Shape-Based Image Retrieval Using Spatio-Temporal Relation Computations

Huan-Chao Keh and Timothy K. Shih

Multimedia Information NEtwork (MINE) Lab
Department of Computer Science and Information Engineering
Tamkang University
Tamsui, Taiwan, R.O.C.
E-mail: tshih@cs.tku.edu.tw

Abstract

Content-based retrieval of multimedia information is one of the most difficult research topics in multimedia computing and information retrieval. In this paper, we present a visual system which allows content-based retrieval of still image. The recognition algorithms we used are based on spatio-temporal relations. Two approaches and algorithms were developed based on the similarity between polygons. The system is incorporated with a visual interface which allows the user to specify polygons as the shape specification of pictures. The preliminary experience shows that, within an image database containing about 300 bitmapped images, the system is able to retrieve correct information of a high satisfaction.

Key Words **:** Temporal Interval Relations, Content-Based Retrieval, Multimedia, Spatio-Temporal Model, Polygon Matching, Information Retrieval

1. Introduction

Content-based information retrieval has become one of the most interesting research topic as the importance of multimedia database development realized by many researchers in the multimedia community [15, 12]. However, due to the difficulty of developing effective and efficient methods to match a database query against its target image or video object, yet, there remains many unsolved problems [19]. Content-based retrieval of still image relies on the color, the texture, or the shape of an object, as well as relies on the relative positions of a number of objects [11, 4, 12, 5, 19, 16, 13, 10]. If the retrieval is to consider objects in a video [4, 19], which can be treated as a sequence of pictures and sound, the timing of object movement is also an important factor. Content-based retrieval researches also include finding information in sound or music files.

The general concepts of Content-Based Image Retrieval (CBIR) can be found in [5]. Many research projects are also found in the literature. The QBIC system [4] is an image/picture retrieval system allowing the user to search for objects in a large database, based on color, shape, texture, and sketches. Image or picture objects are pre-processed by a feature extraction module (i.e., the database population module). Therefore, the query interface module is able to retrieve objects based on their contents. Another paper [12] proposes a multimedia database platform (named GOLS) which adopts image or real-time video understanding techniques for video content-based retrieval. The model of recognition is implemented by the Video Scene Description Language (VSDL) and a state transition mechanism. The implemented system is able to assist content-based editing of TV drama. The informedia project [19] proposes an on-line video library with a retrieval mechanism based on video contents and knowledge. Several techniques used in the project include speech understanding, image processing and natural language understanding. Unsolved problems with respect to each of these techniques were discussed. A content-based retrieval system for human faces is proposed in [16]. Based on a photo and its caption, the Piction system is able to analyze image content and generate constraints, such as spatial, characteristic, and contextual information. The system is tested on a set of pictures and captions from newspaper photographs and performs reasonably well.

In the development of a multimedia information system, we have proposed several sub-systems including a authoring system, a multimedia database, a multimedia network framework, and several other tools. Among these sub-systems, the multimedia database project [15] serves as the core system for its many applications. Recently, we have developed a set of algorithms and a system on the top of our database system. The algorithms rely on temporal interval relations [1] for automatic picture searching. Based on the algorithms, our image content-based retrieval system serves as the front end interface tool to our multimedia database project.

The importance of knowledge underlying temporal interval relations was found in many disciplines. As pointed out in [1], researchers of artificial intelligence, linguistics, and information science use temporal intervals as a time model for knowledge analysis. For instance, in a robot planning program, the outside world is constantly changed according to a robot's actions. The notion of “number three box is on the left of number two box" is true only within a temporal interval (assuming that the robot is given a command to move a number of boxes to a destination). As other examples, linguisticians use temporal intervals to model tense information of a sentence. And information process center uses temporal intervals to process queries such as “who is the president of company X within the period of March 1, 1994 to January 31, 1995?" The work discussed in [1] analyzes the relations among temporal intervals. This research contribution [1] has been used in many temporal modeling of multimedia systems including ours [14, 18]. However, the work [1] only states temporal interval relations. No spatial relations were discussed. We found that these temporal relations can be generalized for spatial modeling.

