Object Localization Based on Directional Information: Case of 2D Raster Data

Pascal Matsakis, JingBo Ni, Xin Wang

Dept. of Computing and Information Science, University of Guelph, Canada
matsakis@cis.uoguelph.ca, jni@uoguelph.ca, xin@uoguelph.ca

Abstract

A directional spatial relationship to a reference object (e.g., "east of the post office") can be represented by a spatial template. The template partitions the space into regions where the relationship holds (to various extents) and regions where it does not hold. The objects for which the relationship holds can then be located. A template can be easily modeled. Computationally, however, exact calculation of the model in case of 2D raster data is prohibitively expensive, and a tractable approximation algorithm was proposed. Here, we introduce a new concept: the concept of the F-template. It leads to a new approximation algorithm, which is faster, gives better results, and is more flexible.

1. Introduction

Space plays a fundamental role in human cognition. In everyday situations, it is often viewed as a construct induced by spatial relationships, rather than as a container that exists independently of the objects located in it. Spatial relationships, therefore, have been thoroughly investigated in many disciplines, including cognitive science, psychology, linguistics, geography and artificial intelligence. They act as a connecting link between visually perceived data and natural language, and an important part of research naturally deals with two types of tasks: those related to the translation of visual information into linguistic expressions (e.g., automatic digital image analysis and description), and those related to the translation of linguistic expressions into visual information (e.g., query processing in spatial database systems). In this paper, we consider only directional (also called projective) relationships (e.g., front, south, above). The past ten to fifteen years have seen significant advancements in the development of mathematical and computational models of these relationships [1] [2] [3] [4] [5]. Tasks of the first type require from such models the capability to identify which relationships hold best between any two objects. Tasks of the second type require different capabilities.

Given a directional relationship to a reference object (e.g., "east of the post office"), the models should be able to identify the objects for which the relationship holds best, and also to distinguish regions where it holds from regions where it does not hold. These regions, of course, blend into one another. They form what Bloch calls a "fuzzy landscape" [6], Logan and Sadler a "spatial template" [7], Gapp an "applicability structure" [8], Olivier and Tsuji a "potential field" [9]. Here, we will use the term "directional spatial template" (or "template", for short). There exists a simple and yet cognitively plausible way to model a template. Computationally, however, exact calculation of the model in case of 2D raster data is prohibitively expensive, and a tractable approximation algorithm was proposed in [6]. In Section 3, we present a new approximation algorithm. A comparative experimental study, summarized in Section 4, shows that it is faster, gives better results, and is more flexible. Conclusions and directions of future work are given in Section 5. First, in Section 2, we introduce the concept this new algorithm is based on.

2. Basic templates and F-templates

In the following, $\Re$ denotes the set of real numbers and $\mathcal{P}$ the Cartesian plane. μ is a mapping from $\Re$ into [0,1], periodic with period 2π, even, decreasing on $[0,\pi]$, and such that $\mu(0)=1$ and $\mu(\pi/2)=0$ (Fig. 1). An *object* is a non-empty subset of $\mathcal{P}$. For any two points q and p in $\mathcal{P}$, with q$\neq$p, the expression $\angle(q,p)$ represents the direction of the vector $\overrightarrow{qp}$. It is a value that belongs to the interval $(-\pi,\pi]$. Directional spatial relationships defy precise definitions, and Freeman proposed that fuzzy set theory should be applied [10]. The idea has been widely accepted. For any direction δ and any two objects A and B, the expression $d(\delta,A,B)$ represents the degree of truth of the proposition "A is in direction δ of B". It is a value that belongs to [0,1]. For instance, $d(0,A,B)=1$ might express the fact that the object A is perfectly to the right of B, the equality $d(-\pi/2,A,B)=0.5$ that A is somewhat behind B, and $d(\pi/2,A,B)=0$ that A is not at all to the north of B (Fig. 2).

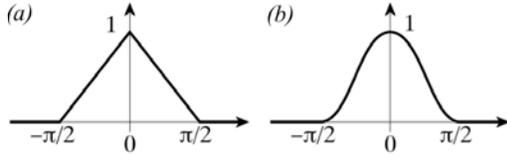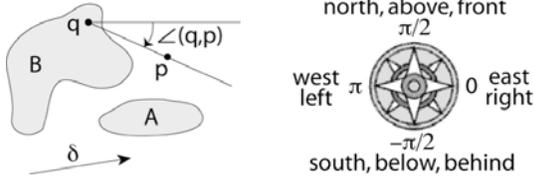**Fig. 1.** Two possible functions μ.

