DIRECTIONAL RELATIONSHIPS TO A REFERENCE OBJECT: A QUANTITATIVE APPROACH BASED ON FORCE FIELDS

Pascal Matsakis, JingBo Ni, Melanie Veltman

Computing and Information Science, University of Guelph, Canada, {pmatsaki, jni}@uoguelph.ca

ABSTRACT

A directional map (or spatial template, fuzzy landscape) is an image where the value of each pixel represents the degree to which the pixel satisfies some directional relationship (e.g., right, left, above, below) to some reference object (i.e., a given set of pixels). There exists a simple quantitative model of such relationships, and the computation of directional maps is usually based on algorithmic implementations of this model. We show here that the model has important flaws, and we respond to the issue with a new, promising approach: all directional maps induced by the reference object are generated from a force field that the object (which is seen as a physical entity) creates around itself. Preliminary experiments illustrate and show the interest of the approach.

Index Terms — Directional relationships, object localization, directional maps, force fields, spatial templates

1. INTRODUCTION

The modeling of spatial relations [1,2,3,4] involves two fundamental questions: (a) How to identify the spatial relations between two given objects? (b) How to identify the object that best satisfies a given relation to a reference object? The second question defines an object localization task (and has received less attention). Cognitive experiments show that people accomplish this task by parsing space around the reference object into *good* regions (where the object being looked for is more likely to be), *acceptable* and *unacceptable* regions [5,6]. These regions blend into one another and define a *spatial template* [6]: each point in space is assigned a value between 0 (unacceptable region) and 1 (good region).

In this paper, we focus on *directional* (also called *projective* [7] or *cardinal* [2]) relationships (e.g., front, south, above). A spatial template (also called *fuzzy landscape* [1]) can then be referred to as a *directional map*, which is in fact an image where the value of each pixel represents the degree to which the pixel satisfies some directional relation to some reference object (i.e., to a given set of pixels). Directional maps can be used for, e.g., spatial reasoning, object localization and identification, structural and model-based pattern recognition [8,9,10]. Let us illustrate this with a simple example. Assume you are familiar with the two reference
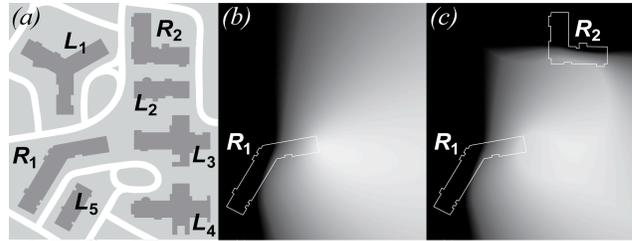
Fig. 1. Directional maps: *(a)* Buildings; *(b)* The directional map "east of R_1". The brighter the region, the more it is considered that the region is east of R_1. Black corresponds to 0 and white to 1; *(c)* The map "east of R_1 and south of R_2".

buildings R_1 and R_2 (Fig. 1a). You need to go to Escher Hall, which you do not know. You are told, however, that it is the building east of R_1 and south of R_2. Which one among L_1 to L_5 do you think is Escher Hall? Let us find it together. First, compute the directional maps $M_1 =$ "east of R_1" (Fig. 1b) and $M_2 =$ "south of R_2" (not shown here). We will expand on this task in the rest of the paper. Then, combine M_1 and M_2 into a single map $M =$ "east of R_1 and south of R_2" (Fig. 1c); e.g., assign each pixel p the value $M(p)=\min\{M_1(p),M_2(p)\}$. Now, let us rate each candidate building L on a scale from 0 (worst) to 1 (best). A pessimistic soul might rate it $\min_{p\in L} M(p)$, a temperate soul $\text{ave}_{p\in L} M(p)$ and an optimistic one $\max_{p\in L} M(p)$. Our decision can actually be based on all three values, i.e., on the triple

$$d(L) = (\min_{p\in L} M(p), \text{ave}_{p\in L} M(p), \max_{p\in L} M(p)). \quad (1)$$

For instance, L can only be Escher Hall if $d(L)=(1,1,1)$. It cannot be Escher Hall if $d(L)=(0,0,0)$. There is no certainty at all but it is totally possible that L is Escher Hall if $d(L)=(0,0.5,1)$. There exists a simple quantitative model of the directional relationships to a reference object, and the computation of directional maps is usually based on algorithmic implementations of this model (Section 2). We show here that the simplicity of the model has eclipsed important flaws. We respond to the issue with a new, promising approach: it relies on the idea of considering the reference object as a physical entity that creates a force field around itself (Section 3). All directional maps induced by the object can be generated from the force field in negligible time (Section 4). Preliminary experiments illustrate and show the interest of the approach (Section 5).

