Retrieval by Spatial Similarity in Image Databases

Mariana Stoeva[1], Slava Jordanova[2]

Abstract – **This paper presents our work in the area of image retrieval in Image Databases for images saved by spatial similarity of domain objects location. We propose a geometric based structure which makes possible the extraction of directional spatial relations among domain objects that are invariant with respect to transformations. We introduce an algorithm for spatial similarity retrieval in Image Database.**

Keywords – **Image databases, Query by index content, Spatial reasoning, Similarity retrieval**

I. Introduction

Similarity based retrieval of images is an important task in many image database application. A major class of users' requests is the one that requires retrieval of those images in databases that are spatially similar to the query images. A survey of existing approaches to the problem of spatial image retrieval and their limitations is described in [9,16]. The Query by Pictorial Example (QPE) [4] 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 witch use objects and spatial relationships to describe the visual content of an image.

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. 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 of symbolic projections [11,12,9,16]. However, in real application, it would also to be able to find the images in the database that present a given pattern, even if it appears mirror reflected.

In this paper we propose a geometric based structure which makes possible the extraction of spatial relations among domain objects that are invariant with respect to transformation such as translation, rotation, scaling, reflection, view point change as well as arbitrary compositions of these transformations. We introduce an algorithm for spatial similarity retrieval I Image Database, presented as SIMR$_{R\theta}$ that recognizes transformed images and sub-images.

[1]Mariana Stoeva is with the Technical University Varna, Department of Computer Systems and Technologies, Bulgaria, E-mail: mariana_stoeva@hotmail.com

[2]Slava Jordanova is with the Technical University Varna, Department of Computer Systems and Technologies, Bulgaria, Email: sl@windmail.net

The paper is organized as follows: In Section 2 we present the symbol description used by us as well as the approximations used for computing the spatial relationships among objects. In Section 3, for the purpose of similarity, we introduce descriptions of spatial relations, of spatial relation similarities, as well as similarity measure between two images and also an algorithm presented as SIMR$_{R\theta}$ that recognizes transformed images and sub-images. We publish a part of the experiments made and their results in Section 4. In Section 5 some conclusions are shown and further work aims are planned.

II. Image Description and Approximations Used

In most applications the objects in the images have not an exactly fixed shape. The performance of any spatial investigation on the exact location and shape of each object domain is very expensive. Due to this an initial approximation or filtration is used. The most commonly used objects shape approximation to the purpose of image retrieval in Image Databases (ID) is the Minimal Boundary Rectangle (MBR – the smallest rectangle that bounds the object). Whatever the approximation approach is, it cannot influence the method of indexing in ID [15].

The use of MBR for shape representation allows obtaining of spatial relations that are invariant with respect to translation and scaling, but variable with respect to rotation and reflection transformations.

In our work we are searching such an approximate objects shape representation that would be invariant with respect to arbitrary compositions of transformations and at the same time would store the information that is necessary for accounting the spatial relations among extended objects in an image.

We assume that an image I consists of n domain objects denoted as O_j. The symbolic description of image I is stored in Image Database. Let C_j is an object centroid and C_I is the centroid of the image obtained from the centroids of the objects contained in the image. The description presents an image as $I = ((O_j, (O_j.x_{cj}, O_j.y_{cj}), ((O_j.x_{jl}, O_j.y_{jl}), 1 \le l \le 4), 1 \le j \le n)$, where O_j is a name of the object, $(O_j.x_{cj}, O_j.y_{cj})$ are coordinates of the centroid C_j of object O_j, $((O_j.x_{jl}, O_j.y_{jl}), 1 \le l \le 4)$ are the coordinates of the typical for the object shape coordinates of four points from the object external contour in Cartesian coordinate system with initial point C_I. The typical points are vertices of an approximating object shape tetragon. We assume that the domain objects are stored in lexicographic /alphabetic/ order on the object names. The notation O_I is used to refer to a set of objects in image I.

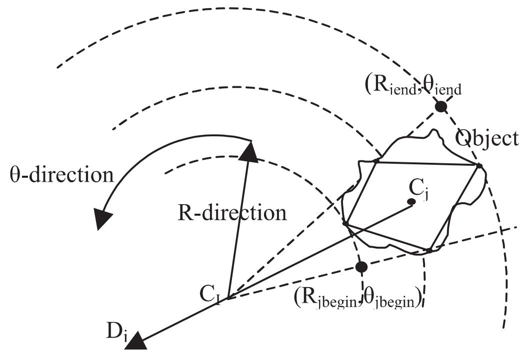

Fig. 1. Illustration of the approximations used

The domain objects of each image are stored in ID by their symbol names, centroid coordinates, and four characterizing the object shape points, whose determination is invariant with reference to transformations. The determination of the four typical for the object shape points used in approximate representation of each object is submitted to the consistent execution of the next requirements in the corresponding priority order" to be independent of transformations, to represent the shape in the form of tetragon whose centroid is identical with the object centroid, and whose area is maximally close to that of the object.

