^2$. We defined a first correspondence function that estimates the similarity between two images, one being the query.

1 Introduction

Generally speaking, two different formats of physical images are referred to by the term image in the literature, the raster and the vector formats. In the raster format an image is considered as a matrix of pixels, each pixel being represented by its colour. This kind of images includes photographs, paintings, ..., and are produced by digitalisation. In the vector format an image is represented as a set of mathematical equations defining n-dimensional objects with explicit knowledge on their structure and locations. This format is typically used to store line art and CAD information, and is produced by computer drawing tools. In this paper the term image designates only the first category, the raster format representation.

A lot of image content representations have been defined in the literature. The image descriptions are classified in three main categories. Most information retrieval systems use a combination of different descriptions to capture all aspects of image content.

- The basic image representation considers the image as a physical object (pixel matrix) without considering semantics interpretation of its content. The indexing of this kind of representation is mainly based on the colour distribution in the image, textures, etc.[4]
- The second category considers an image as a graphics. All the elements recognised in the image are represented using a spatial description. Two general approaches are used in defining a spatial representation, the object oriented and the relation oriented. In the first approach the image content is considered as a set of geometric objects, each defined by a set of points in an Euclidean space. In the second approach the image is represented by a set of objects linked together by a particular set of spatial relationships (topological, metric, ...).[18, 21, 8, 5]
- The third image representation category includes all kind of semantic interpretation of images. A wide range of models have been proposed and used for image retrieval. This category includes the list of external attributes of the image, e.g. the date, the author name, the size, etc. It includes the classical textual descriptions and their indexing using list of terms defining the elements considered as relevant in the image [12]. Some rich semantic descriptions, developed in AI for knowledge representation, are also used to capture the complex image content, e.g. complex objects [15], first order logic [6], Shank's conceptual dependencies [2], etc.

We think that an image model suitable for IR should insure the best precision and exhaustivity in image modelling. This model should be based on a combination of semantic descriptions of image content and spatial information linking semantic descriptors to the two-dimensional entities (sub-images) present in the image. This semantic description is symbolic and is mainly used in the query evaluation process, since users are more skilled in defining their information needs using symbolic descriptors (mainly natural language based). The spatial information, inherent to image data, concerns the shape of objects and their relative positions inside the image. It gives precise information about objects and permits geometric operations inherent to bi-dimensional nature of objects in an image.

An application really needs the possibility to manipulate both spatial and semantic information for retrieval of images, and it is necessary to represent any of these two aspects in a single model. Considering, for example, a cartography application. It is as much interesting to know that a certain number of lines is observed in an image as to know that a line corresponds to a river, and an intersection of two lines shows the meeting place of a river and a road. Furthermore, considering a medical example, a radiological image might be split up into two sub-images represented by polygons, and it is also interesting to know that one shows a cancerous tumour of the right lung and that the other one shows the extension of this tumour through metastasis of the left lung.

The semantic description of an image is an interpretation of its content and is user dependent (subjective). Each different user produces a particular interpretation of the same image. The model should provide mechanisms to associate different interpretations of the same image, and different descriptions to each entity identified in the image.

The entities in an image are real world objects and they are seldom atomic, e.g. house, car, man, etc. The model should consider the representation of complex entities, independently from the semantic description associated to them.

The model should be open to user specific descriptions. The model should offer a generic structure that can be instantiated to fulfil application specificity. It is obvious that all the semantic knowledge must not and can not be stored inside the model: the domain representation of the application must be considered only at the application level. On the other hand, the model must provide abstractions for the manipulation and the interpretation of these contextual data. Hence the model must also be adaptable enough to allow different levels of description for the semantics of an image.

We divided the presentation of $EMIR^2$ in three main parts : the first introduces the basic ideas of the model and the vocabulary; the second part presents the formalisation of these ideas using a general mathematical formalism; and the third part discusses the query language and the correspondence function principles. We conclude by describing the implementation of this model and the experiments currently conducted and the future work to enhance the model.

2 The image model basics

The process of image description construction consists in recognising basic entities relevant to the image content, and assigning a particular semantics to these entities. This process produces a set of completely defined objects according to a semantic model, which can be general or specific to a particular application or domain. We will present our proposition for an information retrieval model adapted to images.

2.1 The basic notions

In $EMIR^2$ the basic description of an image is a particular interpretation of it. An interpretation defines the semantics of the image objects considered as representative of the image in a particular context. To build the best description of the image we combine a set of interpretations, each interpretation corresponds to a particular view of the image, and the image is said to be a multi-viewed object. The two principal views are the physical view and the logical view. The logical view groups all aspects of image contents and its general context. This view is an aggregation of different basic views, the spatial, the structural, the perceptive, and the symbolic views.

The notion of complexity is inherent to the image content, and is considered in $EMIR^2$ by representing the image content as a set of concrete objects identified as relevant in the image, and their interrelationships. Here an image is partitioned into a set of sub-images, each corresponding to some relevant object, which is designated by the term image object. An image object corresponds to the real word objects of the scene whose projection in a two-dimensional space is the described image. The multi-view aspect of the image description is extended to the sub-images and their representative, the image objects. The Fig. 1 below identifies an image and an image object by their basic descriptive views.

The principal views, mainly the components of the logical view of an image and image object, are described below.

2.2 The physical view

The physical view of an image is the corresponding pixel matrix. Four main image types are defined in our model: the bitmap images, where image pixels can be black or white; the grey scale images, where a pixel is one the 256 grey levels; the palette colour images where a pixel has a colour among a set of 256 possible colours; and the true colour images, where a pixel can have a colour among a set of 224 different colours. We defined two categories of operations to manipulate the physical view of images. The first category includes general image processing functions, like the zoom, the scaling, the edge detection, ... The second category includes binary operations to produce new images by combining existing ones by means of , three operators AND, OR, NOT.

2.3 The logical view

Each particular view describing the image content is partial, and the integration of a set of views in the same model leads to a more complete representation of the image. These different views are combined in a global view, the logical view, which integrates all aspects of the image content and can be used as a faithful representation of it. The schema of the Fig. 1 below illustrates how do we combine these partial views to get a global one.

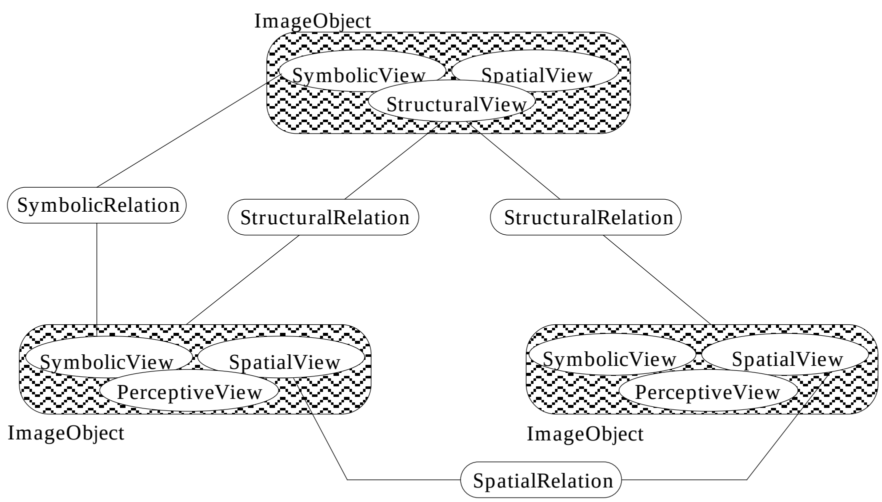

Figure 1: Logical view of an image

We identified four main complementary view types of the image. Each view is generally based on a set of descriptors linked together by relationships specific to the view. In the following we present each view, by identifying its basic components and the relationships between them.

