Spatial Reasoning and Retrieval of Pictures Using Relative Distance

Been-Chian Chien and Peir-Cheng Chieng

Department of Information Engineering
I - Shou University
1, Section 1, Hsueh-Cheng RD. Ta-Hsu Hsiang,
Kaohsiung, Taiwan, R.O.C.
E-Mail: cbc@csa500.isu.edu.tw

Abstract
In this paper, we present a new spatial reasoning method based on relative distance notations of spatial data. The new representation considers the spatial relationships including direction and distance. However, the previous researches provided directional relationships only. We represent an image with relative distance notations and produce a vector-like format. Some simple spatial reasoning rules can be obtained from our representation. It is easy to be understood in theoretical and implemented in practical. Furthermore, an efficient similarity retrieval algorithm can be developed. It provides a better precision than 2D-PIR and uses less storage space than the previous other representations.

1 Introduction
In multimedia database systems, an image information system is one of the most important issues. Applications using image database consist of office automation, computer-aided design, computer vision, remote sensing of earth resources and medical pictorial archiving, etc. An intelligent image database system allows for efficient, flexible retrieval of vast amounts of pictorial information and supports spatial reasoning. The image retrieval problem is concerned with retrieving images that are relevant to users’ requests from an image database. We assume that there is a database containing pictorial images and each image has been associated with some symbolic representation describing the contents of the image. The symbolic representation is generated by an image analysis algorithm, or manually, or by a combination of both in advance. The representation contains information about the objects in the image, such as the properties and the spatial relationships among the objects. These symbolic representing techniques help us that query no longer need to be specified in a textual form. We can take or build an image, and find the image with the best “similarity” from image databases. A query processing mechanism to determine which images will be retrieved from a database by the symbolic representations of images. There are many approaches had been proposed. The most famous representation of symbolic image is 2D-String[1] proposed by Chang et al. Recently, Nabil et al proposed an another representation called 2D-Projection Interval Representation (2D-PIR)[7,8].

In this paper, we will review the two previous approaches briefly and propose a new relationship representation based on relative distance. In our representation, a relative distance between two objects is used to describe the spatial relationships between two objects in an image instead of the spatial operators or direct descriptions of relationships. There are many advantages while using a relative distance as objects’ spatial relationship. First, the distance relationship can be embedded in the representation. The length of objects and the distance between two objects are obtained easily. Second, the relative distance representation can support all of the functions of spatial operators in 2D C-String and the relationships in 2D-PIR. Third, a simpler spatial reasoning algorithm and an efficient similarity retrieval algorithm can be developed. At last, the new notation uses less storage space than the previous other representations. We will show all above advantages in this paper.

The contents of this paper are organized as follows. Section 2 reviews the representations of 2D-String and 2D-PIR. The relative distance notations and the corresponding properties of new representation are introduced in Section 3. Then, Section 4 will show the spatial reasoning algorithm and similarity retrieval method based on the new representation. Finally, a short summary and future work are outlined.
2 Spatial Relationships and Representations of Iconic Images

The representation of 2D-String uses a symbolic projection of a picture along the x-axis and the y-axis. The first string is obtained by projecting all the symbols vertically denoted by u string, while the second one is obtained by projecting the same symbols horizontally denoted by v string. Three spatial operators “<”, “=” and “>” are provided in 2D-String. The operator “<” denotes the “left-right” or “below-above” spatial relationship. The operator “=” denotes the “at the same spatial location as” relationship. The last operator “>” stands for “in the same set as” relationship. Spatial relationships reasoned by 2D-String only 9 types are obtained, they are “east”, “west”, “south”, “north”, “southeast”, “northeast”, “southwest”, “northwest” and “the same location.” Similarity retrieval of images encoded as 2D-String is based on subsequence matching techniques. There are three types of subsequences, type-0, type-1 and type-2 are defined.

