Linguistic Description of Relative Positions in Images

Pascal Matsakis, James M. Keller, *Fellow, IEEE*, Laurent Wendling, Jonathan Marjamaa, and Ozy Sjahputera

***Abstract*—Fuzzy set methods have been used to model and manage uncertainty in various aspects of image processing, pattern recognition, and computer vision. High-level computer vision applications hold a great potential for fuzzy set theory because of its links to natural language. Linguistic scene description, a language-based interpretation of regions and their relationships, is one such application that is starting to bear the fruits of fuzzy set theoretic involvement. In this paper, we are expanding on two earlier endeavors. We introduce new families of fuzzy directional relations that rely on the computation of histograms of forces. These families preserve important relative position properties. They provide inputs to a fuzzy rule base that produces logical linguistic descriptions along with assessments as to the validity of the descriptions. Each linguistic output uses hedges from a dictionary of about 30 adverbs and other terms that can be tailored to individual users. Excellent results from several synthetic and real image examples show the applicability of this approach.**

***Index Terms*—Force histograms, fuzzy logic, linguistic descriptions, relative positions, scene understanding, spatial relations.**

I. INTRODUCTION

DESCRIPTION of natural scenes is one of the most important tasks in an image understanding system. Over the years, it has received considerable attention. The ACRONYM system by Brooks [1] was an early approach to model-based image understanding that identified object instances in the image by matching from a picture graph and an observability graph. Constraint networks [2], [3] have been used to identify where objects might be located in a scene. Andress and Kak [4] used the Dempster-Shafer belief framework to interpret images from knowledge consisting of line drawings of the expected scene. Walker *et al.* [5] developed a system for reasoning about lines, planes, and polygons in two and three dimensions. This work was extended to incorporate fuzzy set theoretic operations to control perceptual grouping of primitive elements [6]. Antony described an object oriented database management system to support spatial and temporal reasoning [7]. He constructed a framework within which spatial, temporal and hierarchical scene reasoning can take place, although no actual imagery was analyzed. He discussed the use of fuzzy set theory for expressing constraints such as “near,” and used quadtree representations to determine (crisp) areas of an image that would correspond to spatial concepts like “northeast.” In [8], Medasani and Krishnapuram described a fuzzy approach to content-based image retrieval that can handle linguistic queries involving region labels, attribute values, and spatial relations.

Given the importance of spatial relations in the description of a scene, many methods have been created to define them for digital image objects. Winston [9] was interested in quantifying spatial relations to create a program that could learn to recognize line drawing representations of structures by building an abstract representation of a given line drawing and examining the applicability of various internalized structure descriptions. He used rules to generate descriptions for line drawings of three-dimensional (3-D) scenes. These descriptions were crisp and context sensitive by nature. The relations involved were ABOVE, SUPPORTS, IN-FRONT-OF, LEFT, RIGHT, and MARRIES. A few years later, Freeman [10] proposed that the relative position of objects be described in terms of 13 primitive spatial relations: 1) LEFT OF, 2) RIGHT OF, 3) ABOVE, 4) BELOW, 5) BEHIND, 6) IN FRONT OF, 7) BESIDE, 8) NEAR, 9) FAR, 10) TOUCHING, 11) BETWEEN, 12) INSIDE, and 13) OUTSIDE. The first six are called the primitive directional relations. While humans seem capable of ascertaining them, they are exceedingly difficult to define precisely. “All-or-nothing” standard mathematical relations are clearly not suitable, and Freeman proposed that fuzzy relations be used. However, computers have not been able to effectively model these vital spatial concepts. For instance, many authors assimilated two-dimensional (2-D) objects to very elementary entities such as a point (centroid) or a (bounding) rectangle [11]–[14]. The procedure is practical, but cannot be hoped to give a satisfactory modeling.