2. DIRECTIONAL MAPS

In the rest of the paper, $\mathbb{P}$ is the Euclidean plane. An *object* R is a nonempty bounded subset of $\mathbb{P}$. For any point p of $\mathbb{P}$ we have $R(p)=1$ if $p \in R$ and $R(p)=0$ if $p \notin R$. The symbol δ denotes a unit vector. The vector from p to q is pq. Its norm (i.e., length) is $|pq|$. The radian measure in $[0,\pi]$ of the angle between two nonzero vectors u and v is denoted by $\angle(u,v)$. μ is a non-increasing function from $[0,\pi]$ into $[0,1]$ such that $\mu(0)=1$ and $\mu(\pi/2)=0$. Two possible candidates for $\mu(x)$ are $\mu_{lin}(x)=\max\{0,1-2x/\pi\}$ and $\mu_{cos}(x)=\max\{0,\cos x\}$.

Given a point p, the function that assigns the value $\mu(\angle(\delta,pq))$ to each point q is the *directional map* induced by p in direction δ. There is no consensus on the terminology, but there is clear consensus on the definition. Directional maps induced by R (instead of p) are not, however, easy to define. The simplest way is to replace R with its centroid p_R and assign the value $\mu(\angle(\delta,p_Rq))$ to each q. We then get the *centroid-based* map $C^{\delta R}$. There exists, however, a quantitative model of the directional relationships to a reference object, which is simple, does not sacrifice the geometry of the object, and whose basic principle is supported by cognitive studies. The *standard map* $S^{\delta R}$ depends essentially on angular deviation (in accordance with [5,6,7]) and is defined by

$$S^{\delta R}(q) = \sup\nolimits_{p \in R} \mu(\angle(\delta,pq)). \qquad (2)$$

In practice, exact calculation of $S^{\delta R}$ is computationally expensive. Two approximation algorithms (one based on a morphological approach [1] and one based on the partitioning of the image into parallel raster lines [11]) have been proposed and are commonly used [8,9,10]. Standard maps, unfortunately, have important flaws, as will be shown in Section 5.

3. FORCE FIELDS

Consider two point-like particles p and q of mass 1. According to Newton's law of gravity, q exerts on p a force $pq/|pq|^{r+1}$ of magnitude $1/|pq|^r$, where $r = 2$. Now, imagine R is a flat metal plate of uniform area density 1. The particle q exerts on R the force

$$\Phi_r^R(q) = \iint_{p \in R} \frac{pq}{|pq|^{r+1}} \, dp = \iint_{p \in \mathbb{P}} R(p) \frac{pq}{|pq|^{r+1}} \, dp. \qquad (3)$$

In this paper, we are not bound to physical laws, and r can be any real number. The function Φ_r^R is called the *force field* induced by R. We introduce here two algorithms for computing force fields in the case of 2D raster data. First, let us rewrite (3) using the Cartesian coordinates (u,v) and (x,y) of q and p:

$$\Phi_r^R(u,v) = \int_x \int_y R(x,y) \frac{(u-x,v-y)}{|(u-x,v-y)|^{r+1}} \, dx \, dy. \qquad (4)$$

In the discrete domain of an $N=m \times n$ image, (4) becomes
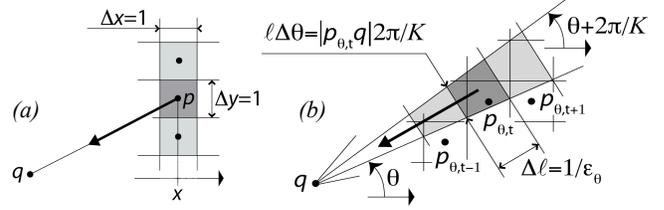
Fig. 2. Force field computation: *(a)* Using Cartesian coordinates; *(b)* Using polar coordinates.

$$\Phi_r^R(u,v) = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} R(x,y) \frac{(u-x,v-y)}{|(u-x,v-y)|^{r+1}} \Delta x \Delta y. \qquad (5)$$

The surface elements the plane is divided into are the pixels of the image, with width $\Delta x=1$ and height $\Delta y=1$ (Fig.2a). Note that when $(x,y)=(u,v)$, all forces in the surface element cancel each other out for symmetry reasons. This results in the convention that $0/0=0$. Equation (5) calculates $\Phi_r^R(u,v)$ in $\mathcal{O}(N)$ time, i.e., it calculates Φ_r^R in $\mathcal{O}(N^2)$ time.