Fig. 2. Points, objects and directions. Is A in direction δ of B?

2.1. Basic templates

If you were told that some unknown object A was perfectly (or somewhat, or not at all) in direction δ (e.g., west, above-right) of some reference object B, where in space would you look for A? Cognitive experiments suggest that you would mentally build a spatial template [7] [11] [12]. Using essentially angular deviation, you would partition the space into regions where "in direction δ of B" holds to various degrees. It therefore makes sense to model this template by the mapping $S^{\delta B}$ from $\mathcal{P}$ into $\Re$ defined as follows:

$$\forall\, p \in \mathcal{P},\ [\, p \in B \Rightarrow S^{\delta B}(p) = 1\,] \tag{1}$$

$$\forall\, p \in \mathcal{P},\ [\, p \notin B \Rightarrow S^{\delta B}(p) = \sup_{q \in B} \mu(\angle(q,p) - \delta)\,] \tag{2}$$

We say that $S^{\delta B}$ is the *basic directional spatial template* induced by B in direction δ. As an example, Fig. 3b shows the basic template induced by some reference building in direction north. The brighter the area, the higher $S^{\delta B}(p)$, i.e., the more it is considered that the area is north of the reference building. Equation (1) is rather arbitrary. It is, however, consistent with Eq. (2), since most points p of a surface B are such that $\sup_{q \in B-\{p\}} \mu(\angle(q,p)-\delta) = 1$. Equation (2) can be compared to the aggregation method described in [1], where d(δ,A,B) is set to a combination of the $\mu(\angle(q,p)-\delta)$ values, for all p in A and q in B. The authors, who consider finite sets, suggest using the min, max, or mean operator. In Eq. (2), sup is a logical choice. Other operators would make it impossible for $S^{\delta B}$ to reach 1 outside of B (a restriction which would not be cognitively plausible).

Note that $S^{\delta B}$ allows the proposition "A is in direction δ of B" to be readily assessed for any object A. For instance, d(δ,A,B) can be set to $\sup_{p \in A} S^{\delta B}(p)$, or to $\inf_{p \in A} S^{\delta B}(p)$. These two values correspond to the most optimistic and most pessimistic points of view (Fig. 3c). They can also be interpreted as a possibility degree and a necessity degree [6].
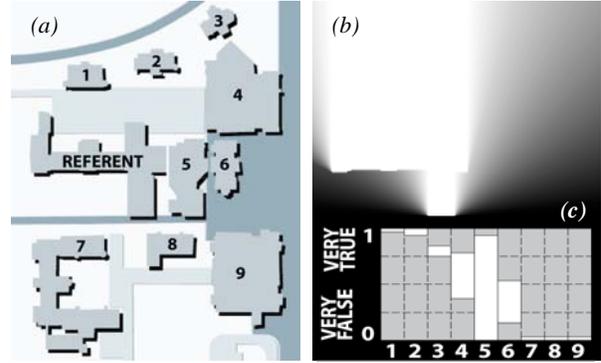

Fig. 3. *(a)* **Campus map.** *(b)* **Basic spatial template: show me where the north is (relative to the reference building).** *(c)* **Are the buildings 1 to 9 north of the reference one? The white bars represent all possible points of view.**