By introducing the notion of the histogram of angles, Miyajima and Ralescu [15] developed the idea that the relative position between two objects can have a representation of its own and can thus be described in terms other than spatial relationships. Actually, angle histograms have generally been used to assess directional relations [15]–[18]. In [19], Keller and Wang presented a fuzzy rule-based approach for linguistic scene description, where directional relationship values were generated from neural networks fed by angle histograms. These networks were trained on aggregate responses from a panel of people [18], and the spatial relationship values were combined with other world knowledge encoded in fuzzy logic rules to produce a final linguistic analysis. However, the “language” used in [19] was very coarse, additional metric features were needed, and the methods of generating the basic spatial relation membership functions did not satisfy certain “reasonable” criteria, such

Manuscript received July 7, 2000; revised February 3, 2001. The work of P. Matsakis, J. Keller, J. Marjamaa, and O. Sjahputera was supported by the Office of Naval Research under Grant N00014-96-0439. The work of J. Keller was supported in part by the Humanitarian De-mining MURI Program from the Army Research Office under Contract DAAG55-97-1-0014. This paper was recommended by Associate Editor A. Bensaid.

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors. This includes six multimedia AVI format movie clips (see Section V-A), which show natural scenes with linguistic descriptions. This material is 2.1 MB in size.

P. Matsakis, J. M. Keller, J. Marjamaa, and O. Sjahputera are with the Department of Computer Science and Computer Engineering, University of Missouri, Columbia, MO 65211 USA (e-mail: pmatsakis@cecs.missouri.edu; keller@cecs.missouri.edu).

L. Wendling is with the LORIA, BP 239, 54506 Vandoeuvre Cedex, France.

Publisher Item Identifier S 1083-4419(01)05664-3.

1083–4419/01$10.00 ©2001 IEEE

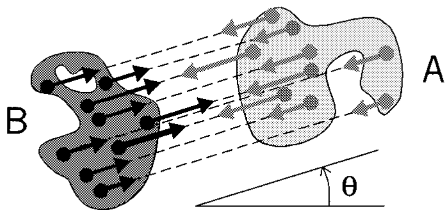


Fig. 1. Computation of $F^{AB}(\theta)$. It is the scalar resultant of forces (black arrows). Each one tends to move B in direction θ.

as the semantic inverse principle [10]. Better descriptive terminology needs to be created and matched to the spatial relationship definitions, particularly if the description language is to be tailored to particular individual experts.

In [20]–[22], Matsakis and Wendling introduced the notion of the histogram of forces. It generalizes and supersedes that of the histogram of angles. It ensures rapid processing of raster data as well as of vector data, and of crisp objects as well as of fuzzy objects. It offers solid theoretical guarantees, and allows explicit and variable accounting of metric information. We briefly present this notion in Section II, and in Section III we describe new families of fuzzy directional relations relying on the computation of force histograms. These families are interfaced with a set of fuzzy rules to generate linguistic descriptions of relative positions, as explained in Section IV. The descriptions contain a richer language than that found in [19]. We tune our system from a group of simple images, and we demonstrate it in Section V on a large number of synthetic images as well as Laser Radar (LADAR) range images of a complex power plant scene.

II. F-Histograms

We represent the relative position of a 2-D object A with regard to another object B by a function F^{AB} from $\Re$ into $\Re_+$. For any direction θ, the value $F^{AB}(\theta)$ is the total weight of the arguments that can be found in order to support the proposition "A is in direction θ of B." More precisely, it is the scalar resultant of elementary forces. These forces are exerted by the points of A on those of B, and each tends to move B in direction θ (Fig. 1). If F^{AB} is defined on $\Re$, i.e., if for any θ the scalar resultant $F^{AB}(\theta)$ is finite, then the pair (A, B) is termed *F-assessable* and F^{AB} is called the *histogram of forces associated with (A, B) via F*, or the *F-histogram associated with (A, B)*. The object A is the *argument*, and the object B the *referent*. Note that in the figures throughout this paper, the referent is always drawn darker than the argument.

Actually, the letter F denotes a numerical function. Let r be a real. If the elementary forces are in inverse ratio to d^r, where d represents the distance between the points considered, then F is denoted by F_r. For any r, any pair of disjoint objects is F_r-assessable. The F_0-histogram (histogram of constant forces) and F_2-histogram (histogram of gravitational forces) have very different and very interesting characteristics as shown in Fig. 2. The former provides a global view of the situation. It considers the closest parts and the farthest parts of the objects equally, whereas the F_2-histogram focuses on the closest parts. Details can be found in [20]–[22].