Now, let us move the origin of the coordinate system to q and let us rewrite (3) using the polar coordinates (θ,ℓ) of p. Since the Jacobian determinant of the coordinate conversion formula is ℓ, i.e., since $dxdy=\ell d\theta d\ell$, we have

$$\Phi_r^R(0,0) = \int_\theta \int_\ell R(\theta,\ell) \frac{(-\ell\cos\theta,-\ell\sin\theta)}{\ell^{r+1}} \ell \, d\theta \, d\ell. \qquad (6)$$

In the discrete domain, (6) becomes

$$\Phi_r^R(0,0) = \sum_\theta \sum_\ell R(\theta,\ell) \frac{(-\ell\cos\theta,-\ell\sin\theta)}{\ell^{r+1}} \ell\Delta\theta\Delta\ell. \qquad (7)$$

θ belongs to a set $\{2\pi k/K\}_{k \in 0..K-1}$ of K reference directions. $\Delta\theta$, therefore, is $2\pi/K$. Starting from q, and using, e.g., Bresenham's algorithm [12], draw a line in direction θ. The pixels $p_{\theta,0}=q$, $p_{\theta,1}$, $p_{\theta,2}$, etc., are successively encountered. Each one is attached to a surface element of width $\Delta\ell=1/\varepsilon_\theta$, where $\varepsilon_\theta=|\cos\theta|$ if $\theta \in [0,\pi/4] \cup [3\pi/4,5\pi/4] \cup [7\pi/4,2\pi)$ and $\varepsilon_\theta=|\sin\theta|$ otherwise (Fig.2b). In the end, (7) is rewritten as

$$\Phi_r^R(q) = -\frac{2\pi}{K} \sum_\theta \left\{ \sum_t \frac{R(p_{\theta,t})}{|p_{\theta,t}q|^{r-1}} \right\} \frac{(\cos\theta,\sin\theta)}{\varepsilon_\theta}. \qquad (8)$$

This equation calculates $\Phi_r^R(q)$ in $\mathcal{O}(K\sqrt{N})$ time, i.e., it calculates Φ_r^R in $\mathcal{O}(KN\sqrt{N})$ time.

4. FORCE FIELD-BASED MAPS

Force fields can be used to generate a variety of directional maps. Consider a given direction (unit vector) δ. Let us show how to derive a map $\Phi_r^{\delta R}$ from the field Φ_r^R. The inequality $\Phi_r^R(q) \cdot \delta \leq 0$ (where "$\cdot$" denotes the dot product) is seen as an indication that the point q is not at all in direction δ of R. The value of $\Phi_r^{\delta R}(q)$ is then set to 0. What value should $\Phi_r^{\delta R}(q)$ take when $\Phi_r^R(q) \cdot \delta > 0$? Two directional transformations are proposed. In the discrete domain of an image of size N, the $\Phi_r^{\delta R}$ described below are generated from Φ_r^R in $\mathcal{O}(N)$ time, and processing times are negligible compared to those for force field computation.

4.1. First directional transformation

Here, $\quad \Phi_r^R(q)\cdot\delta > 0 \;\Rightarrow\; \Phi_r^{\delta R}(q) = \mu(\angle(\delta,\Phi_r^R(q)))$ (9)

μ is as in Section 2, and $\angle(\delta,\Phi_r^R(q))$ is the radian measure in $[0,\pi]$ of the angle between the vectors δ and $\Phi_r^R(q)$.

4.2. Second directional transformation

In this section, the idea is to assign q a directional value $\Phi_r^{\delta R}(q)$ that is a non-decreasing function of $\Phi_r^R(q)\cdot\delta$. The equation below represents an obvious way to proceed:

$$\Phi_r^{\delta R}(q) = \max\left\{0, \frac{\Phi_r^R(q)\cdot\delta}{\sup_{p\in\mathbb{P}}\Phi_r^R(p)\cdot\delta}\right\}. \quad (10)$$

We can show that $|\Phi_r^R|$ is bounded when $r\in[0,2)$. However, $\sup_{p\in\mathbb{P}}\Phi_r^R(p)\cdot\delta$ cannot be easily determined. In practice, therefore, (10) is replaced with

$$\Phi_r^{\delta R}(q) = \max\left\{0, \min\left\{1, \frac{\Phi_r^R(q)\cdot\delta}{\tilde{\Phi}_r^{\delta R}}\right\}\right\}, \quad (11)$$

where $\tilde{\Phi}_r^{\delta R}$ is an approximation of $\sup_{p\in\mathbb{P}}\Phi_r^R(p)\cdot\delta$.