2D G-String and 2D C-String are the two main extensions of 2D-String. Generally speaking, 2D C-String provides a more efficient notation than 2D G-String and preserves the same spatial relationships. 2D C-String can support 13 spatial operators including “equal” (=), “before” (<), “meet” (|), “start” ([), “finish” (]), “overlap” (/), “during” (\), “contain” ([), “inside” (]), “touch” (\), “cover” (\), “overlap” (\), “after” (>) and “before” (<). The symbol “*” stands for “reverse”. That is, if A < B, it means B < A. Since the results of “equal” operator A = B and A = B are the same, it is denoted only one “=”. The spatial reasoning and similarity retrieval methods of 2D C-String are developed by the above 13 spatial operators. Of course, the corresponding algorithms are much more complicated than 2D-String, but there are 169 spatial relationships can be derived from 2D C-String, as shown in [4].

2D-PIR is a symbolic representation proposed by Nabil et al. in 1995[7] and 1996[8]. It adapts three existing representing formalisms (Allen’s temporal intervals, 2D-String, and topological relationships), and produces a unified representation. A 2D-PIR is defined as a triple (δ, χ, ψ). δ is a topological relationship; χ and ψ are the interval relationships projected along the x-axis and the y-axis respectively. A picture can be modeled by a 2D-PIR graph (V, R), where V is a finite non-empty set of symbols representing the objects in a picture, and R is a set of edges labeled by the 2D-PIR relationships. For δ, 8 topological relationships are possible, they are “disjoint” (δt), “equal” (δq), “touch” (δo), “contain” (ct), “inside” (in), “cover” (co), “overlap” (ov) and “covered by” (cb). For χ and ψ, the 13 interval relationships are similar to the definitions of 13 spatial operators in 2D String except for their representing symbols. We rewrite the interval relationships by the notations of 2D-PIR: “equal” (=), “before” (<), “after” (>), “meet” (m), “meet-inverse” (mi), “start” (s), “start-inverse” (si), “finish” (f), “finish-inverse” (fi), “overlap” (o), “overlap-inverse” (oi), “during” (di), “during-inverse” (di). The “after” is the inverse of “before”. Hence, the similarity matching of 2D-PIR becomes the graph matching problem. The similarity retrieval is measured by “degree of similarity” instead of the definition of type-i. The degree of similarity is dependent on the relationship distance between the two corresponding components in two distinct images. The 2D-PIR’s similarity retrieving approach can behave a better precision than the 2D C-String’s.

3 Relative Distance Notation and Image Representation

We assume that an image P consists of n objects identified by O1, O2, ..., On. After preprocessing by image processing and pattern recognition techniques, the n objects in the original image are recognized. Each object is enclosed by a minimum bounding rectangle(MBR) with boundaries parallel to horizontal (x) and vertical (y) axis. Let pxi and qxi be the two coordinates for left and right boundary of Oi’s MBR projected along x-axis, pxi < qxi, pyi and qyi is the coordinates for up and down boundary of Oi’s MBR projected along y-axis and pyi < qyi, where 1 ≤ i ≤ n. The coordinates of objects’ MBR in an image can use a popular measure to be a basic computing unit; for example, a pixel. The relative distance notations for a picture P containing n objects will be defined by the following definitions proposed in [2].

**Definition 1.** The x- and y-axis projection interval of Oi are IXi and IYi defined as IXi = qxi - pxi and IYi = qyi - pyi.

**Definition 2.** Let Oi and Oj are arbitrary two objects in a picture P. We define

\[
D_{x} = pxj - pxi \quad \text{and} \quad D_{y} = pyi - pvy.
\]

Dx and Dy are called the x- and y-axis distance between Oi and Oj referring to Oi, respectively.

**Definition 3.** Xij is the x-axis relative distance between Oi and Oj referring to Oi, i < j and

\[
Xij = (D_{x} / IXi, IYi / IXi).
\]

Yij is the y-axis relative distance between Oi and Oj referring to Oi, i < j and...
\[ Y_{ij} = (Dy_{ij}/Iy_i, Iy_j/Iy_i). \]

Above 3 definitions have some important properties. We describe and prove such properties in the following lemmas.

**Lemma 1.** The x- and y-axis distance are reversible; that is, \( Dx_{ij} = -Dx_{ji} \) and \( Dy_{ij} = -Dy_{ji} \).

**Proof.** From Definition 2, it is trivial that \( Dx_{ij} = px_{ij} - px_i = -(px_i - px_{ij}) = -Dx_{ji} \). \( Dy_{ij} = -Dy_{ji} \) can be derived in a similar way.

