Communications in Information and Systems
Vol. 4, No. 2, pp. 181-190, March 2005

004

ON INTRACTABILITY OF SPATIAL RELATIONSHIPS IN CONTENT-BASED IMAGE DATABASE SYSTEMS*

QING-LONG ZHANG† AND STEPHEN S.-T. YAU‡

Abstract. An image stored in image database systems is assumed to be associated with some content-based meta-data about that image, that is, information about objects in the image and absolute/relative spatial relationships among them. An image query for such an image database system can generally be handled in two ways: exact picture matching and approximate picture matching. In this paper we show the intractability of matching of spatial relationships between a query image and an image stored in the database. In particular, our results suggest that one would not expect to have polynomial-time algorithms for finding the exact picture-matching and computing the maximal similarity between a query picture and a database picture, unless $P = NP$.

1. Introduction. Image database systems have been very active over the past 20 years. With the explosive interest for the last 10 years in multimedia systems, content-based image retrieval has attracted the attention of researchers across several disciplines, including computer vision, pattern recognition, human-computer interaction, and image databases. A recent survey on content-based image retrieval was given by Smeulders et al. [5]. We [8, 9, 10, 11, 12] have proposed to develop a mathematical and algorithmic approach to modelling content-based image database systems.

In this short paper, we intend to show the intractability of matching of spatial relationships between a query image and an image stored in the database. In particular, one would not expect to have polynomial-time algorithms for finding the exact picture-matching and computing the maximal similarity between a query picture and a database picture, unless $P = NP$.

2. Specifying a General User Query. We first show how to specify a general user query, discussed in [12].

A real picture is assumed to be associated with some content-based meta-data about that picture, that is, information about objects in the picture and absolute/relative spatial relationships among them. An object in a real picture corresponds to a significant element of the image. Depending on the application, the significant elements of the image can be pixels, lines, regions, etc. A spatial relationship among objects is relative if it is determined by the position of the centroid of its objects. A spatial

*Received on November, 13, 2003; accepted for publication on May 8, 2004.
†Member,IEEE. Department of Information Technology, University of South Florida, Lakeland Campus, 3433 Winter Lake Road, Lakeland, Florida 33803, USA. Current E-mail contact: zhangq@math.uic.edu
‡Fellow, IEEE. Institute of Mathematics, East China Normal University, Shanghai, China; and Control and Information Laboratory, Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 322 Science and Engineering Offices, 851 South Morgan Street, Chicago, Illinois 60607, USA. E-mail: yau@uic.edu

relationship is absolute if it is determined by the absolute position of its objects in the image. The following various absolute spatial relationships are of common interest in pictorial databases: *left-of, right-of, in-front-of, behind, above, below, inside, outside,* and *overlaps*. Only the first six spatial operators are considered for relative spatial relationships, since *inside, outside,* and *overlaps* operators are not applicable. Note that the first six spatial operators are directional and the last three spatial operators are topological.

We will use x^a and x^r to indicate the absolute spatial operator x and the relative spatial operator x respectively. Note that *right-of* and *above* are dual spatial operators of *left-of* and *below* respectively, and *in-front-of* and *behind* spatial operators are only applicable for three dimensional pictures. Let

$$\boldsymbol{R} = \{\textit{left-of}^{\,a}, \textit{left-of}^{\,r}, \textit{below}^{\,a}, \textit{below}^{\,r}, \textit{inside, outside, overlaps}\}. \tag{1}$$

Clearly we can just use these seven spatial operators in $\boldsymbol{R}$ to specify spatial constraints among objects in a planar (i.e., two-dimensional) picture. Then an image stored in the image database is assumed to be represented by objects in the image and the complete information about absolute/relative spatial relationships of $\boldsymbol{R}$ among them.

Now a user query is of the following form:

(2) **An image query** Q: Find images containing a nonempty finite set $\boldsymbol{O}^Q$ of objects and another set (possibly null) $\boldsymbol{F}^Q$ of absolute/relative spatial relationships of $\boldsymbol{R}$ among them.