5. EXPERIMENTS

In Sections 5.2 to 5.5, the symbols $\Phi_{r,1}$ and $\Phi_{r,2}$ denote the map $\Phi_r^{\delta R}$ induced by R in direction δ as defined in Sections 4.1 and 4.2 respectively. $d_{r,1}(L)$ and $d_{r,2}(L)$ denote triples, as in (1), calculated from $\Phi_{r,1}$ and $\Phi_{r,2}$.

5.1. Comparing the two algorithms

Let $^{\mathrm{CART}}\Phi_r^R$ and $^K\Phi_r^R$ be the force fields calculated according to (5) and (8) respectively. The difference between $^{\mathrm{CART}}\Phi_r^R$ and $^K\Phi_r^R$ is measured by the difference ratio

$$DR = \frac{\sum_x\sum_y\left|\,|^{\mathrm{CART}}\Phi_r^R(x,y)| - |^K\Phi_r^R(x,y)|\,\right|}{\sum_x\sum_y\max\left\{|^{\mathrm{CART}}\Phi_r^R(x,y)|,|^K\Phi_r^R(x,y)|\right\}}, \quad (12)$$

which takes values between 0 and 1, and is 0 if and only if $^{\mathrm{CART}}\Phi_r^R = {}^K\Phi_r^R$. The force field $^K\Phi_r^R$ approximates $^{\mathrm{CART}}\Phi_r^R$. Higher K means higher accuracy (lower DR) (Fig. 3d), but also longer processing time (Fig. 3b). The accuracy depends mostly on K, and is quite high ($DR<1\%$) even when K is relatively small ($K=90$), but it also depends a bit on r and on the image size N. In particular, it gets higher when r gets closer to 1—although the variation is small when r is confined to the interval $[0,2]$ (Fig. 3c). The processing time increases significantly with K and N but does not depend on r.

5.2. Sensitivity to outliers

Standard maps are overly sensitive to outliers and can change drastically because of one single pixel. According to the map in Fig. 4b, the small disk L is indubitably to the right of the object R defined by Fig. 4a ($d_S(L)=(0.95,0.98,1.00)$). This goes against intuition. Centroid and force field-based maps (Fig. 4c) are not affected by the outlier pixel and consider that L is not at all to the right of R ($d_{0,2}(L)=(0,0,0)$).
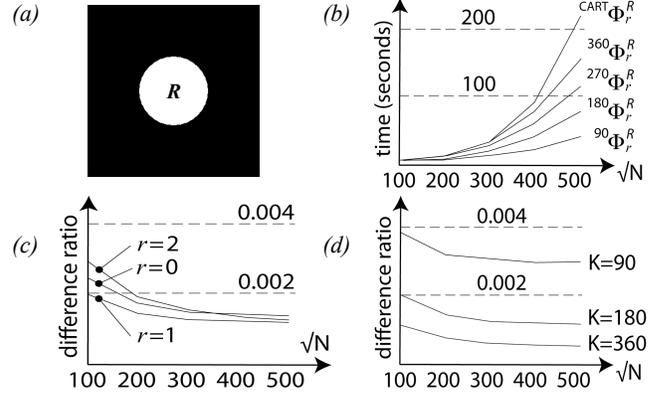
Fig. 3. Comparison of $^{\mathrm{CART}}\Phi_r^R$ and $^K\Phi_r^R$: *(a)* Test image, with N pixels; *(b)* Processing times; *(c)* DR values when $K=180$; *(d)* DR values when $r=1$. The algorithms were implemented in C and run on a machine with Intel Pentium D CPU 3.0GHz and 1GB memory.
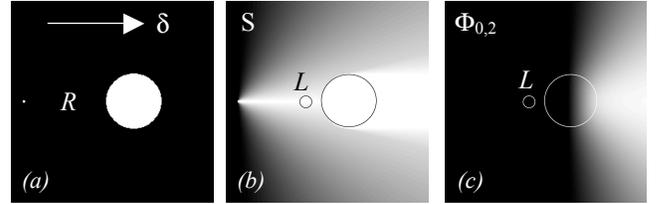
Fig. 4. Sensitivity to outliers: *(a)* R includes the disk, but it also contains the pixel (barely visible) close to the left edge of the image; *(b)* Standard map; *(c)* Force field-based map.
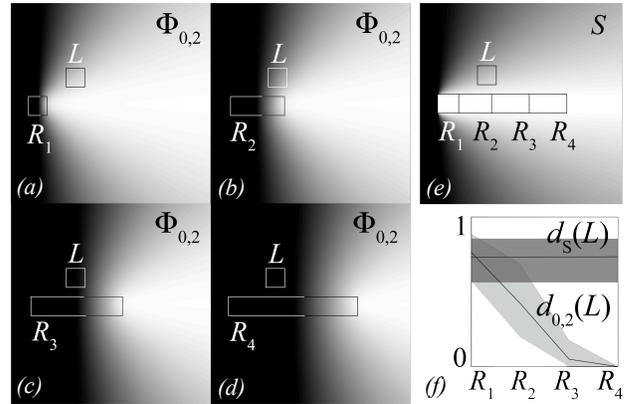
Fig. 5. Case of elongated objects. *(a)-(d)* A force field-based map induced in one direction changes when the reference object grows in that direction. *(e)* This is not the case of standard maps. *(f)* Is L to the right of R_1, R_2, R_3, R_4? Two different views.