Many researchers propose temporal modeling of multimedia objects. The work discussed in [3] presents a framework for data modeling and semantic abstraction of image/video data. Seven generalized n-ary relations were used to describe the temporal relationships among n objects. The authors also defined spatial events in terms of these n-ary relations. Temporal events were then specified in terms of these spatial events. However, there was no discussion of the conflict situation among relations. A functional model extends media segments to include executable programs, live media streams and the links among them was proposed in [6]. The work also provides a set of operators to combine intervals. Based on Allen's temporal interval algebra, a set of directional and topological spatial relations were addressed in [7]. The authors also provided a set of spatial inference rules for automatic deduction. A methodology for spatial and temporal object composition under a distributed multimedia environment was proposed in [8]. A set of n-ary temporal relations with their temporal constraints were discussed in [9], which is an early result of the work addressed in [3]. The temporal model of reverse play and partial interval evaluation for midpoint suspension and resumption were also discussed. Algorithms for accessing objects in a database were presented. The work in [18] introduced a spatial and temporal model for actors of multimedia applications. A hundred and sixty nine directional relationships between two 2-D object bounding boxes were proposed. A number of spatial and temporal operators were used to compose a presentation. The composition mechanism was defined in an extended BNF formal syntax definition. However, no discussion of the conflict situation among relations were found. Efficient indexing schemes based on the R-tree spatial data structure were proposed in [17]. The mechanism handles 1-D and 2-D objects, as well as 3-D objects which treat a time line as the third axis. The discussion in [20] identified various temporal interaction forms and discussed their temporal dependencies. The work also integrated some interaction forms into a temporal model. The 2D Projection Interval Relationships [11] (2D-PIR) is based on both directional and topological temporal relations. Image retrieval algorithms were discussed based on PIR. The system is able to handle rotated and reflected images. However, the matching mechanism focuses on the relative positions among a set of objects instead of searching on a single object of a particular shape.

We have surveyed many researches related to the spatio-temporal semantics of multimedia objects. We found that, the use of spatio-temporal relations serves as a reasonable semantic tool for the underlying representation of objects in many multimedia applications. In this paper, we extend Allen's work by means of a complete analysis of all temporal relation domains. These domains are also extended for spatio computation. We then propose that, the similarity between spatio relations can be used to compare the shapes of two images. Therefore, our image content-based information retrieval system is developed.

This paper is organized as the following. We propose a generalized temporal relation model based on Allen's work [1] in section 2. Interested readers should look at the 18 relations proposed in the section in order to understand the definition of relation similarity, which is discussed in section 3. In section 4, based on relation similarity, we propose two algorithms to compute polygon similarity. These algorithms are used to construct a picture retrieval system, which is discussed in section 5. Finally, we summarize our contributions and future directions in section 6.

2. A Generalized Model of Temporal/Spatial Relations

James F. Allen [1] introduced 13 binary relations for temporal intervals. These relations cover all possible cases of the bibary relation between an arbitrary pair of line segments lie on a 2-D line. However, considering that a line segment might have a length of zero, Allen's temporal interval relations can be generalized. We introduce five new relations. To analyze a generalized model of spatio-temporal relations, we consider the following situations. For an arbitrary pair of points, *A* and *B*, located on a 1-dimensional line, there are three possibilities: $A < B$, $A = B$, or $A > B$, for *A* is before *B*, *A* is at the same position as *B*, and *A* is after *B*, respectively. If these two points are located on a 2-dimensional plan, there exists nine (i.e., 3*3) cases. The X and the Y coordinates of these two points on the plan are independent. The possible relations between these two points on a plan can be denoted as $A(<, <)B$, $A(<, =)B$, $A(<, >)B$, ..., $A(>,>)B$, where the first element in in the pair representing a relation denotes the order on the X coordinate while the second is for the Y coordinate. Now, considering two line segments located on a 1-dimensional line, or on a 2-dimensional plan, the situation becomes complicated. Since each line segment has a starting point and an ending point, we analyze the temporal/spatial relations of two line segments according to these points. Allen's research is a special case of two line segments on the 1-dimensional line, with each line segment of length greater than zero.

Suppose *As* and *Ae* are the starting and ending points of the line segment *A*. And, *Bs* and *Be* are those of *B*. We define a binary relation, <.>, (either <, =, or > for "*A* is before *B*", "*A* is the same as *B*", or "*A* is after *B*") of two points. The 13 relations introduced by Allen make the binary relations hold in the first part of the following table:

Table 1. Starting and ending point relations

As <.> Bs	As <.> Be	Ae <.> Bs	Ae <.> Be	Relations
<	<	<	<	<
>	>	>	>	>
>	<	>	<	d
<	<	>	>	di
<	<	>	<	o
>	<	>	>	oi
<	<	=	<	m
>	=	>	>	mi
=	<	>	<	s
=	<	>	>	si
>	<	>	=	f
<	<	>	=	fi
=	<	>	=	e
=	=	>	>	10s
<	<	=	=	10e
=	<	=	<	01s
>	=	>	=	01e
=	=	=	=	00

The second part of the table has five special cases. For instance, we use *A*10*sB* to represent *A* is a line and *B* is a point (i.e., the starting and the ending points are located at the same position), where *A* and *B* meets at the starting point of *A*. These five special cases were not considered in [1]. The situations of points of two line segments could have upto $3^4 = 81$rows in the above table. However, except for the 18 cases, others are conflict situations (i.e., it is physically impossible for the situation to occur).