2.3.1 The structural view

The structural view of an image defines the set of image objects that have been considered by the indexer as the most relevant to the image description. Each image object can be simple or complex, i.e. described by a structural view. The structural view is not a complete partition of the image, only relevant image objects are considered in the decomposition.

The structural elements of the image representation form a connex oriented graph whose nodes are the structural objects of the model (image and image objects) and the arcs correspond to the composition relation between these basic objects. The Fig. 2 shows an example of structural description of an image. In this example two image objects have been identified, a tree and a house, and considered in the structural view of the image. These image objects are described using structural views and decomposed into simpler image objects: foliage and trunk for the tree, and façade and roof for the house image object.

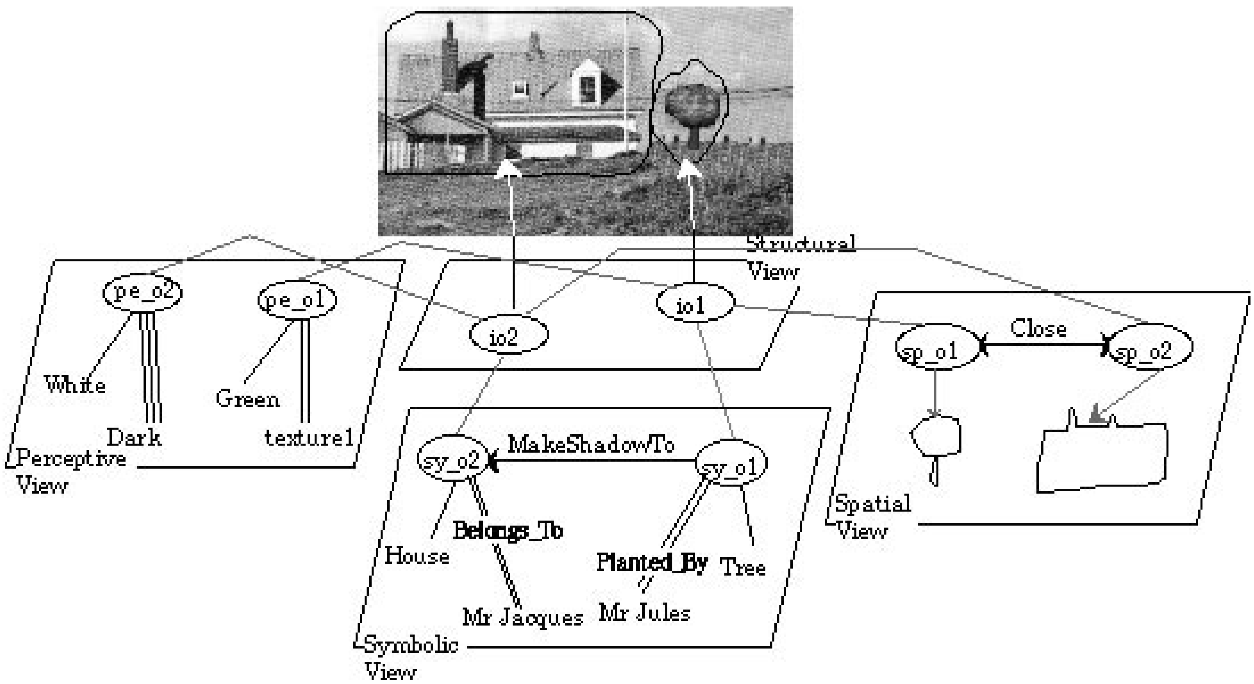

Figure 2: Example of an image description in EMIR^2

2.3.2 *The spatial view*

The spatial view of an image object represents the shape of the image objects (polygon, segment, ...) and the spatial relationships (far, north, overlap, ...) that indicate their relative positions inside the image. We define the spatial view of an image object as a combination of a set of modelling spaces, generally used in the literature.

According to [17] only four modelling spaces, topological, metric, vector and Euclidean, are relevant to spatial reasoning. Topological spaces are the most relevant, and they include only concepts of connectedness and continuity. Metric structures involve notions of distances. Vector spaces are well known; coordinates, directions, dimensions are typically vectorial. The more realistic structures, the Euclidean ones, admit notions of scalar products, orthogonality, angle and norm. We consider in EMIR^2 a complete mathematical model that represents all aspect of spatial knowledge about objects in an image.

The four modelling spaces are then considered in our model.

- The Euclidean space is used to describe the object shapes. Three basic categories of spatial objects are considered, the point, the segment and the polygon. A spatial object is defined given a list of points, and a point is identified by its Cartesian coordinates.
- The metric space is reduced to two spatial relationships based on the object distances. These relations are far and close.
- In the vector space we consider the four direction relations north, south, east and west.
- In the topological space we chose a relevant set of topological relations defined in [9]. This five relations set has two principal advantages. Its completeness, i.e. a couple of spatial objects are related at least by a relation, and its exclusiveness, i.e. a couple of spatial objects are related at most by a relation of the set. These relations are cross, overlap, disjoint, in, and touch.

One should keep in mind that the spatial relations do not relate image objects to each other, but relate their corresponding spatial views, which corresponds to their projection on the bi-dimensional plan of the image. This constraint implies that when a spatial view is inside another spatial view, the corresponding real world objects may not be included one in another.

The relations considered in EMIR^2 are computed using the Euclidean space. We provide a set of procedures that determines the topological, metric, and vector space descriptions of an image given the Euclidean space description of the image objects that compose it. In order to compute the vector and metric spaces we substitute to each spatial object its equidistant barycentre.

The spatial view of an image object based on the four modelling spaces combines the two classical approaches in image content modelling [7]. The object oriented approach is represented by the Euclidean space, and the relation oriented approach is represented by the three other modelling spaces.

The spatial view of an image can be seen as a graph whose nodes represents the spatial objects and whose arcs corresponds to spatial relations linking the spatial objects. The presence of a spatial view in the description of an image object is interpreted as the visibility of the image object. So, an image object with no spatial view is considered as not visible in the physical view of the image. In EMIR^2 we can represent all elements relevant to image content even if they are not visible (hidden elements).

Combining this set of modelling spaces induces a set of dependencies between the spatial views. In fact, when two objects intersect we can deduce they are close to each other, and when two objects are far from each other we can deduce they are disjoint.

2.3.3 The perceptive view

The perceptive view includes all the visual attributes of the image and/or image objects. It describes the appearance of the image components as perceived by an observer. In EMIR^2 we consider mainly three basic visual attributes, the colour, the brightness, and the texture.

- The Colour attribute captures the colour distribution in the image. The representation considered in EMIR^2 is the colour histogram, in which we consider the set of dominant colours and the ratio of the object surface represented by the colour. A colour value can be represented in different colour spaces, we consider in EMIR^2 for the moment the RGB colour space.

- The texture of an object is represented by the regular pattern that fills all the surface of the object [22]. For the moment we consider a set of basic textures, and the texture attribute of an object is instantiated by a value from this set.

- The brightness attribute is represented by a value corresponding to the average light in an object or surface.

An important advantage of using the perceptive view is the possibility to use automatic procedures to compute the image description according to this view. The same procedures can be used for image objects once their spatial views are defined.