That is, an image f stored in the image database is matched by an image query Q if f contains a set $\boldsymbol{O}^Q$ of objects satisfying spatial constraints $\boldsymbol{F}^Q$ among these objects in $\boldsymbol{O}^Q$. We call it *an exact picture-matching* between Q and f. Note that the set $\boldsymbol{F}^Q$ in the image query Q implicitly indicates the conjunction of all absolute/relative spatial relationships in $\boldsymbol{F}^Q$ using the $\wedge$ (i.e., AND) logical operator. A more general user query is indeed the disjunction of the above user queries in (2) using $\vee$ (i.e., OR) logical operators. Therefore, a more general user query can always be handled via its user subqueries in (2); that is, the query outcome of a more general user query consists of all query outcomes of its user subqueries in (2).

3. Intractability of Spatial Constraints in an Image Query. In this section we show the intractability of matching of spatial relationships between a query image and an image stored in the database.

A real picture is represented by objects in the picture and spatial relationships among them for storage and retrieval. Spatial relationships may be absolute or relative, and directional or topological. The 2D string approach developed by Chang et al. [1] is based on (relative) directional spatial relationships: *left-of, right-of, above*, and *below*. Spatial relationships used in [4] are (absolute) directional or (absolute) topological. Spatial relationships proposed in our work [8, 9] are more general, can be (absolute) directional, (relative) directional, or (absolute) topological. Tucci et al. [6] proved that the type-1 symbolic picture matching, developed in [1], is *NP*-complete. Zhang [7] formulated a small class of picture retrieval problem in picture retrieval systems [4], and proved that the exact picture matching problem is *NP*-complete; then showed that, as a corollary, if there exists a polynomial-time algorithm to compute the maximal similarity between a query picture and a database picture, then $P = NP$. In fact, Zhang's results [7] are also valid in our framework of content-based image database systems [11]. Here we follow [7] to briefly present the proofs.

3.1. The Exact Picture Matching Problem. Let *PLOB* be a collection of planar (i.e., 2-dimensional) pictures, in which each object has only its name and no other information such as its description of properties (e.g., attributes), and each spatial relationship is only directional (i.e., $\textit{left-of}^a$, $\textit{left-of}^r$, $\textit{below}^a$, or $\textit{below}^r$), and O $\textit{left-of}^a$ O' (O $\textit{below}^a$ O', respectively) is in the picture if and only if O $\textit{left-of}^r$ O' (O $\textit{below}^r$ O') is in the same picture too.

One can use the simple transitive rule, for each $x \in \{\textit{left-of}^a, \textit{left-of}^r, \textit{below}^a, \textit{below}^r\}$, to deduce the directional spatial relationship AxC from the directional spatial relationships AxB and BxC. Let $\boldsymbol{F}$ be a set of directional spatial relationships involving only $\textit{left-of}^a$, $\textit{left-of}^r$, $\textit{below}^a$, and $\textit{below}^r$. Then the transitive closure of $\boldsymbol{F}$, denoted by $CL\{\boldsymbol{F}\}$, is the biggest set of directional spatial relationships, deducible from $\boldsymbol{F}$ using the transitive rule, including all directional spatial relationships in $\boldsymbol{F}$.

Given a picture f, we will use $\boldsymbol{O}^f$ and $\boldsymbol{F}^f$, respectively, to denote a collection of objects in f and a set of spatial relationships among objects in f. Given a set $\boldsymbol{F}$ of spatial relationships among objects in $\boldsymbol{O}$, we will also use $\boldsymbol{F}(\boldsymbol{O'})$ to denote the set of spatial relationships in $\boldsymbol{F}$ among objects only in $\boldsymbol{O'} \subseteq \boldsymbol{O}$.

In the following Definitions 3.1 and 3.2, any two objects O and O' in *PLOB* are the same if they have exactly the same name; and any two directional spatial relationships O_1xO_2 and $O'_1yO'_2$ in *PLOB* are the same if $x = y$ (i.e., both are $\textit{left-of}^a$, $\textit{left-of}^r$, $\textit{below}^a$, or $\textit{below}^r$), and $O_1 = O'_1$, $O_2 = O'_2$. Then any two sets $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ of directional spatial relationships in *PLOB* are the same if the objects and the directional spatial relationships in both sets are the same respectively.

