

Using Spatial Relations for Graphical Symbol Description

Santosh K.C., Laurent Wendling, Bart Lamiroy

► **To cite this version:**

Santosh K.C., Laurent Wendling, Bart Lamiroy. Using Spatial Relations for Graphical Symbol Description. 20th International Conference on Pattern Recognition - ICPR 2010, Aug 2010, Istanbul, Turkey. IEEE, pp.2041 - 2044, 2010, ICPR 2010 - 20th International Conference on Pattern Recognition. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5595915>. <10.1109/ICPR.2010.503>. <inria-00516725>

HAL Id: inria-00516725

<https://hal.inria.fr/inria-00516725>

Submitted on 11 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Using Spatial Relations for Graphical Symbol Description

Santosh K.C.^{*}, Laurent Wendling[†] and Bart Lamiroy[‡]

^{*}INRIA Nancy-Grand Est, [‡]Nancy Université INPL

LORIA - Campus Scientifique, BP 239 - 54506 Vandoeuvre-lès-Nancy Cedex, France

Email: Santosh.KC, Bart.Lamiroy@loria.fr

[†]LIPADE, Université Paris Descartes, 75270 Paris Cedex 06, France

Email: Laurent.Wendling@math-info.univ-paris5.fr

Abstract—In this paper, we address the use of unified spatial relations for symbol description. We present a topologically guided directional relation signature. It references a unique point set instead of one entity in a pair, thus avoiding problems related to erroneous choices of reference entities and preserves symmetry.

We experimentally validate our method on showing its ability to serve in a symbol retrieval application, based only on a spatial relational descriptor that represents the links between the decomposed structural patterns called “*vocabulary*” in a spatial relational graph.

Keywords-Unified Spatial Relations; Vocabulary; Spatial Relation Graphs; Symbol Retrieval

I. INTRODUCTION

Use of pairwise spatial relations can greatly ease image understanding, scene analysis and pattern recognition tasks. However, it is difficult to automatically organise and obtain them [1], [2].

One of the approaches to fully exploit the information embedded in a symbol is to extract/decompose structural components and formalise the links that exist between them. In this paper, we develop a new spatial relation model to express the links between the structural descriptions of patterns. Spatial relations can be either topological [3], [4], [5], [6] or directional [7], [8], [9], [10] in nature, their choice depending on the kind of application. Building on [1], we unify both topological and directional information into one descriptor [11]. The proposed spatial reasoning approach is applied to the visual vocabulary illustrated in [12] and we experimentally validate its expressiveness for symbol retrieval.

The rest of the paper is organised as follows. Section II reviews the existing models. The proposed approach is explained in section III which is followed by experimental results and discussions in section IV. Section V concludes the paper.

II. OVERVIEW OF EXISTING MODELS

In general, there is no particular spatial reasoning approach that can adapt to any type of application. It is possible, however, to identify three main levels of information: topological relations, quantitised/metric refinement and symmetry.

Topological relations are invariant to topological transformations. These encompass, but are not restricted to rigid transforms as rotation, scaling, and translation. It is to be noted that no existing model fully integrates topology. They rather have various degrees of sensitivity to or awareness of topological relations.

Region Connection Calculus-8 (RCC-8) [6], is one way to express basic topological predicates closely related to human understanding like *Disconnected*, *Externally Connected*, *Overlap*, *Contain/Inside*, *Cover/Covered by* and *Equal* as shown in Fig. 1. Based on the intersections of the boundaries, interiors and exteriors of two sets \mathbb{A} and \mathbb{B} , we can use the 9-intersection model [4], to express these topological relations between them in a 9-dimensional binary space. However, this lacks metric information, and is therefore not sufficiently discriminant. Most models integrate quantitised information, but the level of detail in the expression of the spatial predicates (like *Left*, *Right etc.*) can vary widely [13]. The introduction of this metric information most often introduces asymmetry, thus rendering them subject to erroneous choices of reference objects that, on their turn, influence the global positioning semantics. We took care, in our method, to integrate the idea of semantic inverse theory [1] and to preserve symmetry.