2.3.4 The symbolic views

A symbolic view associates a semantic description to an image or to an image object. A wide range of possible descriptions can be used as symbolic views, we will limit ourselves to the ones generally used in IR. The simplest descriptions are terms, database attributes, and complex terms (compound nouns). More sophisticated representations like first order logic, terminological logic, conceptual graphs, semantic nets, have been used as well. We think the use of such symbolic views, with rich semantics, is necessary to achieve the best effectiveness in information retrieval, but using simple descriptions leads to efficient IRS.

Three general types of symbolic views are considered in EMIR^2 : classes, properties, and symbolic relations.

- A class defines the semantic category of an image object. For example we can describe an image containing a tree and a house using two image objects each described by a symbolic view of type class. The first object is described as an element of the class Tree and the second is described as an element of the class House. The set of possible classes corresponds to all possible concrete objects and is organised in an ontology by the IS-A (specificity/genericity) relationship, and is part of the image model.

- A property corresponds to an attribute defined by a couple of elements, representing the property identification and the domain value of the property. For example we can describe an image by a property called Author whose values are a subset of the string data type. We consider here two subsets, properties associated to images, e.g. size, date, author, etc., and those associated to image objects, e.g. identifier, name, etc.

- A symbolic relation corresponds to different elements of the image content involving the image objects : an action involving one or more objects in the image, the states of the objects, etc. For example an image showing two persons who are fighting can be described using two image objects whose symbolic views indicate the persons' names and a symbolic relation between these two objects corresponding to the fight event. A set of constraint rules have to be defined to control symbolic views construction. A rule defines for each relation the possible symbolic objects that can be linked using it. The RIME semantic model [2] is an example of such symbolic view.

2.3.5 Inter-View dependencies

Being partial descriptions of the same image object is not the only link between the different views of an image object. A particular property follows from this definition of the model. The structural relation between two image objects implies that the spatial view of the component object is inside the spatial view of the composed object :

The Fig. 2 shows a complete example of a description of an image. Here, two image objects are considered relevant to the image content description, the tree and the house. The image object corresponding to the tree is represented using a term symbolic view (tree), a general spatial view is used to state that the tree is visible in the image but its shape is not relevant, and the third component of the object is its structural view which states that it is composed f two simple entities, corresponding to its foliage and its trunk, and which are represented in the same way. In this example some spatial relations are used to link the spatial views of image objects, the foliage touches and is north of the trunk. No perceptive view has been used to describe the image in this example.

3 Formalisation of the image model

We present in this section the formalisation of the image model described above using a general mathematical formalism. This formal description considers all the elements presented above, plus some contextual elements necessary for the semantic interpretation of the image.

3.1 The physical view model

3.1.1 Definition

A physical view model is defined the following tuple :

$\mathcal{M}_{\rm ph}$ = ($I_{\rm ph}$, POINT, EC, TYPE, h, w, tc, $pixels$, $type$)

- $I_{\rm ph}$ is the set of physical view identifiers in $EMIR^2$.

- POINT is the set of natural number pairs representing the cartesian coordinates of possible points : POINT = $N^+ \times N^+$.

- EC is the colour set defined in a particular colour space. We consider in $EMIR^2$ the RGB colour space defined by : EC = $\{0, 1, ..., 255\} \times \{0, 1, ..., 255\} \times \{0, 1, ..., 255\}$.

- TYPE is the set of physical view types. For the moment it contains four elements : TYPE = {BW, GS, PC, TC}, with BW = Black & White, GS = Grey Scale, PC = Palette Colour, and TC = True Colour.

- $h : I_{\rm ph} \longrightarrow N^+$, h is a function that associates with each physical view identifier a positive number corresponding to the image height.

- $w : I_{\rm ph} \longrightarrow N^+$, w is a function that associates with each physical view identifier a positive number corresponding to the image width.

- *tc* : $I_{ph} \longrightarrow \mathcal{P}$(EC), *tc* is a function that associates with each physical view identifier the set of colours used in the corresponding image. $\mathcal{P}$(s) stands for the set of subsets of s.

- *pixels* : $I_{ph} \longrightarrow \mathcal{P}$(POINT $\times$ EC), this function associates with each physical view identifier the set of pixels of the image. Each pixel being defined as the association of a point and a colour.

- *type* : $I_{ph} \longrightarrow$ TYPE, this function associates with each physical view identifier the type of the corresponding image.

3.1.2 Constraints on the physical view model

A given physical view model is said to be coherent if the following constraints hold :

- The colour table of a physical view is linked to its type by the following rules :

  - $\forall$ i $\in I_{ph}$, if *type*(i_{ph}) = BW then *tc*(i) = $\{$(0,0,0), (255,255,255)$\}$.
  - $\forall$ i $\in I_{ph}$, if *type*(i_{ph}) = GS then *tc*(i) = $\{$(n,n,n), with $0 \leq n \leq 255\}$.
  - $\forall$ i $\in I_{ph}$, if *type*(i_{ph}) = PC then $\|$ *tc*(i) $\| \leq 256$.

- The height and length of the image gives the pixel numbers :

  $\forall$ i $\in I_{ph}$, $\|$ pixels(i) $\|$ = *w*(i) * *h*(i).

- We can not associate with the same point two different colours :

  $\forall$ i $\in I_{ph}$, $\forall$ (p_1, c_1) $\in$ *pixels*(i), $\forall$ (p_2,c_2) $\in$ *pixels*(i), if $c_1 \neq c_2$ then $p_1 \neq p_2$.

- The coordinates of the points of the pixels are limited to the image dimensions, and the colours belongs to the image colour table :

  $\forall$ i $\in I_{ph}$, $\forall$ (p,c) $\in$ *pixels*(i), p = (x,y), then $1 \leq x \leq$ *w*(i) and $1 \leq y \leq$ *h*(i) and c $\in$ *tc*(i).

3.2 The structural view model

3.2.1 Definition

An image structural view model (Mst) is defined by a set of image object identifiers and the composition relation between the objects.

$\mathcal{M}_{st}$ = (I_{io}, *CONT*)

- I_{io} is the set of possible image object identifiers in the structural view.

- *CONT* is the composition relation between image objects, *CONT* $\subseteq I_{io} \times I_{io}$. This composition relation depends on the semantics associated to the image objects.

3.2.2 Constraints on the structural view model

A structural view model is coherent if it respects the following set of constraints:

- The relation *CONT* is anti-symmetrical and transitive.

- An image object is component of only one image object.

  $\forall$ (io_1, io_2), (io_3, io_4) $\in$ *CONT*, if $io_2 = io_4$ then $io_3 = io_1$.

3.3 The perceptive view

3.3.1 Definition

The perceptive view is defined by :
$\mathcal{M}_{\mathrm{pe}}$ = (I_{pe}, TX, BR, CL, *tx*, *br*, *cl*)

- I_{pe} is the set of perceptive object identifiers.
- TX is the set of possible textures in the model.
- BR is the set of possible brightness values.
- CL is the set of possible colour values, CL $\subseteq$ EC.
- *tx* : $\mathrm{I}_{\mathrm{pe}} \longrightarrow$ TX, a function that associates with a perceptive object identifier a texture from TX.
- *br* : $\mathrm{I}_{\mathrm{pe}} \longrightarrow$ BR, a function that associates with a perceptive object identifier a brightness from BR.
- *cl* : $\mathrm{I}_{\mathrm{pe}} \longrightarrow$ CL, a function that associates with a perceptive object identifier a colour from CL.

Each basic set (TX, BR, CL) is augmented by a null value ($\mathrm{tx}_{\emptyset} \in$ TX, $\mathrm{br}_{\emptyset} \in$ BR, $\mathrm{cl}_{\emptyset} \in$ CL) to be used when the value is unknown or undefined.