**Lemma 2.** If \( X_{ij} \) and \( Y_{ij} \) are defined as Definition 3, and representing \( X_{ij} = (\alpha_{ij}, \chi_{ij}) \), \( Y_{ij} = (\beta_{ij}, \psi_{ij}) \). Let \( X_{ij} = (\alpha_{ij}, \chi_{ij}) \) and \( Y_{ij} = (\beta_{ij}, \psi_{ij}) \), we have \( \alpha_{ij} = -\alpha_j/\chi_{ij}, \chi_{ij} = 1/\chi_{ij}, \beta_{ij} = -\beta_j/\psi_{ij}, \) and \( \psi_{ij} = 1/\psi_j \).

**Proof.** Since \( X_{ij} = (Dx_{ij}/Ix_i, Ix_j/Ix_i) \) and \( Y_{ij} = (Dy_{ij}/Iy_i, Iy_j/Iy_i) \), \( \alpha_{ij} = Dx_{ij}/Ix_i \), \( \chi_{ij} = Ix_j/Ix_i \). By Lemma 1, for \( X_{ij} \), we have \( \alpha_{ij} = Dx_{ij}/Ix_j \) and \( \chi_{ij} = Ix_j/Ix_i \). Hence,

\[
\alpha_{ij} = \frac{Dx_{ij}}{Ix_j} = 1/\chi_{ij}, \quad \chi_{ij} = \frac{Ix_j}{Ix_i}.
\]

It is the same for obtaining \( \beta_{ij} = -\beta_j/\psi_{ij} \), and \( \psi_{ij} = 1/\psi_j \).

Lemma 1 shows that \( Dx_{ij} \) are the reverse direction of \( Dx_{ji} \). Hence, the direction characteristics are embedded in the notation implicitly. The reversible property in Lemma 2 indicates that relationships between object \( O_i \) and \( O_j \) referred to \( O_i \) can be reasoned from the relationships referring to \( O_j \).

**Lemma 3.** For \( n \) objects, let \( 1 \leq i, j, k \leq n \), if \( X_{ij} = (\alpha_{ij}, \chi_{ij}) \), \( Y_{ij} = (\beta_{ij}, \psi_{ij}) \) and \( X_{ik} = (\alpha_{ik}, \chi_{ik}) \), \( Y_{ik} = (\beta_{ik}, \psi_{ik}) \), then \( X_{jk} = (\alpha_{jk}, \chi_{jk}) \), \( Y_{jk} = (\beta_{jk}, \psi_{jk}) \).

**Proof.** For \( X_{jk} \) from Definition 3, we can obtain \( \alpha_{jk} = Dx_{jk}/Ix_j, \chi_{jk} = Ix_j/Ix_k \) and \( \alpha_{ik} = Dx_{ik}/Ix_k, \chi_{ik} = Ix_i/Ix_k \).

By Definition 2 and 3,

\[
\alpha_{jk} = \frac{Dx_{jk}}{Ix_j} = \frac{Dx_{ik} - px_i}{Ix_j} = \frac{(px_i - px_i + px_i)}{Ix_j} = \frac{\{(Dx_{ik} - Dx_{ik})/Ix_j\}}{Ix_j/Ix_k} = \frac{\{\alpha_{ik} - \alpha_{ik}\}}{\chi_{jk}} = \frac{\{\alpha_{ik} - \alpha_{ik}\} \chi_{jk}}{\chi_{jk}}.
\]

It is the same for obtaining \( \beta_{jk} = (\beta_{jk} - \beta_{jk})/\psi_{jk} \) and \( \psi_{jk} = \psi_k/\psi_j \).

Lemma 3 gives an important transitive property that we can derive new relationships from two indirect relationships belonging to different objects. By this property, we will reduce the quantity of stored information. The next lemma shows our image representation.

**Lemma 4.** The relative distance for an image \( P \) containing \( n \) objects can be represented by \( (O_i, rx_i, ry_i) \), where \( rx_i = Ix_j, \) \( ry_i = Iy_j, 2 \leq j \leq n \) and \( 1 \leq i \leq n \).