2.2. F-templates

The concept of the F-template is a new concept, dual to that of the F-histogram. Let us briefly describe F-histograms [13] [4]. Consider two objects A and B. An *F-histogram* associated with the pair (A,B) is a mapping F^{AB} from $\Re$ into $\Re$. It is one possible quantitative representation of the position of A with respect to B. In "F^{AB}", the letter "F" actually denotes a function. For any θ in $\Re$ and any p in $\mathcal{P}$, let $A_\theta(p)$ and $B_\theta(p)$ be the intersections of A and B with the line in the direction θ that passes through p. These intersections are *longitudinal sections* of A and B (Fig. 4a). The function F is from $\Re \times \mathcal{L} \times \mathcal{L}$ into $\Re$, where $\mathcal{L}$ denotes the set of all possible longitudinal sections. The histogram value $F^{AB}(\theta)$ is a combination of the $F(\theta, A_\theta(p), B_\theta(p))$ values, for all p (Fig. 4b). Different functions F and combination operators lead to different families of F-histograms, like Allen F-histograms [14] and force histograms [4], with applications in many areas (e.g., pattern recognition [15], scene description [16], human-robot communication [17]). The fundamental property of a force histogram F^{AB} is that it allows the propositions "A is in direction δ of B" to be readily assessed, for any δ. There are many ways, of course, to define the degree of truth d(δ,A,B) and extract it from F^{AB}. See, e.g., [16].
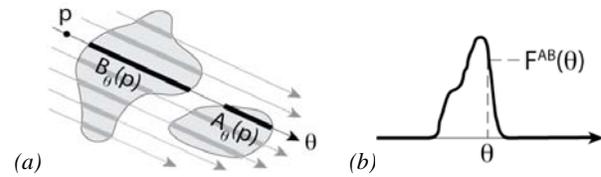

Fig. 4. *(a)* **Longitudinal sections. The value $F(\theta, A_\theta(p), B_\theta(p))$ is associated with each line.** *(b)* **F-histogram. $F^{AB}(\theta)$ is a combination of the $F(\theta, A_\theta(p), B_\theta(p))$ values, for all p.**

Now, consider the description of F-histograms that we have just provided, "replace" the set A of points with directions, and directions with points. You then get a description of F-templates. Let δ be a direction and B an object. An *F-template* associated with the pair (δ,B) (or induced by B in direction δ) is a mapping $F^{\delta B}$ from $\mathcal{P}$ into $\mathfrak{R}$. In "$F^{\delta B}$", the letter "F" denotes a function from $\mathcal{P} \times \mathfrak{R} \times \mathcal{L}$ into $\mathfrak{R}$. Consider an element p of $\mathcal{P}$. The value $F^{\delta B}(p)$ is a combination of the $F(p,\delta,B_p(\theta))$ values, for all θ (here, to be true to the duality mechanism, longitudinal sections of B are denoted by $B_p(\theta)$ instead of $B_\theta(p)$). Compare Figs. 5a and 5b. In the rest of this paper, we assume that for any p

$$F^{\delta B}(p) = \sup_{\theta \in (-\pi,\pi]} F(p,\delta,B_p(\theta)),$$

with $F(p,\delta,B_p(\theta))$ defined as follows. If $B_p(\theta)=\varnothing$ then $F(p,\delta,B_p(\theta))=0$. If $p \in B_p(\theta)$ then $F(p,\delta,B_p(\theta))=1$. Assume $B_p(\theta) \neq \varnothing$ and $p \notin B_p(\theta)$. If $\angle(q,p)=\theta$, for all $q \in B_p(\theta)$, then $F(p,\delta,B_p(\theta))=\mu(\theta-\delta)$. If $\angle(q,p)=\theta+\pi$ (or $\theta-\pi$), for all $q \in B_p(\theta)$ (and this is the case shown in Fig. 5b), then $F(p,\delta,B_p(\theta))=\mu((\theta+\pi)-\delta)$. Otherwise (imagine, in Fig. 5b, that p is between the two segments whose union is $B_p(\theta)$), then $F(p,\delta,B_p(\theta))=\max\{\mu(\theta-\delta),\mu((\theta+\pi)-\delta)\}$. It is easy to show that the F-template $F^{\delta B}$ so defined is equal to the basic directional spatial template $S^{\delta B}$ (Section 2.1). Its fundamental property, therefore, is that it allows the propositions "A is in direction δ of B" to be readily assessed, for any A. Compare Figs. 6a and 6b.
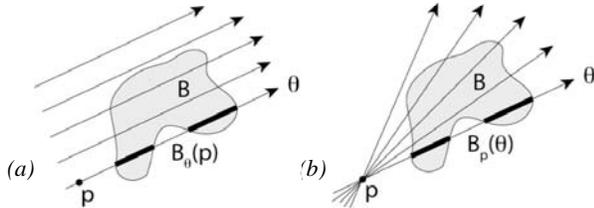

Fig. 5. (a) F-histograms. p varies, θ does not.
(b) F-templates. θ varies, p does not.

