Virtual Images for Similarity Retrieval in Image Databases

Gennaro Petraglia, Monica Sebillo, Maurizio Tucci, and Genoveffa Tortora, *Senior Member*, *IEEE*

Abstract—We introduce the virtual image, an iconic index suited for pictorial information access in a pictorial database, and a similarity retrieval approach based on virtual images to perform content-based retrieval. A virtual image represents the spatial information contained in a real image in explicit form by means of a set of spatial relations. This is useful to efficiently compute the similarity between a query and an image in the database. We also show that virtual images support real-world applications that require translation, reflection, and/or rotation invariance of image representation.

Index Terms—Image databases, query by image content, iconic indexing, spatial reasoning, similarity retrieval.

1 Introduction

Recently, important efforts in the design of image database systems have been stimulated by growing interest toward the image as a basic component of information interchange. A great number of applications based on image processing and management have been implemented and applied to office automation, computer-aided design, medical pictorial archiving, and several industrial domains.

In general, an Image Database System is a combination of three basic components. The first component deals with image processing for the extraction of information from physical images, the second one is responsible for the storage and management of the original images and the extracted information, and the third component concerns querying the database.

In traditional database systems, the use of indexing to allow database accessing has been well-established. Analogously, image indexing techniques have been studied during the last decade to support pictorial information retrieval from an image database. An image can be associated with two kinds of descriptors: information about its content (in textual form) and information related to the shape and to the spatial arrangement of its pictorial elements. To make an image database flexible, the spatial knowledge embedded in images should be preserved by the data structures used to store them. The use of icons as indexes is the basic issue of the iconic indexing methodologies developed to this aim [8], [45], [46].

Retrieving pictures that satisfy high level spatial queries is also an important issue of image database systems. An example of high-level pictorial information retrieval could be "find all pictures showing a tree to the left of a house." With the Query by Pictorial Example [7] approach, the index of an image is an iconic image itself, which represents the visual information contained in the image in a form suitable for different levels of abstraction and management [10], [12].

In this paper, we define a virtual description of an image, called *virtual image*. The virtual image of a real image consists of a set of objects and a set of binary spatial relations over the objects. Then, the spatial information embedded in a real image is represented in a relational format and the spatial relations among its objects are expressed explicitly. This capability is a fundamental requirement in many advanced applications where it is required that image objects and their implicit information are converted into a pictorial knowledge structure able to support spatial reasoning and flexible information retrieval. The interest in using virtual images as an iconic index is also motivated by the possibility of exploiting traditional database techniques for the solution of pictorial queries. Moreover, the virtual image is suited for defining a similarity measure between images for similarity retrieval that is flexible and easy to compute. The similarity degree between a query Q and a virtual image im_{V} takes into account the fact that pairs of objects of Q possibly have different spatial relations in im_{vi}, although Q and im_{vi} have similar visual contents. Based on such an index, we present a simple and efficient matching criterion for similarity retrieval.

Then, we show that virtual images are effective means for retrieval and browsing. As a matter of fact, a query can be simply specified by drawing an iconic sketch from which a virtual image is easily obtained. Given a query, the set of images satisfying the condition expressed in the query is selected by computing the similarity degree between each image and the query by means of their corresponding virtual images. Then, the query is solved by matching its virtual image against those of the images in the database. On the other hand, the virtual image associated with an image of the database can be easily transformed into an iconic sketch, which can be efficiently used for a visual interface in place of the real image for browsing.

- *The authors are with the Dipartimento di Matematica e Informatica, University of Salerno, Via S. Allende, I-84081 Baronissi, Salerno, Italy. E-mail: {petragli, monica, mautuc, jentor}@unisa.it.*

Manuscript received 28 May 1998; revised 1 Mar. 2000; accepted 13 Mar. 2000; posted to Digital Library 6 Apr. 2001.
For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org, and reference IEEECS Log Number 106914.

The virtual image index and the similarity degree function have been applied to the implementation of a prototype database system for diagnostic images that provides users with content-based retrieval [1], [2]. Experimental results performed on this system have shown that virtual images can be used profitably in terms of efficiency and effectiveness.

In real applications, it would also be important to enable the database system to solve queries of the type: "find the images in the database that present a given pattern even if it appears as in a mirror reflection." This capability is required in specific applications, such as in virtual reality or robotic scenes, where the observer viewpoint is continuously moved and the visual content representation of an image is object-centered.

However, an image cannot be retrieved unless the query is issued according to the storing direction because the conventional methodologies based on orthogonal projections [38] do not provide simple ways to recognize similarities between two indexes corresponding to an image and one of its possible manipulations, say rotation, translation, reflection, change of point of view, projection from 3D to 2D views, etc.

It is evident that the spatial relations of an image do not change if it is translated along one or both directions and it can also be observed that the spatial relations of an image reflected along a given direction can be easily obtained from those of the original one. In other words, there is an immediate correspondence between the virtual image of a translated or reflected real image and the virtual image of the original one. In this sense, the virtual image is independent of translation and reflection.

However, the indexing methods based on orthogonal projections are sensitive to rotations. In fact, each image is indexed according to a specific orientation, which should be taken into account in storing and accessing images in the database. When no implicit orientation can be assumed, as, for example, in the management of nuclear images, the iconic index should be able to capture the visual information contained in the images independently of any orientation. To obtain the rotation invariance, we introduce the *R-virtual image* iconic index, which is constructed on analogy with virtual images, by using a system of polar axes.

The R-virtual image describes the spatial relationship between each pair of objects in terms of conditions over the beginning points and the ending points of the objects. These conditions are the same as for virtual images, with the only difference being the use of a polar system. Then, one scanning direction corresponds to concentric circles, moving from the origin to outside (c-direction), and the other corresponding to a trace swiping clockwise around the origin (s-direction).

The paper is organized as follows: In Section 2, we give a brief overview of existing iconic indexing methods. The concept of virtual image is given in Section 3, together with the Atomic Relation Extraction Method (AREM). In Section 4, we introduce a similarity retrieval method based on the virtual images and give a formula to compute a similarity degree among images. An experimental application and corresponding experimental results are reported in Section 5. Finally, the reflection and rotation invariant indexing methodologies are provided in Sections 6 and 7, respectively. Some conclusions and further works are drawn in Section 8.

2 Iconic Indexing for Pictorial Retrieval in Image Databases

One of the most important issues in the design of Image Database Systems is a representation of image contents that allows for efficient storage and retrieval of pictorial data through a user-friendly interface. An image contains two types of information: information regarding its objects and information related to the shape and to the spatial arrangement of its pictorial elements. Then, to make an image database flexible, this spatial knowledge embedded in images should be preserved by the data structures used to store them.

The use of icons as indexes is the basic issue of the iconic indexing methodologies developed to this aim. Following the Query by Pictorial Example (QPE) [7] approach, the index of an image is an iconic image itself which represents the visual information contained in the image in a form suitable for different levels of abstraction and management [10], [12], [13]. In particular, the QPE philosophy expresses the objects and the spatial relations to be retrieved through a symbolic image which serves as a query and which is matched against the images in the database. Then, a query is an iconic image itself, represented by the same method used to describe an iconic index. As a result, a variety of approaches have been proposed which use objects and spatial relationships to describe the visual content of an image. Their main difference consists of what category of spatial relationships they handle. As a matter of fact, spatial relationships can be decomposed into three categories of relations, namely topological, directional, distance.

The application of the iconic indexing methodologies to the direction relationships has been widely discussed and, among several approaches, a great deal of interest has been devoted to the development of the QPE philosophy.

One of the most remarkable iconic indexing methodologies, proposed by Chang et al., is based on the theory of the Symbolic Projections, which models the distribution of image objects using orthogonal spatial relationships [9], [11], [46]. Such a formalism segments the plane with respect to a system of axes, called *symbolic projections*, and translates exact metric information into a qualitative form, thus allowing reasoning about the spatial relationships among objects in a 2D plane. According to their proposal, the initial step to obtain an iconic index from a real image is the recognition of its logical content in terms of the objects it contains. Each object of the real image is symbolically represented by an icon, which can have either a fixed size or the same size as the minimum bounding rectangle (MBR) of the object. Then, the iconic index for the image, called 2D string, is derived from the symbolic picture by projecting its objects along the x- and y-directions and applying the basic set of spatial operators.

Successively, the QPE approach and the basic set of spatial relations and operators have been further developed

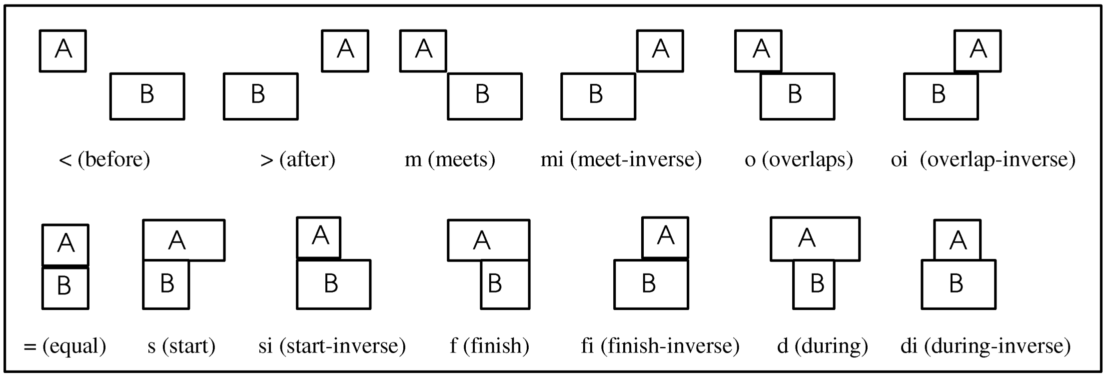


Fig. 1. The 13 types of spatial relations in one dimension (horizontal projection).

in several ways [27], [28], [31]. The common viewpoint of the several existing extensions of 2D strings is that any images can be described by symbolic projections, although the set of basic spatial relational operators can differ depending on efficiency and/or accuracy requirements. In particular, Allen's 13 types of relationships between two intervals in one dimension [3], shown in Fig. 1, are widely adopted to specify spatial relationships. If x- and y-direction are considered independently, a total of 169 spatial relationships arise between two objects in two dimensions through which the spatial content of an image can be suitably represented.

Along this line, Lee and Hsu proposed the basic set of spatial operators {"<," "|," "=," "[," "]," "%," "/"} and a set of algebraic laws which can be used both to identify the spatial relations between a small number of objects local to a given area and to guarantee global reasoning [31], [32], [33]. In Table 1, we recall the meaning of the operators proposed by Lee and Hsu through their definition in terms of the beginning and ending points of the objects involved.

Other extensions to the indexing methodology based on 2D strings have been studied in order to match different requirements, both in two dimensions and three or more dimensions, among them we recall [5], [13], [22], [23], [24], [25], [35]. In [46], Chang and Jungert summarized the fundamental results of the literature on the symbolic projections theory and its further extensions, while, in [14], the current indexing techniques by content are recalled.