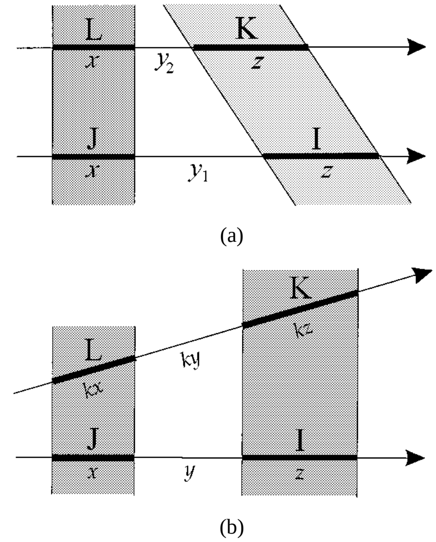


Fig. 2. Main characteristics of the F_0 and F_2-histograms. (a) Independence from distance (F_0-histograms): the force exerted by K on L is equal to the force exerted by I on J. (b) Independence from scale (F_2-histograms): the force exerted by K on L is equal to the force exerted by I on J.

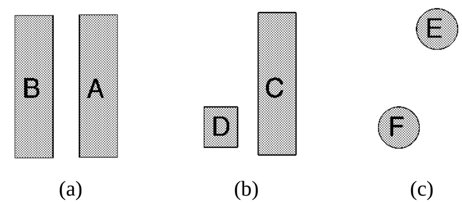


Fig. 3. No existing family of directional relations fits with this perception of the world: (a) A is perfectly to the right of B, (b) C is more to the right of D than above it, and (c) E is more above F than to the right of it.

III. New Families of Fuzzy Directional Relations

A. The Why

The linguistic descriptions generated in Section V make use of spatial prepositions related to the four primitive directional relationships: "to the right of," "above," "to the left of," and "below." Many families of fuzzy directional relations rely on the construction of angle histograms [15]–[18]. Some do not [23]–[25]. The former can be advantageously redefined using histograms of constant forces [20], [26]. This stems from the fact that F_0-histograms coincide with angle histograms [20]–[22], but without their weaknesses (anisotropy, requirement for raster data, etc.) However, none of the above families fits with the set of descriptions presented in Fig. 3. We could argue whether the "natural" perception illustrated by this figure is the most intuitive, but it is one possible perception, and we will use it for our discussion. Many families of directional relations consider that the object A is not perfectly to the right of the object B (one may wonder when such an event occurs). Others consider that C is more above D than to the right of it. No one is able to reconcile the two last descriptions, because no one really takes metric information into account. Moreover, an object is often assessed to be in many directions with respect to another: both to the right and somewhat to the left, and/or both above and a little below. This feature is questionable. It runs counter to the fact that, generally, people do not combine more than two spatial prepositions when translating visual

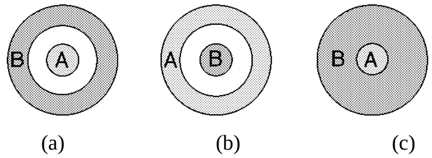


Fig. 4. Directional relations cannot substitute for the spatial relation "surround". (a) A is surrounded by B. (b) A surrounds B. (c) A is included in B. In each case, according to the existing families of directional relations, object A is somewhat above, below, to the right and to the left of object B as well. It does not mean that A surrounds B.

information into natural language descriptions [27], [28] . Some authors [15], [23] support the idea that it allows more complex relationships—like "surround"—to be derived. For instance, knowing that an object A is somewhat above, below, to the right, and to the left of an object B as well, one could conclude that A surrounds B. In our opinion, drawing such a conclusion is not reasonable (see Fig. 4). The directional relations are not the only spatial relations [10], and they cannot represent the relative position of an object with regard to another all by themselves. In particular, they cannot substitute for the spatial relation "surround."