3. Case of raster objects in 2D space

In this section, objects are non-empty subsets of the set $\mathcal{I} = (0..m-1) \times (0..n-1) \subset \mathcal{P}$, with m and n two positive integers (image width and height). For any direction δ and object B, the basic directional spatial template $S^{\delta B}$ is the mapping from $\mathcal{I}$ into $\mathfrak{R}$ defined as in Section 2.1:

$$\forall \, p \in \mathcal{I}, \; [\, p \in B \Rightarrow S^{\delta B}(p) = 1 \,]$$
$$\forall \, p \in \mathcal{I}, \; [\, p \notin B \Rightarrow S^{\delta B}(p) = \max_{q \in B} \mu(\angle(q,p)-\delta) \,]$$

The algorithm that corresponds to these equations is straightforward but computationally expensive, since it is in $\mathcal{O}(m^2 n^2)$ time. In 1999, Bloch proposed an approximation algorithm based on a mathematical morphology approach [6]. Here, we introduce a new approximation algorithm based on F-templates. Let Δ be a positive

integer divisible by 4. The set $\{2\pi i/\Delta\}_{i \in 0..(\Delta-1)}$ is the set of *reference directions*. They are evenly distributed in the Cartesian plane (Fig. 7a). The F-template $F^{\delta B}$ defined as in Section 2.2 is computed very much like F-histograms, using the duality between the two concepts. The algorithm is presented next page and illustrated by Fig. 7b. Its complexity is linear in the number of pixels in the image (like the complexity of the algorithm described in [6]) and also linear in the number Δ of reference directions. The initialization step $F^{\delta B} \leftarrow S^{\delta\{q_0\}}$ is preferred to $F^{\delta B} \leftarrow 0$ because it enables efficient handling of the pixels that are far from B. This way, Δ (and processing time) can be kept low even when the reference object is small compared to the image. $S^{\delta\{q_0\}}$ is computed in $\mathcal{O}(mn)$ time.
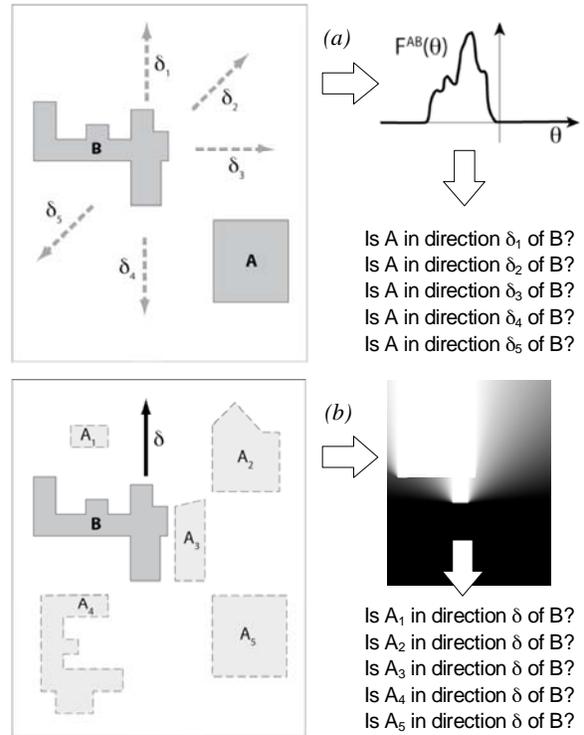

Fig. 6. (a) F-histograms. Tell me where object A is (relative to B). (b) F-templates. Show me where direction δ is (relative to B).
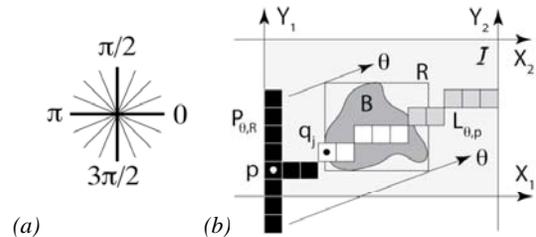

Fig. 7. (a) Reference directions. Example. (b) Algorithm. For the pixels in white (they belong to B): $F^{\delta B}(q_i) \leftarrow 1$. For the pixels in grey: $F^{\delta B}(q_i) \leftarrow \max (F^{\delta B}(q_i) , \mu(\theta-\delta))$.