TABLE 1
The Definition of the Characteristic Spatial Operators by Lee and Hsu

Notation	Meaning	Condition
A<B	A disjoins B	end(A)<begin(B)
A=B	A is the same as B	begin(A)=begin(B) end(A)=end(B)
A I B	A is edge to edge with B	end(A)=begin(B)
A%B	A and B have not the same bound and A contains B	begin(A)<begin(B) end(A)>end(B)
A(B	A and B have the same begin bound and A contains B	begin(A)=begin(B) end(A)>end(B)
A)B	A and B have the same end bound and A contains B	begin(A)<begin(B) end(A)=end(B)
A/B	A is partly overlapping with B	begin(A)<begin(B) < <end(A)<end(B)

3 The Virtual Image as an Iconic Index

As we have seen in the previous section, the 2D strings are compact iconic representations of images suitable for indexing an image database. However, it is difficult to define and efficiently perform types of matching between 2D strings that reflect the similarity between the corresponding images. This is due to the fact that the iconic indexes are influenced by the methodology used to obtain it, i.e., by the choice of the segmentation algorithm.

On the other hand, an iconic index should be able to explicitly describe the information contained in the image independently of any method used to extract it, thus providing a virtual description of the image. In [5], Chang and Lee introduced a data structure, called the Relative Coordinates Oriented Symbolic String (RCOS string), which does not require any segmentation algorithm because it describes each object by the x- and y-projection of its beginning and ending points. In particular, a symbolic picture is represented by the set of the objects contained in the real image and by a set of pairs of coordinates for those objects. Then, the spatial relation between any two objects of the image can be derived from the RCOS string according to the definition of spatial operators given in Table 1, thus obtaining one among 169 possible types of spatial relations in a two-dimensional space.

As an example, let us consider the image of Fig. 2 that shows the minimum enclosing rectangles corresponding to three symbolic objects. The corresponding RCOS string is:

$$[\mathrm{abc}, (1,1)(2,2), (3,3)(4,4), (5,2)(6,4)],$$

where the three objects ob_1, ob_2, and ob_3 in the image are represented by the list of symbols "abc" and each couple of pairs $(\mathrm{x}_{i1}, \mathrm{y}_{i1})(\mathrm{x}_{i2}, \mathrm{y}_{i2})$ denotes the relative coordinates of the left-bottom and right-up corners of the object ob_i.

This representation is independent of any cutting mechanism, but the spatial relationships are not explicitly

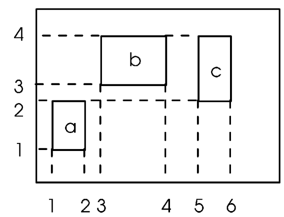


Fig. 2. The relative coordinates of the objects.

expressed and require an additional computational cost if they are needed in explicit form.

In this section, we introduce the virtual image iconic index. It is based on the theory of Symbolic Projections and describes a real image in terms of objects and binary spatial relations over the objects. The spatial information embedded in a real image is represented in a relational format and the spatial relations among the objects are expressed explicitly. Moreover, the virtual image allows us to define a similarity measure between images for similarity retrieval that is flexible and easy to compute.

Although, in general, we can assume that every kind of binary relation can be used, in the following, we restrict ourselves to the set of spatial relations defined in Table 1.

Definition 1. *Given a real image im, the* virtual image *P associated with im is a pair* $(\mathrm{Ob}, \mathrm{Rel})$, *where:*

- $\mathrm{Ob} = \{\mathrm{ob_1}, \dots, \mathrm{ob_n}\}$ *is a set of objects*
- $\mathrm{Rel} = (\mathrm{Rel_x}, \mathrm{Rel_y})$ *is a pair of sets of binary spatial relations over* Ob, *where* $\mathrm{Rel_x}$ *(resp.,* $\mathrm{Rel_y}$*) contains the mutually disjoint subsets of* $\mathrm{Ob} \times \mathrm{Ob}$ *that express the relationships "<," "|," "=," "[," "]," "/," "%" holding between pairs of objects of im along the x-projection (resp., y-projection).*

For simplicity, we use the notation $\mathrm{ob_i}\gamma\mathrm{ob_w}$ *to indicate that the pair* $(\mathrm{ob_i}, \mathrm{ob_w})$ *belongs to the relation* γ, *where* $\mathrm{ob_i}, \mathrm{ob_w} \in \mathrm{Ob}$ *and* $\gamma \in \{<, |, =, [,], /, \%\}$. *A triple like* $\mathrm{ob_i}\gamma\mathrm{ob_w}$ *is called an* atomic relation *in the following. We also say that the atomic relation* $\mathrm{ob_i}\gamma\mathrm{ob_w}$ belongs to $\mathrm{Rel_x}$ *(resp.,* $\mathrm{Rel_y}$*) if the spatial relation holding between* $\mathrm{ob_i}$ *and* $\mathrm{ob_w}$ *along the x-projection (resp., y-projection) is* γ. *Then, we can regard both* $\mathrm{Rel_x}$ *and* $\mathrm{Rel_y}$ *simply as sets of atomic relations.*

It is worth noting that the objects in the image are identified manually or automatically and labeled according to classes or names in a certain domain.

As an example, let us consider the picture in Fig. 3a and its symbolic picture corresponding to Fig. 3b; the corresponding virtual image is:

$$\mathrm{Ob} = \{\mathrm{A, B, C, D, E, F}\};$$
$$\mathrm{Rel_x} = \{\mathrm{A < B, A < C, A < F, A < E, A < D, B/C, B < F,}$$
$$\mathrm{B < E, B < D, C < F, C < D, C < E, F|D, F \% E, E < D}\}$$
$$\mathrm{Rel_y} = \{\mathrm{A/B, A < C, A < D, A < E, A < F, B < C, B < D,}$$
$$\mathrm{B < E, B < F, C < D, C < E, C < F, F < D, D \% E, F < E}\}.$$

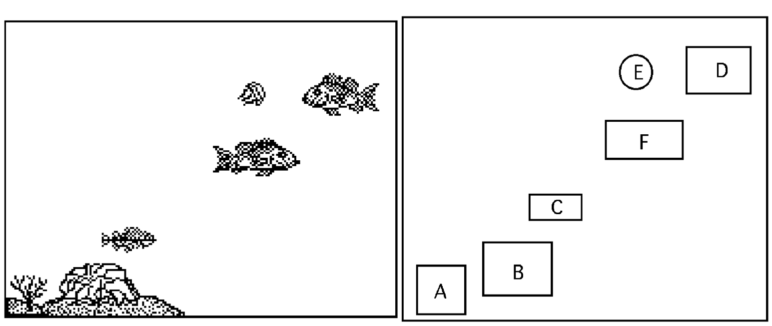
Then, it is easily seen that the virtual image of a real image can be obtained by applying the definition of the spatial operators given in Table 1 to each pair of objects in the image. Of course, this implies a computational cost that is quadratic in the number of objects. The following Atomic Relation Extraction Method (AREM) algorithm derives the virtual image from a given real image.

Algorithm AREM

Input: A real image *im*.
Output: A virtual image $(\mathrm{Ob}, (\mathrm{Rel_x}, \mathrm{Rel_y}))$.
Method: Let Ob be the set of the objects contained in *im*: perform the following steps to compute each of the sets $\mathrm{Rel_x}$ and $\mathrm{Rel_y}$:
Step 1: Let $\mathrm{Rel_x}$ (resp., $\mathrm{Rel_y}$) be the empty set;
Step 2: Scan the image *im* along the x-direction (resp., y-direction) to compute the values begin(A) and end(A) for every $\mathrm{A} \in \mathrm{Ob}$;
Step 3: For each pair of objects $\mathrm{A, B} \in \mathrm{Ob}$, add to $\mathrm{Rel_x}$ (resp., $\mathrm{Rel_y}$) the relation obtained by the following case-statement:

Case:

end(A) < begin(B)	:	A < B
end(B) < begin(A)	:	B < A
begin(A) = begin(B) and end(A) = end(B)	:	A = B
end(A) = begin(B)	:	A | B
end(B) = begin(A)	:	B | A
begin(A) < begin(B) and end(A) > end(B)	:	A % B
begin(B) < begin(A) and end(B) > end(A)	:	B % A
begin(A) = begin(B) and end(A) > end(B)	:	A [B
begin(A) = begin(B) and end(A) < end(B)	:	B [A
begin(A) < begin(B) and end(A) = end(B)	:	A] B
begin(A) > begin(B) and end(A) = end(B)	:	B] A
begin(A) < begin(B) < end(A) < end(B)	:	A / B



(a) (b)

Fig. 3. An image and its corresponding symbolic picture.

begin(B) < begin(A) < end(B) < end(A) : B / A
endcase

A virtual image describes its corresponding real image in terms of objects and spatial relationships and preserves the spatial knowledge embedded in the real image.

We briefly discuss storage space and access time efficiency of virtual image indexing. Storing the atomic relations derived from N objects takes $O(N^2)$ space, while other existing indexing methodologies only require a linear space. For example, the length of a 2D string is bounded by the number of subparts of objects and an RCOS string takes a space linear with the number of objects. However, a virtual image represents the spatial information embedded in a real image in explicit form, whereas a 2D string also contains implicit information. The difference is relevant for similarity retrieval purposes. In the next section, we present a simple matching criterion for similarity retrieval which exploits the explicit information stored in virtual images. To work on 2D strings, this method should be applied after a preprocessing phase to extract explicit spatial information from them [32], [33], thus affecting the time efficiency of retrieval.

On the other hand, some spatial operators defined in Table 1 satisfy several transitivity properties. In particular, the spatial relations {<, =, [,], %} are transitive. Therefore, if, for example, A < B and B < C, then it follows that A < C so that storing the latter relationship can be avoided. Moreover, some spatial relations can be inferred from one another. As an example, A | B and B | C imply A < C. It is thereby possible to reduce the number of atomic relations needed to express the spatial knowledge embedded in an image and store only a minimal set of atomic relations from which all the others can be obtained by transitivity.

By taking into account the transitivity of the spatial relation "<," the following more compact virtual image corresponds to the virtual image of Fig. 3:

$$\begin{aligned}
&\mathrm{Ob} = \{\mathrm{A, B, C, D, E, F}\}\\
&\mathrm{Rel_x} = \{\mathrm{E < D, F|D, A < C, A < B, B < F, B/C, C < F,}\\
&\qquad\quad \mathrm{C < E, B < E, F \,\%\, E}\}\\
&\mathrm{Rel_y} = \{\mathrm{A/B, A < C, B < C, F < D, C < F, D \,\%\, E, F < E}\}.
\end{aligned}$$

Of course, when a minimal set of atomic relations is stored, then an additional cost of retrieval must be added, as discussed before, in order to enhance the algorithm and generate inference. Then, the advantage of compactness would be balanced again by a higher cost of retrieval. A complete study on spatial relations transitivity and the methods to encode/decode them, and the relative costs, is out of the scope of this paper and will not be discussed further.

4 Similarity Retrieval by Virtual Images

Virtual images are a reasonably compact representation, in the sense discussed earlier, that can be used to index real images in a database. In the following, we show that virtual images are also efficient means for browsing and retrieval.