**Proof.** An image representation should be able to describe all relationships among the objects in the image. That is, we can find the relationships between two arbitrary objects in the representation without information lost. There are \( n(n+1)/2 \) relationships totally in an image containing \( n \) objects. The information of an image we stored in the representation is \( (O_i, rx_i, ry_i), (O_2, rx_2, ry_2), (O_3, rx_3, ry_3), \ldots, (O_n, rx_n, ry_n) \). The detailed replacements with the symbols of definitions are \( (O_i, Ix_j, Iy_j), (O_2, X_{12}, Y_{12}), (O_3, X_{13}, Y_{13}), \ldots, (O_n, rx_n, ry_n) \). Thus, we can derive any relationships among the \( n \) objects by Lemma 2 and Lemma 3 based on object \( O_j \) in two computing steps at most.

### 4 Spatial Reasoning and Similar Image Retrieving

Our representation in Section 3 also can be used to stand for the 13 relationships illustrated in 2D C-String and 2D-PIR. We describe the corresponding relationships in Table 1. The boundary of objects in our representation is...
represented by $\alpha$, $\alpha + \chi$, $\beta$ and $\beta + \psi$. The 169 types of spatial relationships in 2D space will be obtained by comparing the values of $\alpha$, $\beta$, $\chi$ and $\psi$. All queries supported by 2D C-String and 2D-PIR can also be provided by our representation. Furthermore, the performance of reasoning is more effective than the two previous representations. In addition to the spatial queries, our representation can support the distant queries, such as “find the pictures with closest distance between $O_i$ and $O_j$”, “find the pictures with the farthest distance between $O_i$ and $O_j$”, and “find the pictures with the same distance between $O_i$ and $O_j”, etc. The distance reasoning method is similar to the spatial reasoning. We define the distance between two objects $O_i$ and $O_j$ in an image as the following formula:

$$Distance = \sqrt{(Dx_i - Dx_j)^2 + (Dy_i - Dy_j)^2}.$$  

For supporting similarity retrieval, we provide a similarity matching algorithm as 2D-PIR’s. Let $O_i$ and $O_j$ be two arbitrary objects in a picture $P$, $r_{ij}$ is the relative distance between $O_i$ and $O_j$. The similarity degree of two different pictures is dependent on the value of difference between two pictures[2]. The value of difference for the same objects $O_i$ and $O_j$ in two different pictures $P_1$ and $P_2$ can be defined as the measurement functions in [3]. The total difference is computed by the following equation:

$$D(P_1, P_2) = \sum_{i=1}^{n} \sum_{j=1}^{n} D(r_{ij}^{P_1}, r_{ij}^{P_2}).$$  

The value of $D(P_1, P_2)$ is the difference of pictures $P_1$ and $P_2$. A smaller value of $D$ means that two pictures are more similar. It supports a better precision than 2D C-String and 2D-PIR in similarity retrieval. The matching algorithm is as follows:

**Algorithm:** Picture similarity matching  
**Input:** two representing pictures $P_1$ and $P_2$.  
**Output:** the difference value between $P_1$ and $P_2$  
**begin**  
set $D = 0$ /* different value = 0 */  
if $P_1$ is a subset of $P_2$ then  
for each object $O_i \in P_1$ and $P_2$  
compute the difference $D(P_1, P_2)$  
end for  
return $D$  
else  
return $\infty$ /* different pictures */  
end if  
**end**

5 Conclusions

The access and management of image data is an important task in multimedia systems. Spatial relationship is one of the main features in an image. The directional and distant relationships among objects in an image are the two key spatial relationships. Although the 2D-String and 2D-PIR had discussed the directional relationships, the effect of distant factor has not proposed yet so far. We propose a new representation for spatial relationships based on relative distance notation. We prove that our new representation can support spatial reasoning and similarity retrieval easily without loss the generality. The new method also saves the storage space of iconic information and improves the performance of image retrieval.

