High Expressive Spatio-temporal Relations

R. CHBEIR, Y. AMGHAR, A. FLORY

LISI – INSA de Lyon
20, Avenue A. Einstein
F-69621 VILLEURBANNE - FRANCE
Phone: (+33) 4 72 43 85 88 - Fax: (+33) 4 72 43 87 13
E-mail: {rchbeir, amghar, flory}@lisi.insa-lyon.fr

Abstract
The management of images remains a complex task that is currently a cause for several research works in various domains such as GIS, medicine, etc. In line with this, we are interested in this work with the problem of image retrieval that is mainly related to the complexity of image representation. In this paper, we address the problem of relation representation and we propose a novel method for identifying and indexing relations between salient objects. Our method is based on defining several intersection sets for each feature like shape, time, etc. Relations are then computed using a binary intersection matrix between sets. Several types of relations can be calculated using our method in particular spatial and temporal relations. We show how our method is able to provide high expressive power and homogeneous representation to relations in comparison to the traditional methods.

Introduction
During the last decade, a lot of work has been done in information technology in order to integrate image retrieval in the standard data processing environments. Image retrieval is involved in several domains [(Yoshitaka and Ichikawa 1999), (Rui, Huang, and Chang 1999), Grosky 1997], (Smeulders, Gevers, and Kersten, 1998)] such as Geographic Information Systems, Medicine, Surveillance, etc. where queries criteria are based on different types of features. Principally, spatio-temporal relations are very important in image retrieval and need to be considered in several domains. In medicine, for instance, the spatial data in surgical or radiation therapy of brain tumors is decisive because the location and the temporal evolution of a tumor has profound implications on a therapeutic decision (Chbeir and Favetta 2000).

Hence, it is crucial to provide a precise and powerful system to express spatio-temporal relations required in image retrieval. In this paper, we address this issue and propose a novel method that can easily compute several types of relations with better expressions.

The rest of this paper is organized as follows. In section 2, we present the related work. Section 3 is devoted to define our method for calculating relations. Finally, conclusions and future orientations are given in section 4.

Related Work
Relations between either salient objects, shapes, points of interests, etc. have been widely used in image indexing such as R-tree (Guttman 1984), R+-tree (Sellis, Roussopoulos, and Faloutsos 1987), R*-tree (Beckmann 1990), hB-tree (Lomet and Salzberg 1990), ss-tree (White and Jain 1996), TV-tree (Lin, Jagadish, and Faloutsos 1994), 2D-String (Chang, Shi, and Yan 1987), 2D-G String (Chang and Jungert 1991), 2D-C+ String (Huang and Jean 1994), θR-String (Gudivada 1995), σ-tree (Chang and Jungert 1997), etc. Temporal and spatial relations are the most used relations in image indexing.

To calculate temporal relations, Allen relations (Allen 1983) are often used. Allen proposes 13 temporal relations (Before, After, Meet, Touched By, Started by, Overlapped By, Start With, Started With, Finish wish, Finishes, Contain, During, Equal) in which 6 are symmetrical.

On the other hand, three major types of spatial relations are generally proposed in image representation (Egenhofer, Frank, and Jackson 1989):

- Metric relations: measure the distance between salient objects (Peuquet 1986). For instance, the metric relation “far” between two objects A and B indicates that each pair of points A_i and B_j has a distance greater than a certain value δ.
- Directional relations: describe the order between two salient objects according to a direction, or the localisation of salient object inside images (El-kwae and Kabuka 1999). In the literature, fourteen directional relations are considered: north, south, east, west, north-east, north-west, south-east, south-west, left, right, up, down, front and behind.
- Topological relations: describe the intersection and the incidence between objects [(Egenhofer and Herring 1991), (Egenhofer 1997)]. Egenhofer (Egenhofer and
Herring 1991) has identified six basic relations: Disjoint, Touch, Overlap, Cover, Contain, and Equal.