Intuitively, the virtual image associated with an image of the database can be easily transformed into an iconic sketch, which can be efficiently used for a visual interface in place of the real image for browsing. On the other hand, an end user can specify a query either as a real image (or a subpart of it) or as an iconic picture composed on the screen by means of a given set of graphic objects, taken from a palette of icons. Spatial operators are deduced from the relative positions of the objects themselves, following the Query by Pictorial Example philosophy. Then, the query is solved by matching its virtual image against those of the images in the database.

Regarding this task, the virtual image allows both an exact matching by comparing the sets of objects and atomic relations and an extension of the conventional retrieval methodologies. As a matter of fact, a typical requirement to an image database could be: "I would like to retrieve an image like this, even if it has fewer objects or the objects have a different spatial arrangement." In this scenario, the image used as a search parameter does not match the query, but it could be considered similar to it in a perceptual sense, i.e., there could be an object to be recognized through its model, or there could be a set of objects with a significant spatial arrangement in terms of after/before, top/down relationships [42].

This observation is the basic issue faced in [43], [44], where the authors' goal is both to present the psychological findings regarding similarity, which can be useful to computer scientists, and to discuss the issues connected to the specification and management of complex queries, where distinct similarity measures can be applied. Further investigations concerning similarity retrieval methodologies can be found in [6], [24], [33], [34], [35], [46].

In the following, we introduce a similarity retrieval methodology based on the virtual image which performs a comparison among images based on qualitative features.

The core problem for similarity retrieval is determining a similarity metrics that is efficient to compute and capture the essential aspects of similarity that humans use [35]. To this aim, we introduce a formula that measures the similarity degree between an image and a query in terms of objects and spatial relationships and expresses it as a value in the range $[0, 1]$. Moreover, we also introduce a minimum required similarity degree: An image satisfies a query if it has a similarity measure with the query that is greater than or equal to a minimum required threshold.

We first consider that our goal is to retrieve from a database those images that are similar to a given query, in the sense that they contain *almost* the same objects with similar spatial arrangement. Thus, our similarity measure has to take into account the similarities among spatial operators. As an example, the spatial operators | and < can be considered very similar since both A < B and A | B indicate that the object B is on the right of the object A. On the contrary, the *totally contain* operator % and the *disjoint* operator < are intuitively very different.

Then, we adopt the interval neighborhood graph as in [35], which formally defines the distances among spatial operators. According to the definition of the interval neighborhood graph given by Freksa in [19], two projection relationships are neighbors if they can be transformed into one another by continuously deforming (shortening,

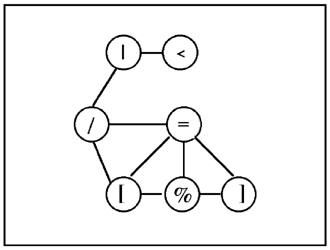


Fig. 4. Interval neighborhood graph.

lengthening moving) the projections. Fig. 4 shows a neighborhood graph for the projection relationships corresponding to the spatial operators in Table 1, where the edges represent the neighborhood relation [35]. The distance between two projection relationships, denoted by $distance(\gamma_1, \gamma_2)$, is defined as the length of the shortest path between their corresponding nodes in the graph.

Since the maximum distance on the interval neighborhood graph is 4 and the minimum one is 0, we define the similarity value $sim(\gamma_1, \gamma_2) = 1 - (distance(\gamma_1, \gamma_2)/4)$ for each pair (γ_1, γ_2) of spatial relations. Table 2 illustrates the values of similarity that are derived from the interval neighborhood graph and that we consider in our example. It should be understood that these values may be appropriately tuned in specific applications according to different requirements.

Now, we are ready to formally define a similarity retrieval method based on virtual images, where the similarity between two images represented by their virtual images is computed as a similarity degree value that depends on the similarity between their corresponding atomic relations. The result of a query is the set of images satisfying the condition expressed in it, i.e., the images that have a similarity degree with respect to the query greater than a specified threshold. Formally, the notions of *query* and *similarity degree* between images are described in the following definitions:

Definition 2. *A* query *Q is a 4-tuple* $(\mathrm{F}, \mathrm{G}, \mathrm{Rel}, t)$, *where* $(\{\mathrm{F} \cup \mathrm{G}\}, \mathrm{Rel})$ *is a virtual image and* $t \in [0, 1]$ *is a similarity degree threshold.*

In the definition above, F is the set of *mandatory* objects: An image *im* of the database satisfies the query Q only if its virtual image im_{vi} contains all the objects of F, with the same atomic relations as in Q. G is the set of *optional* objects:

TABLE 2
The Values of Similarity between Spatial Relations

SIM	<	=	|	/	%	[	]
<	1	0.25	0.75	0.5	0	0.25	0
=	0.25	1	0.5	0.75	0.75	0.75	0.75
|	0.75	0.5	1	0.75	0.25	0.5	0.25
/	0.5	0.75	0.75	1	0.5	0.75	0.5
%	0	0.75	0.25	0.5	1	0.75	0.75
[	0.25	0.75	0.5	0.75	0.75	1	0.5
]	0	0.75	0.25	0.5	0.75	0.5	1

Roughly speaking, the more objects of G are contained in im_{vi}, the higher the value of similarity degree between Q and *im* is. Finally, *t* is the minimum required similarity degree between Q and *im* in order for *im* to satisfy Q.

As a sample query, let us consider the image in Fig. 5, showing a query where the user has specified the objects in the set F (the fish and the leftmost sponge) and the ones in the set G (the octopus and the rightmost sponge) by means of a pointing device. Then, he/she specifies the minimum required similarity degree *t*, let us say t = 0, 6. Assigning symbolic names to these objects, as in Fig. 5b, the resulting query is $\mathrm{Q} = (\mathrm{F}, \mathrm{G}, \mathrm{Rel}, t)$, with

$$\begin{aligned} &\mathrm{F} = \{\mathrm{a}, \mathrm{c}\} \quad \mathrm{G} = \{\mathrm{a}, \mathrm{d}\} \\ &\mathrm{Rel_x} = \{\mathrm{c}]\mathrm{a}, \mathrm{c} < \mathrm{a}, \mathrm{a} < \mathrm{a}, \mathrm{c} < \mathrm{d}, \mathrm{a} < \mathrm{d}\} \\ &\mathrm{Rel_y} = \{\mathrm{a} \,\%\, \mathrm{a}, \mathrm{a}/\mathrm{d}, \mathrm{a}|\mathrm{d}, \mathrm{a} < \mathrm{c}, \mathrm{d} < \mathrm{c}\}, \end{aligned}$$

where the atomic relations have been derived by applying AREM.

The similarity degree between a query Q and a virtual image $\mathrm{im_{vi}}$ is computed by considering which portion of F and G is contained in im_{vi} and how similar spatial relations are found in im_{vi} with respect to those required in Q. In particular, such a formula takes into account that it is possible to derive different spatial relationships between the same pair of objects since the objects required in G could appear in im_{vi} more than once. Then, the highest value of the similarity degree is chosen. This task is performed by computing the sets X and Y according to the formula given below.

Definition 3. *Let* $\mathrm{Q} = (\mathrm{F}, \mathrm{G}, \mathrm{Rel}, t)$ *be a query with*

$$\begin{aligned} &\mathrm{F} = \{\mathrm{q_1}, \ldots, \mathrm{q_n}\}, \mathrm{G} = \{\mathrm{q_{n+1}}, \ldots, \mathrm{q_{n+m}}\}, \\ &\text{and } \mathrm{Rel} = (\mathrm{Rel_x}, \mathrm{Rel_y}) \end{aligned}$$

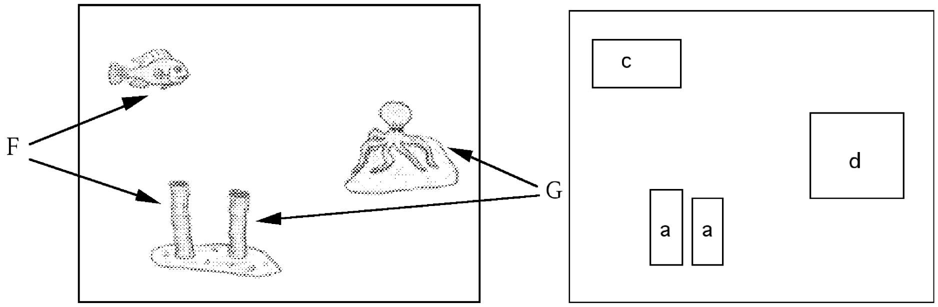
and let $im_{\mathrm{vi}} = (\mathrm{Ob}_{im}, \mathrm{Rel}_{im})$ *be the virtual image of an image im. By* $\mathrm{Rel_{F_x}}$ *(resp.,* $\mathrm{Rel_{F_y}}$*) we denote the subset of* $\mathrm{Rel_x}$ *(resp.,* $\mathrm{Rel_y}$*) consisting of all the atomic relations between pairs of objects in F along the x- (resp., y-) projection. Then, the sets X and Y consider the highest value of the similarity between each pair of atomic relations as follows:*

$$\begin{aligned} \mathrm{X} = \{&(\mathrm{q}\gamma\mathrm{q}', \mathrm{s} \mid \mathrm{q}\gamma\mathrm{q}' \in (\mathrm{Rel_X} - \mathrm{Rel_{F_X}}) \\ &\text{and } \mathrm{s} = \mathrm{Max}_{\mathrm{q}\gamma'\mathrm{q}' \in \mathrm{Rel_{im_X}}}\{\mathrm{sim}(\gamma, \gamma')\}\} \\ \mathrm{Y} = \{&(\mathrm{q}\gamma\mathrm{q}', \mathrm{s}) \mid \mathrm{q}\gamma\mathrm{q}' \in (\mathrm{Rel_y} - \mathrm{Rel_{F_X}}) \\ &\text{and } \mathrm{s} = \mathrm{Max}_{\mathrm{q}\gamma'\mathrm{q}' \in \mathrm{Rel_{im_y}}}\{\mathrm{sim}(\gamma, \gamma')\}\}. \end{aligned}$$

As an example, let us compute the sets X and Y for the query Q of Fig. 5 and the virtual image $P_{\mathrm{vi}} = (\mathrm{Ob_P}, \mathrm{Rel_P})$ corresponding to the image *P* of Fig. 6, where:

$$\begin{aligned} \mathrm{Ob_P} = \{&\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\} \text{ and a, b, c, and d represent the sponge,} \\ &\text{the jelly-fish, the fish and the octopus, respectively;} \\ \mathrm{Rel_{P_x}} = \{&\mathrm{c} \,\%\, \mathrm{a}, \mathrm{c}]\mathrm{a}, \mathrm{c} < \mathrm{d}, \mathrm{c} < \mathrm{b}, \mathrm{a} < \mathrm{a}, \mathrm{a} < \mathrm{b}, \mathrm{a} < \mathrm{d}, \mathrm{b} = \mathrm{d}\} \\ \mathrm{Rel_{P_y}} = \{&\mathrm{a}/\mathrm{d}, \mathrm{a} < \mathrm{c}, \mathrm{a} < \mathrm{b}, \mathrm{a} \,\%\, \mathrm{a}, \mathrm{d} < \mathrm{c}, \mathrm{d} < \mathrm{b}, \mathrm{c}/\mathrm{b}\}. \end{aligned}$$

According to Definition 3, the $\mathrm{Rel_{F_x}}$, $\mathrm{Rel_{F_y}}$, X, and Y sets are:



(a) (b)

Fig. 5. A query specification and its corresponding symbolic picture.

$$\mathrm{Rel}_{\mathrm{F_x}} = \{\mathrm{c}]\mathrm{a}\} \qquad \mathrm{Rel}_{\mathrm{F_y}} = \{\mathrm{a} < \mathrm{c}\}$$
$$\mathrm{X} = \{(\mathrm{a} < \mathrm{a}, 1), (\mathrm{c} < \mathrm{a}, 0), (\mathrm{c} < \mathrm{d}, 1), (\mathrm{a} < \mathrm{d}, 1)\}$$
$$\mathrm{Y} = \{(\mathrm{a} \,\%\, \mathrm{a}, 1), (\mathrm{a}|\mathrm{d}, 0.75), (\mathrm{d} < \mathrm{c}, 1), (\mathrm{a}/\mathrm{d}, 1)\}.$$

For example, the value associated with the atomic relation (a | d) along the y-direction is 0.75 since the image contains (a/d) and Table 2 assigns sim(|, /) = 0.75.