3.4 The spatial view model

3.4.1 Definition

$\mathcal{M}_{\mathrm{sp}}$ = (I_{sp}, POINT, OS, RSPA, *shape*, R_{sp})

- I_{sp} is the set of spatial objects identifiers.
- POINT is the set of integer pairs that represent the cartesian coordinates of all possible points, POINT = $\mathrm{N}^{+} \times \mathrm{N}^{+}$.
- OS is the set of basic image objects that can be used to represent the shape of the object in an image. Three basic types are used in EMIR^2 for the moment, the point, the segment and the polygon, and they are defined as follows :

  SEGMENT $\subseteq$ POINT $\times$ POINT, the points being the segment extremities,

  POLYGON $\subseteq \mathcal{P}$(SEGMENT), each segment being a side of the polygon,

  OS = POINT $\cup$ SEGMENT $\cup$ POLYGON.

- RSPA represents the set of spatial relations defined in EMIR^2, and which is : RSPA = {Far, Close, East, West, North, South, In, Disjoint, Touch, Overlap, Cross}.
- *shape* : $\mathrm{I}_{\mathrm{sp}} \longrightarrow$ OS is a function that associates with each spatial object identifier its shape which is defined by a subset of OS.
- $R_{sp} \subseteq$ RSPA $\times\ \mathrm{I}_{\mathrm{sp}} \times \mathrm{I}_{\mathrm{sp}}$, is the relation that represents all possible spatial relations linking the spatial objects of the spatial view.

3.4.2 Constraints on the spatial view model

- The extremities of a segment are disjoint.

  $\forall$ (p_1, p_2) $\in$ SEGMENT, then $p_1 \neq p_2$.

- Each point of a segment in a polygon definition is shared by two and only two distinct segments in the same polygon definition.

  $\forall$ p $\in$ POLYGON, $\forall$ (p_{11}, p_{12}) $\in$ p, then $\exists!$ (p_{21}, p_{22}) $\in$ p, such that ($p_{11} = p_{21}$ and $p_{12} \neq p_{22}$) xor ($p_{11} = p_{22}$ and $p_{12} \neq p_{21}$) xor ($p_{12} = p_{21}$ and $p_{11} \neq p_{22}$) xor ($p_{12} = p_{22}$ and $p_{11} \neq p_{21}$).

- Each element of OS is used only once in the definition of the shape of a spatial object.

  $\forall$ o_sp_1, $o\_sp_2 \in I_{sp}$, *shape*(o_sp_1) $\neq$ *shape*(o_sp_2).

- The spatial relations between spatial objects are computed using the shape of the objects.

  (sr, o_sp_1, o_sp_2) $\in R_{sp}$ iff HOLDS(sr, *shape*(o_sp_1), *shape*(o_sp_2)).

  HOLDS(sr, so_1, so_2) is a boolean function that checks if the spatial relation sr holds between the two elements so_1 and $so_2 \in$ OS. The definition of HOLDS for each spatial relation is given in the annex A of this paper.

3.5 The symbolic view

As presented previously the symbolic view is specific to a particular application and can not be defined independently from the application specificity. We will define the symbolic view model as the association between an application semantic model and a set of abstractions representing the symbolic view.

3.5.1 The application semantic model

The application semantic model includes the object class ontology, the definition of the composition relation between object classes, the symbolic relations definitions, and the properties definition.

$\mathcal{M}_{\rm app}$ = (ID_{cl}, *IS-A*, ID_{pr}, ID_{rs}, VAL_PROP, *PROP*, RSYMB, *COMP*, *domain*)

- ID_{cl} is the set of class identifiers. This set is organised as a lattice by the *IS-A* relation with a minimal and maximal element ($\perp$ and $\top$).

- ID_{pr} is the set of property identifiers.

- ID_{rs} is the set of symbolic relation identifiers.

- VAL_PROP is the set of possible values of the properties, VAL_PROP = Real $\cup$ Integer $\cup$ String $\cup$ Boolean.

- *domain* : $ID_{pr} \longrightarrow \mathcal{P}$(VAL_PROP), is the function that defines for each property the set of its possible values.

- *PROP* is the set of property definitions.

  PROP $\subseteq ID_{pr} \times ID_{cl} \times \mathcal{P}$(VAL_PROP)

- RSYMB is the set of symbolic relation definitions.

  RSYMB $\subseteq ID_{rs} \times ID_{cl} \times ID_{cl}$.

- *COMP* $\subseteq ID_{cl} \times ID_{cl}$, is the composition relation between classes. (c_1, c_2) $\in$ *COMP* means that objects of the class c_1 can be components of objects of the class c_2. This relation is mainly used to control the validity of the structural view (object decomposition) of an image.

3.5.2 The symbolic view model definition

The symbolic view model is defined relative to an application semantic model. It associates with the set of symbolic objects their semantic interpretation, and is defined by :

$\mathcal{M}_{\rm sy}$ = ($\mathcal{M}_{\rm app}$, $I_{\rm sy}$, *cl*, *RI*, *PI*)

- $I_{\rm sy}$ is the set of symbolic objects identifiers.
- *cl* : $I_{\rm sy} \longrightarrow ID_{\rm cl}$, is the function that associates with a symbolic object identifier its class.
- *RI* $\subseteq ID_{\rm rs} \times I_{\rm sy} \times I_{\rm sy}$, is the relation that represents the symbolic relations between the symbolic object identifiers.
- *PI* $\subseteq ID_{\rm pr} \times I_{\rm sy} \times$ VAL_PROP, is the relation that represents all the properties associated with symbolic objects.

3.5.3 Constraints on the symbolic view model

- The elements of RI are instances of the symbolic relation definitions in the application semantic model.

  $\forall$ sy_o$_1$, sy_o$_2$ $\in I_{\rm sy}$, (rs, sy_o$_1$, sy_o$_2$) $\in$ *RI*, iff $\exists$ rs $\in ID_{\rm rs}$, and c_1, $c_2 \in ID_{\rm cl}$, and (rs, c_1, c_2) $\in$ RSYMB, such that *cl*(sy_o$_1$) IS_A c_1, and *cl*(sy_o$_2$) IS_A c_2.

- The elements of *PI* are instances of the property definitions in the application semantic model.

  $\forall$ sy_o$_1$ $\in I_{\rm sy}$, (pr, sy_o$_1$, v) $\in$ *PI*, iff $\exists$ pr $\in ID_{\rm pr}$, and $c_1 \in ID_{\rm cl}$, and (pr, c_1, $ID_{\rm val}$) $\in$ *PROP*, such that *cl*(sy_o$_1$) $\leq c_1$, and v $\in$ *domain*(pr).