Now, considering two line segments on a 2-dimensional plan, according to the above table and since the position of these two line segments are independent at the X and the Y coordinates, there exists $18^2 = 324$ possible relations between these two line segments on a plan. These relations, similar to those of two points on a plan, are denotes by pairs as: (<, <), (<, >), (<, *d*), (<, *di*),..., (00, 01*e*), and (00, 00). We use these 324 binary relations to model our temporal/spatial relations of multimedia objects. Conclusively, the generalized model of temporal/spatial relations include four cases:

- two points on a line
- two points on a plan
- two line segments on a line
- two line segments on a plan

The first two cases are not used in our project simply because they are trivial and they do not efficiently express intervals. The third case can represent the relations of two temporal intervals. The fourth case is the semantic tool that we rely on to develop our temporal/spatial computation mechanism.

3. Relation Similarity

Relations are similar to each other in certain degree. For example, "during" and "starts" are similar since the only difference is the starting points of the two intervals are different. However, "before" and the inverse of "meets" are not quite the same. If we consider two line segments on the plan, based on the relative positions of two lines, it is feasible to construct a mechanism to compare the similarity between two polygons since polygons are made of line segments. Therefore, an evaluation mechanism is necessary to compute *relation similarity*.

Table 2. Similarity relation sets of indices from 1 to 4

Rel.	SRS of index 1	SRS of index 2
<	{ o,m }	{ d,di,s,fi,10e,01s }
>	{ oi,mi }	{ d,di,si,f,10s,01e }
d	{ o,oi,s,f }	{ <,>,di,m,mi,si,fi,e,01s,01e }
di	{ o,oi,si,fi }	{ <,>,d,m,mi,s,f,e,10e,10s }
o	{ <,d,di,m,s,fi }	{ oi,si,f,e,10e,01s }
oi	{ >,d,di,mi,si,f }	{ o,s,fi,e,10s,01e }
m	{ <,o,10e,01s }	{ d,di,s,fi }
mi	{ >,oi,10s,01e }	{ d,di,si,f }
s	{ d,o,si,e,01s }	{ <,di,oi,m,f,fi,10s }
si	{ di,oi,s,e,10s }	{ >,d,o,mi,f,fi,01s }
f	{ d,oi,fi,e,01e }	{ >,di,o,mi,s,si,10e }
fi	{ di,o,f,e,10e }	{ <,d,oi,m,s,si,01e }
e	{ s,si,f,fi }	{ d,di,o,oi,10e,01s,00,10s,01e }
10s	{ si,mi }	{ di,s,e,00,oi,01e,> }
10e	{ m,fi }	{ <,o,di,01s,e,00,f }
01s	{ m,s }	{ <,10e,o,e,si,00,d }
01e	{ f,mi }	{ fi,e,00,10s,d,oi,> }
00	{ }	{ 10e,01s,e,10s,01e }

Rel.	SRS of index 3	SRS of index 4
<	{ oi,si,f,e }	{ >,mi,00,10s,01e }
>	{ o,s,fi,e }	{ <,m,10e,01s,00 }
d	{ 10e,10s }	{ 00 }
di	{ 01s,01e }	{ 00 }
o	{ >,mi,10s,01e }	{ 00 }
oi	{ <,m,01e,01s }	{ 00 }
m	{ oi,si,f,e,00 }	{ >,mi,10s,01e }
mi	{ o,s,fi,e,00 }	{ <,m,10e,01s }
s	{ >,mi,10e,00,01e }	{ }
si	{ <,m,10e,00,01e }	{ }
f	{ <,m,01s,00,10s }	{ }
fi	{ >,mi,01s,00,10s }	{ }
e	{ <,>,m,mi }	{ }
10s	{ o,fi,01s,d,f }	{ <,10e,m }
10e	{ s,si,d,oi,01e }	{ 10s,mi,> }
01s	{ fi,di,10s,f,oi }	{ 01e,mi,> }
01e	{ 10e,o,di,s,si }	{ <,m,01s }
00	{ m,fi,s,si,f,mi }	{ <,o,di,d,oi,> }

In section 2, each of the 18 temporal relations is defined by four “< . >” relations. These relations can be used as a base of our evaluation criterion. Let's starts with some definitions. An incompatible difference of two “< . >” relations belong to two different temporal relations occurs if those two temporal relations hold different relations in the same column of table 1. A similarity relation set (SRS) defined with respect to a temporal relation *r* of index *n* is a set containing temporal relations which have *n* incompatible differences from *r*. Since there are four < . > relations used in each temporal relation, the index of a SRS is from 1 to 4[1]. Note that, the lower the index, the closer the relation to the relations in its SRS. The following table shows the SRSs of different indices:

The index of SRS with respect to each temporal relation in table 1 can also be retrieved from the length of a shortest path in a distance graph (see figure 1). For example, relations *d* and *mi* have a SRS index (i.e., the number of incompatible differences) of 2. Note that, in the distance graph, the shortest path between an arbitrary pair of nodes (i.e., relations) has a length between 1 and 4 since the distance graph is to represent the similarity of relations of two line segments on a line. Note that, in the distance graph, there are five special edges illustrated in thick curved lines (i.e., from relation 00 to 10*e*, 01*s*, *e*, 10*s*, and 01*e*). This is due to the fact that, there exists no SRS element of index 1 of the 00 relation. A thick curved line has a weight of 2, instead of 1 as to the regular cases.

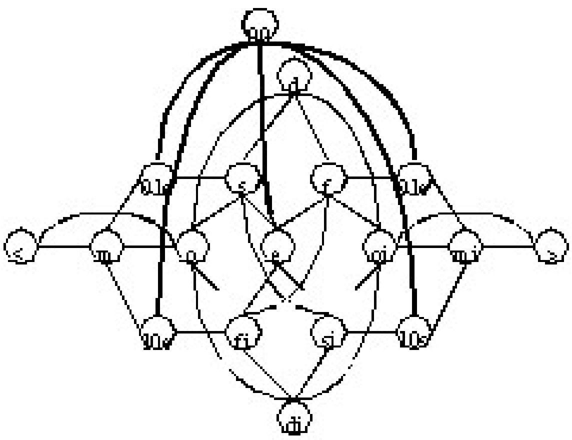

Figure 1. Distance graph of temporal relations

Table 1 describes only the SRS of two line segments on a line. The index of SRS of relations between two line segments on a plan (which hold a spatial relation) ranges from 1 to 8, since the X and the Y coordinates are independent from each other and upto 8 “< . >” relations are used (i.e., four for the X and another four for the Y coordinates). To compute the similarity relation sets of two line segments on a plan is not difficult at all. The algorithm relies on the SRSs of relations of two line segments on a line. Let 18*REL* denote the set of 18 relations of two line segments on a line, and 324*REL* denote the one of two line segments on a plan. Obviously, 324*REL* = 18*REL* × 18*REL*. Figure 2 illustrates the relation between 18*REL* and 324*REL*.

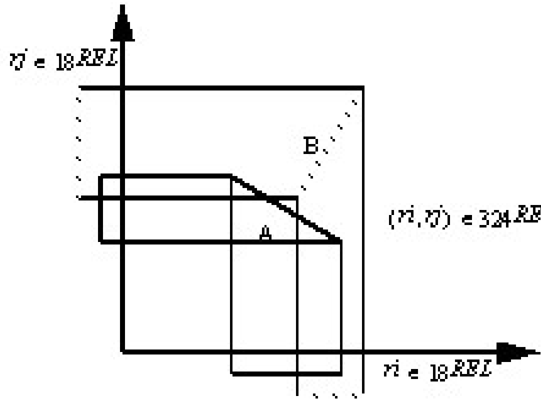

Figure 2. Relation between 18*REL* and 324*REL*

Also, let $ind18(r_i\ ,\ r_j)$ be a SRS index function takes as input two relations, $r_i\ ,\ r_j \in 18REL$, and returns a similarity index from 0 to 4. And, $SRS18(r_i;\ d)$is the SRS of r_i of index d. We have:

$$d = ind18(r_1\ ,\ r_2) \Leftrightarrow,$$
$$r_2 \in SRS18\ (r_1\ ,\ d) \vee r_1 \in SRS18\ (r_2\ ,\ d)$$

Let $ind324\ ((r_i\ ,\ r_j)\ ,\ (r_k\ ,\ r_l))$ and $SRS324((r_i\ ,\ r_k)\ ,\ d)$ be those of 324*REL*, we have:

$$\begin{aligned}
&\forall\ r_1, r_2, r_3, r_4 \in 18REL \bullet\\
&\quad (r_1\ ,\ r_3)\ ,\ (r_2\ ,\ r_4) \in 324REL \bullet\\
&\quad d = ind324\ ((r_1\ ,\ r_3)\ ,\ (r_2\ ,\ r_4))\\
&\qquad = ind18(r_1\ ,\ r_2) + ind18(r_3\ ,\ r_4)\\
&\qquad \Leftrightarrow\\
&\quad (r_2 \in SRS18\ (r_1) \wedge r_4 \in SRS18\ (r_3) \Rightarrow\\
&\qquad (r_2\ ,\ r_4) \in SRS324\ ((r_1\ ,\ r_3)\ ,\ d) \vee\\
&\quad r_1 \in SRS18\ (r_2) \wedge r_4 \in SRS18\ (r_3) \Rightarrow\\
&\qquad (r_1\ ,\ r_4) \in SRS324\ ((r_2\ ,\ r_3)\ ,\ d) \vee\\
&\quad r_2 \in SRS18\ (r_1) \wedge r_3 \in SRS18\ (r_4) \Rightarrow\\
&\qquad (r_2\ ,\ r_3) \in SRS324\ ((r_1\ ,\ r_4)\ ,\ d) \vee\\
&\quad r_1 \in SRS18\ (r_2) \wedge r_3 \in SRS18\ (r_4) \Rightarrow\\
&\qquad (r_1\ ,\ r_3) \in SRS324\ ((r_2\ ,\ r_4)\ ,\ d))
\end{aligned}$$

[1] A SRS of relation *r* of index 0 is an atomic set containing relation *r* itself.

Similarity relation sets are important concepts used in the polygon matching mechanism we developed in this paper.

3.1 Representation of Relation Computation

Even it is easy to compare the similarity of two relations by looking at the SRSs, however, it is too time consuming to search an element in the SRSs. We have developed a fast computation mechanism to solve this problem. Using a bit-slicing representation of the relations, the index of polygon similarity can be computed in a few operations.

We use 10, 00, and 01 to represent the comparisons of end points of a line segment, >, =, and <, respectively. Eight bits are used to represent a relation of two line segments on a line. And 16 bits are used to represent a relation of two line segments on a plan. For instance, the two relations, *d* and *mi*, in the starting and ending point relation table in section 2 can be represented as 10 01 10 01 and 10 00 10 10, respectively. An exclusive or logical operation is used to compute the similarity syndrome, which represents the difference:

Temporal Relations	Starting and Ending Point Relations	Bit Codes
d	> < > <	10 01 10 01
mi	> = > >	10 00 10 10
		---- XOR ----
		00 01 00 11
		(similarity syndrome)

An integer array, A, of 2^8 or 2^{16} elements (for either temporal or spatial relation, respectively) is used to simulate the SRS index function (i.e., $ind18(r_i, r_j)$, or $ind324((r_i, r_j), (r_k, r_l))$). The similarity syndromes (represented as unsigned integers) are indices to the array. For instance, A(00 01 00 11) = 2, which represents that the two relations, *d* and *mi*, have two incompatible differences at the second and the fourth "< . >" relation. Therefore, the index of similarity is computed by one logical operator and one direct table lookup. The efficiency is very important in computing polygon similarity, as discussed in the next section.

4. Polygon Similarity

Polygon similarity can be derived by comparing spatial relations of edges belong to two different polygons. Since polygons are the fundamental element to represent computer graphics images, it is feasible that, polygon similarity can be used as an evaluation criterion for image content-based retrieval. The matching mechanism developed based on SRSs has the advantage of precisely detecting objects that are operated by the combinations of the following computer graphics operations: 1. rotation, 2. shape intact scaling, and 3. translation. However, the following operations make an object hard to be recognized: 1. metamorphic transformation, 2. slanting, and 3. shape altered scaling.

In a multimedia resource database that we built [15], pictures are bitmapped images associated with shap representation polygons (SRPs). An image processing mechanism is used to compute the shapes of objects in a picture. The outcome is then adjusted by a database administrator. Each picture has a set of SRPs. It is this SRP sets that we based on to compute polygon similarity between a query polygon (QP) and those in the SRP sets. In this section, two polygon matching mechanisms based on SRSs are discussed. The first focuses on the analysis of two consecutive sides of a polygon. The second focuses on the analysis of the interior structures of polygons.

4.1 Similarity on Consecutive Edges

Suppose the query polygon has *n* sides and a candidate shape representation polygon has *m* sides, where *n* and *m* are not necessarily equal. We have

$$QP_1\ QP_{r.1}\ QP_2\ ,\ QP_2\ QP_{r.2}\ QP_3\ ,\ldots,\ QP_{n\text{-}1}\ QP_{r.n\text{-}1}\ QP_n\ ,\ \text{and}\ QP_n\ QP_{r.n}\ QP_1$$

where QP_is, $1 \le i \le n$, are the *n* sides of the query polygon, and $QP_{r.i}$s, $1 \le i \le n$, are relations of two line segments on a plan. In the same way, we have

$$SRP_1\ SRP_{r.1}\ SRP_2\ ,\ SRP_2\ SRP_{r.2}\ SRP_3\ ,\ldots,\ SRP_{m\text{-}1}\ SRP_{r.m\text{-}1}\ SRP_m\ ,\ \text{and}\ SRP_m\ SRP_{r.m}\ SRP_1$$