Once the most similar atomic relations have been selected, it is necessary to derive the similarity degree between the query and the image in order to decide if the image belongs to the answer set.

Generally speaking, our own experience is that the user discriminates the images according to some basic features: the objects, their shape, their color, and their spatial arrangement. On the other hand, the accuracy of an image retrieval algorithm has to be based on both how successful the system is in satisfying the user expectations and how independent of the human interpretation the retrieval method is. With this observation in mind, we propose a similarity retrieval criterion based on a formula which evaluates the similarity between an image and a query by comparing the objects they share and their corresponding atomic relations. Such a formula returns a value in the range $[0, 1]$ by computing the ratio between the amount of objects retrieved together with the similarity values of their atomic relations and the amount of objects required by the user in terms of cardinality of F and G, say | F | and | G |, and their atomic relations. When the candidate image does not satisfy the user's requirements in terms of mandatory objects and minimum similarity threshold, it is discarded; otherwise, it belongs to the answer set with the corresponding similarity degree. Formally, the similarity degree is derived according to the following definition:

Definition 4. *Let* Q *and* $\mathrm{im_{vi}}$ *be defined as in Definition 3, then the* similarity degree *between* Q *and* im_{vi}*, denoted by* $Sim\_deg(\mathrm{Q}, im_{\mathrm{vi}})$*, is defined by the formula:*

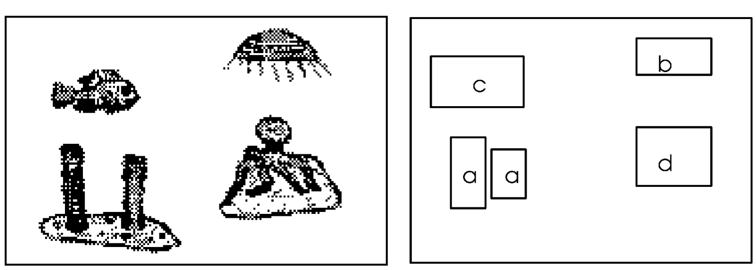
$$Sim\_\mathrm{deg}(\mathrm{Q}, \mathrm{im_{vi}}) = \begin{cases} \text{if } (\mathrm{F} \subseteq \mathrm{Ob_{im}}) \text{ and } (\mathrm{Rel_{F_x}} \subseteq \mathrm{Rel_{im_x}}) \\ \text{and } (\mathrm{Rel_{F_y}} \subseteq \mathrm{Rel_{im_y}}) \text{ then} \\ \dfrac{|\mathrm{F}|+|\mathrm{Rel_{F_x}}|+|\mathrm{Rel_{F_y}}|+|\mathrm{G}\cap\mathrm{Ob_{im}}|+\sum_{(\mathrm{q}\gamma\mathrm{q}',\mathrm{s})\in\mathrm{X}}\mathrm{s}+\sum_{(\mathrm{q}\gamma\mathrm{q}',\mathrm{s})\in\mathrm{Y}}\mathrm{s}}{|\mathrm{F}|+|\mathrm{G}|+|\mathrm{Rel_x}|+|\mathrm{Rel_y}|\sum_{(\mathrm{q}\gamma\mathrm{q}',\mathrm{s})\in\mathrm{X}}} \\ \text{otherwise} \qquad 0. \end{cases}$$

As an example, let us compute the similarity degree between the query of Fig. 5 and the virtual image corresponding to the image of Fig. 6. We obtain:

$$Sim\_deg(\mathrm{Q}, P\mathrm{vi}) = \frac{2+1+1+2+3+3.75}{2+2+5+5} = \frac{12.75}{14} = 0.91.$$

Since $Sim\_deg(\mathrm{Q}, P\mathrm{vi}) \geq t$ (the minimum required similarity degree), the image in Fig. 6 satisfies the user's requirements and belongs to the set of the images retrieved.

The Sim_deg function has been effectively used as a metrics for similarity of images in the implementation of a prototype database system for diagnostic images that provides users with content-based retrieval [1], [2]. Experimental results obtained by applying this similarity measure to the field of medicine have shown that virtual images can be used profitably in terms of efficiency and effectiveness of the performances.



(a) (b)

Fig. 6. An image P similar to the query of Fig. 5 and its symbolic image.

5 An Experimental Image Information Retrieval Application

Recently, there has been much interest in various kinds of database management systems for nontextual information such as multimedia, spatial, object-oriented, and image databases [20]. In image databases, special effort has been devoted to efficient retrieval of iconic information by contents [16], [18], [21], [39], [40]. In particular, in [16], an indexing methodology is proposed which uses the image content, such as color, texture, and shape, as the basis of the queries. This methodology is part of the Query by Image Content (QBIC™) project [17], whose techniques are used in some commercial products (IBM's Multimedia Manager, IBM's Digital Library, and the DB2 series of products). The same technology for content-based retrieval has been adopted for the commercial product ILLUSTRA™.

5.1 The IME Image Management Environment

The goal of this section is to provide readers with tools to evaluate the efficiency and effectiveness of the virtual image and the similarity degree function. To this aim we illustrate the Image Management Environment (IME), an experimental system for the management of images with content-based access system [1], [2]. It supports any application domain where the spatial relationships among objects are the primary keys to index the images. The current prototype implementation of the IME system deals with medical images and the set of the objects that images contain is determined by this application domain. To achieve user-friendliness, it incorporates a visual interface that merges the query-by-pictorial-example and query-by-sketch paradigms. The user interface was conceived as a typical Windows application. A palette includes a set of anatomic objects—e.g., right lung, left lung, spine, aorta, nodule, etc.—for each specific human body district. Each object in the palette is taken from a standard CT scan and is represented as an icon by means of its contour.

In imaging diagnosis, segmentation and edge detection are very helpful for physicians [15], [26] since some features, such as spatial locations, opacities, shapes, and geometrical measures, can be automatically produced. In order to obtain such features, many segmentation and edge detection techniques are available [15], [26]. In contrast, several applications require manual annotation or at least an expert control over the automatic feature extraction. For example, in clinical radiology, a physician is faced with the task of determining the presence or absence of disorders in images from different modalities, such as CT scan, when the correct answer is unknown. In particular, for a given image, a physician has to identify abnormalities, consider their spatial relationships with certain organs, and evaluate their morphological and geometrical features like opacity, shape, symmetry, roundness, area, etc. In other words, a physician has to evaluate the semantic contents of the image based on his/her personal knowledge to formulate a diagnosis and a treatment plan for the patient. The personal knowledge of a physician is typically based on remembering similar features of previously examined patients. In order to formulate a diagnosis, the physician has to perform three complex operations:

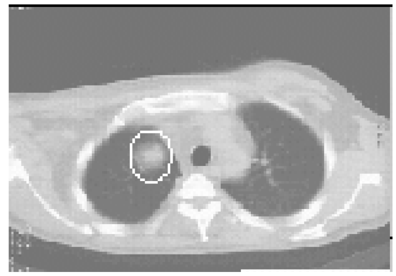


Fig. 7. A lung CT scan: The circle focuses on a suspect cancer.

1. Identify the hot spots;
2. Retrieve related images by similarity from his/her mnemonic data base (experience) or from a very large archive of analogic films;
3. Formulate a diagnosis.

Therefore, it would be helpful for him/her to have more support in accessing images and related data of patients with similar abnormalities. As an example, consider Fig. 7 in which a lung CT scan is illustrated. The circle focuses on a suspect cancer. Such a pathology was diagnosed by the radiologist, considering its morphological and geometrical features and its position with respect to other objects, e.g., lungs, spine, etc.

Unlike alphanumeric text, medical images usually require a long time to be accessed, processed, stored, retrieved, and distributed because of their size and internal complexity. Thus, an electronic database should meet the requirement of effectively handling the above time-consuming operations [28]. The image processing tools integrated in the IME system allow hot spots and canonical objects extraction by means of an entropy-based method for segmentation and edge detection [15]. Applying such a method to the lung CT scan of Fig. 7, the image of Fig. 8 is produced, which includes some canonical organs (left and right lungs, aorta, and thorax) and a hot spot, pointed to by an arrow.

Before storing the image in the database, it must be processed to extract the key features to be used for indexing. In this phase, the visual interaction with the physician is particularly useful to assign the correct meaning to the patterns in the image and to select the significant ones. Contrarily, the physician can invoke several processing tools in order to outline the features of interest in the original image. Fig. 9 shows a typical IME window where the CT scan of Fig. 7, namely the ANC10.BMP subwindow, is loaded in the left-up corner.

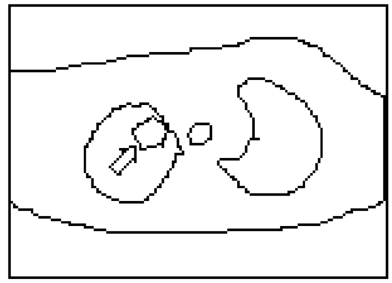


Fig. 8. Edge detection of the lung CT scan illustrated in Fig. 7.