5.3. Case of elongated objects

Consider Fig. 5abcd. When R_1 lengthens and becomes R_2, then R_3, and then R_4, the force field-based map $\Phi_{0,2}$ changes accordingly. Therefore, although L is somewhat to the right of R_1 ($d_{0,2}(L)=(0.6,0.8,0.9)$), this becomes less and less true as R_1 grows (Fig. 5f), and L is not at all to the right of R_4 ($d_{0,2}(L)=(0,0,0)$). Comparable results are obtained with other force field-based maps, and with centroid-based maps. They are intuitively sound. On the other hand, the standard maps induced by the four objects are exactly the same (Fig.5e),

and L is to the right of R_4 as much as it is to the right of R_1 ($d_S(L)$=(0.6,0.8,0.9)). This goes against common sense.

5.4. Case of concave objects

Consider the concentric shell R and the disks L_1 to L_6 in Fig. 6a. The standard map S (Fig. 6b) assigns 1 to every pixel in the region enclosed by R and considers that every disk L_i is perfectly to the right of R ($d_S(L_i)$=(1,1,1)). This is another counterintuitive result. On the other hand, according to the force field-based map $\Phi_{0.5,1}$ (Fig. 6c), L_1 and L_2 are not at all to the right of R; L_3 and L_4 are somewhat to the right of R; L_5 and L_6 are to the right of R. See Fig. 6f. The centroid-based map C (not shown here) shares exactly the same view about L_1, L_2, L_5, L_6, but is more positive about L_3, L_4 (e.g., $d_C(L_3)$=(0.6,0.8,1.0) while $d_{0.5,1}(L_3)$=(0.4,0.7,1.0)).

$\Phi_{1.0,1}$ and $\Phi_{1.5,1}$ (Fig. 6de) are way different from S and C and show unique characteristics. According to $\Phi_{1.0,1}$ (Fig. 6g), L_1 to L_4 are not at all to the right of R. The map $\Phi_{1.5,1}$ takes a more local view of the situation and considers that L_1, L_2 are to the right of R because they definitely are to the right of the closest part of R (Fig. 6h). Note that if L_6 was drifting to the right, away from the shell, $\Phi_{1.5,1}$ would still consider (with the same confidence) that L_6 is to the right of the shell. This shows the main difference between $\Phi_{r,1}$ and $\Phi_{r,2}$ when r>0: the triple $d_{r,1}(L_6)$ would remain pretty much equal to (1,1,1), while $d_{r,2}(L_6)$ would decrease and tend towards (0,0,0).

5.5. Escher Hall

According to C, S, $\Phi_{r,1}$, $\Phi_{0,2}$, the triple $d(L)$ cannot decrease when L is moving away from R in direction δ. In most cases, it actually increases: the farther L, the more it is in direction δ of R. For example, all these approaches believe that the building L_4 in Fig. 1a is a better candidate than L_3 for "south of R_2"—and a building on the other side of town would be an even better candidate. $\Phi_{r,2}$ with r>0 breaks this rule. The directional maps in Fig. 1bc are actually $\Phi_{0.1,2}$ maps. Contrary to S maps, they consider that L_3 is a better candidate than L_4 for "south of R_2", and is the best candidate for "east of R_1 and south of R_2". We have found Escher Hall.