where SRP_is and $SRP_{r.i}$s are those for the candidate shape representation polygon. Moreover, assuming that spatial relation $r_i = (r_{ix}, r_{iy}) \in 324REL$ holds for line segments *A* and *B*, and A_s, A_e, B_s, and B_e are the starting and ending points of the two line segments on a plan, we want to define a length ratio function, $LR(r_{ix})$ and $LR(r_{iy})$:

$$LR(r_{ix}) = (A_e.x - A_s.x) / (B_e.x - B_s.x) \wedge$$
$$LR(r_{iy}) = (A_e.y - A_s.y) / (B_e.y - B_s.y)$$

where $A_s.x$ and $A_s.y$ are the X and the Y coordinates of the starting point of line segment A (other notations are prepresented in a similar manner). Let function $sim(r_i\ ,\ r_j)$ be a similarity function, which takes as input two relations, r_i , and r_j , and returns a similarity:

$$r_i = r_j \Rightarrow sim(r_i , r_j) = (LR(r_{ix}) + LR(r_{iy})) - (LR(r_{jx}) + LR(r_{jy})) \vee$$
$$r_i \neq r_j \Rightarrow sim(r_i , r_j) = ind324(r_i\ ,\ r_j) * ((LR(r_{ix}) + LR(r_{iy})) - (LR(r_{jx}) + LR(r_{jy}))$$

The similarity function, $sim(r_i , r_j)$, estimates the similarity between two line segments on a plan based on similarity index and the length ratio function. Based on the similarity function, we construct the polygon similarity function. Polygons are represented as sets of line segments (polygon sides):

$$QP = \{ QP_1 , QP_2 , \ldots, QP_n \}$$
$$SRP = \{ SRP_1 , SRP_2 , \ldots, SRP_m \}$$

The polygon similarity function, $psim(QP,SRP)$, takes as input the query polygon and the candidate shap representation polygons, and returns an integer:

$$n \geq m \Rightarrow psim(QP, SRP) = \sum_{i=1}^{m} sim(QP_{r.i} , SRP_{r.i}) + (n - m) * 8 \vee$$
$$n < m \Rightarrow psim(QP, SRP) = \sum_{i=1}^{m} sim(QP_{r.i} , SRP_{r.i}) + (m - n) * 8$$

Given the coordinates of points, it is easy to compute the relations of two consecutive sides of a polygon. Therefore, the sets of $QP_{r.i}$ s and $SRP_{r.i}$ s are computed. However, the mechanism decides the first relation in a sequence of n or m relations plays an important role in the computation accuracy of polygon similarity. This is due to that, a candidate shap representation polygon might be rotated in the image database. A brute force approach to this problem is to fix a relation of the query polygon, and to iterate through each relation in the candidate shap representation polygon as the first relation. A branch-and-bound strategy can be used to select the minimum similarity out of the iterations. Another approach is to apply a heuristic mechanism to detect an "irregular" adjacency of two sides of a polygon, such as the one creates a sharp convex or concave angle. The relation between these two adjacent sides can be chosen as the first relation.

If the number of sides of the candidate shap representation polygon and the query polygon are different, the extra sides are completely mismatch (i.e., the similarity index is equal to 8). However, it is possible to improve the accuracy of matching. Assuming that, polygon P is either the query polygon or the candidate shap representation polygon, whichever has a larger number of sides. If the number of extra sides is n, we can simplify the first n pairs of consecutive sides with a low slope difference. That is, suppose the two consecutive sides are A and B, and the slope difference between A and B is low, we can eliminate the adjacent vertex of A and B, by connecting the other two points of A and B. Therefore, A and B are merged. After the first n pairs of consecutive sides are simplified, the two polygons have the same number of sides. The preliminary experience shows that, the accuracy of matching is improved.

4.2 Similarity on Structures

The first algorithm considers relations between two consecutive sides. However, the interior structure (i.e., the relations between each pair of sides) is not considered. It is possible that a query polygon is sheared, or scaled with shape altered. But the index of change does not affect the relations between pairs of consecutive sides. This situation makes the first matching algorithm erroneous.