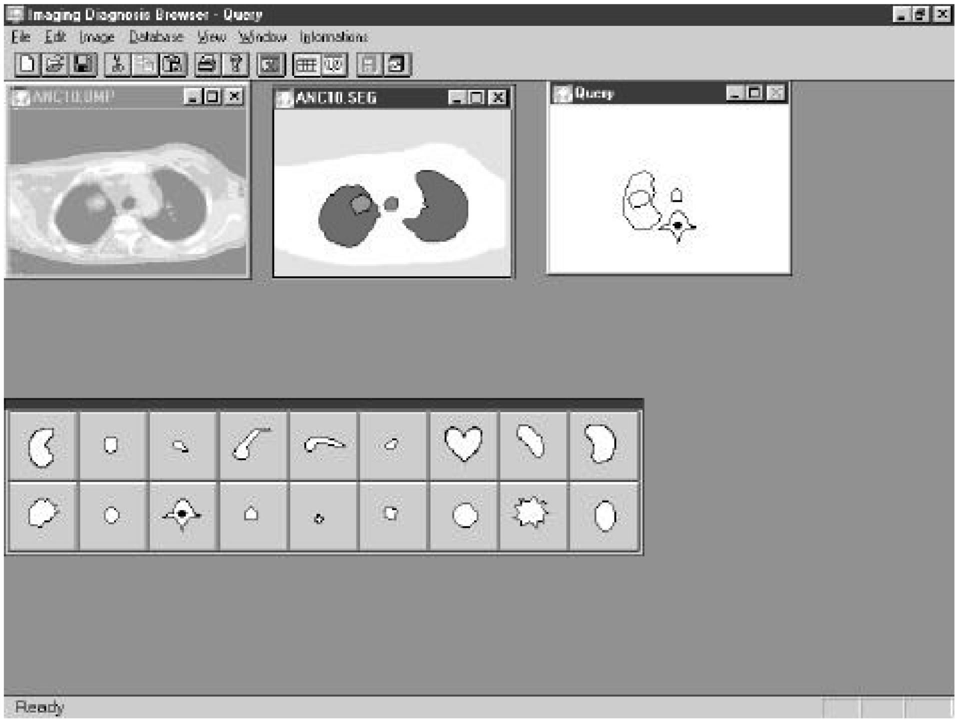


Fig. 9. A query by example.

The results of its segmentation is shown in the ANC10.SEG subwindow. Usually, the physician is supposed to choose a subset of the edges to be the representative ones. He/she has to add more information to each of them, such as which organs they represent, whether they are normal or not, which abnormalities he/she has detected, and so on. Besides these descriptors, a virtual representation of the image is automatically computed to describe the spatial contents of the real image and to allow content-based similarity retrieval. All this collected information is, in fact, the index of the image. Fig. 9 also shows an IME subwindow, namely the Query window, where a user is composing a query starting from a segmented image and using the iconic palette of the environment. The latter contains some thorax CT scan iconic features, such as spine, left and right lung, etc. As previously stated, a record of alphanumeric descriptors is automatically extracted from the image, such as spatial relationships, morphological and geometrical measures of each hot spot. When considering a diagnostic image, a physician may want to retrieve some similar images from the database and compare them to the first one in order to be supported in formulating his/her diagnosis. Therefore, he/she first asks for edge detection from the image and then selects the objects by which he/she wants to formulate a query. In the example, the query consists of spine and right lung with a star-shaped hot spot inside the aorta. All the images in the database with a similar abnormality in approximately the same position will be retrieved.

Fig. 10 shows the IME window that displays the results of the previous query, that is, a set of retrieved images, and their associated similarity degrees with respect to the query. As expected, each image has a similar abnormality in approximately the same location as in the image of Fig. 7 so that the similarity degrees are all greater than 90 percent. It is worth noting that no false alarms were retrieved, i.e., images that do not show any abnormalities or images with abnormalities in different positions (e.g., within the left lung). Besides comparing their contents with those of the original one in order to formulate a diagnosis, the retrieved images can, in turn, be used to formulate new queries.

A complete description about the morphological—geometrical measures and the k-d tree search used in IME to manage the index for images can be found in [2].

5.2 Experimental Results

This section recalls experimental results and performance evaluation from [2] obtained by testing the IME system on a database with a large number of lung CT scan images chosen from digital archive of the Institute of Radiology of the Second University of Naples. It should be noted that such a set of images is relative to the same human body district, namely the lung district, and then all contain approximately the same objects in the same positions. The only relevant difference between two images is related to the presence/absence of a similar abnormality, owing to several possible lung pathologies.

Frequently used evaluation criteria to measure effectiveness of retrieval system are the recall and the precision [41]:

- The recall is the capability of system to retrieve all relevant images;
- The precision is the capability of system to retrieve only relevant images.

In order to assess the performances of IME, an extension of the former, namely the Normalized Recall (NR) [41], has been chosen. In particular, NR reflects how close is the set of

Fig. 10. The set of images resulting from the query of Fig. 9.

the retrieved images to an ideal retrieval in which the most REL relevant images appear in the first REL positions.

Formally, relevant images are ranked 1, 2, ..., REL, where REL is the number of relevant images and Ideal Rank (IR) is given by

$$\frac{\sum_{r=1}^{REL} r}{REL}.$$

Now, let

$$\frac{\sum_{r=1}^{REL} Rank_r}{REL}$$

be the Average Rank (AR) over the set of relevant images retrieved by the system, where $\mathrm{Rank_i}$ represents the rank of relevant images. The difference between AR and IR, given by $\mathrm{AR} - \mathrm{IR}$, represents a measure of the effectiveness of the system. This difference can range from 0, for the perfect retrieval $(\mathrm{AR} = \mathrm{IR})$, to $(\mathrm{TOT} - \mathrm{REL})$, for the worst case, where TOT is the number of images in the collection. Hence, the above expression can be normalized by dividing it by $(\mathrm{TOT} - \mathrm{REL})$ and then by subtracting the result from 1. Finally, the Normalized Recall (NR) is

$$1 - \frac{AR - IR}{(TOT - REL)}.$$

This measure ranges from 1 for the best case (perfect retrieval) to 0 for the worst case.

In order to evaluate the performances of IME with respect to human perceptual similarity, a sample of 15 heterogeneous images has been selected to be used as queries. For each of them, a manual selection of the 25 most similar images has been first made from the database, that is, a set of 25 images containing a similar abnormality in the same position as in the query. Then, they have been contrasted with the corresponding images automatically retrieved. For each query, the corresponding NR has been computed. In [2], the authors run experiments on a set of 2,000 images, selected from an analogic medical archive, which had already been classified with respect to the different pathologies. The results are shown in Table 3 and demonstrate the effectiveness of IME. The value of NR is averaged over the 15 test queries.

In this test, the resulting NR value is equal to 0.975, which is very close to 1. This reflects the fact that the ranks of relevant images deviate very little in average from the ideal case.

In order to achieve faster than sequential search, IME uses k-d-trees as spatial access structures. To show how efficient this choice is in terms of computing time, in the following, the performance gains it offers with respect to sequential scanning are illustrated.

The IME system has been tested considering an IDB of 20,000 images and 50 typical queries randomly repeated with different values of the parameters *t* and ε. Fig. 11 plots

TABLE 3
Effectiveness of IME Evaluated by Normalized Recall

Measure	Value
Size of idb	2000
Number of queries	15
Normalized Recall	0,975

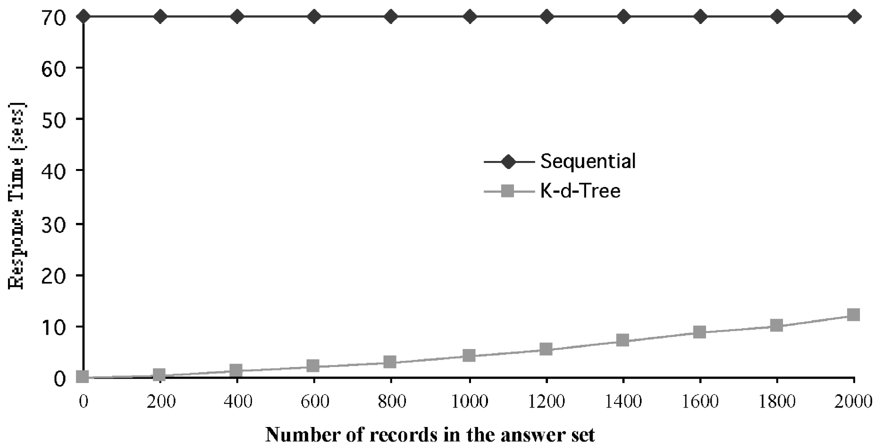


Fig. 11. The response time as a function of the retrieved images set size.

the average response times of sequential and k-d-tree searches as a function of the size of the retrieved set of images. Notice that the size of the retrieved set of images grows as t decreases and ε increases. Fig. 11 shows an evident speed up of k-d-tree search over sequential search. Moreover, the k-d-tree search never requires more than 20 seconds.

Another test considers the response time as a function of the IDB size. In this experiment the authors set $t = 0.8$ and issued each query with four different values for ε. Fig. 12 plots the average retrieval response times compared to sequential search. Notice the k-d-tree achieves considerably better computing time than sequential search and that the performance gap widens as the IDB grows.

6 Transformation Invariance of Iconic Indexes

In Section 3, we have introduced the virtual image and discussed its usefulness as an iconic index to perform a flexible retrieval in the Image Databases. However, in real applications, it would also be important to enable the database system to be able to find the images in the database that present a given pattern, even if it appears reflected.

This capability of recognizing the similarity between an image and its mirror reflection is important in specific applications, such as in virtual reality or robotic scenes, where the observer viewpoint is continuously moved and the visual content representation of an image is object-centered. In such cases, the basic requirement is both to maintain a unique index derived with respect to the left-bottom viewpoint and to be able to derive the indexes associated either with different subpictures or with different viewpoints, i.e., to derive the atomic relations according to different viewpoints.

As an example, let us suppose that O' is the observer viewpoint in Fig. 13. Then, the image can be ideally divided into four subpictures, P1, P2, P3, and P4, which will have different scanning directions, from the origin O' to outside.

However, this image cannot be retrieved unless the query is issued according to the direction through which it has been stored because the conventional indexing methods based on orthogonal projections do not provide simple ways to recognize similarity between two indexes corresponding to an image and one of its possible manipulations, say translation, reflection, change of point of view, etc.

Then, to deal with this issue, let us observe that the spatial operators in Table 1 are defined in terms of conditions over their borders (beginning point and ending

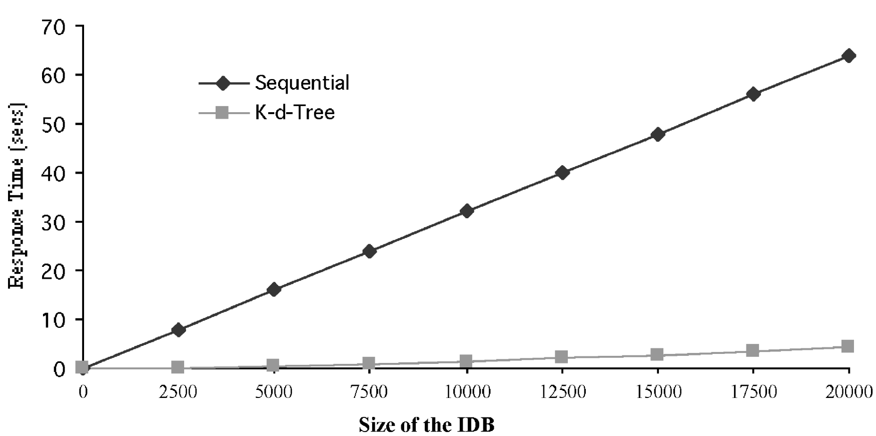


Fig. 12. The average retrieval response time as a function of the IDB.