3.5.4 Example

In an image base representing photographs, each image is described by a set of attributes : author name, place, etc. The main subject of images is landscapes and houses. We define, for this particular application, two properties of the image (author and place) and only one symbolic relation MakeShadowTo. The semantic model of the application is then defined by :

$\mathcal{M}_{\rm app\_ex}$ = ($ID_{\rm cl}$, $ID_{\rm pr}$, $ID_{\rm rs}$, VAL_PROP , *PROP*, RSYMB, *COMP*)

- $ID_{\rm cl}$ = {Trunk, Tree, Foliage, Door, House}.
- $ID_{\rm pr}$ = {Author, Place},
- $ID_{\rm rs}$ = {MakeShadowTo},
- VAL_PROP = Real $\cup$ Integer $\cup$ String $\cup$ Boolean.
- *domain*(Author) = NPhoto $\subset$ VAL_PROP, *domain*(Place) = NPlace $\subset$ VAL_PROP, NPhoto = {"NADAR PAUL", ...} is the set of the author names, and NPlace = {"Paris", ... } is the set of possible places.
- *PROP* = { (Author $\times$ Image $\times$ NPhoto), (Place $\times$ Image $\times$ NPlace)}
- RSYMB = {({"MakeShadowTo"} $\times$ Tree $\times$ House)}
- *COMP* = {(Trunk, Tree), (Foliage, Tree), (Door, House)}

The symbolic view of a particular image taken by the famous French photograph NADAR in Paris and containing a tree and a house and the shadow of the tree covering the house is represented by the expression SyViewImage$_1$ which is an instance of the symbolic view model :

SyViewImage$_1$ = ($\mathcal{M}_{\rm app\_ex}$, {sy_o$_0$, sy_o$_1$, sy_o$_2$}, *cl*, *PI*, *RI*)

sy_o$_0$ corresponds to the entire image, and sy_o$_1$, sy_o$_2$ to two symbolic objects.

- *cl*(sy_o$_0$) = Image, *cl*(sy_o$_1$) = Tree, *cl*(sy_o$_2$) = House.
- *PI* = {(Place, sy_o$_0$, "Paris"), (Author, sy_o$_0$, "NADAR PAUL")}
- *RI* = { ("MakeShadowTo", sy_o$_1$, sy_o$_2$)}.

3.6 The image model

3.6.1 Definition

An image model $\mathcal{M}_{\rm im}$ is defined as an aggregation of a the basic coherent EMIR^2 view models and a set of relations that represent the inter-view dependencies :

$\mathcal{M}_{\rm im}$ = ($\rm I_{im}$, $\mathcal{M}_{\rm ph}$, $\mathcal{M}_{\rm st}$, $\mathcal{M}_{\rm pe}$, $\mathcal{M}_{\rm sp}$, $\mathcal{M}_{\rm sy}$, $L_{\rm sp}$, $L_{\rm sy}$, $L_{\rm pe}$)

- $\rm I_{im}$ is the set of EMIR^2 image identifiers.
- $\mathcal{M}_{\rm ph}$ is a coherent physical view model.
- $\mathcal{M}_{\rm st}$ is a coherent structural view model.
- $\mathcal{M}_{\rm pe}$ is a coherent perceptive view model.
- $\mathcal{M}_{\rm sp}$ is a coherent spatial view model.
- $\mathcal{M}_{\rm sy}$ is a coherent symbolic view model.
- $L_{\rm sp} \subseteq \rm I_{io} \times I_{sp}$, is the relation that associates with an image object a spatial object from the spatial view.
- $L_{\rm pe} \subseteq \rm I_{io} \times I_{pe}$, is the relation that associates with an image object a perceptive object from the perceptive view.
- $L_{\rm sy} \subseteq \rm I_{io} \times I_{sy}$, is the relation that associates with an image object a symbolic object from the symbolic view.

3.6.2 Constraints on the image model

An instance of the image model is noted i and each element e of i is noted i.e, for example $\rm i.i_{ph}$ is the identifier of the physical view of the image i.

An image model Mim is coherent if it respects the following constraints :

- The relation $L_{\rm sp}$ (res. $L_{\rm sy}$, $L_{\rm pe}$) associates at most one spatial object (res. symbolic, perceptive) with an image object.
  - $\forall$ sp_o $\in \rm i.I_{sp}$, $\exists$ io $\in \rm i.I_{io}$, such that (io, sp_o) $\in$ i.$L_{\rm sp}$, and $\forall$ ($\rm io_1$, $\rm sp\_o_1$), ($\rm io_2$, $\rm sp\_o_2$) $\in$ i.$L_{\rm sp}$, $\rm io_1 = io_2 \longleftrightarrow sp\_o_1 = sp\_o_2$.
  - $\forall$ sy_o $\in \rm i.I_{sy}$, $\exists$ io $\in \rm i.I_{io}$, such that (io, sy_o) $\in$ i.$L_{\rm sy}$, and $\forall$ ($\rm io_1$, $\rm sy\_o_1$), ($\rm io_2$, $\rm sy\_o_2$) $\in$ i.$L_{\rm sy}$, $\rm io_1 = io_2 \Longleftrightarrow sy\_o_1 = sy\_o_2$.
  - $\forall$ pe_o $\in \rm i.I_{pe}$, $\exists$ io $\in \rm i.I_{io}$, such that (io, pe_o) $\in$ i.$L_{\rm pe}$, and $\forall$ ($\rm io_1$, $\rm pe\_o_1$), ($\rm io_2$, $\rm pe\_o_2$) $\in$ i.$L_{\rm pe}$, $\rm io_1 = io_2 \Longleftrightarrow pe\_o_1 = pe\_o_2$.
- The cartesian coordinates of the points used in the spatial object shape definition are included within the limits of the physical view of the image :

  $\forall$ e $\in$ OS, $\forall$ (x, y) $\in$ *pts*(e), $1 \leq x \leq$ i.*w*($\rm i.i_{ph}$) and $1 \leq y \leq$ i.*h*($\rm i.i_{ph}$).

  pts : OS $\longrightarrow \mathcal{P}$(POINT), is a function that gives the points used in the definition of a spatial object.
- The composition relation between image objects in the structural view is an instantiation of the composition relation defined on the object classes in the symbolic view :

  $\forall$ ($\rm io_1$,$\rm io_2$) $\in$ i.*CONT*, $\exists$ ($\rm c_1$, $\rm c_2$) $\in$ i.*COMP*, *cl*($\rm sy\_o_1$) IS_A $\rm c_1$, *cl*($\rm sy\_o_2$) IS_A $\rm c_2$, ($\rm io_1$, $\rm sy\_o_1$), and ($\rm io_2$, $\rm sy\_o_2$) $\in$ i.$L_{\rm sp}$.
- The colours used in the perceptive view are included in the colour table of the physical view :

  $\forall$ pe_o $\in \rm i.I_{pe}$, *cl*(pe_o) $\in$ i.*tc*($\rm i.i_{ph}$) $\cup \{\rm cl_{\emptyset}\}$.

3.7 The image base

An EMIR^2 image base is defined as a collection of instances of a coherent EMIR^2 image model.
$\mathcal{EMIR}_{\rm base}$ = ($\mathcal{M}_{\rm im}$, $\rm I_{im}$).

- $\mathcal{M}_{\rm im}$ is a coherent EMIR^2 image model.
- $\rm I_{im}$ is a set of instances of the image model $\mathcal{M}_{\rm im}$.

4 The query language and the correspondence function

We present in this section the elements of the correspondence model intended for EMIR^2. We will give the general guidelines for the query language, and the query definition, thereafter the list of selection criteria to be considered in comparing an EMIR^2 query and an EMIR^2 image.

4.1 The query language

A query is an instance of the image model with some new possibilities :

- We can use generic identifiers instead of real identifiers to represent all objects : image objects, symbolic, perceptive, and spatial objects, image identifier and physical view identifier. For more convenience we can use the undefined identifier * for image and physical view identifiers.
- Fuzzy values for the perceptive views Colour and Brightness can be used in a query. These fuzzy values correspond to a set of basic values from the domains Colour and Br defined in the image base context.
  - The colour of an object (image or image object) in a query can be represented by an identifier that correspond to a subset of the colour space defined in the image base context. This sub-set is denoted by the function dom. For example the colour Green, does not correspond to a single colour, but to a set of colours that can be perceived as green by a human being, dom(Green) = $\{\rm cc_1, cc_2, ..., cc_n\}$. We define a set VAL_CL that includes the set of terms representing the fuzzy colours that can be used in an EMIR^2 query, and with each we associated a set of colours from Colour.