The fact is that integrating both high level metric precision and topologically sound descriptions is computationally expensive. Existing approaches present a trade-off between these factors: The *Cone-Shaped Model* reduces relative positioning to the discretised angle [10] of the sole centroids. It is robust to small variations of shape and size, and separation. However, in cases where the centroids coincide it cannot produce any measure. Extensions, like [14], do not lift this ambiguity. *Angle Histograms* tend to be more capable of dealing with overlapping. However, they consider all pixels, instead of just the centroids, which severely increases computational time, while not offering any substantial benefit when two objects are at a sufficiently large distance. It also does not completely lift possible ambiguities for describing very different pairs of objects, since it may give identical histograms [8]. Another trade-off is to approximate objects by their *Minimum Bounding Rectangle* (MBR). This generally gives more interesting relations as the

approximation captures shape and size [15], [16], [17], [5]. However, its quality, and therefore precision, largely depends on the compactness of the underlying shapes. These methods only approximate topological relations, and may generate wrong connection/overlapping information. *F-Histograms*, pairs of longitudinal sections instead of pairs of points, give coherent results [9] at the cost of high time complexity. They do not cover basic topological relations such as *Inside* and *Overlap* nor give them metric information. Another well-known approach uses fuzzy landscapes [7], and is based on fuzzy morphological operators.

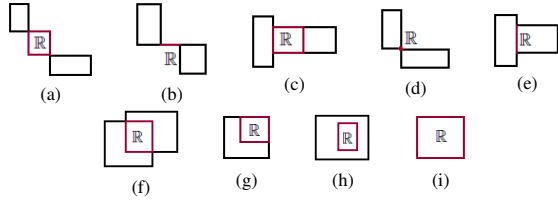


Figure 1. Topologically guided \mathbb{R} : (a-c) Disconnected, (d-e) Externally Connected (f) Overlap (g) Cover/Covered By, (h) Contain/Inside and (i) Equal

III. SPATIAL RELATION ATTRIBUTE

Unlike the previously mentioned approaches, our approach integrates both topology and quantised directional relations between two objects. We first extract a unique reference region \mathbb{R} based on the topology of their respective MBRs by using the 9-intersection model [4]. As shown in Fig. 1 and in connection with [6], \mathbb{R} is derived as either the common region of two neighbouring sides in the case of *disconnected* MBRs or the intersection of MBRs in the case of *overlapping*, *equal* or otherwise *connected* MBRs. In what follows we shall use the characteristic points $\mathbb{R}p_i$ composed of the extrema and the centroid of $\mathbb{R} = \{\mathbb{R}p_i\}_{i=1\dots 2n+1}$, where n is the dimensionality of the region. The dimension of \mathbb{R} changes with the topological relations (Fig. 1).

Let \mathbb{X} be one of the initial objects \mathbb{A} or \mathbb{B} , and let their reference region be \mathbb{R} . From a set of key points $\mathbb{R}p_i$, defining \mathbb{R} , we cover the surrounding space at regular radial intervals of $\Theta = 2\pi/m$. The line rotates over a cycle, and intersecting with object \mathbb{X} , generates a boolean histogram of angular coverage \mathcal{H} ,

$$\mathcal{H}(\mathbb{X}, \mathbb{R}p_i) = [I(\mathbb{R}p_i, j\Theta)]_{j=0\dots m}$$

$$\text{where, } I(\mathbb{R}p_i, \theta_j) = \begin{cases} 1 & \text{if } \text{line}(\mathbb{R}p_i, \theta_j) \cap \mathbb{X} \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

This is extended *wlog* to the sector defined by two successive angle values and is normalised with respect to the total area of the studied object [18] such that $\sum \mathcal{H}(\cdot) = 1$. The process is repeated from every $\mathbb{R}p_i$, and the resulting averaged histogram is called *quantised directional relation signature* $\mathcal{H}(\mathbb{X}, \mathbb{R})$.