The implementation of these requirements in the typical points determination imposes the use of criteria treating the point distance from the external contour to the object centroid C_j, the length of the maximal segments determined by the cross-section of the external contour by axes passing through the centroid C_j, as well as the angle between these axes. These criteria are presented in details in [17].

The extended object O_j shape approximation built on stored information for each object has the view of a ring sector obtained from concentric circles around image C_I centroid. The ring boundaries are determined by the minima and maxima of the polar coordinates of the points from the tetragon determined by the stored four typical object points $((O_j.x_{jl}, O_j.y_{jl}),\ 1 \leq l \leq 4)$. The ring sector initial point has polar coordinates $(R_{\text{begin}}, \theta_{\text{begin}})$ and the final point has polar coordinates. $(R_{\text{end}}, \theta_{\text{end}})$. The description of object O_j takes the form $(O_j, (R_{cj}, \theta_{cj}),\ (R_{j\text{begin}}, \theta_{j\text{begin}}),\ (R_{j\text{end}}, \theta_{j\text{end}}),\ 1 \leq j \leq n))$, where (R_{cj}, θ_{cj}) are the polar coordinates of the centroid C_j, $(R_{j\text{begin}}, \theta_{j\text{begin}})$, $(R_{j\text{end}}, \theta_{j\text{end}})$ – the origin and the end of the object approximating ring sector. An example of ring sector approximating object O_j shape and described by its four typical points is depicted in Fig. 1. Objects centroids polar coordinates and the polar coordinates of the initial and the final points of the approximating them ring sectors are used for obtaining the spatial relations of each object with the other image objects.

III. Spatial Similarity Image Retrieval in Image Database

Following the QPE philosophy the query is processed by matching the obtained from its description spatial relations with these of the stored in the database images. Investigations concerning similarity retrieval methodologies can be found in [3,10]. We use a directional approach to determine the spatial relationships between each match of domain objects in an image. We use two directions corresponding to the directions in a polar coordinate system whose centre is the image centroid. The linear scanning R-direction begins from the image centroid C_I and corresponds to the concentric circles going from the origin to outside. The circle θ-direction corresponds to a trace swiping anticlockwise around the origin. The directions are shown in Fig. 1. Allen's well-known 13 types of spatial relationships [1] whose defining is adapted to a polar coordinate system are obtained in determining the relationships between each object match in the sense of conditions for their initial and final points in each separate direction. In both directions 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. The spatial relations among the image objects are computed by using the polar coordinates of the initial $(R_{\text{begin}}, \theta_{\text{begin}})$ and the final $(R_{\text{end}}, \theta_{\text{end}})$ points of object approximating ring segments whose centre is the object centroid.

Definition 1. A triple like $(O_j\, \gamma\, O_i)$ is called an atomic spatial relation, where $O_j,\ O_i \in O_I$ are object names and $\gamma \in \{<, >, |,$ ", =, [, (,],), /, \, %, #$\}$ is an operator.

We use the notation $(O_j\, \gamma_R\, O_i)$ to indicate that the pair of objects (O_j, O_i) belongs to relation γ in R-direction and $(O_j\, \gamma_\theta\, O_i)$ to indicate that the pair of objects (O_j, O_i) belongs to relation γ in θ-direction.

The defining of binary spatial relations is identical to those in [1] adapted to a polar system and it is presented in [16]. The 7 known symbols for spatial operators used by Lee and Hsu [13,14] are utilized and four new symbols are introduced. In the linear direction R the analogy of spatial relations defining with the one used in orthogonal model is complete due to the linear type of the direction. To achieve invariance in determining the spatial relations for each object match with reference to the transformations rotation and reflection in circle direction θ it is necessary the angle coordinates to be recomputed. This is necessary due to the existing peculiarity of the circle direction (its relative beginning and end are identical). The re-computation stores the proximity of points and corresponds to the real ordering of the initial and final object points in θ-direction.