For a polygon with n sides, there are $C(n , 2)$ (i.e., the combination) such relations can be used to describe the interior structure. Instead of using consecutive sides, our second algorithm uses the $C(n , 2)$ relations. Let the query polygon and the shape representation polygon have n and m sides, respectively:

$$QP = \{ QP_1 , QP_2 , \ldots, QP_n \}$$
$$SRP = \{ SRP_1 , SRP_2 , \ldots, SRP_m \}$$

We want to redefine the polygon similarity function, $psim$. Firstly, we need a function to decide the order of relations. Suppose polygon P has n sides, we define the lexicographical order function, $LO(P)$, to return an ordered list which contains $C(n , 2)$ elements:

$$(r_{1,2} , r_{1,3} ,\ldots, r_{1,n} , r_{2,3} , r_{2,4} ,\ldots, r_{2,n} ,\ldots, r_{n\text{-}1,n})$$

Each $r_{i,j}$, $1 \leq i , j \leq n$, is a relation of two line segments, where $r_{i,j} \in LO(P)$. The polygon similarity function can be redefined as:

$$n \geq m \Rightarrow psim(QP, SRP) = \sum_{\forall r_i \in LO(QP), \forall r_j \in LO(SRP)} sim(r_i, r_j) + (C(n, 2) - C(m, 2)) * 8 \vee$$
$$n < m \Rightarrow psim(QP, SRP) = \sum_{\forall r_i \in LO(QP), \forall r_j \in LO(SRP)} sim(r_i, r_j) + (C(m, 2) - C(n, 2)) * 8$$

Note that, in the process of summation, r_i and r_j are in the corresponding positions in $LO(QP)$ and $LO(SRP)$, respectively. The extra elements in either $LO(QP)$ or $LO(SRP)$ are also considered as complete mismatch relations. It is possible to use the mechanism discussed in the previous section to simplify a polygon so that the number of sides of the candidate shap representation polygon and the one of query polygon are the same.

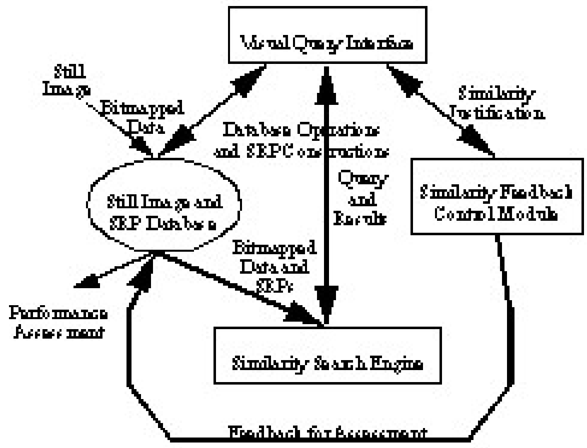

Figure 3. The architecture of our image content-based retrieval system

5. An Image Content Based Retrieval System

Figure 3 illustrates the overview of the content-based retrieval system developed based on our polygon matching algorithms. The system is incorporated with a still image and SRP (i.e., shape representation polygons) database. The database operations and the construction of SRPs are performed through a visual query interface, which provides a simple image processing function and a number of polygon drawing functions. Each picture is associated with a SRP edited by the user and stored in the database. The user can issue a query, which is a query polygon (QP), and the similarity search engine takes as input the QP and searches polygons in the database. Candidate pictures are displayed. The system will ask the user to justify the intensity of matching, in terms of degrees from 0 to 10. The justification (as a feedback) is saved in the system for the purpose of evaluating the accuracy of our algorithms.

The system is implemented in *Visual C++* with its visual interface constructed in *Visual Basic*. We use *MS Access* to build the database. Three graduate students spent about six months in developing the database, the interface, and the similarity search engine. A number of undergraduate students tested the system for another two months.

6. Conclusions

This paper proposes a generalized temporal relation model based on the 13 temporal interval relations proposed by James F. Allen [1]. Similarity between two relations can be computed according to the < . > relation defined in section 2. The most important contribution of this paper is the two algorithms to compute polygon similarity discussed in section 4. We believe that, the newly proposed algorithms bring a new mechanism for estimating the similarity between two images. Therefore, a new approach to image content-based information retrieval is possible. We implemented a system to test our approach. The experience is satisfiable.

There are several extensions that can be made. The proposed mechanism is to match the shape of two singular objects. We are developing mechanisms to match two sets of multiple polygons. It is possible to use spatial relations to estimate the relative positions among a set of polygons. However, the computation load is high. A normalization scheme is essential to bring down the complexity. However, it is difficult to define an efficient scheme and to maintain the accuracy at the same time. Moreover, we did not consider properties such as colors or textures of pictures. But, it is possible to apply many existing technologies in the literature of content-based retrieval to our system. Even there exists no best solution yet for content-based information retrieval, we believe that, our newly proposed mechanism based on polygon similarity points out a new way of solving the problem.

References

[1] Allen, James F., "Maintaining Knowledge about Temporal Intervals," Communications of the ACM, Vol. 26, No. 11 (1983).

[2] Chung, Chi-Ming, Shih, Timothy K., Huang, Jiung-Yao, Wang, Ying-Hong, and Kuo,

Tsu-Feng, "An Object-Oriented Approach and System for Intelligent Multimedia Presentation Designs," in proceeding of the International Conference on Multimedia Computing and Systems (ICMCS'95), Washington DC, U.S.A., pp. 278 – 281, May 15 - 18 (1995).