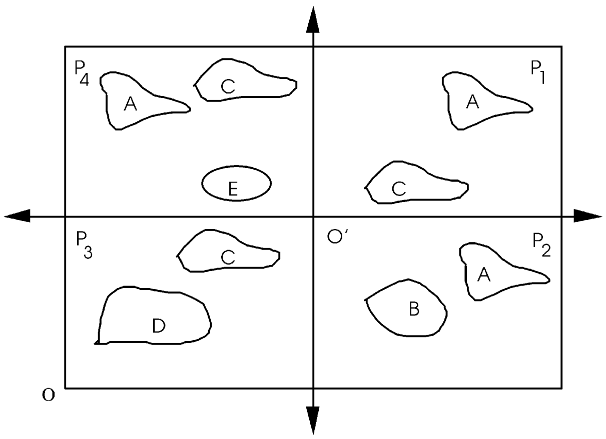


Fig. 13. An image P with observer viewpoint O'.

point); when reversing the scanning direction, the beginning point of an object is exchanged with its ending point. For this reason, when evaluating the usual iconic indexes of a reversed image, the algorithm for the string representation would determine different relationships, and then different indexes. On the contrary, AREM derives the atomic relations in terms of relations between couple of objects and, so, they are simple to transform because this knowledge is directly used as an iconic index. Moreover, the atomic relations of a picture reversed along a given direction can be easily obtained from those of the original picture. Table 4, called TX-6 transformation law, summarizes the correspondences existing between the atomic relations of a picture *f* along the x-direction (resp., y-direction) and the corresponding atomic relations in a picture *g* obtained by reversing *f* along the y-axis (resp., x-axis).

Then, in order to verify whether a picture *f* is obtained by reversing a picture *g*, it is sufficient to compare the virtual image of *f*, say f_{vi}, to the one obtained by transforming the virtual image of *g*, say g_{vi} via TX-6, as asserted by the following theorem.

Theorem 1. *Let f and g be two pictures such that g is obtained by reversing f along the x-axis (resp., y-axis) and let* $\mathrm{f}_{vi} = (\mathrm{Ob_f}, \mathrm{Rel_f})$ *with* $\mathrm{Rel_f} = (\mathrm{Rel_{f_x}}, \mathrm{Rel_{f_y}})$ *and* $\mathrm{g_{vi}} = (\mathrm{Ob_g}, \mathrm{Rel_g})$ *with* $\mathrm{Rel_g} = (\mathrm{Rel_{g_x}}, \mathrm{Rel_{g_y}})$ *be their corresponding virtual images. Then, the virtual image* g_{vi} *is given by*

$$\mathrm{g}_{vi} = (\mathrm{Ob_g}, \mathrm{Rel_g}) \text{ with } \mathrm{Rel_g} = (\mathrm{Rel_{f_x}}, \mathrm{TX\text{-}6}(\mathrm{Rel_{f_y}}))$$
$$(\text{resp., } \mathrm{g}_{vi} = (\mathrm{Ob_g}, \mathrm{Rel_g}) \text{ with } \mathrm{Rel_g} = (\mathrm{TX\text{-}6}(\mathrm{Rel_{f_x}}), \mathrm{Rel_{f_y}})).$$

Proof. Let us observe that, when reversing a picture along the x-axis, the beginning point and the ending point of any object in the y-direction are mutually exchanged; then, the thesis follows from Table 4. □

TABLE 4
The TX-6 Transformation Law

atomic relation in f	atomic relation in g
A<B	B<A
A=B	B=A
A|B	B|A
A%B	A%B
A[B	A]B
A]B	A[B
A/B	B/A

Then, given the virtual image P of the picture in Fig. 13, the virtual images of the subpictures P_1, P_2, P_3, and P_4 can be obtained by proper applications of the TX-6 transformation law and vice versa.

For example, let us consider the subpicture P_4 whose virtual image with respect to the O' viewpoint is $(\mathrm{Ob}, \mathrm{Rel})$, where

$$\begin{aligned}
&\mathrm{Ob} = \{\mathrm{A}, \mathrm{C}, \mathrm{E}\};\\
&\mathrm{Rel}' = (\mathrm{Rel}'_{\mathrm{x}}, \mathrm{Rel}'_{\mathrm{y}}) \text{ with}\\
&\mathrm{Rel}'_{\mathrm{x}} = \{\mathrm{C}|\mathrm{A}, \mathrm{E} < \mathrm{A}, \mathrm{C}\ \%\ \mathrm{E}\}\\
&\mathrm{Rel}'_{\mathrm{y}} = \{\mathrm{A}/\mathrm{C}, \mathrm{E} < \mathrm{A}, \mathrm{E} < \mathrm{C}\}.
\end{aligned}$$

When moving the viewpoint from O' toward O, the virtual image corresponding to the subpicture P_4 can be obtained by applying the TX-6 transformation law to the $\mathrm{Rel_x}$ set as follows:

$$\begin{aligned}
&\mathrm{Ob} = \{\mathrm{A}, \mathrm{C}, \mathrm{E}\};\\
&\mathrm{Rel} = (\mathrm{Rel_x}, \mathrm{Rel_y}) \text{ with}\\
&\mathrm{Rel_x} = \mathrm{TX\text{-}6}\ (\mathrm{Rel}'_{\mathrm{x}}) = \mathrm{TX\text{-}6}\ (\{\mathrm{C}|\mathrm{A}, \mathrm{E} < \mathrm{A}, \mathrm{C}\ \%\ \mathrm{E}\})\\
&\qquad = \{\mathrm{A}|\mathrm{C}, \mathrm{A} < \mathrm{E}, \mathrm{C}\ \%\ \mathrm{E}\}\\
&\mathrm{Rel_y} = \{\mathrm{A}/\mathrm{C}, \mathrm{E} < \mathrm{A}, \mathrm{E} < \mathrm{C}\}.
\end{aligned}$$

Analogously, it is possible to derive the virtual image P of the picture in Fig. 13 by composing the virtual images of the subpictures P_1, P_2, P_3, and P_4 after proper application of the TX-6 transformation law.

Therefore, it is possible to maintain the virtual image as the unique index associated with an image, being able to navigate in the image and derive information about the relative positions of the objects by inspecting the corresponding virtual image.

7 ROTATION INVARIANT ICONIC INDEXING

Since each image is indexed according to a specific orientation, the virtual image index defined in the previous sections is suitable for all the database applications where images have an implicit orientation, which has to be taken into account when accessing the database. On the other hand, there are real-world applications in which no implicit orientation can be assumed, as, for example, the management of nuclear images, in which objects and their spatial relations do not refer to any implicit scanning direction. In these cases, a good iconic index should be able to capture the visual information contained in the images independently of any orientation, i.e., in these cases, the index should be rotation invariant.

It is well-known that the iconic indexing methodologies based on topological and distance relationships are rotation invariant; on the contrary, the iconic indexes based on

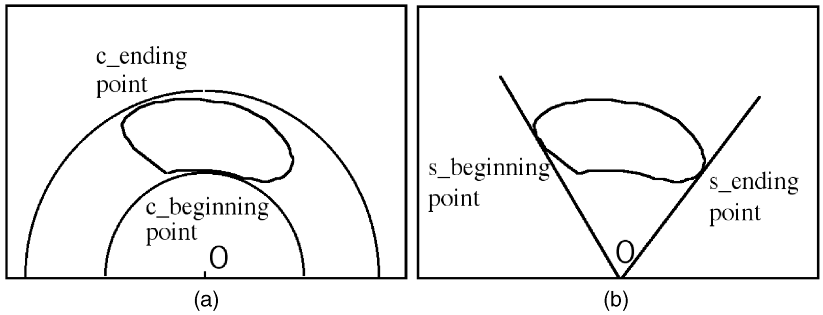


Fig. 14. The beginning points and the ending points along the c- and s-direction.

orthogonal projections are sensitive to rotations. However, if direction relationships are also based on object viewpoint, rather than viewer-based, then even directional relationships are rotation invariant. This scenario is typical of virtual reality scenes, where all the objects move together with the ones which represents the viewpoint of the whole scene. In this case, a rotated image is indexed according to a viewpoint which moves with the image itself, then spatial relations don't change. Differently, in real-world database applications, the rotation invariance is a basic issue because each image is captured and stored in agreement with a viewpoint which is implicitly dependent on an outside viewer who establishes a fixed scanning direction. Any translation of this viewpoint and/or any rotation of the image affects the direction relations between each pair of objects and, therefore, the whole index.

In the recent literature, several approaches can be found whose aim is to provide a solution to the rotation invariance of the conventional indexing methodologies based on symbolic projections [29], [30].

The aim of this section is to introduce an iconic indexing methodology which guarantees the rotation invariance of the image visual content, where it is described by direction relationships, which are viewer-based.

To obtain the rotation invariance, we introduce the *R-virtual image* iconic index, which is similar to the virtual image, but is computed with respect to a system of polar axes and inherits important properties from orthogonal indexes. Preliminary results can be found in [36], [37], where the 2R-string was presented as the rotation invariant version of the 2D C-string.

Generally speaking, the R-virtual image describes the spatial relationships between each pair of objects in terms of conditions over the beginning points and the ending points of the objects. These conditions are the same as for the virtual image, with the only difference being that they refer to a system of polar axes, where the first scanning direction is determined by concentric circles, moving from the origin to outside (c-direction), and the second direction corresponds to a trace sweeping clockwise around the origin (s-direction).

Without loss of generality, we choose the centroid, that is, the closest object to the center of the image, as the origin of the polar axes system. This object is called *the observer* and its centroid is called *the rotation center*. Let us notice that the atomic relations between the observer and each of the remaining objects in the image are unnecessary to describe the spatial information related to the observer since this information is implicitly inferred from the fact that the observer is in the origin of the polar system. Therefore, an R-virtual image consists of a set of objects, including the observer, and a set of atomic relations over the other objects.

Fig. 14 shows the beginning and ending points of an object with respect to the rotation center O; they have been recognized along the c- and s-direction as previously discussed. Table 5 represents the definition of the R-virtual image operators; in particular, the R-virtual image operators and the virtual image operators use the same notations, but their meaning refers to the polar system/Cartesian system, respectively.

As a sample, the *disjoint* relation "<" between A and B is satisfied along the c-direction, i.e., $A < B$ holds, if the ending point of A is encountered before the beginning point of B when moving from the rotation center to outside, as stated in Table 5.

Let us note that the AREM method is independent of the coordinate system, thus the algorithm introduced in Section 3 works also for building the R-virtual image.