We provide the independence of the spatial relations in θ-direction of the leading object O_j and the other objects O_i by re-computing their initial coordinates $(\theta_{i\text{begin}})$ with reference to an origin point D_j. The point D_j is at a distance of π from the value θ_{cj}. For object O_j the relative circle origin in θ-direction is point D_j $(\theta_{cj} \pm \pi)$. Thus the coordinate θ_{cj} of object O_j centroid remains always in the relative middle of the numerical axis. The coordinates of the end points $(\theta_{i\text{end}})$ are computed with reference to the recomputed origin and the stored distance between the initial $(\theta_{i\text{begin}})$ and the final $(\theta_{i\text{end}})$ coordinates. The relative origin for computing the spatial relations of object O_j in circle direction θ in the example at Fig. 1 is denoted by D_j. We hold without formal proof, that with thus recomputed initial and final points of the ring sector that describes the domain object, the relations remain invariant with reference to rotation transformation. In case of reflection transformation the relations in θ-direction may be easily obtained by changing each operator with corresponding inverse operator.

Since the query processing goal is to retrieve those images from the ID that contain almost the same objects as contained in the query, and have similar spatial ordering. Our similarity measure has to take into account the similarities between the spatial operators for each relation.

Definition 2. Let the objects O_j and O_i belong to image $Q(O_j, O_i \in O_Q)$ and to image $I(O_j, O_i \in O_I)$. Let the object match $((O_j, O_i))$ in image Q belongs to the spatial relation $\gamma(O_j\, \gamma\, O_i)$ and the object match (O_j, O_i) in image I belongs to the spatial relation $\gamma'(O_j\, \gamma'\, O_i)$. We define the similarity between the spatial relations in both images Q and I for the object match (O_j, O_i) as $\text{sim}_{ji}(\gamma, \gamma') = t$, where $t \in [0, 1]$, is the similarity between the spatial operators γ and γ'. We denote the similarity between the relations in both the images Q and I for the object match (O_j, O_i) in R-direction as $\text{sim}_{ji}(\gamma_R, {\gamma'}_R)$ and as $\text{sim}_{ji}(\gamma_\theta, \gamma'_\theta)$ in θ-direction.

We adopt the interval neighborhood graph as in [16], that defines the distances among spatial operators. According the definition of the interval neighborhood graph given by Freksa in [7], two projection relationships are neighbors if they can be transformed into one another by continuous deforming the projections. The similarity values $t = \text{sim}(\gamma, \gamma')$ between each operator match we use, are shown in [16]. In case of multiple instances of one object, the greatest value of similarity $\max\{\text{sim}_{ji}(\gamma, \gamma')\}$ between the operators that describe it is used. 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 [0,2]. According our understanding for similarity and our desire the method to be independent of human interpretation, we put forward a formula for similarity evaluation that assesses the similarity between the common for both the images objects and their corresponding atomic relations.

Definition 3. Let the query image be Q and the Image Database image is I. We define the similarity distance between Q and I by equation (1),

$$\text{sim}(Q, I) = \frac{1}{m}\left(n + \frac{1}{2n}\sum_{j}^{n}\sum_{i}^{n}\text{sim}_{ji}(\gamma_R, \gamma'_R) + \text{sim}_{ji}(\gamma_\theta, \gamma'_\theta)\right) \tag{1}$$

where $m = ||O_Q||$ is the number of objects in the query image, $n = ||O_Q \cap O_I||$ is the number of the common for both the images objects, $\text{sim}_{ji}(\gamma_R, \gamma'_R)$ is the spatial similarity between the images Q and I for the object match (O_j, O_i) in R-direction, and $\text{sim}_{ji}(\gamma_\theta, \gamma'_\theta)$ is the spatial similarity between the images Q and I for the object match (O_j, O_i) in θ-direction.

An image I from the Database answers the query Q only if it contains all the query objects with the same relations as in image Q. The more objects from Q are contained in image I, the higher the similarity degree. This formula returns a value of the order of [0,2] by computing the correlation between the number of objects that are retrieved together with the similarity value of their atomic relations and the number of objects demanded by the user and their atomic relations. We expect a value of 2 when the input images are identical, otherwise the function will take a lower values. These values are proportional to the degree of disagreement in the spatial relationships between the corresponding objects in the input images.

We present a spatial similarity algorithm, that not only recognizes transformation variants of an image but also recognizes sub-images (or transformation variants of sub-images) of the query image in the Image Database.

⋄ Q and I are the query image and the Database image.