In spite of all the proposed work to represent complex visual situations, several shortcomings exist in the methods of spatial relation computations. For instance, Figure 1 shows two different spatial situations of three salient objects that are described by the same spatial relations in both cases, in particular topological relations: a1 Touch a2, a1 Touch a3, a2 Touch a3; and directional relations: a1 Above a3, a2 Above a3, a1 Left a2.

<table>
<thead>
<tr>
<th>a1</th>
<th>a2</th>
<th>a3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Two different spatial situations.

In this paper, we present our method able to provide a high expressive power to represent several types of relations in particular spatial and temporal ones.

Proposition

Our proposal represents an extension of the 9-Intersection model (Egenhofer and Herring 1991). It provides a general method for computing not only topological relations but also other types of relations in particular temporal and spatial relations with higher precision. The idea is to identify relations in function of two values of a feature such as shape, position, time, etc. The shape feature gives spatial relations, the time feature gives temporal relations, and so on. To identify a relation between two values of the same feature, we propose the use of an intersection matrix between several sets defined in function of each feature:

Definition

Let us first consider a feature F. We define its intersection sets as follows:

- F\textsuperscript{\text{int}} (or inferior): contains elements of F that do not belong to any other intersection set and inferior to ∂F elements on the basis of i dimension.
- F\textsuperscript{\text{sup}} (or superior): contains elements of F that do not belong to any other intersection set and superior to ∂F elements on the basis of i dimension.
- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{sup}} (or inferior): contains elements of F that do not belong to any other intersection set and inferior to ∂F elements on the basis of i dimension.
- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{sup}}: contains elements of F that do not belong to any other intersection set and superior to ∂F elements on the basis of i dimension.

If we consider a feature of 2 dimensions i and j (as the shape of a salient object in a 2D space), we can define 4 intersection subsets:

- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{sup}} (or inferior): contains elements of F that do not belong to any other intersection set and inferior to F\textsuperscript{\text{in}} and ∂F elements on the basis of i and j dimensions.
- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{sup}}: contains elements of F that do not belong to any other intersection set and superior to F\textsuperscript{\text{in}} and ∂F elements on the basis of i and j dimensions.
- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{in}} (or superior): contains elements of F that do not belong to any other intersection set and superior to F\textsuperscript{\text{in}} and ∂F elements on the basis of i and j dimensions.
- F\textsuperscript{\text{in}} ∩ F\textsuperscript{\text{in}}: contains elements of F that do not belong to any other intersection set and superior to F\textsuperscript{\text{in}} and ∂F elements on the basis of i and j dimensions.

More generally, we can determine intersection sets (2^n) of n dimensional feature.

In addition, we use a tolerance degree in the feature intersection sets definition in order to represent separations between sets and to provide a simple certainty parameter for the end-user. For this purpose, we use two tolerance thresholds:

- Internal threshold ε\textsuperscript{\text{in}}: defines the distance between F\textsuperscript{\text{in}} and ∂F,
- External threshold ε\textsuperscript{\text{out}}: defines the distance between subsets of F.

To calculate relation between two features, we establish an intersection matrix of their corresponding feature intersection sets. Matrix cells have binary values:

- 0 whenever intersection between sets is empty,
- 1 otherwise

For one-dimensional feature (such as the acquisition time), the following intersection matrix is used to compute relations between two values A and B:
For two-dimensional feature (such as the shape) of two values A and B, we obtain the following intersection matrix:

\[
\begin{array}{cccc}
A \cap B & A \cup B & A \cap \neg B & A \cup \neg B \\
\neg A \cap B & \neg A \cup B & \neg A \cap \neg B & \neg A \cup \neg B \\
A \cap \neg A \cap B & A \cup \neg A \cup B & A \cap \neg A \cap \neg B & A \cup \neg A \cup \neg B \\
\neg A \cap \neg A \cap B & \neg A \cap \neg A \cup B & \neg A \cap \neg A \cap \neg B & \neg A \cap \neg A \cup \neg B \\
\end{array}
\]

Figure 4: Intersection matrix of two values A and B on the basis of two-dimensional feature.