\mathbb{R} thus avoids problems related to erroneous choices of reference entities and will guarantee uniqueness of subsequent spatial relations. For symmetry preserving reasons, we use $\mathbb{R}(\star, \ast) = \{\mathcal{H}(\star, \mathbb{R}), \mathcal{H}(\ast, \mathbb{R})\}$. In this model, the parameter Θ defines the resolution of \mathcal{H} that accounts the trade-off between precision and time complexity. Besides, for all *inclusion* topological relations like *Cover/Covered By*, *Contain/Inside* and *Equal*, $\mathcal{H}(\mathbb{X}, \mathbb{R}) = \emptyset$. Therefore, only one part of \mathbb{R} needs to be computed. This eventually reduces time complexity. Fig. 3 shows an example where $\mathcal{H}(\text{thick}, \mathbb{R}) = \emptyset$ since *CoveredBy*(*thick*, \mathbb{R}).

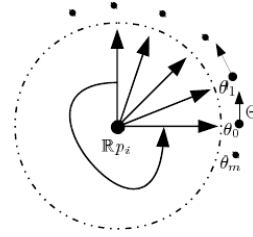


Figure 2. Radial line $\text{line}(\mathbb{R}p_i, \theta_j)$ rotation

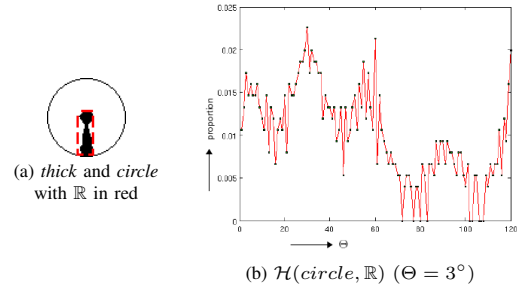


Figure 3. Average relation signature

IV. EXPERIMENTAL SETUP

In order to prove the usefulness of our approach, we show experimentally that it can be used as a basis for position based descriptors of symbols, and can be used in a recognition and retrieval framework. We build upon the method presented in [11] and apply it to the visual vocabulary in [12].

A. Symbol Description

We use a set of well controlled elementary part detectors to define a visual *vocabulary*. In our case, it consists of: *circles*, *corners*, *loose ends/extremities* and *thick/filled* components. Rather than using the detected elements as a basis for expressing and computing spatial relations, we group them together into “classes” of elements having the same type, as shown in Fig. 4 (a). The whole symbol is then expressed as a complete Spatial Relational Graph in which each node represents a distinct class of elements and the arcs are labeled with a numerical expression of the spatial relation \mathbb{R} between the connected nodes. In other words, we construct a graph $G = (V, E, F_V, F_E)$ with V the set of vertices, E

the set of graph arcs: $E \subseteq V \times V$, and $F_V : V \rightarrow P_V$ and $F_E : E \rightarrow R_E$ the attribute functions on nodes and arcs.

$$\begin{aligned} V &= \{A, B, C, \dots\}, \quad E = \{(A, B), (A, C), \dots\} \\ F_V &= \{(A, \text{thick}), (B, \text{circle}) \dots\} \\ F_E &= \{((A, B), \mathfrak{R}(A, B)), ((A, C), \mathfrak{R}(A, C)) \dots\} \end{aligned}$$

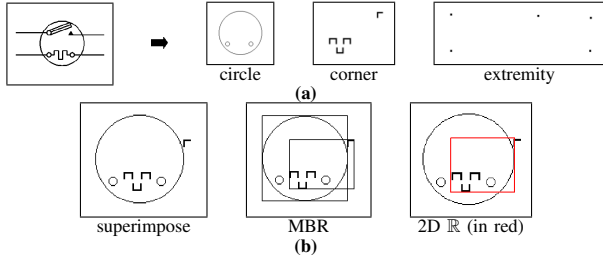


Figure 4. (a) Types of vocabulary [12] and (b) \mathbb{R} generation from a pair: circle and corner – an example

B. Symbol Retrieval

The previously described setup can now allow us to evaluate the expressive power of our spatial relational attribute for symbol retrieval.