Algorithm for F-template computation.

Calculate R, the minimum-bounding rectangle of B.
Pick any pixel $q_0 \in B$ and initialize: $\mathbf{F^{\delta B} \leftarrow S^{\delta(q_0)}}$.

FOR each reference direction θ **such that** $\mathbf{\mu(\theta-\delta)\neq 0}$ DO

IF $\theta \in [\pi/4, 3\pi/4)$ THEN project R along θ on the X_1-axis
ELSE IF $\theta \in [3\pi/4, 5\pi/4)$ THEN project it on the Y_2-axis
ELSE IF $\theta \in [5\pi/4, 7\pi/4)$ THEN project it on the X_2-axis
ELSE project R on the Y_1-axis.
/* Let $P_{\theta,R}$ be the segment (i.e., set of pixels)
that results from this projection. */

FOR each pixel p in $P_{\theta,R}$ DO

From p, grow a rasterized line $L_{\theta,p}$ along direction θ.
/* (Use Bresenham's algorithm [18].) */
/* When growing the line, the pixels q_1, q_2, ...q_k
are successively encountered in I. Let q_j be the
first pixel encountered in B. */

FOR all i in j..k DO

IF q_i belongs to B THEN $\mathbf{F^{\delta B}(q_i) \leftarrow 1}$
ELSE $\mathbf{F^{\delta B}(q_i) \leftarrow max\ (\ F^{\delta B}(q_i)\ ,\ \mu(\theta-\delta)\ ).}$

4. Comparative experimental study

In this section, the basic template $S^{\delta B}$ is compared with its approximations $F^{\delta B}$ (computed as in Section 3) and $M^{\delta B}$ (computed as in [6], using a morphological approach). Experiments were conducted on a 2.4GHz P4 with 1024MB memory, running *Windows 2000*.
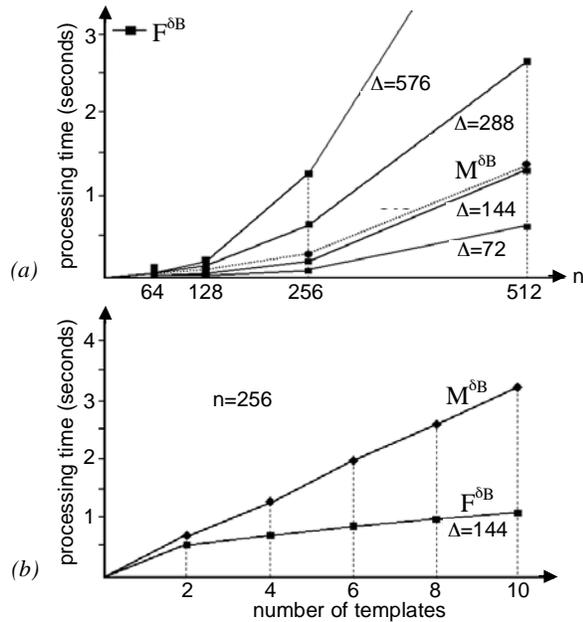

(a)

(b)

Fig. 8. Efficiency analysis. *(a)* **Processing time depends on n and, for $F^{\delta B}$, on Δ. Processing time for $S^{\delta B}$ (not shown) is about 75 seconds when n=256 and 1200 seconds when n=512.** *(b)* **F-templates induced by the same reference object in different directions can be batch-processed.**

The implementation language was $C++$. The function μ was as in Fig. 1a. The images were 8-bit m×n images, with n=m. Figures 8 and 9 illustrate well our findings, which are summarized below.