    VAL_CL = {Green, Red, Blue, ...}.

    dom(Green), dom(Red), dom(Blue) $\in \mathcal{P}$(Colour).
  - The fuzzy values for the brightness are terms representing subintervals of the possible brightness values, [0 .. 1], this interval is denoted by the function dom. For example the term Dark corresponds to the interval dom(Dark) = [0 .. 0.1], and the term bright corresponds to the interval dom(Bright) = [0.9 .. 1]. We define a set VAL_BR that includes all possible fuzzy brightness terms that can be used in an EMIR^2 query.

    VAL_BR = {Dark, Bright, Mat, ...}

    dom(Dark), dom(Bright), dom(Mat) $\subset$ [0 .. 1]

4.2 The correspondence function

We will list here the basic selection criteria to be respected by a query (q) and an image (d) represented in the EMIR^2 model such that the image can be considered as relevant to the query.

The image d and the query q are defined as instances of the image model $\mathcal{M}_{\rm im}$, with the possible extensions to the query described in the section above.

The image d is considered as answering the query q iff we can find a surjective application, denoted $\mathcal{A}$, from the set $\rm d.I_{io}$ in the set $\rm q.I_{io}$ that respects the following constraints:

$\mathcal{A} \subseteq \rm d.I_{io} \times q.I_{io}$

The application $\mathcal{A}$ has the following properties :

(c1) It is surjective, i.e. $\forall\ \rm io_{q1} \in q.I_{io}$, $\exists\ \rm io_{d1} \in d.I_{io}$, such that $(\rm io_{d1}, io_{q1}) \in \mathcal{A}$.

(c2) The antecedent of an element of $\rm q.I_{io}$ is unique : $\forall\ (\rm io_{d1}, io_{q1}) \in \mathcal{A}$ if $\exists\ (\rm io_{d2}, io_{q1}) \in \mathcal{A}$ then $\rm io_{d1} = io_{d2}$. This constraint is introduced so that all the constrains on the views of an object of the query are verified by the views of the same object from the image d.

4.2.1 Constraints on the physical view

The type, width, and height of the query physical view are similar to image physical view definition.

(c3) $w(\mathrm{d.i_{ph}})$ similar_to $w(\mathrm{q.i_{ph}})$,

(c4) $h(\mathrm{d.i_{ph}})$ similar_to $h(\mathrm{q.i_{ph}})$,

(c5) $type(\mathrm{d.i_{ph}})$ similar_to $type(\mathrm{q.i_{ph}})$.

4.2.2 Constraints on the structural view

If a composition relation between two image objects from $\mathrm{q.I_{io}}$ holds, then a composition relation must hold between their corresponding objects from $\mathrm{d.I_{io}}$, and considering the transitivity of the composition relation. We define a function Components : $\mathrm{I_{io}} \longrightarrow \mathcal{P}(\mathrm{I_{io}})$, as the set of image objects directly or indirectly composing the object io.

Components(io) = $\{\mathrm{io_1} \in \mathrm{I_{io}}, (\mathrm{io}, \mathrm{io_1}) \in \mathit{CONT}$ or $\exists\ \mathrm{io_2} \in \mathrm{I_{io}}, \mathrm{io_1} \in$ Components($\mathrm{io_2}$) and $(\mathrm{io}, \mathrm{io_2}) \in \mathit{CONT}\}$.

(c6) $\forall\ (\mathrm{io_{d1}}, \mathrm{io_{q1}}) \in \mathcal{A}$, if $\exists\ \mathrm{io_{q2}}$ such that $(\mathrm{io_{q1}}, \mathrm{io_{q2}}) \in \mathrm{q.}\mathit{CONT}$ then $\exists\ \mathrm{io_{d2}} \in \mathrm{d.I_{io}}$, such that $(\mathrm{io_{d2}}, \mathrm{io_{q2}}) \in \mathcal{A}$ and $\mathrm{io_{d2}} \in$ Components($\mathrm{io_{d1}}$).

4.2.3 Constraints on the symbolic view

(c7) Compatibility of corresponding image object classes.

$\forall\ (\mathrm{io_{d1}}, \mathrm{io_{q1}}) \in \mathcal{A}$, such that $\exists\ (\mathrm{io_{q1}}, \mathrm{sy\_o_{q1}}) \in \mathrm{q.}L_{\mathrm{sy}}$, and $\exists\ (\mathrm{io_{d1}}, \mathrm{sy\_o_{d1}}) \in \mathrm{d.}L_{\mathrm{sy}}$, then $\mathrm{d.}cl(\mathrm{sy\_o_{d1}})$ IS_A $\mathrm{q.}cl(\mathrm{sy\_o_{q1}})$.

(c8) Compatibility of corresponding image object properties.

$\forall\ (\mathrm{io_{d1}}, \mathrm{io_{q1}}) \in \mathcal{A}$, such that $\exists\ (\mathrm{io_{q1}}, \mathrm{sy\_o_{q1}}) \in \mathrm{q.}L_{\mathrm{sy}}$, and $\exists\ (\mathrm{io_{d1}}, \mathrm{sy\_o_{d1}}) \in \mathrm{d.}L_{\mathrm{sy}}$, and $\exists\ (\mathrm{pr_1}, \mathrm{sy\_o_{q1}}, \mathrm{v_q}) \in \mathrm{q.}\mathit{PI}$, then $\exists\ (\mathrm{pr_1}, \mathrm{sy\_o_{d1}}, \mathrm{v_d}) \in \mathrm{d.}\mathit{PI}$, such that $\mathrm{v_q}$ similar_to $\mathrm{v_d}$.

(c9) Compatibility of corresponding image object symbolic relations.

$\forall\ (\mathrm{io_{d1}}, \mathrm{io_{q1}}), (\mathrm{io_{d2}}, \mathrm{io_{d2}}) \in \mathcal{A}$, such that $\exists\ (\mathrm{io_{q1}}, \mathrm{sy\_o_{q1}}), (\mathrm{io_{d2}}, \mathrm{sy\_o_{d2}}) \in \mathrm{q.}L_{\mathrm{sy}}$, and $\exists\ (\mathrm{io_{d1}}, \mathrm{sy\_o_{d1}}), (\mathrm{io_{d1}}, \mathrm{sy\_o_{d1}}) \in \mathrm{d.}L_{\mathrm{sy}}$, and $\exists\ (\mathrm{r_1}, \mathrm{io_{q1}}, \mathrm{io_{d2}}) \in \mathrm{q.}\mathit{RI}$, then $\exists\ (\mathrm{r_1}, \mathrm{io_{d1}}, \mathrm{io_{d2}}) \in \mathrm{d.}\mathit{RI}$.

4.2.4 Constraints on the perceptive view

$\forall\ (\mathrm{io_d}, \mathrm{io_q}) \in \mathcal{A}$, such that $\exists\ (\mathrm{io_q}, \mathrm{pe\_o_q}) \in \mathrm{q.}L_{\mathrm{pe}}$, and $\exists\ (\mathrm{io_d}, \mathrm{pe\_o_d}) \in \mathrm{d.}L_{\mathrm{pe}}$.

(c10) Compatibility of corresponding image object colours.

$\mathrm{d.}cl(\mathrm{pe\_o_d}) = \mathrm{q.}cl(\mathrm{pe\_o_q})$ or $\mathrm{d.}cl(\mathrm{pe\_o_d}) \in \mathrm{q.}cl\ (\mathrm{pe\_o_q})$ if $\mathrm{q.}cl(\mathrm{pe\_o_q}) \in$ VAL_CL.