[3] Day, Young Francis, et. al., "Spatio-Temporal Modeling of Viedo Data for On-Line Object-Oriented Query Processing," in proceedings of the International Conference on Multimedia Computing and Systems, Washington DC, U.S.A., pp. 98 – 105, May 15 - 18 (1995).

[4] Flickner, Myron, et. al., "Query by Image and Video Content: The QBIC System," IEEE Computer, pp. 23 – 32, September (1995).

[5] Gudivada, Venkat N., and Raghavan, Vijay V., "Content-Based Image Retrieval Systems," IEEE Computer, pp. 18 – 22, September (1995).

[6] Keramane, Cherif and Duda, Andrzej, "Interval Expressions - a Functional Model for Interactive Dynamic Multimedia Presentations," in proceedings of the 1996 International Conference on Multimedia Computing and Systems, Hiroshima, Japan, pp. 283 –286, June 17 - 23 (1996).

[7] Li, John Z., TamerOzsu, M., and Szafron, Duane, "Spatial Reasoning Rules in Multimedia Management Systems," in proceedings of the 1996 Multimedia Modeling international conference (MMM'96), Toulouse, France, pp. 119 – 133, November 12 - 15 (1996).

[8] Little, Thomas D. C. and Ghafoor, Arif, "Spatio-Temporal Composition of Distributed Multimedia Objects for Value-Added Networks," IEEE Computer, pp. 42 – 50, October (1991).

[9] Little, Thomas D. C. and Ghafoor, Arif, "Interval-Based Conceptual Models for Time-Dependent Multimedia Data," IEEE transactions on knowledge and data engineering, Vol. 5, No. 4, pp. 551 – 563 (1993).

[10] Manjunath, B. S., and Ma, W. Y., "Texture Features for Browsing and Retrieval of Image Data," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 8, pp. 837 – 842, August (1996).

[11] Nabil, Mohammad, Ngu, Anne H. H., and Shepherd, John, "Picture Similarity Retrieval Using 2D Projection Interval Representation," IEEE Transactions on Knowledge and Data Engineering, Vol. 8, No. 4, pp. 533 – 539, August (1996).

[12] Sakauchi, Masao, Satou, Takashi, and Yaginuma, Yoshitomo, "Multimedia Database Systems for the Contents Mediator," IEICE Trans. Inf. and Syst., Vol. E79-D, No. 6, pp 641 – 646, June (1996).

[13] Samet, Hanan and Soffer, Aya, "MARCO: MAp Retrieval by COntent," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 8, pp. 783 –798, August (1996).

[14] Shih, Timothy K., Lo, Steven K. C., Fu, Szu-Jan, and Chang, Julian B., "Using Interval Temporal Logic and Inference Rules for the Automatic Generation of Multimedia Presentations," in Proceedings of the IEEE International Conference on Multimedia Computing and Systems, Hiroshima, Japan, pp. 425 – 428, June 17 - 23 (1996).

[15] Shih, Timothy K., Kuo, Chin-Hwa, Keh, Huan-Chao, Chao, T. Fang-Tsou, and An, Kuan-Shen, "An Object-Oriented Database for Intelligent Multimedia Presentations," in proceedings of the 1996 IEEE International Conference on Systems, Man and Cybernetics, Beijing, China, October 14 - 17 (1996).

[16] Srihari, Rohini K., "Automatic Indexing and Content-Based Retrieval of Captioned Images," IEEE Computer, pp. 49 – 56, September (1995).

[17] Theodoridis, Yannis , Vazirgiannis, Michael , and Sellis, Timos, "Spatial Temporal Indexing for Large Multimedia Applications," in proceedings of the 1996 International Conference on Multimedia Computing and Systems, Hiroshima, Japan, pp. 441 – 448, June 17 - 23 (1996).

[18] Vazirgiannis, Michael, Theodoridis, Yannis, and Sellis, Timos, "Spatio - Temporal Composition in Multimedia Applications," in proceedings of the International Workshop on Multimedia Software Development, Berlin, Germany, pp. 120 – 127, March 25 – 26 (1996).

[19] Wactlar, Howard D., Kanade, Takeo, Smith, Michael A., and Stevens, Scott M., "Intelligent Access to Digital Video: Informedia Project," IEEE Computer, pp. 46 – 52, May (1996).

[20] Wahl, Thomas, et. al., "TIEMPO: Temporal Modeling and Authoring of Interactive Multimedia" in proceedings of the international conference on multimedia computing and systems, Washington DC, U.S.A., pp. 274 – 277, May15 – 18 (1995).

Manuscript Received: Feb. 8, 2001
and Accepted: Feb. 21, 2001