B. The How

The previous analysis leads us to introduce alternative families of fuzzy directional relations. These families will be used in Section V to generate the linguistic descriptions of relative positions. They stem from a new way to exploit the force histograms. The idea is to impose physical considerations on the histograms. The broad outline of the method has been given in [29], and we now present it in detail. The French-speaking reader is also invited to consult [26] (or [20]), in which it is shown that the new families satisfy the four basic axiomatic properties [21], [22]:

1) two objects can be assimilated to points if they are distant enough;
2) the directional relations are not sensitive to scale changes;
3) neither a space dimension nor a direction are preferred;
4) the semantic inverse [10] principle is respected (e.g., object A is to the left of object B as much as B is to the right of A).

Let r be a real and (A, B) an F_r-assessable pair of objects. Our goal is to assess the degree of truth of a proposition like "A is in direction α of B," where α represents any angle. In this section, we will only consider the proposition "A is in direction 0 of B," *i.e.* "A is to the right of B." For another value of α, you can simply perform the computations described below on the shifted histogram $F_r^{AB}(\theta+\alpha)$. The forces exerted on B can be classified in different types. First, the set of directions is divided into four quadrants, as shown in Fig. 5. The forces $F_r^{AB}(\theta)$ of the outer quadrants ($\theta \in [-\pi, -\pi/2]$ or $\theta \in [\pi/2, \pi]$) are elements which, to various degrees, weaken the proposition "A is to the right of B." The forces of the inner quadrants ($\theta \in [-\pi/2, 0]$ or $\theta \in [0, \pi/2]$) are elements which support the proposition. Some forces of the third quadrant are used to compensate—as much as possible—the contradictory forces of the fourth one. The proportion of these compensatory forces is defined by an angle θ_+, to which we will return in Section III-C. Forces of the second

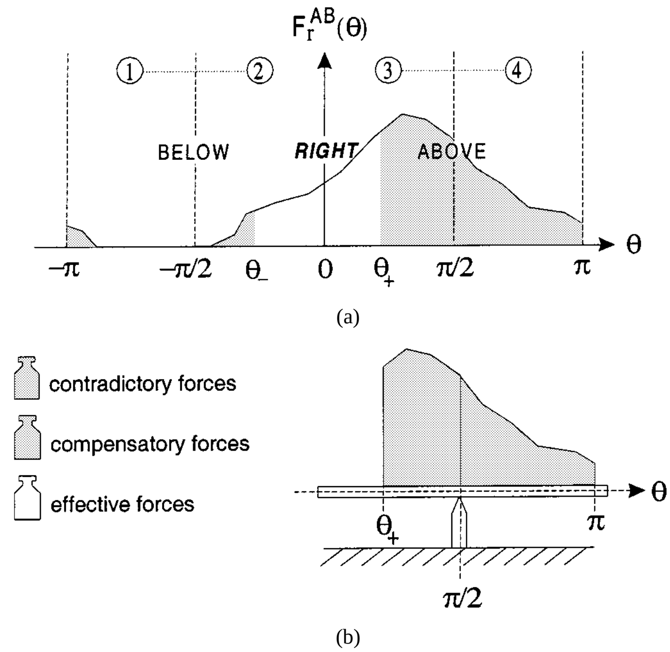


Fig. 5. Contradictory (CTF), compensatory (CPF) and effective (EF) forces. (a) The set of directions is divided into four quadrants. (b) θ_+ is chosen such that the 2-D material system depicted above is balanced. The percentage of the effective forces is denoted $b_r(RIGHT)$: $b_r(RIGHT) = EF/(CTF + CPF + EF)$.

quadrant are used in a similar way to compensate the contradictory forces of the first one. The amount of these compensatory forces is defined by θ_-. The remaining forces are called the effective forces.

Thanks to this first classification, the new families of directional relations will not run counter to the fact that, generally, people do not combine more than two spatial prepositions when translating visual information into natural language descriptions. In Fig. 6, for instance, the object A is not considered to the right of, above and below object B as well, but only to the right of it. When assessing the degree of truth of "A is above B," the contradictory and compensatory forces cancel each other out, and the proposition is found completely false. The same applies to "A is below B."