(c11) Compatibility of corresponding image object brightness values

$\mathrm{d.}br(\mathrm{pe\_o_d}) = \mathrm{q.}br(\mathrm{pe\_o_q})$ or $\mathrm{d.}br(\mathrm{pe\_o_d}) \in \mathrm{q.}br(\mathrm{pe\_o_q})$, if $\mathrm{q.}br(\mathrm{pe\_o_q}) \in$ VAL_BR.

(c12) Compatibility of corresponding image object textures

$\mathrm{d.}tx(\mathrm{pe\_o_d})$ similar_to $\mathrm{q.}tx(\mathrm{pe\_o_q})$.

4.2.5 Constraints on the spatial view

(c13) The shape of corresponding image objects, represented by the spatial objects, should be similar.

$\forall\ (\mathrm{io_d}, \mathrm{io_q}) \in \mathcal{A}$, such that $\exists\ (\mathrm{io_d}, \mathrm{sp\_o_d}) \in \mathrm{d.}L_{\mathrm{sp}}$, and $\exists\ (\mathrm{io_{q1}}, \mathrm{sp\_o_q}) \in \mathrm{q.}L_{\mathrm{sp}}$, then $\mathrm{q.}shape(\mathrm{sp\_o_q})$ similar_to $\mathrm{d.}shape(\mathrm{sp\_o_d})$.

(c14) The spatial relations defined in Rsp on the spatial views of the query should be respected by spatial views of the corresponding image objects in the image d. Only these spatial relations are considered in comparing a spatial objects of the query with those of the image.

$\forall\ (\mathrm{io_{d1}}, \mathrm{io_{q1}}), (\mathrm{io_{d2}}, \mathrm{io_{d2}}) \in \mathcal{A}$,

such that $\exists\ (\mathrm{io_{d1}}, \mathrm{sp\_o_{d1}}), (\mathrm{io_{d2}}, \mathrm{sp\_o_{d2}}) \in \mathrm{d.}L_{\mathrm{sp}}$,

and $\exists\ (\mathrm{io_{q1}}, \mathrm{sp\_o_{q1}}), (\mathrm{io_{d2}}, \mathrm{sp\_o_{d2}}) \in \mathrm{q.}L_{\mathrm{sp}}$,

and $\exists\ (\mathrm{rs}, \mathrm{sp\_o_{q1}}, \mathrm{sp\_o_{d2}}) \in \mathrm{q.}R_{\mathrm{sp}}$,

then $\exists\ (\mathrm{rs}, \mathrm{sp\_o_{d1}}, \mathrm{sp\_o_{d2}}) \in \mathrm{d.}R_{\mathrm{sp}}$.

rs $\in$ RSPA = {north, south, west, east, far, close, cross, in, disjoint, intersect, overlap, touch}.

Important : the function similar_to is used to estimate the similarity between two elements, its expression depends on the elements type : spatial objects, image width and height, property values, etc. In the first implementation of this model using Sowa's conceptual graph formalism we use a simple form for this function, which is the equality.

5 Conclusion and future work

We presented in this paper our approach for an extended content based representation and retrieval of images. $EMIR^2$ is a formal model that integrates all aspects considered as relevant to image content description for effective information retrieval. In this model we combine different types of image representations to get the most precise and the most exhaustive image content description. These different representations are identified as particular views and an abstraction to combine them is defined. A general mathematical formalism has been used to state the model elements, the query language and the selection criteria to be used for image-query similarity estimation.

An operational model $EMIR^2$-CG, based on Sowa's conceptual graph formalism, is defined to implement the concepts of the model $EMIR^2$, and the similarity function to be used in the retrieval engine. We are currently experimenting $EMIR^2$-CG using a collection of images of the old Paris areas. The retrieval engine, based on a conceptual graph framework, has been developed on top of the object oriented database system O2. The test collection is composed of two main parts: the indexing of the images, which has been done by specialists, using a sophisticated term based symbolic view, and the modelling of domain dependent knowledge which includes the concept type lattices corresponding to the different views, mainly the class type symbolic view (thesaurus of the domain), and a set of image properties [20].

$EMIR^2$ is open to integrate other media description in the same framework. The symbolic view associated with images was inspired from textual data representation, and according to that a text can be easily represented in $EMIR^2$ using a particular symbolic view, and then a comparison between an image and a text could be based on this symbolic description.

The future work in $EMIR^2$ is conducted in three directions. First one we try to introduce some uncertainty and/or relevance measures in the image representation, since the image interpretation process, depending on its nature (manual or automatic), produces descriptions which are far from being perfect: they can be partial, ambiguous, uncertain, more or less relevant, etc. The second work axe concerns the definition of a complete graphics model to be used as a spatial view and mainly to get an effective function for comparing object shapes. The third work axe concerns the use of an operational model more suitable for IR, since Conceptual Graphs does not deal with logical inferences. Terminological logic based models seems to be the more promising for the moment and we will start soon working on this point.

6 Annex A. Definition of the spatial relations

6.1 Metric modeling space relations

The definition of the metric relations is based on the normalised distance function ndist : ndist (so_1,so_2) = , where distmax corresponds to the diagonal of the image, and mdist(so_1, so_2) is the minimal distance between the objects so_1 and so_2.

- HOLDS(close, so_1, so_2) $\Longleftrightarrow$ ndist(so_1, so_2) $\leq d_{\min}$, $d_{\min} \in [0 .. 1]$.
- HOLDS(far, so_1, so_2) $\Longleftrightarrow$ ndist(so_1,so_2) $\geq d_{\max}$, $d_{\max} \in [0 .. 1]$.

$d_{\min}$ and $d_{\max}$ are two parameters of the model, that are dependent on the properties of the relations far and close.

6.2 Vector modeling space relations

Let Bo_1 and Bo_2 be the barycentres of the spatial objects so_1 and so_2, α is the angle between the line defined by Bo_2 and which is parallel to Y-axis, and the line defined by the points Bo_1 and Bo_2.

- HOLDS(north, so_1, so_2) $\Longleftrightarrow 0 \leq \alpha \leq \pi$ and $b_{\min} \leq \sin(\alpha) \leq b_{\max}$.
- HOLDS(south, so_1, so_2) $\Longleftrightarrow -\pi \leq \alpha \leq 0$ and $b_{\min} \leq -\sin(\alpha) \leq b_{\max}$.
- HOLDS(east, so_1, so_2) $\Longleftrightarrow -\pi/2 \leq \alpha \leq \pi/2$ and $b_{\min} \leq \cos(\alpha) \leq b_{\max}$.
- HOLDS(west, so_1, so_2) $\Longleftrightarrow \pi/2 \leq \alpha \leq 3\pi/2$ and $b_{\min} \leq -\cos(\alpha) \leq b_{\max}$.

With $b_{\min}$, $b_{\max} \in [0 .. 1]$, being two parameters of the model.

6.3 Topological modeling space relations

The topological modeling space relations are taken from [9]. Their definition is based upon three functions, the boundary (∂o), the interior (o$^\circ$) of the objects, and Dim.