6. CONCLUSION

We have shown that the standard quantitative model of the directional relationships to a reference object has important flaws. In response to the issue, we have designed new models. The reference object R is seen as a physical entity that creates a force field around itself. The directional maps induced by the object can be generated from the force field. These maps show unique characteristics. Distance is explicitly taken into account. The closest parts of the object can receive more or less attention. Note that all the equations presented here still hold when R is fuzzy (i.e., $R(p)$ belongs to [0,1] instead of {0,1}). In future work, we will introduce an efficient algorithm for force field computation and we will extend the approach to vector and 3D data.
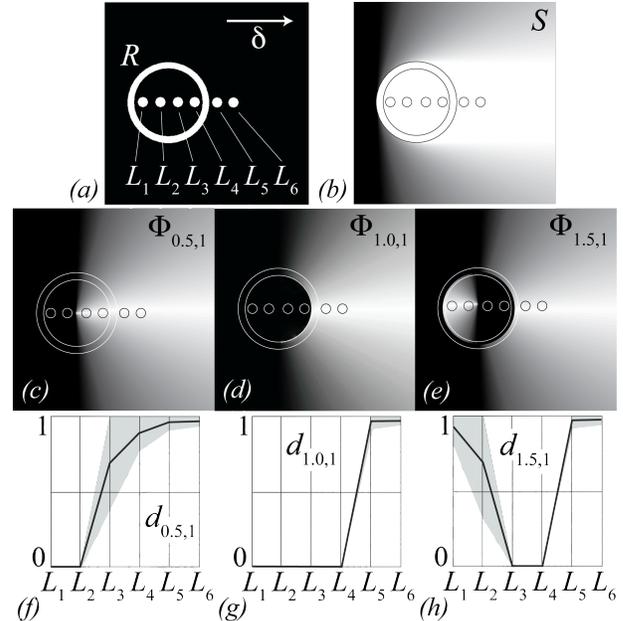
Fig. 6. Case of concave objects: *(a)* Direction, reference and located objects; *(b)* The standard map; *(c)(d)(e)* Various force field-based maps and *(f)(g)(h)* corresponding $d(L)$ values.

7. REFERENCES

[1] I. Bloch, "Fuzzy relative position between objects in image processing: a morphological approach", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(7), pp. 657-64, 1999.

[2] A. U. Frank, "Qualitative spatial reasoning: Cardinal directions as an example", *Int. J. of Geographical Information Systems*, 3(10), pp. 269-90, 1996.

[3] R. Krishnapuram, J. M. Keller, Y. Ma, "Quantitative analysis of properties and spatial relations of fuzzy image regions", *IEEE Trans. on Fuzzy Systems*, 1(3), pp. 222-33, 1993.

[4] P. Matsakis, L. Wendling, "A new way to represent the relative position between areal objects", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(7), pp. 634-43, 1999.

[5] N. Franklin, L. Henkel, T. Zangas, "Parsing surrounding space into regions", *Memory & Cognition*, 23(4), pp. 397-407, 1995.

[6] G. D. Logan, D. D. Sadler, "A computational analysis of the apprehension of spatial relations", *Language and Space*, Cambridge, MA: MIT Press, pp. 493-529, 1996.

[7] K.-P. Gapp, "Angle, distance, shape, and their relationship to projective relations", *Proceedings of the 17th Annual Conf. of the Cognitive Science Society*, pp. 112-7, 1995.

[8] Colliot, O. Camara, I. Bloch, "Integration of fuzzy spatial relations in deformable models: Application to brain MRI segmentation", *Pattern Recognition*, pp. 1401-14, 2006.

[9] R. Krishnapuram, S. Medasani, S.-H.Jung, Y.-S. Choi, R. Balasubramaniam, "Content-based image retrieval based on a fuzzy approach", *IEEE Trans. on Knowledge and Data Engineering*, 16(10), pp. 1185-99, 2004.

[10] Smith, G. Brannon, S. Bridges, "Fuzzy spatial data mining". *North American Fuzzy Information Processing Society Conf.*, New Orleans, LA, pp. 184-89, 2002.

[11] P. Matsakis, J. Ni, X. Wang, "Object localization based on directional information: Case of 2D raster data", *Proceedings of the 18th Int. Conf. on Pattern Recognition*, 2, pp. 142-6, 2006.

[12] J. E. Bresenham, "Algorithm for computer control of a digital plotter", *IBM Systems Journal*, 4(1), pp. 25-30, 1965.