**Temporal relations**

To calculate temporal relations, we define the following intersection sets for each interval T (Figure 5):

- TR1: The interval \( \Delta T_1 \) starts before \( \Delta T_2 \) and ends almost before \( \Delta T_2 \).
- TR2: The interval \( \Delta T_1 \) starts almost before \( \Delta T_2 \) and finishes almost after the beginning of \( \Delta T_2 \).
- TR3: The interval \( \Delta T_1 \) starts before the beginning of \( \Delta T_2 \) and finishes during \( \Delta T_2 \).
- TR4: The interval \( \Delta T_1 \) starts before the beginning of \( \Delta T_2 \) and finishes almost before the beginning of \( \Delta T_2 \).
- TR5: The interval \( \Delta T_1 \) starts almost before \( \Delta T_2 \) and finishes during \( \Delta T_2 \).

Using our method, 33 (32+1) temporal relations between two intervals are identified instead of 13 (Allen 1983). We show here below 7 of these principal temporal relations (non symmetrical) between two intervals \( \Delta T_1 \) and \( \Delta T_2 \):

1. TR1: The interval \( \Delta T_1 \) starts and ends before \( \Delta T_2 \):

   - In this paper, we do not consider intersections sets between instants.
Spatial relations
In this paper, we consider only spatial relations calculated between polygonal shapes. Intersections sets of a polygon \( P \) are defined as follows (Figure 6):

- \( P^\Omega \): represents the surface of \( P \). It contains all points situated in the internal part of \( P \).
- \( \partial P \): is the frontier of \( P \).
- \( P_x \cap P_y \): contains all points that do not belong to \( P^\Omega \) nor to \( \partial P \), where coordinates are inferior to barycentre’s.
- \( P_x \cup P_y \): contains all points that do not belong to \( P^\Omega \) nor to \( \partial P \), where coordinates are superior to barycentre’s.
- \( P_x \cap P_y \): contains all points that do not belong to \( P^\Omega \) nor to \( \partial P \), where abscises are strictly inferior to barycentre’s and ordinates are strictly superior to barycentre’s.
- \( P_x \cup P_y \): contains all points that do not belong to \( P^\Omega \) nor to \( \partial P \), where abscises are strictly superior to barycentre’s and ordinates are strictly inferior to barycentre’s.

Using our method, we are able to provide a high expression power to spatial relations that can be applied to describe images and formulate complex visual queries in several domains. For example, for Figure 1 that shows two different spatial situations between three salient objects \( a_1 \), \( a_2 \), and \( a_3 \), our method expresses the spatial relations as shown in Figure 3. The relations \( R(a_1, a_2) \) and \( R'(a_1, a_2) \) are equal but the relations \( R(a_1, a_3) \) and \( R'(a_1, a_3) \) are clearly distinguished. Similarly, we can express relations between \( a_2 \) and \( a_3 \) in both situations (Figure 7).

Moreover, our method allows combining both directional and topological relation into one binary relation, which is very important for indexing purposes. There are no directional and topological relations between two salient objects but only one spatial relation. Hence, we can propose a 1D-String to index images instead of 2D-Strings (Chang, Shi, and Yan 1987).

Discussion
Using our method, we can provide other type of relations. For instance, we can calculate spatio-temporal relations into one representation. For that, we define a new feature with 3 dimensions: 1 dimension for time representation, and 2 dimensions for 2D shape’s. Intersections sets are defined exactly as time and shape features’. We obtain 10 \((2+2)^3\) intersection sets that conduct to 10x10 intersection matrix cells in order to calculate spatio-temporal relation between two features.

Conclusion
In this paper, we presented our method to identify relations in image representation. We used spatial and temporal relations as a support to show how relations can be powerfully expressed within our method. This work aims to homogenize, reduce and optimize the representation of relations. We currently experiment its integration in our prototype MIMS (Chbeir, Amghar, and Flory 2001).

However, in order to study its efficiency and limits, our method requires more intense experiments in complex environment where great number of feature dimensions and salient objects exist. Furthermore, we will examine it in other type of media like videos.

References


Yoshitaka A. and Ichikawa T., “A Survey on Content-Based Retrieval for Multimedia Databases”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 11, No. 1, 1999, P. 81-93.