$F^{\delta B} \geq S^{\delta B}$, and the error $F^{\delta B}(p) - S^{\delta B}(p)$ is maximum (up to 0.5) for 8-neighbors of B (Fig. 9d). This error, however, decreases rapidly when getting only a few pixels further from B. When the number Δ of reference directions increases, the average error drops, then increases a bit and stabilizes (Fig. 9h). By simply adjusting Δ, users can finely control the balance between quality and processing time (Fig. 8a). When more than a hundred reference directions are considered, it becomes difficult to visually distinguish $F^{\delta B}$ (Fig. 9f) from $S^{\delta B}$ (Fig. 9b). On the other hand, $M^{\delta B}$ shows flaws (Fig. 9g). $M^{\delta B} \leq S^{\delta B}$. The error $S^{\delta B}(p) - M^{\delta B}(p)$ does not get higher than $F^{\delta B}(p) - S^{\delta B}(p)$ (up to 0.4), but it gets high for way more pixels. Moreover, these pixels tend to form noticeable patterns (Figs. 9eg).
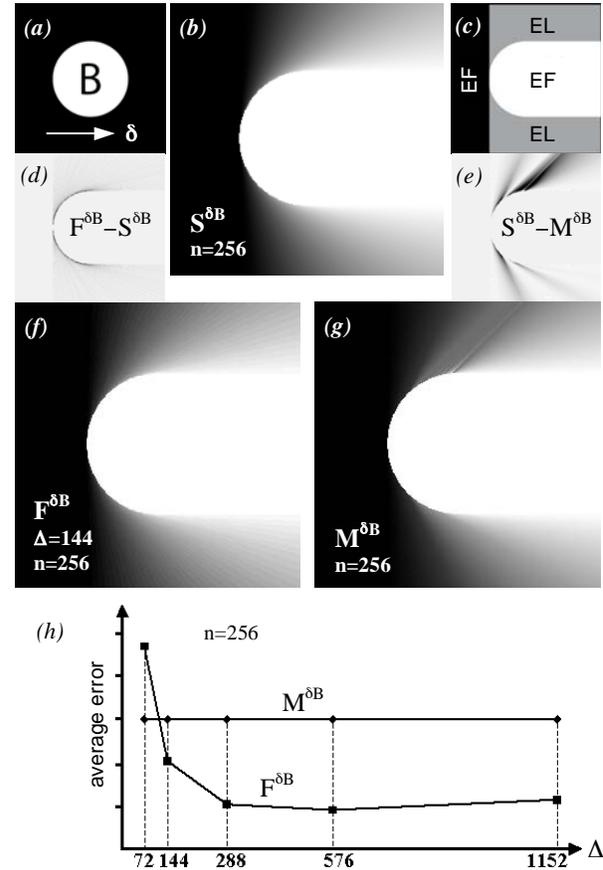

Fig. 9. Quality analysis. *(b)* $\mathbf{S^{\delta B}}$ **is approximated by** *(f)* $\mathbf{F^{\delta B}}$ **and** *(g)* $\mathbf{M^{\delta B}}$**, which are divided into** *(c)* **Error-Likely areas (in grey) and Error-Free areas.** *(d)(e)* **Contrast-enhanced error images. The darker, the higher the error.** *(h)* **The average error of $M^{\delta B}$ (around 0.02 in the error-likely areas) is about three times the average error of $F^{\delta B}$.**

In the end, basic directional spatial templates can be approximated faster and better by F-templates. Also note that F-templates induced by the same reference object B in different directions δ_1, δ_2, δ_3, … can be batch-processed (Fig. 8b). The algorithm described in Section 3 requires little modification. Only the four statements in bold are affected.

5. Conclusions and Future Work

Directional spatial templates (which are given different names in the literature) play an important role in object localization tasks. A template can be easily modeled through some mapping called *basic template*. Computationally, however, exact calculation of this model in case of 2D raster data is prohibitively expensive, and an approximation algorithm was proposed. In this paper, we have introduced a new concept, the concept of the F-template. We have shown that basic directional spatial templates can be seen as F-templates and we have presented a new approximation algorithm, which is faster, more flexible, and gives better results. The concept of the F-template is dual to that of the F-histogram. F-histograms can embed not only angular information, but also metric information. They can handle not only 2D crisp objects in raster form, but also 3D objects, fuzzy objects, and vector objects [13] [4] [19] [17]. We will show, in future publication, that F-templates have the same capabilities.

Acknowledgments

References

[1] R. Krishnapuram, J.M. Keller, and Y. Ma, "Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions", *IEEE Trans. on Fuzzy Systems*, 1993, vol. 1, no. 3, pp. 222-233.

[2] J. Sharma, and D.M. Flewelling, "Inferences from Combined Knowledge about Topology and Direction", *SSD'95 (Advances in Spatial Databases: 42nd Symposium)*, 1995, *Proceedings*, pp. 279-291.