Definition 3.1. *Given a user query Q with $\boldsymbol{O}^Q = \{O_1, O_2, \ldots, O_q\}$ and a picture f in* PLOB*, f is called a type*-1 *satisfied picture of the query Q if there is a subset $\boldsymbol{O'}^f = \{O'_1, O'_2, \ldots, O'_q\} \subseteq \boldsymbol{O}^f$ such that $O_j = O'_j (1 \leq j \leq q)$ and* $\mathrm{CL}\{\boldsymbol{F}^Q\} = \mathrm{CL}\{\boldsymbol{F}^f\}(\boldsymbol{O'}^f)$.

Definition 3.2. *Given a user query Q with $\boldsymbol{O}^Q = \{O_1, O_2, \ldots, O_q\}$ and a picture f in* PLOB*, f is called a type*-0 *satisfied picture of the query Q if there is a subset $\boldsymbol{O'}^f = \{O'_1, O'_2, \ldots, O'_q\} \subseteq \boldsymbol{O}^f$ such that $O_j = O'_j (1 \leq j \leq q)$ and* $\mathrm{CL}\{\boldsymbol{F}^Q\} \subseteq \mathrm{CL}\{\boldsymbol{F}^f\}(\boldsymbol{O'}^f)$.

Given a user query Q and a picture f in *PLOB*, determining whether f is a type-1 (type-0, respectively) satisfied picture of Q is called the type-1 (type-0) picture matching problem (in *PLOB* of our framework). We also call, both the type-1 and type-0 picture matchings, *the exact picture matching*. Now we can prove as in [7] that the type-1 picture matching problem is an *NP*-complete problem.

3.1.1. The NP-completeness of Type-1 Symbolic Picture Matching. A picture or a pattern is represented by an $m \times n$ matrix containing objects/entries which are denoted by a set of symbols V, for symbolic picture retrieval [1].

DEFINITION 3.3. ([*1*]) *Given a set of symbols* V*, a symbolic picture* f *on* V *is a mapping* $M \times N \longrightarrow W$*, where* $M = \{1, 2, \ldots, m\}$*,* $N = \{1, 2, \ldots, n\}$*, and* W *is the power set of* V*.*

It is proved in [6] that determining whether a pattern p is a type-1 subpicture of a picture f is an NP-complete problem.

The type-1 symbolic picture matching (T-1M) problem is formulated as follows:
Instance: An $m \times n$ matrix F (picture) and an $s \times t$ matrix P (pattern) of symbols from an alphabet A.
Question: Is P a type-1 subpicture of F, i.e., do there exist two ascending sequences of indices $\mathbf{r} = (r_1, r_2, \ldots, r_s)$ and $\mathbf{c} = (c_1, c_2, \ldots, c_t)$ such that $F(r_i, c_j) = P(i, j)$ for $1 \le i \le s$, and $1 \le j \le t$.

THEOREM 3.4. ([*6*]) *The type*-1 *symbolic picture matching* (*T-1M*) *is* NP*-complete.*

In the proof of Theorem 3.4 given in [6], the number of elements in each entry of a symbolic picture f is bounded by 1. So, we can modify the above Definition 3.3 of a symbolic picture as follows: a symbolic picture f on V is a mapping $M \times N \longrightarrow V \cup \{blank\}$. Furthermore, when a row or a column is empty, it can be removed. Thus, in the following Section 3.1.2, we shall assume that f maps $M \times N$ to $V \cup \{blank\}$, and each row and each column of f must contain at least one element.

3.1.2. The NP-completeness of Type-1 Picture Matching in PLOB. Now we want to prove that the type-1 picture matching problem in *PLOB* of our framework is an *NP*-complete problem.

The type-1 picture matching (T-1EPM) problem in *PLOB* of our framework is formulated as follows:
Instance: A user query Q with $|\boldsymbol{O}^Q| = q$ and a picture f with $|\boldsymbol{O}^f| = N$ in *PLOB*, and a set C of all object names from which the name of each object in Q and f is chosen.
Question: Is f a type-1 satisfied picture of Q?

THEOREM 3.5. *The type*-1 *picture matching in* PLOB *of our framework is* NP*-complete.*

Proof. It follows from a similar proof in [7].