The extensions and applications of our work are vast. 2D image notations can be extended to 3D environments easily. The three dimensional version should have the similar representation but consist of more spatial relationships. An another extension is in video data. Video data are organized by a sequence of related images called frames. Our representation is also suitable to describe the change of a sequence of images in temporal. Other future works include indexing structure, precision of similarity retrieval and applications.
<table>
<thead>
<tr>
<th>operators</th>
<th>Relationships</th>
<th>$X_{ij}$</th>
<th>$Y_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;</td>
<td></td>
<td>$\alpha_{ij} &gt; 1$</td>
<td>$\beta_{ij} &gt; 1$</td>
</tr>
<tr>
<td>=</td>
<td></td>
<td>$\alpha_{ij} = 0$ and $\alpha_{ij} + \chi_{ij} = 1$</td>
<td>$\beta_{ij} = 0$ and $\beta_{ij} + \psi_{ij} = 1$</td>
</tr>
<tr>
<td>/</td>
<td></td>
<td>$\alpha_{ij} = 1$</td>
<td>$\beta_{ij} = 1$</td>
</tr>
<tr>
<td>%</td>
<td></td>
<td>$0 &lt; \alpha_{ij} &lt; 1$ and $\alpha_{ij} + \chi_{ij} &gt; 1$</td>
<td>$0 &lt; \beta_{ij} &lt; 1$ and $\beta_{ij} + \psi_{ij} &gt; 1$</td>
</tr>
<tr>
<td>[</td>
<td></td>
<td>$\alpha_{ij} = 0$ and $\alpha_{ij} + \chi_{ij} &gt; 1$</td>
<td>$\beta_{ij} = 0$ and $\beta_{ij} + \psi_{ij} &gt; 1$</td>
</tr>
<tr>
<td>]</td>
<td></td>
<td>$\alpha_{ij} &lt; 0$ and $\alpha_{ij} + \chi_{ij} = 1$</td>
<td>$\beta_{ij} &lt; 0$ and $\beta_{ij} + \psi_{ij} = 1$</td>
</tr>
<tr>
<td>)*</td>
<td></td>
<td>$\alpha_{ij} &lt; 0$ and $\alpha_{ij} + \chi_{ij} &lt; 0$</td>
<td>$\beta_{ij} &lt; 0$ and $\beta_{ij} + \psi_{ij} &lt; 0$</td>
</tr>
<tr>
<td>)*</td>
<td></td>
<td>$\alpha_{ij} = 0$ and $\alpha_{ij} + \chi_{ij} = 1$</td>
<td>$\beta_{ij} = 0$ and $\beta_{ij} + \psi_{ij} = 1$</td>
</tr>
<tr>
<td>)*</td>
<td></td>
<td>$\alpha_{ij} &lt; 0$ and $\alpha_{ij} + \chi_{ij} = 0$</td>
<td>$\beta_{ij} &lt; 0$ and $\beta_{ij} + \psi_{ij} = 0$</td>
</tr>
<tr>
<td>*/</td>
<td></td>
<td>$\alpha_{ij} &lt; 0$ and $0 &lt; \alpha_{ij} + \chi_{ij} &lt; 1$</td>
<td>$\beta_{ij} &lt; 0$ and $0 &lt; \beta_{ij} + \psi_{ij} &lt; 1$</td>
</tr>
<tr>
<td>%*</td>
<td></td>
<td>$0 &lt; \alpha_{ij} &lt; 1$ and $\alpha_{ij} + \chi_{ij} &lt; 1$</td>
<td>$0 &lt; \beta_{ij} &lt; 1$ and $\beta_{ij} + \psi_{ij} &lt; 1$</td>
</tr>
<tr>
<td>[*</td>
<td></td>
<td>$\alpha_{ij} = 0$ and $\alpha_{ij} + \chi_{ij} &lt; 1$</td>
<td>$\beta_{ij} = 0$ and $\beta_{ij} + \psi_{ij} &lt; 1$</td>
</tr>
<tr>
<td>]*</td>
<td></td>
<td>$0 &lt; \alpha_{ij} &lt; 1$ and $\alpha_{ij} + \chi_{ij} = 1$</td>
<td>$0 &lt; \beta_{ij} &lt; 1$ and $\beta_{ij} + \psi_{ij} = 1$</td>
</tr>
</tbody>
</table>

Table 1: The 13 basic relationships for spatial reasoning.

6 References