Now, as shown in Fig. 7, a threshold τ is employed to divide the effective forces into optimal and suboptimal components. The optimal components support the idea that A is "perfectly" to the right of B: whatever their direction, they are regarded as horizontal and pointing to the right. The "average" direction $\alpha_r(RIGHT)$ of the effective forces is then computed, in conformity with this agreement. We will return to τ and $\alpha_r(RIGHT)$ in Section III-C. Due to the distinction between optimal and suboptimal components, the new families of directional relations will be able to fit with the linguistic descriptions given in Fig. 3(b) and (c), and to ensure a coherent and continuous handling of intermediate configurations. Fig. 8 depicts the complete classification of the forces exerted on the reference object.

Finally, the degree of truth $a_r(RIGHT)$ of "A is to the right of B" is set to

$$a_r(RIGHT) = \mu(\alpha_r(RIGHT)) \times b_r(RIGHT).$$

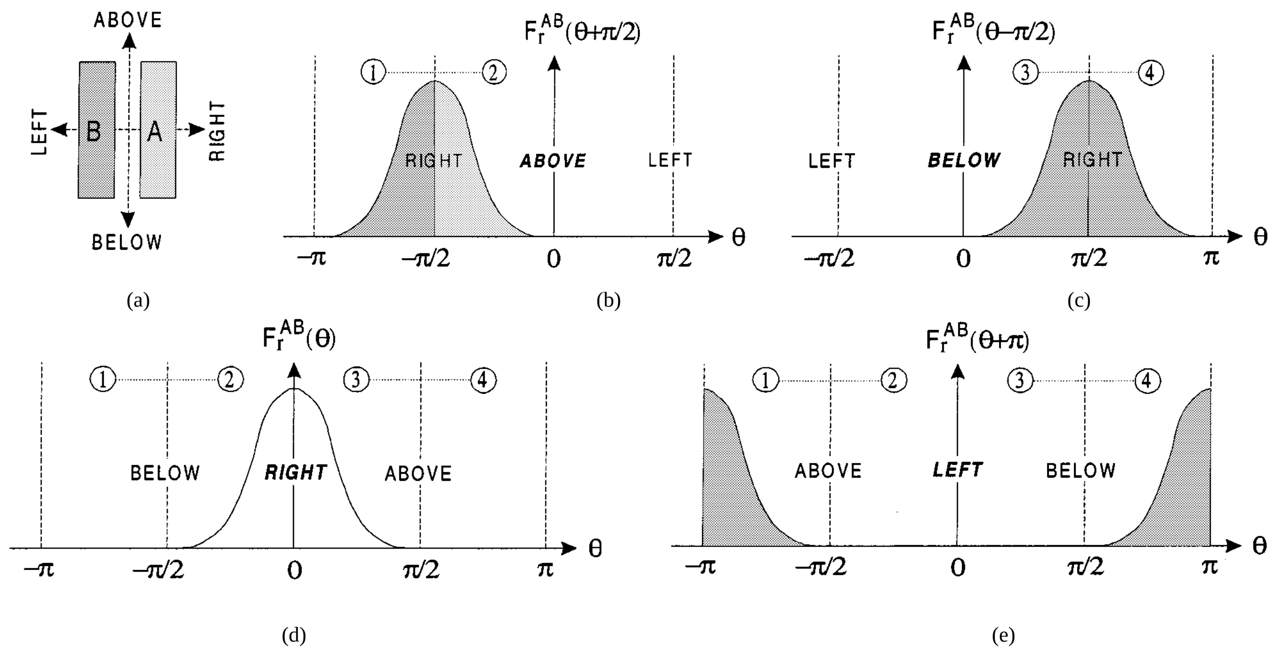


Fig. 6. According to the new families of directional relations, an object cannot be both to the right and to the left of another, or both above and below. (a) What is the relative position of A with regard to B? (b) All forces are either contradictory or compensatory. "A is above B" is completely false. (c) All forces are either contradictory or compensatory. "A is below B" is completely false. (d) All forces are effective and the histogram is symmetric. "A is to the right of B" is completely true. (e) All forces are contradictory. "A is to the left of B" is completely false.