It is easy to see that T-1EPM $\in NP$. Since a nondeterministic algorithm needs only to guess a subset $\boldsymbol{O'}^f \subseteq \boldsymbol{O}^f$ with $|\boldsymbol{O'}^f| = q$, and checks whether $\boldsymbol{O}^Q = \boldsymbol{O'}^f$ and $CL\{\boldsymbol{F}^Q\} = CL\{\boldsymbol{F}^f\}(\boldsymbol{O'}^f)$. The check could be done in polynomial-time (Note that the polynomial with order 3 is enough, since we can have such an efficient algorithm with order 3 to calculate the transitive closure $CL\{\boldsymbol{F}\}$ of a given set $\boldsymbol{F}$ of spatial relationships).

We now transform the above type-1 symbolic picture matching (T-1M) in Section 3.1.1 into our T-1EPM.

Given an $m \times n$ matrix F (picture) and an $s \times t$ matrix P (pattern) of symbols from an alphabet A, the corresponding instance of T-1EPM can be obtained by setting:

$\boldsymbol{O}^f$: A collection of objects f_i ($1 \le i \le N$) such that each f_i has only its name and no attributes, and its name is the same nonblank symbol as some entry in matrix F. Each nonblank entry in matrix F corresponds to one object f_i, $1 \le i \le N$, where N is the total number of nonblank entries.

$\boldsymbol{F}^f$: A collection of all directional spatial relationships among f_i's involving only *left-of*^{a}, *left-of*^{r}, *below*^{a}, and *below*^{r}. If object f_i and object f_j, $1 \le i, j \le N$, are in the k-th row (column, respectively) and l-th row (column) of matrix F, respectively, and $k < l$, then $\boldsymbol{F}^f$ contains the relationships f_j *below*^{a} f_i and f_j *below*^{r} f_i (f_i *left-of*^{a} f_j and f_i *left-of*^{r} f_j).

$\boldsymbol{O}^Q$ $(= \{O_1, O_2, \ldots, O_q\})$ and $\boldsymbol{F}^Q$ for the corresponding user query Q can be defined similarly as the picture f. And

C: Just the alphabet A.

Clearly the transformation is polynomial-time in the size of the instance of T-1M (Note that the polynomial with order 2 is enough, since $\boldsymbol{F}^f$ and $\boldsymbol{F}^Q$ have at most $4 * \frac{N*(N-1)}{2}$ and $4 * \frac{q*(q-1)}{2}$ relationships respectively).

EXAMPLE. Suppose that we have a symbolic picture F (3×3 matrix) and a pattern P (2×2 matrix), where all nonblank entries in F and P are: $F(1,1) = d$, $F(2,2) = b$, $F(2,3) = c$, $F(3,1) = a$; and $P(1,2) = c$, $P(2,1) = a$; and $A = \{a, b, c, d\}$. Then,

$$\begin{aligned}
\boldsymbol{O}^f &= \{f_1(a), f_2(d), f_3(b), f_4(c)\},\\
\boldsymbol{F}^f &= \{f_1 \textit{ left-of}^a\ f_3, f_1 \textit{ left-of}^r\ f_3, f_1 \textit{ left-of}^a\ f_4, f_1 \textit{ left-of}^r\ f_4, f_2 \textit{ left-of}^a\ f_3,\\
&\qquad f_2 \textit{ left-of}^r\ f_3, f_2 \textit{ left-of}^a\ f_4, f_2 \textit{ left-of}^r\ f_4, f_3 \textit{ left-of}^a\ f_4, f_3 \textit{ left-of}^r\ f_4,\\
&\qquad f_1 \textit{ below}^a\ f_3, f_1 \textit{ below}^r\ f_3, f_1 \textit{ below}^a\ f_4, f_1 \textit{ below}^r\ f_4, f_1 \textit{ below}^a\ f_2,\\
&\qquad f_1 \textit{ below}^r\ f_2, f_3 \textit{ below}^a\ f_2, f_3 \textit{ below}^r\ f_2, f_4 \textit{ below}^a\ f_2, f_4 \textit{ below}^r\ f_2\},\\
\boldsymbol{O}^Q &= \{O_1(a), O_2(c)\},\\
\boldsymbol{F}^Q &= \{O_1 \textit{ left-of}^a\ O_2, O_1 \textit{ left-of}^r\ O_2, O_1 \textit{ below}^a\ O_2, O_1 \textit{ below}^r\ O_2\}, \text{and}\\
C &= A = \{a, b, c, d\}.
\end{aligned}$$

Note that we have $CL\{\boldsymbol{F}^f\} = \boldsymbol{F}^f$ and $CL\{\boldsymbol{F}^Q\} = \boldsymbol{F}^Q$ for the above transformation.