Algorithm $\text{SIM}_{R\theta}(Q = ((O_j, (O_j.x_{cj}, O_j.y_{cj}), ((O_j.x_{jl}, O_j.y_{jl}), 1 \leq l \leq 4)), 1 \leq j \leq m),\ I = ((O_j, (O_j.x_{cj}, O_j.y_{cj}), ((O_j.x_{jl}, O_j.y_{jl}), 1 \leq l \leq 4)), 1 \leq j \leq v))$

$Q' \leftarrow ((Q.O_j, (Q.O_j.x_{cj}, Q.O_j.y_{cj}), ((Q.O_j.x_{jl}, Q.O_j.y_{jl}), 1 \leq l \leq 4)), 1 \leq j \leq m) | O_j \in O_Q \cap O_I$
$I' \leftarrow ((I.O_j, (I.O_j.x_{cj}, I.O_j.y_{cj}), ((I.O_j.x_{jl}, I.O_j.y_{jl}), 1 \leq l \leq 4)), 1 \leq j \leq m) | O_j \in O_Q \cup O_I$
$n \leftarrow |O_{Q'}| = |O_{I'}|$
$m \leftarrow |O_Q|$
$Q'_{R\theta} \leftarrow ((O_j, (R_{cj}, \theta_{cj}), (R_{j\text{begin}}, \theta_{j\text{begin}}), (R_{j\text{end}}, \theta_{j\text{end}}), 1 \leq j \leq n)$
$I'_{R\theta} \leftarrow ((O_j, (R_{cj}, \theta_{cj}), (R_{j\text{begin}}, \theta_{j\text{begin}}), (R_{j\text{end}}, \theta_{j\text{end}}), 1 \leq j \leq n)$
sim $\leftarrow 0$
for $j \leftarrow 1$ **to** n **do**

⋄ Accumulating object factor

sim $\leftarrow$ sim $+ 1$

⋄ Accumulating spatial factor

for $i \leftarrow 1$ **to** n **do**

⋄ The atomic spatial relations between objects O_j and O_i in image Q'

$O_j \gamma_R O_i \leftarrow (Q'_{R\theta})$
$O_i \gamma_\theta O_j \leftarrow (Q'_{R\theta})$

⋄ The atomic spatial relations between objects O_j and O_i in image I'

$O_j \gamma'_R O_i \leftarrow (I'_{R\theta})$
$O_i \gamma\_\theta O_j \leftarrow (I'_{R\theta})$

⋄ Defining the spatial similarity values between the atomic relations of images Q' and I' for the object match (O_j, O_i) in R-direction and θ-direction

$\text{sim}_{ji}(\gamma_R, \gamma'_R) \leftarrow ((O_j \gamma_R O_i), (O_j \gamma'_R O_i))$
$\text{sim}_{ji}(\gamma_\theta, \gamma'_\theta) \leftarrow ((O_j \gamma_\theta O_i), (O_j \gamma'_\theta O_i))$
sim $\leftarrow$ sim $+ (\text{sim}_{ji}(\gamma_R, \gamma'_R) + \text{sim}_{ji}(\gamma_\theta, \gamma'_\theta))/(2n)$
endfor
endfor
sim $\leftarrow$ sim$/m$
return sim
end $\text{SIM}_{R\theta}$

Fig. 2. Spatial similarity algorithm

The proposed algorithm $\text{SIM}_{R\theta}$ is further described. The formal definition of $\text{SIM}_{R\theta} : Q \times I \rightarrow [0, 2]$, where Q and I are the symbol presentations of the query and the Database images, respectively. The algorithm is shown in Fig. 2. Lines beginning with the symbol ⋄ indicate comments. Certain lines are numbered for comment convenience.

The algorithm has computational complexity n^2, where n is the number of the common for Q' and I' objects. The computational complexity is determined by the time necessary for the performance of the body of the imbedded for loops (lines $8 \div 17$). This complexity is the same as the complexity of algorithms that take in account orthogonal relations of domain objects [16].

IV. Experiments

For the purpose of the experiment we use a test image collection of 10 original images and 9 variants of each of these images. Variants 1-6 are generated by applying transformation operators to all domain objects, while for variants 8-9 the operators are applied to a subset of domain objects. The transformation operators used in producing the variants of the original images include compositions of the transformations (translation, scaling, rotation, reflection) with different length.

The experimental evaluation is done in two parts. In the first part the focus is on understanding the correspondence between the similarity computed by $\mathrm{SIM}_{R\theta}$ and that to the intuitively expected. In the second part the robust behavior of $\mathrm{SIM}_{R\theta}$ algorithm in combined transformations is studied. The idea of using expert-provided rank ordering of images, elative to each query in given set of test queries, in quantifying the retrieval quality is presented in [8] and [9]. This quantification is based on a measure referred to as the R_{norm} [8].