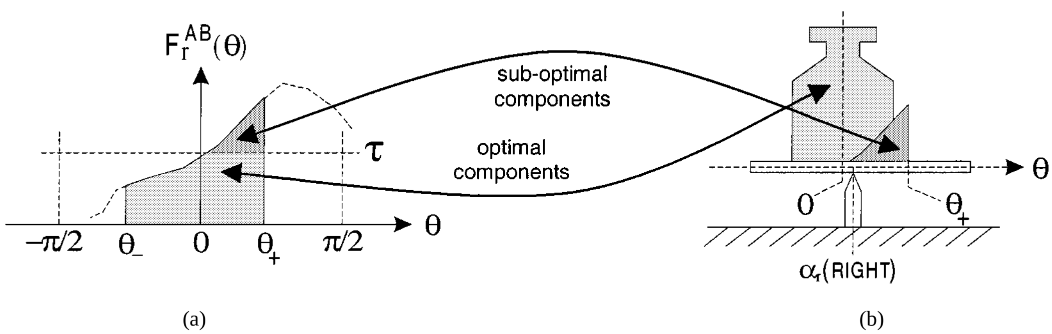


Fig. 7. Optimal and suboptimal components. (a) τ divides the effective forces into optimal and suboptimal components. (b) $\alpha_r(RIGHT)$ is such that the 2-D material system depicted above is balanced. The area of the light gray weight in (b) is equal to the area of the light gray region in (a).

In this expression, $b_r(RIGHT)$ denotes the percentage of the effective forces (Fig. 5). Here, μ is the membership function of a fuzzy set on $[-\pi, \pi]$ that can be employed to define a family of fuzzy directional relations between points [22]. In our experiments, we used the typical triangular function graphed in Fig. 9(a). Let us note that the most optimistic point of view consists in saying that any effective force is optimal. Then, $\alpha_r(RIGHT)$ is always equal to 0, and $\mu(\alpha_r(RIGHT))$ is 1. The value $b_r(RIGHT)$ therefore corresponds to the maximum degree of truth that can reasonably be attached to the proposition "A is to the right of B."

C. Computation of the Different Variables

On an oriented straight line, a material point P is defined by its abscissa x (location), and its mass m (weight). The barycenter of a discrete weighted system like $\{P_i\}_i$, or $\{(x_i, m_i)\}_i$, is the material point $((\sum_i m_i x_i)/(\sum_i m_i), \sum_i m_i)$. The value of the angle θ_+ is the element of $[-\pi/2, \pi/2]$ such that $\left(\pi/2, \int_{\theta_+}^{\pi} F_r^{AB}(\theta)d\theta\right)$ corresponds—as far as possible—to the barycenter of the system $\{(\theta, F_r^{AB}(\theta))\}_{\theta\in[\theta_+,\pi]}$. The physical interpretation of θ_+ is illustrated by Fig. 5(b). The barycenter of $\{(\theta, F_r^{AB}(\theta))\}_{\theta\in[\theta_+,\pi]}$ is

$$\left(\frac{\int_{\theta_+}^{\pi} \theta F_r^{AB}(\theta)d\theta}{\int_{\theta_+}^{\pi} F_r^{AB}(\theta)d\theta}, \quad \int_{\theta_+}^{\pi} F_r^{AB}(\theta)d\theta\right).$$

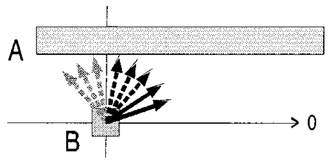
Therefore, θ_+ should be chosen such that

$$\frac{\pi}{2} = \frac{\int_{\theta_+}^{\pi} \theta F_r^{AB}(\theta)d\theta}{\int_{\theta_+}^{\pi} F_r^{AB}(\theta)d\theta}.$$

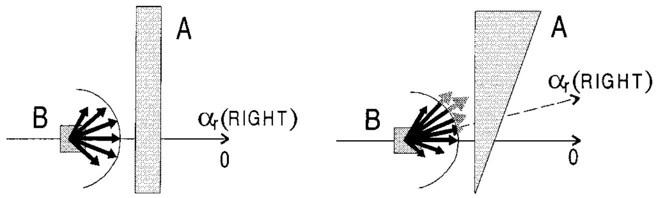
i.e., it should be chosen such that

$$\int_{\theta_+}^{\pi} \left(\theta - \frac{\pi}{2}\right) F_r^{AB}(\theta)d\theta = 0.$$