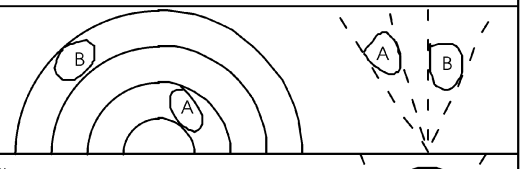
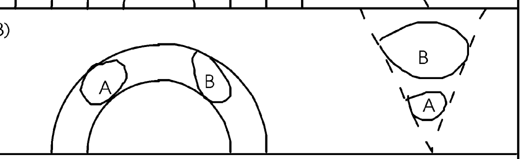
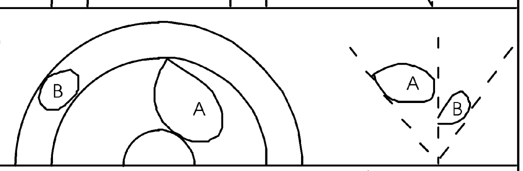
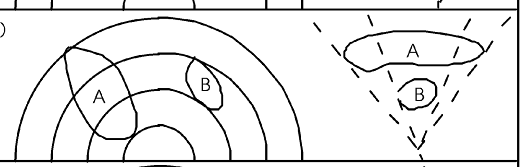
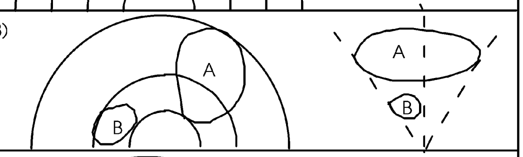
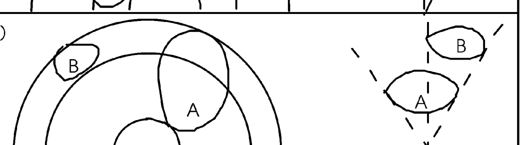
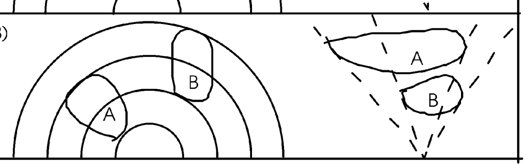
Then, in order to derive the R-virtual image from a symbolic image, we can apply the AREM algorithm since it is independent of the choice of the coordinate system; the algorithm introduced in Section 3 works also for building the R-virtual image.

However, let us make some considerations regarding the extraction of the atomic relations along the s-direction. The spatial relation between two disjoint objects along the s-direction depends on the choice of the initial position of the trace. As an example, in Fig. 15b, the initial position c_1 determines the relation $A < B$, while the choice of c_2 as the initial position determines the relation $B < A$.

Moreover, it may happen that, starting from an initial position, the ending point of an object is encountered before its beginning point, as for the object B of Fig. 16b, if the initial position c_1 is assumed. In both examples, the atomic relation $A < B$ alone is not sufficient to describe the spatial relationship between A and B. In fact, the same atomic relation would result from the two different cases shown in Fig. 15b and in Fig. 16b.

To solve these ambiguities, we express the rotation-invariant spatial relationship between a pair of objects A and B by a couple of atomic relations along the s-direction, namely the two atomic relations derived by choosing the beginning point of A and B, respectively, as the initial position. Hence, the R-virtual image corresponding to the picture in Fig. 15 is $(\mathrm{Ob}, (\mathrm{Rel}_c, \mathrm{Rel}_s))$, where:

TABLE 5
The Definition of the 2-R String Spatial Operators

Notation	Condition	Meaning (c-direction, s-direction)
A < B	end(A) < begin(B)	
A = B	begin(A) = begin(B) end(A) = end(B)	
A | B	end(A) = begin(B)	
A % B	begin(A) < begin(B) end(A) > end(B)	
A (B	begin(A) = begin(B) end(A) > end(B)	
A) B	begin(A) < begin(B) end(A) = end(B)	
A / B	begin(A) < begin(B) < end(A) < end(B)	

$$\mathrm{Ob} = \{\mathrm{A}, \mathrm{B}\}$$
$$\mathrm{Rel_c} = \{\mathrm{A}/\mathrm{B}\}$$
$$\mathrm{Rel_s} = \{\mathrm{A} < \mathrm{B}, \mathrm{B} < \mathrm{A}\};$$

whereas the R-virtual image of the picture in Fig. 16 is represented by $(\mathrm{Ob}, (\mathrm{Rel_c}, \mathrm{Rel_s}))$, where:

$$\mathrm{Ob} = \{\mathrm{A}, \mathrm{B}\}$$
$$\mathrm{Rel_c} = \{\mathrm{A}/\mathrm{B}\}$$
$$\mathrm{Rel_s} = \{\mathrm{A} < \mathrm{B}, \mathrm{B}/\mathrm{A}\}.$$

The latter pair of atomic relations, A < B and B/A, corresponds to starting the clockwise scanning of the image *both* from the initial position c_1 *and* from c_2.

Then, the R-virtual image guarantees the rotation invariance with respect to an object chosen as rotation center.

Moreover, the R-virtual image exploits the same spatial operators as the virtual image, by adapting their meaning to the polar system. Therefore, all the transformation laws [31] hold for it, too, since they do not depend on the fact that an orthogonal or a polar system is used. In particular, the reflection invariance is guaranteed since it is possible to extend the transformation law TX-6 to the polar index to recognize the similarity between an image and its reflection when they are represented by the R-virtual images: When an image is reflected along the x- or y-axis, the spatial relations along the c-direction are unchanged, while the spatial relations along the s-direction are different since, also, the trace direction is reversed. Anyhow, let *f* and *g* be a picture and its reflection, the R-virtual image of *g*, say g_{vi}, can be obtained from the R-virtual image of *f*, say f_{vi}, as follows:

For each atomic relation $\mathrm{ob_i}\gamma\mathrm{ob_w}$ in f_{vi}, the corresponding atomic relation in g_{vi} is:

$$\begin{cases} \mathrm{ob_i}\gamma\mathrm{ob_w} & \text{if } \mathrm{ob_i}\gamma\mathrm{ob_w} \text{ has been extracted along the c-direction} \\ \mathrm{TX\text{-}6}(\mathrm{ob_i}\gamma\mathrm{ob_w}) & \text{otherwise.} \end{cases}$$

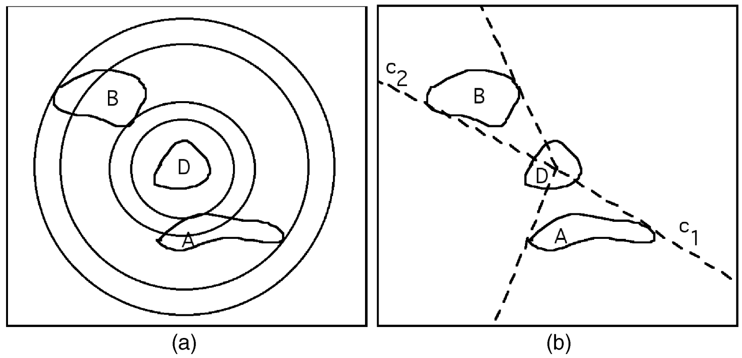


Fig. 15. The beginning points and the ending points along the c- and s-direction.

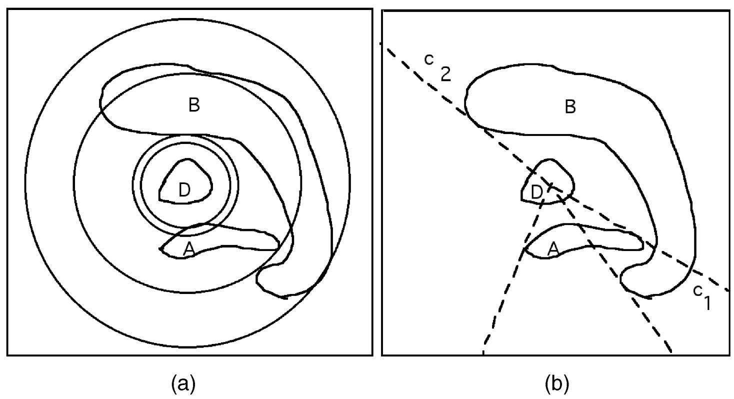


Fig. 16. Ambiguous situations along the s-direction.

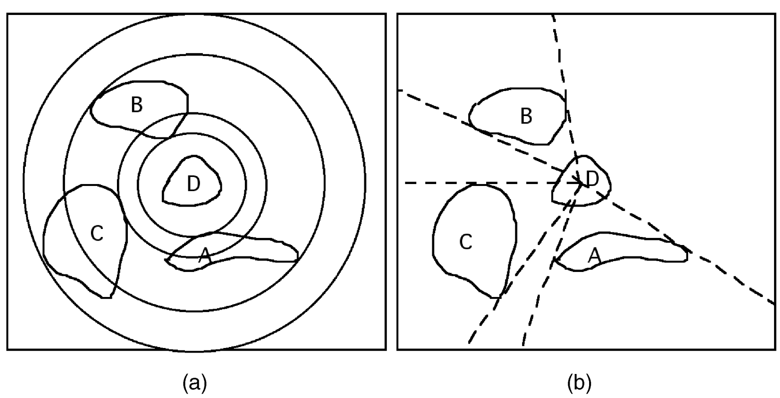


Fig. 17. A picture used as a query.

Then, the same discussion on spatial reasoning as in Section 3 applies to R-virtual images. On the other hand, it is also evident that the same similarity retrieval process introduced in Section 4 can be applied to them. As an example, let us consider the images of Fig. 17 and Fig. 15 as a query image and a candidate image, respectively. The rotation center for both the images corresponds to the object D. Let the sets F and G be {A} and {B, C}, respectively, and let us consider a minimum required similarity degree $\mathrm{t} = 0.7$.

Then, the atomic relations corresponding to the query $\mathrm{Q} = (\{\mathrm{A}\}, \{\mathrm{B}, \mathrm{C}\}, \mathrm{Rel}, 0, 7)$ are:

$$\mathrm{Rel_c} = \{\mathrm{A} = \mathrm{B}, \mathrm{A}/\mathrm{C}, \mathrm{B}/\mathrm{C}\}$$
$$\mathrm{Rel_s} = \{\mathrm{A} < \mathrm{B}, \mathrm{B} < \mathrm{A}, \mathrm{A} < \mathrm{C}, \mathrm{C} < \mathrm{A}, \mathrm{B} < \mathrm{C}, \mathrm{C} < \mathrm{B}\}.$$

According to Definition 3, the sets X and Y are:

$$\mathrm{X} = \{(\mathrm{A} = \mathrm{B}, 0.75)\}$$
$$\mathrm{Y} = \{(\mathrm{A} < \mathrm{B}, 1), (\mathrm{B} < \mathrm{A}, 1)\}.$$

Then, let us compute the similarity degree between the corresponding virtual images; we obtain

$$Sim\_deg(\mathrm{Q}, P\mathrm{vi}) = \frac{1 + 0 + 0 + 1 + 0.75 + 2}{1 + 2 + 3 + 6} = \frac{4.75}{12} = 0.39.$$

Hence, since $Sim\_deg(\mathrm{Q}, P\mathrm{vi}) < t$ (the minimum required similarity degree), the image in Fig. 15 does not satisfy the user's requirements and is discarded.

8 Conclusions and Further Work

In this paper, we have introduced the virtual image as an iconic index and a similarity retrieval method based on virtual images. We have shown that the proposed methodology has several advantages with respect to other existing indexes in terms of explicit representation of the spatial information and efficiency of similarity retrieval. Then, we used the same methodology to obtain a rotation invariant index, called R-virtual image.

In general, virtual images satisfy the property of translation and reflection invariance, but they are sensitive to picture rotation, that is, a rotation can influence the orthogonal spatial relations among pairs the objects. On the contrary, R-virtual images are rotation and reflection invariant but are sensitive to translation, i.e., the iconic index is based on the object-centered representation and varies according to it. The different characteristics of the two introduced indexes correspond to the different requirements of real-world Image Database Systems applications.