Suppose that P is a type-1 subpicture of F, i.e., there exist two ascending sequences of indices $\mathbf{r} = (r_1, r_2, \ldots, r_s)$ and $\mathbf{c} = (c_1, c_2, \ldots, c_t)$ such that $F(r_i, c_j) = P(i, j)$ for $1 \leq i \leq s$ and $1 \leq j \leq t$. Let $\boldsymbol{O'}^f = \{O'_1, O'_2, \ldots, O'_q\}$ and $\boldsymbol{F'}^f$ be the collections of objects and relationships, respectively, defined from the submatrix $F(\mathbf{r}, \mathbf{c})$ formed by s rows $\mathbf{r}$ and t columns $\mathbf{c}$ in matrix F, under the above definition of transformation. Then $\boldsymbol{O}^Q = \boldsymbol{O'}^f \subseteq \boldsymbol{O}^f$ and $\boldsymbol{F}^Q = \boldsymbol{F'}^f = \boldsymbol{F}^f(\boldsymbol{O'}^f)$. So f is a type-1 satisfied picture of Q.

Conversely, suppose that f is a type-1 satisfied picture of Q, i.e., there exists a subset $\boldsymbol{O'}^f \subseteq \boldsymbol{O}^f$ such that $\boldsymbol{O}^Q = \boldsymbol{O'}^f$ and $\boldsymbol{F}^Q = \boldsymbol{F}^f(\boldsymbol{O'}^f)$. This means that both $\boldsymbol{O'}^f$ and $\boldsymbol{F}^f(\boldsymbol{O'}^f)$ together have the same information as both $\boldsymbol{O}^Q$ and $\boldsymbol{F}^Q$ together.

LEMMA. Let H be an $s \times t$ matrix and h be the picture obtained from H under the above transformation, then H can be uniquely determined by $\boldsymbol{O}^h$ and $\boldsymbol{F}^h$.

Proof of Lemma. Firstly let us reconstruct a matrix H' from $\boldsymbol{O}^h$ and $\boldsymbol{F}^h$. Since, the lengths (i.e., the number of relationships in the sequence) of the longest sequences in $\boldsymbol{F}^h$ involving only *left-of*^{a} (*left-of*^{r}, alternatively) and *below*^{a} (*below*^{r}, alternatively), respectively, are just $t-1$ and $s-1$, so H' is an $s \times t$ matrix. Furthermore, for any object h_i in $\boldsymbol{O}^h$, the position of h_i in the matrix H' can be uniquely determined by its relative position given by *left-of*^{a} (*left-of*^{r}, alternatively) and *below*^{a} (*below*^{r}, alternatively) relationships in $\boldsymbol{F}^h$, that is, there exist two longest sequences involving the object h_i and only *left-of*^{a}, respectively, and *below*^{a} relationships in $\boldsymbol{F}^h$,

$$e_1 \textit{ left-of}^a\ e_2 \textit{ left-of}^a \ \ldots \textit{left-of}^a\ e_l = h_i \textit{ left-of}^a \ \ldots \ \textit{left-of}^a\ e_t, \text{ and}$$
$$e'_1 \textit{ below}^a\ e'_2 \textit{ below}^a \ \ldots \textit{below}^a\ e'_k = h_i \textit{ below}^a \ \ldots \ \textit{below}^a\ e'_s,$$

and the positions, l and k, of h_i in any such two longest sequences are uniquely fixed, then h_i is in the $(s-k+1, l)$-entry of the matrix H'. Now, we can easily check $H' = H$ from the construction. This ends the proof of lemma.