1) *Similarity Metric*: We define similarity between two matched graphs G and G' over all edges $e \in E$ matched to their counterparts $e' \in E'$

$$Sim(G, G') = 1 - \sum_{e \in E} \delta(F_E(e), F_E'(e'))$$

where δ expresses the Manhattan distance between histograms $\mathcal{H}(\cdot)$ and normalised to the $[0, 1]$.

It is important to notice, that, by construction, the graph matching problem is straightforward since every single node bears distinct labels (vocabulary). The only matching issue that can occur is that specific labels may or may not be present. Matching can therefore be done in near-constant time.

2) *Retrieval Efficiency*: We use the retrieval efficiency metric η , described in [19] to measure our method,

$$\eta_K = \begin{cases} n/N & \text{if } N \leq K \\ n/K & \text{otherwise} \end{cases}$$

where n is the number of returned relevant symbols, N the total number of relevant symbols and K the number of ranked symbols requested.

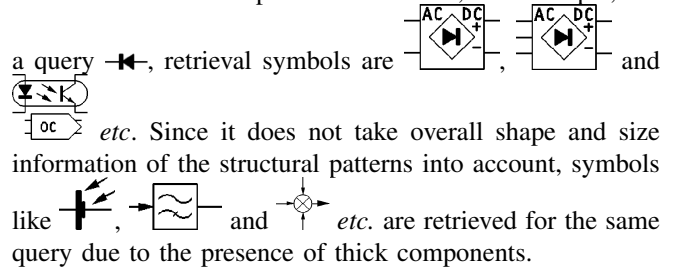
C. Results and Discussions

We have conducted two kinds of queries: one using non-filtered ranking, another by pre-filtering using *Candidate Selection* (CS). With CS, we only retrieve those candidates which share the same set of nodes, with the exact same labels as the query. This reduces the time of matching to symbols that are obviously irrelevant. Fig. 5 (a) shows a part of our used symbols. Query symbols are presented to the database

and results are ranked according to the similarity with the database symbols. Fig. 5 (b) shows a sample of retrieved symbols, ranked in decreasing order of similarity, both with and without CS. Table. I shows comparison between them, based on average results over 30 queries for varying values of K and databases ranging from 100 to 400 symbols.

Overall, we have found that the CS not only optimises retrieval efficiency but also reduces processing time dramatically. However, for certain queries, CS truncated the retrieval scope to the value lower than similar number of symbols in the database and consequently decreased the value of η_K .

Basically our method retrieves the symbols based on spatio-structural description of the structural patterns. Thanks to visual vocabulary, it is possible to retrieve symbols embedded in complex environments, for example, for



V. CONCLUSIONS AND FURTHER WORK

We have presented and applied a new approach to structural symbol description. The proposed approach is simple and flexible, and has an ability to efficiently express spatial relations between any number of components. We have shown that such a representation can be used for symbol retrieval.

Further work consists of using intra-class spatial relations and better candidate selection techniques to increase retrieval performance. To account for shape and size information of the structural components, we plan to integrate our Spatial Relational Graph with other popular descriptors like the Generic Fourier Descriptor, for instance.

REFERENCES

- [1] J. Freeman, "The modelling of spatial relations," *Computer Graphics and Image Processing*, vol. 4, pp. 156–171, 1975.
- [2] A. Rosenfeld and A. Kak, *Digital picture processing*. Academic Press, 1982, vol. 2.
- [3] M. Egenhofer and R. Franzosa, "Point-set Topological Spatial Relations," *International Journal of Geographical Information Systems*, vol. 5, no. 2, pp. 161–174, 1991.
- [4] M. Egenhofer and J. R. Herring, "Categorizing Binary Topological Relations Between Regions, Lines, and Points in Geographic Databases," in *University of Maine, Research Report*, 1991.
- [5] D. Papadias, T. Sellis, Y. Theodoridis, and M. J. Egenhofer, "Topological Relations in the world of Minimum Bounding Rectangles: a Study with R-trees," in *International Conference on Management Data*, 1995, pp. 92–103.