We note that more improvements can be done to establish spatial reasoning laws to recognize the similarity between two images also when their corresponding R-virtual images have been extracted with respect to different rotation centers. As a matter of fact, the dependence of the index on the object-centered representation is an important issue and it needs to be further investigated in order to guarantee that images really similar from the user's point of view are recognized as similar also according to the iconic representation.

Finally, we are currently performing an extension of IME in order to also handle topological relationships which relate the geometric basic objects, point, line, and region to the topological primitives, interior, boundary, and exterior of an object.

REFERENCES

[1] A.F. Abate, M. Nappi, G. Tortora, and M. Tucci, "Assisted Browsing in a Diagnostic Image Database," *Proc. Workshop Advanced Visual Interfaces, AVI '96*, pp. 223-232, May 1996.

[2] A.F. Abate, M. Nappi, G. Tortora, and M. Tucci, "IME: An Image Management Environment with Content-Based Access," *Image and Vision Computing 17*, pp. 967-980, 1999.

[3] J.F. Allen, "Maintaining Knowledge about Temporal Intervals," *Comm. ACM*, vol. 26, pp. 832-843, 1983.

[4] L. Bentley and J.H. Friedman, "Data Structures for Range Searching," *ACM Computing Survey*, vol. 11, no. 4, pp. 397-409, Dec. 1979.

[5] C.C. Chang and C.F. Lee, "Relative Coordinates Oriented Symbolic String for Spatial Relationship Retrieval," *Pattern Recognition*, vol. 28, no. 4, pp. 563-570, 1995.

[6] C.C. Chang and S.Y. Lee, "Retrieval of Similar Pictures in Pictorial Databases," *Pattern Recognition*, vol. 7, pp. 675-680, 1991.

[7] N.S. Chang and K.S. Fu, "Query-by-Pictorial Example," *IEEE Trans. Software Eng.*, vol. 6, pp. 519-524, Nov. 1980.

[8] S.-K. Chang and A. Hsu, "Image Information Systems: Where Do We Go from Here?" *IEEE Trans. Knowledge and Data Eng.*, vol. 4, no. 5, pp. 431-442, Oct. 1992.

[9] S.-K. Chang and E. Jungert, "Pictorial Data Management based upon the Theory of Symbolic Projections," *J. Visual Languages and Computing*, vol. 2, no. 3, pp. 195-215, Sept. 1991.

[10] S.-K. Chang and S.-H. Liu, "Picture Indexing and Abstraction Techniques for Pictorial Database," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 4, pp. 475-484, July 1984.

[11] S.-K. Chang, Q.Y. Shi, and C.W. Yan, "Iconic Indexing by 2D String," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 9, no. 3, pp. 413-428, May 1987.

[12] S.-K. Chang, C.W. Yan, D.C. Dimitroff, and T. Arndt, "An Intelligent Image Database System," *IEEE Trans. Software Eng.*, vol. 14, no. 5, pp. 681-688, May 1988.

[13] G. Costagliola, G. Tortora, and T. Arndt, "A Unifying Approach to Iconic Indexing for 2-D and 3-D Scenes," *IEEE Trans. Knowledge and Data Eng.*, vol. 4, no. 3, pp. 205-222, June 1992.

[14] M. De Marsico, L. Cinque, and S. Levialdi, "Indexing Pictorial Document by Their Content: A Survey of Current Techniques," *Image and Vision Computing*, no. 15, pp. 119-141, 1997.

[15] C. Di Ruberto, M. Nappi, and S. Vitulano, "Different Methods to Segment Biomedical Images," *Pattern Recognition Letters*, vol. 18, nos. 11/13, pp. 1125-1131, Apr. 1998.

[16] C. Faloutsos, W. Equitz, M. Flickner, W. Niblack, D. Petkovic, and R. Barber, "Efficient and Effective Querying by Image Content," *J. Intelligent Information Systems*, vol. 3, nos. 3/4, pp. 231-262, July 1994.

[17] M. Flickner et al., "Query by Image and Video Content: The QBIC System," *Computer*, pp. 23-32, 1995.

[18] R. Flinn, "Querying by Image Content," *IBM Research*, no. 3, pp. 22-25, 1996.

[19] C. Freksa, "Temporal Reasoning Based on Semi-Intervals," *Artificial Intelligence*, vol. 54, nos. 1-2, pp. 199-227, 1992.

[20] Special Issue on Content-Based Image Retrieval Systems, *Computer*, V.N. Gudivada and V.V. Raghavan, guest eds., vol. 28, no. 9, 1995.

[21] K. Hirata and T. Kato, "Query by Visual Example, Content-Based Image Retrieval," *Proc. Advances in Database Technology—EDBT '92*, A. Pirotte, C. Delobel, and G. Gottlob, eds., 1992.

[22] C.C. Hsu, W.W. Chu, and R.K. Taira, "A Knowledge-Based Approach for Retrieval Images by Content," *IEEE Trans. Knowledge and Data Eng.*, vol. 8, no. 4, pp. 522-532, Aug. 1996.

[23] P.W. Huang and Y.R. Jean, "Using 2D C+-Strings as Spatial Knowledge Representation for Image Database Systems," *Pattern Recognition*, vol. 27, no. 9, pp. 1249-1257, 1994.

[24] P.W. Huang and Y.R. Jean, "Reasoning about Pictures and Similarity Retrieval for Image Information Systems Based on SK-SET Knowledge Representation," *Pattern Recognition*, vol. 28, no. 12, pp. 1915-1925, 1995.

[25] P.W. Huang, "Indexing Pictures by Key Objects for Large-Scale Image Database," *Pattern Recognition*, vol. 30, no. 7, pp. 1229-1237, 1997.

[26] M. Kobashi and L.G. Shapiro, "Knowledge-Based Organ Identification from CT Images," *Pattern Recognition*, vol. 28, no. 4, pp. 475-491, 1995.

[27] E. Jungert, "Extended Symbolic Projection Used in a Knowledge Structure for Spatial Reasoning," *Proc. Fourth BPRA Conf. Pattern Recognition*, Mar. 1988.

[28] E. Jungert and S.-K. Chang, "An Image Algebra for Pictorial Data Manipulation," *CVGIP: Image Understanding*, vol. 58, no. 2, pp. 147-160, Sept. 1993.

[29] E. Jungert, "Rotation Invariance in Symbolic Projection as a Means for Determination of Binary Object Relations," *Proc. Workshop Qualitative Reasoning and Decision Technologies (QUARDET '93)*, pp. 503-512, June 1993.

[30] E. Jungert, "Rotation Invariance in Symbolic Slope Projection as a Means to Support Spatial Reasoning," *Proc. Third Int'l Conf. Automation, Robotics and Computer Vision (ICARCV '95)*, pp. 364-368, Nov. 1994.

[31] S.Y. Lee and F.J. Hsu, "2D C-String: A New Spatial Knowledge Representation for Image Database Systems," *Pattern Recognition*, vol. 23, pp. 1077-1088, Oct. 1990.

[32] S.Y. Lee and F.J. Hsu, "Picture Algebra for Spatial Reasoning of Iconic Images Represented in 2D C-String," *Pattern Recognition Letters*, vol. 12, pp. 425-435, July 1991.

[33] S.Y. Lee and F.J. Hsu, "Spatial Reasoning and Similarity Retrieval of Images Using 2D C-String Knowledge Representation," *Pattern Recognition*, vol. 25, pp. 305-318, Mar. 1992.

[34] S.Y. Lee, M.K. Shan, and W.P. Yang, "Similarity Retrieval of Iconic Image Database," *Pattern Recognition*, vol. 22, pp. 675-682, June 1988.

[35] M. Nabil, A.H.H. Ngu, and J. Shepherd, "Picture Similarity Retrieval Using the 2D Projection Interval Representation," *IEEE Trans. Knowledge and Data Eng.*, vol. 8, no. 4, pp. 533-539, Aug. 1996.

[36] G. Petraglia, M. Sebillo, M. Tucci, and G. Tortora, "Towards Normalized Iconic Indexing," *Proc. 1993 IEEE Symp. Visual Language*, pp. 392-394, Aug. 1993.

[37] G. Petraglia, M. Sebillo, M. Tucci, and G. Tortora, "Rotation Invariant Iconic Indexing for Image Database Retrieval," *Progress in Image Analysis and Processing III,* S. Impedovo, ed., pp. 271-278, World Scientific, 1994.

[38] G. Petraglia, M. Sebillo, M. Tucci, and G. Tortora, "A Normalized Index for Image Databases," *Intelligent Image Database Systems,* S.K. Chang, E. Jungert, and G. Tortora, eds., pp. 43-69, Singapore: World Scientific, Sept. 1996.

[39] E.G.M. Petrakis and C. Faloutsos, "Similarity Searching in Medical Image Databases," *IEEE Trans. Knowledge and Data Eng.,* vol. 9, no. 3, pp. 435-447, May/June 1997.

[40] E.G.M. Petrakis and S.C. Orphanoudakis, "Methodology for the Representation, Indexing and Retrieval of Images by Content," *Image and Vision Computing,* vol. 8, pp. 504-521, 1993.

[41] G. Salton and M.J. McGill, *Introduction to Modern Information Retrieval.* McGraw-Hill, 1983.

[42] S. Santini and R. Jain, "Similarity Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence,* submitted.

[43] S. Santini and R. Jain, "Similarity Queries in Image Databases," *Proc. CVPR '96, Int'l IEEE Computer Vision and Pattern Recognition Conf.,* 1996.

[44] S. Santini and R. Jain, "The Graphical Specification of Similarity Queries," *J. Visual Languages and Computing,* no. 7, pp. 403-421, 1996.

[45] S.L. Tanimoto, "An Iconic Symbolic Data Structuring Scheme," *Pattern Recognition and Artificial Intelligence,* pp. 452-471, New York: Academic Press, 1976.

[46] *Symbolic Projection for Image Information Retrieval and Spatial Reasoning,* S.K. Chang and E. Jungert, eds. Academic Press, 1996.

Gennaro Petraglia received the Laurea degree in industrial chemistry from the University of Naples. He is an associate professor of computer science at the University of Salerno. His research interests include database design, image databases, information systems.

Monica Sebillo received the Laurea degree in computer science from the University of Salerno and the PhD degree in applied mathematics and computer science from the University of Naples. Her research interests include image databases, information systems, human-computer interaction, and visual languages.

Maurizio Tucci received the Laurea degree in computer science from the University of Salerno. He is a professor of computer science at the University of Salerno. His research interests include multimedia databases, content-based image retrieval, software-development environments, visual languages, and pictorial information systems.

Genoveffa Tortora received the Laurea degree in computer science from the University of Salerno. She is a professor of computer science and chair of the Department of Mathematics and Computer Science of the University of Salerno. Her research interests include human-computer interaction, image databases, software-development environments, visual languages, geographic information systems, and pictorial information systems. She is author and coauthor of more than 50 papers published in scientific journals or proceedings of refereed conferences, and is coeditor of three books. She is also on the editorial board of both the *International Journal of Software Engineering and Knowledge Engineering* and the *Journal of Visual Languages and Computing*. She is a senior member of the IEEE.