So, by the Lemma, an $m \times n$ matrix F and an $s \times t$ matrix P, respectively, can be uniquely reconstructed from the picture f and the query Q; and both $\boldsymbol{O'}^f$ and $\boldsymbol{F}^f(\boldsymbol{O'}^f)$ together determine an $s \times t$ matrix P', which is just the same as P. Note that every two objects in $\boldsymbol{O'}^f$ have the same relative order in both matrix F and matrix P'. For example, two objects in $\boldsymbol{O'}^f$ are in the same row (column, respectively) of matrix F if and only if they are in the same row (column) of matrix P'; and any number of objects in $\boldsymbol{O'}^f$ are in the same row (column, respectively) of matrix F if and only if they are in the same row (column) of matrix P'. Now, choose one object e_i, $1 \leq i \leq s$, (one object e'_j, $1 \leq j \leq t$, respectively) from the i-th row (j-th column) in P', and let each e_i (e'_j, respectively) be in the r_i-th row (c_j-th column) of F, then $r_1 < r_2 < \ldots < r_s$ and $c_1 < c_2 < \ldots < c_t$. Let $F(\mathbf{r}, \mathbf{c})$ be the submatrix of F formed by s rows $\mathbf{r} = (r_1, r_2, \ldots, r_s)$ and t columns $\mathbf{c} = (c_1, c_2, \ldots, c_t)$ in matrix F. Then we can easily check $F(\mathbf{r}, \mathbf{c}) = P'$ from the construction, since, for any object O'_i in $\boldsymbol{O'}^f$, $1 \leq i \leq q$, let O'_i be in the same k-th row with the object e_k and the same l-th column with the object e'_l in the matrix P', then O'_i is also in the same r_k-th row with the object e_k and the same c_l-th column with the object e'_l in the matrix F, that is, O'_i is in the (k, l)-th entry of P' if and only if it is in the (r_k, c_l)-th entry of F. Hence P is a type-1 subpicture of F.

This concludes the proof of NP-completeness of T-1EPM in PLOB of our framework. □

Theorem 3.5 indicates that even a small class of picture matching problem in *PLOB* of our framework is *NP*-complete. Observe that the type-1 picture matching is a particular part of the type-0 picture matching. One would expect that the exact picture matching (i.e., the type-0 picture matching) in *PLOB* of our framework is *NP*-complete.

3.2. The Maximal Similarity Problem. While the exact picture matching yields the query outcome consisting of only those stored images matched exactly by a user query, it might take much long time to perform the query processing for certain irregular stored images because of *NP*-completeness of the exact picture matching. To address this type of inefficiency, approximate or heuristic picture matching algorithms need to be developed to help improve the performance of pictorial retrieval. Image retrieval based on similarity measures can be found in the literature.

For our framework of content-based image database systems, a real image is represented by its description of objects in the image and absolute/relative spatial relationships among them for storage and retrieval. To develop approximate image retrieval by similarity measures, one may give the notion of similarity $Sim(\cdot, \cdot)$ between two objects, two absolute/relative spatial relationships, and two images. The similarity value is either negative infinity or between 0 and 1. Specifically, the similarity measure in the small class *PLOB* defined above in Section 3.1 should be straightforward.

In *PLOB*, given a user specified object O and a system specified object O', the similarity $Sim(O, O')$ of O and O' is 1 if and only if both O and O' have the same name. Let $r = O_1 x O_2$ and $r' = O'_1 y O'_2$ be user specified and system specified directional spatial relationships respectively. Then the similarity $Sim(r, r')$ of r and r' is 1 if and only if both directional spatial relationship operators x and y are same (i.e., both *left-of*^{a}, both *left-of*^{r}, both *below*^{a}, or both *below*^{r}), and $O_1 = O'_1$ and $O_2 = O'_2$.

Let $Q = (O_1, O_2, \ldots, O_n, r_1, r_2, \ldots, r_m)$ and $f' = (O'_1, O'_2, \ldots, O'_n, r'_1, r'_2, \ldots, r'_m)$ be the descriptions of a user specified picture and a system specified picture respectively. Then the type-1 similarity $Sim(Q, f')$ of Q and f' under the natural order (i.e., O_j and r_k in Q correspond to O'_j and r'_k in f', respectively, for $1 \leq j \leq n$ and $1 \leq k \leq m$) is 1 if and only if $O_j = O'_j$ for $1 \leq j \leq n$, and $r_k = r'_k$ for $1 \leq k \leq m$; and the type-0 similarity $Sim(Q, f')$ of Q and f' under the natural order is 1 if and only if $O_j = O'_j$ for $1 \leq j \leq n$, and either r_k is null spatial relationship or $r_k = r'_k$ for $1 \leq k \leq m$.