Image 1 Image 2 Image 3 Image 4

Fig. 3. . Test images for intuitive understanding of SIMR? similarity computation

To establish whether $\mathrm{SIM}_{R\theta}$ rank ordering of Database images with respect to a query image agrees with our intuitive visual rank ordering, we consider the images shown in Fig. 3. We consider each of these images as the query and compute its similarity with all the other images by $\mathrm{SIM}_{R\theta}$ and by an expert. The $\mathrm{SIM}_{R\theta}$ rank ordering confirms to the intuitively expected rank ordering. The research for robust behavior includes two sets consisting of one original image and its all 9 variants. Then we consider again each image in the set as the query. The experiment shows that the $\mathrm{SIM}_{R\theta}$ algorithm performs our expectations.

V. Conclusions

This paper proposes a geometric based structure which makes possible the extraction of directional spatial relations among domain objects that are invariant with respect to transformations. The presented spatial similarity retrieval distance and algorithm solve a class of tasks namely contain-based spatial similarity research. The similarity distance and the algorithm are not influenced by possible image transformations of the query and database, moreover, they catch also transformed sub-images. The algorithm temporary complexity is n^2, where n is the number of objects that are common for both the query and database images. The algorithm is robust in the sense that it can recognize translation, scale, and rotation variant images and the variants generated by an arbitrary composition of these three geometric transformations. The effectiveness and efficiency of the spatial similarity retrieval algorithm are evaluated by using an expert-provided rank ordering of a test collection with respect to a set of test queries. Our efforts for the future point to improving the algorithm effectiveness and reducing the error inserted by the used approximations.

References

[1] J. F Allen, "Maintaining Knowledge about temporal Intervals", Comm. ACM, vol. 26 (1983) 832-843

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

[3] C.C. Chang, S.Y. Lee, "Retrieval of Similar Pictures in Pictorial Database", Pattern Recognition, vol. 7 (1991) 675-680

[4] N. S. Chang, K.S. Fu, " Query-by-pictorial-example", IEEE Trans. Soft. Eng., vol. 6 (1980) 519-524

[5] S.K. Chang, E. Yungert, "Pictorial Data Management based upon the Theory of symbolic Projection", J. Visual Language and Computing, vol. 2 no3 pp.195-215 1991

[6] N.S. Chang, K.S Fu, "Query-by-pictorial-example", IEEE Trans. Soft. Eng., vol. 6 pp.519-524 1980

[7] C. Freksa, "Temporal Reasoning Based on Semi-Intervals", Artificial Intelligence, vol. 54 no.1-2 pp.199-227 1992

[8] V.Gudivada, V. Raghavan "Design and Evaluation of Algorithms for Image Retrieval by Spatial Similarity", ACM Trans. Information Systems, vol. 13 no.1 pp.115-144 1995

[9] V. Gudivada, "$\theta\Re$-string: A geometry-based representation for efficient and Effective Retrieval of Images by Spatial Similarity", IEEE Trans. KDE, vol. 10, no.3, pp.504-511, 1998

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

[11] E. Jungert, "Rotation Invariance in Symbolic Projection as a Means for Determination of Binary Object Relations", Proc. Workshop (QUARDET '93) pp.503-512 1993

[12] 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 1994

[13] S.Y. Lee, F.J. Hsu, "2D C-string: A new Spatial Knowledge Representation for Image Database System", Pattern Recognition, vol. 23 no.10 pp.1077-1088 1990

[14] S.Y. Lee, F.J. Hsu, " Pictorial Algebra for Spatial Reasoning of Iconic Images Represented 2D C-string", Pattern Recognition, vol. 12 pp.425-435 1991

[15] B. C. Ooi, " Efficient Query Processing in Geographic Information Systems", Edited by G. Goos, G, Hartmanis, J. (eds.) Springer-Velag 1990

[16] G. Petraglia,, M. Sebillo , G. Tucci G. Tortora, "Virtual Images for Similarity Retrieval in Image Databases", IEEE Trans. on knowledge and data engineering vol. 13 no.6 pp.951-967 2001

[17] M. Stoeva, Sl. Jordanova, " Similarity Shape Retrieval from Image, Database Proc. XXXVII International scientific conference on information, communication and energy systems and technologies 1 (ICEST' 2002) pp.283-287 2002