[3] D. Papadias, and Y. Theodoridis, "Spatial Relations, Minimum Bounding Rectangles, and Spatial Data Structures", *Int. J. of Geographical Information Science*, 1997, vol. 11, pp. 111-138.

[4] P. Matsakis, and L. Wendling, "A New Way to Represent the Relative Position of Areal Objects", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1999, vol. 21, no. 7, pp. 634-643.

[5] F. Petry, M. Cobb, D. Ali, R. Angryk, M. Paprzycki, S. Rahimi, L. Wen, and H. Yang, "Fuzzy Spatial Relationships and Mobile Agent Technology in Geospatial Information Systems", in P. Matsakis, and L. Sztandera (Eds.), *Applying Soft Computing in Defining Spatial Relations*, Studies in Fuzziness and Soft Computing, Physica-Verlag, 2002, vol. 106, pp. 123-155.

[6] I. Bloch, "Fuzzy Relative Position Between Objects in Image Processing: A Morphological Approach", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1999, vol. 21, no. 7, pp. 657-664.

[7] G.D. Logan, and D.D. Sadler, "A Computational Analysis of the Apprehension of Spatial Relations", in P. Bloom, M.A. Peterson, L. Nadel, and M.F. Garrett (Eds.), *Language and Space*, Cambridge, MA: MIT Press, 1996.

[8] K.-P. Gapp, "Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space", *12th National Conf. on Artificial intelligence*, Seattle, Washington, 1994, *Proceedings*, vol. 2, pp. 1393-1398.

[9] P. Olivier, and J.-I. Tsuji, "Quantitative Perceptual Representation of Prepositional Semantics", in *Artificial Intelligence Review (Special Issue on Integration of Natural Language and Vision Processing)*, 1994, vol. 8, pp. 147-158.

[10] J. Freeman, "The Modeling of Spatial Relations", *Computer Graphics and Image Processing*, 1975, vol. 4, pp. 156-171.

[11] K.-P. Gapp, "Angle, Distance, Shape, and Their Relationship to Projective Relations", *17th Annual Conf. of the Cognitive Science Society*, Pittsburgh, PA, 1995, *Proceedings*, Mahwah, NJ: Lawrence Erlbaum, pp. 112-117.

[12] N. Franklin, L.A. Henkel, and T. Zengas, "Parsing Surrounding Space Into Regions", *Memory and Cognition*, 1995, vol. 23, pp. 397-407.

[13] P. Matsakis, *Relations Spatiales Structurelles et Interprétation d'images*, PhD Thesis, Institut de Recherche en Informatique de Toulouse, France, 1998.

[14] P. Matsakis, and D. Nikitenko, "Combined Extraction of Directional and Topological Relationship Information from 2D Concave Objects", in M. Cobb, F. Petry, and V. Robinson (Eds.), *Fuzzy Modeling with Spatial Information for Geographic Problems*, Springer-Verlag Publications, 2005, pp. 15-40.

[15] L. Wendling, S. Tabbone, and P. Matsakis, "Fast and Robust Recognition of Orbit and Sinus Drawings Using Histograms of Forces", *Pattern Recognition Letters*, 2002, vol. 23, no. 14, pp. 1687-1693.

[16] P. Matsakis, J. Keller, L. Wendling, J. Marjamaa, and O. Sjahputera, "Linguistic Description of Relative Positions in Images", *TSMC Part B* (*IEEE Trans. on Systems, Man and Cybernetics*), 2001, vol. 31, no. 4, pp. 573-588.

[17] M. Skubic, P. Matsakis, G. Chronis, and J. Keller, "Generating Multi-level Linguistic Spatial Descriptions from Range Sensor Readings Using the Histogram of Forces", *Autonomous Robots*, 2003, vol. 14, no. 1, pp. 51-69.

[18] J.E. Bresenham, "Algorithm for Computer Control of a Digital Plotter", *IBM Systems J.*, 1965, vol. 4, no. 1, pp. 25-30.

[19] J. Ni, P. Matsakis, and L. Wawrzyniak, "Quantitative Representation of the Relative Position Between 3D Objects", *VIIP 2004 (4th IASTED Int. Conf. on Visualization, Imaging, and Image Processing)*, Marbella, Spain, 2004, *Proceedings*, ACTA Press, pp. 452-289.