In the following Definitions 3.6 and 3.7, $\boldsymbol{F}^f$ is assumed to have the complete information about absolute/relative spatial relationships in a picture f. Given a set $\boldsymbol{F}$ of spatial relationships, we use MAX$\{\boldsymbol{F}\}$ to denote the set of all deducible spatial relationships from $\boldsymbol{F}$ under a system of rules (i.e., the maximal set of $\boldsymbol{F}$), defined in [8, Chapter 2] [9]. Now we define the maximal similarity $MaxSim(Q, f)$ between a query picture Q and a database picture f.

Definition 3.6. *Given a user query Q and a picture f, let type-1 $MaxSim(Q, f)$ be the maximal one among all type-1 similarities $Sim(Q, f')$ of Q and subpicture f' in f under the natural order, where $\boldsymbol{O}^Q = \{O_1, O_2, \ldots, O_q\}$ and $\boldsymbol{O}^{f'} = \{O'_1, O'_2, \ldots, O'_q\} \subseteq \boldsymbol{O}^f$ (some O'_j could be null objects); and set $l = max(|MAX\{\boldsymbol{F}^Q\}|, |\boldsymbol{F}^f(\boldsymbol{O}^{f'})|)$, then $(r_1, r_2, \ldots, r_l)$ and $(r'_1, r'_2, \ldots, r'_l)$, respectively, are the enumeration sequences from $MAX\{\boldsymbol{F}^Q\}$ and $\boldsymbol{F}^f(\boldsymbol{O}^{f'})$ (some r_j and r'_k could be null spatial relationships if necessary); now $Q = (O_1, O_2, \ldots, O_q, r_1, r_2, \ldots, r_l)$ and $f' = (O'_1, O'_2, \ldots, O'_q, r'_1, r'_2, \ldots, r'_l)$.*

Definition 3.7. *Given a user query Q and a picture f, let type-0 $MaxSim(Q, f)$ be the maximal one among all type-0 similarities $Sim(Q, f')$ of Q and subpicture f' in f under the natural order, where $\boldsymbol{O}^Q = \{O_1, O_2, \ldots, O_q\}$ and $\boldsymbol{O}^{f'} = \{O'_1, O'_2, \ldots, O'_q\} \subseteq \boldsymbol{O}^f$ (some O'_j could be null objects); and set $l = max(|\boldsymbol{F}^Q|, |\boldsymbol{F}^f(\boldsymbol{O}^{f'})|)$, then $(r_1, r_2, \ldots, r_l)$ and $(r'_1, r'_2, \ldots, r'_l)$, respectively, are the enumeration sequences from $\boldsymbol{F}^Q$ and $\boldsymbol{F}^f(\boldsymbol{O}^{f'})$ (some r_j and r'_k could be null spatial relationships if necessary); now $Q = (O_1, O_2, \ldots, O_q, r_1, r_2, \ldots, r_l)$ and $f' = (O'_1, O'_2, \ldots, O'_q, r'_1, r'_2, \ldots, r'_l)$.*

Theorem 3.8. *If there exists an algorithm with the time complexity* TC *to compute the type-*1 $MaxSim(Q, f)$ *of a user query Q and a picture f, then there is also an algorithm with the time complexity of the same order as TC to answer the type-*1 *picture matching problem in* PLOB.

Proof. Suppose that the assumption is true and $\mathcal{A}$ is the required algorithm with the time complexity *TC*. Now, given a user query Q and a picture f in *PLOB*, we use algorithm $\mathcal{A}$ to compute type-1 $MaxSim(Q, f)$, and also check whether type-1 $MaxSim(Q, f) = 1$. Note that f is a type-1 satisfied picture of Q if and only if type-1 $MaxSim(Q, f) = 1$. Hence, we have this algorithm to answer the type-1 picture matching problem in *PLOB* of our framework. □

Theorem 3.9. *If there exists an algorithm with the time complexity* TC *to compute the type-*0 $MaxSim(Q, f)$ *of a user query Q and a picture f, then there is also an algorithm with the time complexity of the same order as TC to answer the type-*0 *picture matching problem in* PLOB.