6.3.1 Basic functions

- ∂so represents the set of points of the boundary of a spatial object so.

so	∂(so)
point	Ø
segment(p_1, p_2)	$\{p_1, p_2\}$
polygon	$\{s_i$ / $s_i \in$ segment(so)$\}$

- o$^\circ$ represents the interior points of the object so.

so	so$^\circ$
point, p	p
segment, (p_1, p_2)	so - $\{p_1, p_2\}$
polygon	so - $\{$ s_i / $s_i \in$ segment(so)$\}$

- The operator so represents all the points of the object so. So we have :

  so = ∂so $\cup$ so$^\circ$, and ∂so $\cap$ so$^\circ$ = Ø.

- We define the function Dim as the dimension of a set of points (ps).

ps	dim(ps)
Ø	-
ps contains at least a point but no lines nor areas.	0
ps contains at least a line but no areas.	1
ps contains at least an area.	2

6.3.2 Topological relations definition

- HOLDS(touch, so_1, so_2) $\Longleftrightarrow$ ($so_1{}^\circ \cap so_2{}^\circ$ = Ø) $\wedge$ ($so_1 \cap so_2 \neq$ Ø).

  The touch relationship holds if the contours of the two objects intersects. It applies to the pairs poly/poly, segt/segt, segt/poly, point/poly, point/segt, but not to point/point.

- HOLDS(in, so_1, so_2) $\Longleftrightarrow$ ($so_1 \cap so_2 = so_1$) $\wedge$ ($so_1{}^\circ \cap so_2{}^\circ \neq$ Ø).

  The in relationship holds if the first object is included in the second.

- HOLDS(cross, so_1, so_2) $\Longleftrightarrow$ (dim($so_1{}^\circ \cap so_2{}^\circ$) = (MAX (dim($so_1{}^\circ$), dim($so_2{}^\circ$)) - 1)) $\wedge$ ($so_1 \cap so_2 \neq so_1$) $\wedge$ ($so_1 \cap so_2 \neq so_2$).

  The cross relationship applies to segt/segt and segt/poly situations.

- HOLDS(overlap, so_1, so_2) $\Longleftrightarrow$ (dim($so_1{}^\circ$) = dim($so_2{}^\circ$) = dim($so_1{}^\circ \cap so_2{}^\circ$)) $\wedge$ ($so_1 \cap so_2 \neq so_1$) $\wedge$ ($so_1 \cap so_2 \neq so_2$).

  The overlap relationship applies to poly/poly and segt/segt situations.

- HOLDS(disjoint, so_1, so_2) $\Longleftrightarrow$ ($so_1 \cap so_2$ = Ø).

  The disjoint relationship applies to every situation.

- HOLDS(intersect, so_1, so_2) $\Longleftrightarrow$ (so_1, touch, so_2) $\vee$ (so_1, cross, so_2) $\vee$ (so_1, overlap, so_2) $\vee$ (so_1, in, so_2) $\vee$ (so_2, in, so_1).

  The intersect relationship applies to every situation, and represents the union of the relations touch, in, overlap and cross.

References

[1] Yoshitaka A., Kishida S., Hirakawa M., and Ichikawa T. Knowledge-assisted content-based retrieval for multimedia databases. In *Proceedings of the international conference on mulitmedia computing and systems*, pages 131–139, may 14-19 1994.

[2] Catherine Berrut and Yves Chiaramella. Indexing medical reports in a multimedia environment: the rime experimental approach. In *ACM-SIGIR Conference on Research and Development in Information Retrieval, Boston, Massachussetts, USA*, pages 187–197, June 25–28 1989.

[3] Catherine Berrut and Mourad Mechkour. Representation of images in databases: a preliminary study. In *IFIP-WG2.6 2nd Conference on Visual Database Systems, Budapest, Hungary*, September 30 - October 3 1991.

[4] Faloutsos C., Flickner M., Niblack W., Petrovic D., Equitz W., and Barber R. Efficient and effective querying by image content. Research report, IBM Almaden Research Center, 1993.

[5] Goble C., O'docherty M., Crowther P., Ireton M., Oakley J., and Xydeas C. The manchester multimedia information system. In *Proceedings VLDB 1992*, pages 39–55, 1992.

[6] Meghini C. An image retrieval model based on classical logic. In *ACM SIGIR*, July 1995.

[7] Papadias D. and Sellis T. The semantics of relations in 2d space using representative points : spatial indexes. In *Lecture Notes in computer science, 716. Spatial information theory*, pages 234–247. Springer-Verlag, September 1993. Proceedings of the european conference, COSIT'93, Marciana Marina, Elba Island, Italy, September 1993.

[8] Bertino E., Rabitti F., and Thanos C. Multos : a document server for distributed office systems. In *International Conference in Extending Database technology, EDBT'88*, 1988.

[9] Clementini E., Di Felice P., and Van Oostrom P. A small set of formal topological relationships suitable for end-user interaction. In *Lecture notes in computer science. Advances in Spatial Databases*, pages 277–295. Springer-Verlag, June 1993. Proceedings of the third international symposium, SSD'93, Singapore.

[10] Hirabayashi F., Matoba H., and Kasahara Y. Information retrieval using impression documents as a clue. In *ACM-SIGIR*, pages 233–243, 1988.

[11] Sowa J. F. *Conceptual structures : information processing in mind and machine*. Addison-Wesley publishing company, 1984.

[12] Halin G., Crehange M., and Kerekes P. Machine learning and vectorial matching for an image retrieval model : Exprim and the system rivage. In *ACM-SIGIR*, pages 99–114, 1990.

[13] Salton G. and McGill M. J. *Introduction to modern information retrieval*. McGraw-Hill, 1983.

[14] Tamura H. and Yokoya N. Image database systems : a survey. *Pattern Recognition*, 17(1):29–43, 1984.

[15] BACH J., PAUL S., and JAIN R. An interactive image management system for face retrieval. In *Proceedings of the 1st CIKM conference*, pages 3.1–3.22, 1992.

[16] Nie J. An outline of a general model for information retrieval systems. In *Proceedings of the 11th ACM SIGIR conference*, pages 495–506, 1988.

[17] Buisson L. Reasoning on space with object-centred knowledge representations. In Buchmann A., GÇnther O., Smith T. R., and Wang Y. F., editors, *Lecture notes in computer science. Design and implementation of large spatial databases*, pages 325–344. Springer-Verlag, July 1989. First Symposium SSD'89. Santa Barbara, California, July 1989, Proceedings.

[18] Mourad Mechkour. A conceptual graphics model for information retrieval. In *Work Part 3 Deliverable: A Model for the Semantic Content of Multimedia Data*. ESPRIT BRA Project No. 8134 - FERMI, 1995.

[19] Mourad Mechkour. *EMIR2. Un Modèle étendu de représentation et de correspondance d'images pour la recherche d'informations. Application à un corpus d'images historiques.* PhD thesis, Université Joseph Fourier, 1995.

[20] Mourad Mechkour. A multifacet formal image model for information retrieval. In *MIRO final workshop, Glasgow, UK.*, page (to appear), 18-20 september 1995.

[21] Chang N. S. and FU K. S. Picture query languages for pictorial database systems. *IEEE COMPUTER*, November 1981.

[22] Caelli T. and Reye D. On the classification of image regions by colour, texture and shape. *Pattern Recognition*, 26(4):461–470, 1993.

[23] Kato T. and Kurita T. Visual interaction with electronic art gallery. In *DEXA 90. Database and expert system applications*, pages 234–240, 1990.

[24] Gong Y., Zhang H., Chuan H. C., and Sakauchi M. An image database system with content capturing and fast image indexing abilities. In *Proceedings of the international conference on multimedia computing and systems*, pages 121–130, May 14-19 1994.

[25] Takao Y., Itoh S., and Iisaka J. *An image-oriented database system*. Springer-Verlag, 1980.