Proof. Apply the similar proof of Theorem 3.8, using type-0 instead of type-1. □

Theorem 3.10. *If there exists a polynomial-time algorithm to compute the type-*1 $MaxSim(Q, f)$ *of a user query Q and a picture f, then* P = NP.

Proof. Immediately from the above two Theorems 3.5 and 3.8. □

As suggested in the above Section 3.1, we would expect that the exact picture matching (i.e., the type-0 picture matching) in *PLOB* of our framework is *NP*-complete. Thus, in view of Theorem 3.9, we would also expect that the similar result of Theorem 3.10 holds for the case of type-0 $MaxSim(\cdot,\cdot)$; that is, if there exists a polynomial-time algorithm to compute the type-0 $MaxSim(Q, f)$ of a user query Q and a picture f, then $P = NP$.

"Does $P = NP$?" arising in the field of Computer Science has become one of the well-known hardest questions in Mathematics [2, 3]. For example, it is one of the seven million-dollar Millennium Prize Problems listed by the Clay Mathematics Institute (see *www.claymath.org*). As mentioned in [2, 3], people would expect that $P \neq NP$ though it still remains open. That means, in view of Theorems 3.5 and 3.10, one would not expect to have polynomial-time algorithms for finding the exact picture-matching and computing $MaxSim(\cdot,\cdot)$ between a query picture and a picture stored in the database, unless $P = NP$.

In [12], we have addressed the approximate picture matching problem, and have presented a stepwise approximation of intractable spatial constraints in an image query. Especially, this stepwise approximation may be pre-processed on an image query before an advanced picture matching algorithm is invoked. Advanced polynomial-time algorithms for the approximate picture matching need to be much developed to help improve the performance of image retrieval due to the *NP*-completeness of the exact picture matching.

REFERENCES

[1] S.-K. Chang, Q. Y. Shi, and C. W. Yan, *Iconic Indexing by 2-D Strings*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 9:3(1987), pp. 413–428.

[2] S. Cook, *The* P *versus* NP *Problem*, one of the seven million-dollar Millennium Prize Problems in Mathematics, available at *www.claymath.org/Millennium_Prize_Problems/P_vs_NP*.

[3] S. Cook, *The Importance of the* P *versus* NP *Question*, Journal of the ACM, 50:1(2003), pp. 27–29.

[4] A. P. Sistla, C. Yu, and R. Haddad, *Reasoning about Spatial Relationships in Picture Retrieval Systems*, Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile, pp. 570–581, 1994.

[5] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, *Content-Based Image Retrieval at the End of the Early Years*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:12(2000), pp. 1349–1380.

[6] M. Tucci, G. Costagliola, and S.-K. Chang, *A Remark on NP-Completeness of Picture Matching*, Information Processing Letters, 39(1991), pp. 241–243.

[7] Q.-L. Zhang, *A Remark on Intractability of Picture Retrieval by Contents*, Technical Report of Summer Research (Supported by the Research Assistantship of Professor Clement T. Yu), Univ. of Illinois, Chicago, June 1994.

[8] Q.-L. Zhang, *A Unified Framework for Iconic Indexing of Spatial Relationships in Image Databases*, Ph.D. Thesis, Department of Mathematics, Statistics, and Computer Science,

Univ. of Illinois, Chicago, 1996.

[9] Q.-L. Zhang, S.-K. Chang, and S. S.-T. Yau, *A Unified Approach to Iconic Indexing, Retrieval, and Maintenance of Spatial Relationships in Image Databases*, Journal of Visual Communication and Image Representation, Special Issue on Indexing, Storage, Retrieval and Browsing of Images and Video, Academic Press, 7:4(1996), pp. 307–324.

[10] Q.-L. Zhang, S.-K. Chang, and S. S.-T. Yau, *The Consistency Problem on Content-based Pictorial Description in Pictorial Database Systems*, Communications in Information and Systems, International Press, 1:2(2001), pp. 225–240.

[11] Q.-L. Zhang and S. S.-T. Yau, *A General Approach to Indexing and Retrieval of Images in Image Databases*, Communications in Information and Systems, International Press, 3:1(2003), pp. 61–73.

[12] Q.-L. Zhang and S. S.-T. Yau, *A Stepwise Approximation of Intractable Spatial Constraints in Image Queries*, Communications in Information and Systems, International Press, 3:3(2003), pp. 203–221.