The integral $\int_{\pi/2}^{\pi}(\theta-\pi/2)F_r^{AB}(\theta)d\theta$ is greater than or equal to 0. However, $\int_{-\pi/2}^{\pi}(\theta-\pi/2)F_r^{AB}(\theta)d\theta$ is not necessarily less

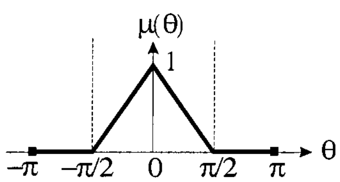


(a)

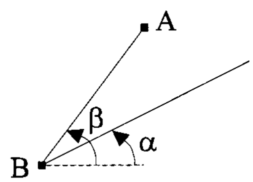


(b)

Fig. 8. Forces exerted on the referent can be classified in four types: contradictory forces, compensatory forces, optimal effective forces and suboptimal effective forces. (a) Gray dotted arrows: contradictory forces. Black dotted arrows: compensatory forces. Black continuous arrows: effective forces. (b) Gray arrows: suboptimal components of the effective forces. Black arrows: optimal components.

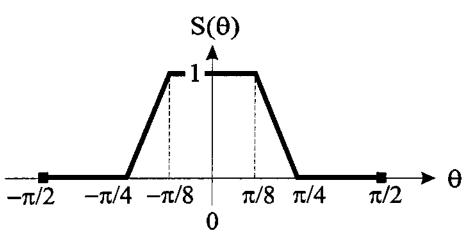


(a)

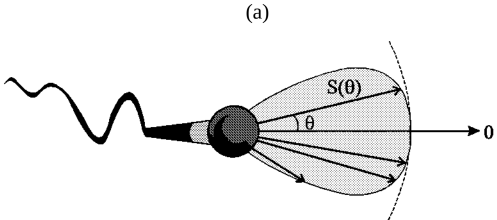


(b)

Fig. 9. Example of directional relations between points. (a) A typical fuzzy set. (b) Is A in direction α of B? The degree of truth of the proposition "A is in direction α of B" is $\mu(\beta-\alpha)$.



(a)



(b)

Fig. 10. Directional sensitivity given to the F-histograms is characterized by a function S. (a) An example of weighting function S. (b) Analogy with the design features of a microphone. Note that the upper horizontal segment in (a) corresponds to an arc in (b).

than or equal to 0, which means that the previous equation may not have any solution. In that case (no solution), all the forces of the third and second quadrants should be used to compensate the contradictory forces of the fourth one (no effective forces). Finally, a rigorous definition of θ_+ can be expressed as follows:

Choose θ_+ so that

$$\text{if } \int_{-\pi/2}^{\pi}\left(\theta-\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta \le 0$$

$$\text{then } \int_{\theta_+}^{\pi}\left(\theta-\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta = 0,$$

$$\text{if } \int_{-\pi/2}^{\pi}\left(\theta-\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta > 0 \text{ then } \theta_+ = -\frac{\pi}{2}.$$

In a similar way:

Choose θ_- so that

$$\text{if } \int_{-\pi}^{\pi/2}\left(\theta+\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta \ge 0$$

$$\text{then } \int_{-\pi}^{\theta_-}\left(\theta+\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta = 0,$$

$$\text{if } \int_{-\pi}^{\pi/2}\left(\theta+\frac{\pi}{2}\right)F_r^{AB}(\theta)d\theta < 0 \text{ then } \theta_- = \frac{\pi}{2}.$$

Hence, as in Fig. 5, the value of $b_r(RIGHT)$ is given by the following.

$$\text{If } \theta_+ > \theta_- \text{ then}$$
$$b_r(RIGHT) = \frac{\int_{\theta_-}^{\theta_+} F_r^{AB}(\theta)d\theta}{\int_{-\pi}^{\pi} F_r^{AB}(\theta)d\theta},$$
$$\text{if } \theta_+ \le \theta_- \text{ then } b_r(RIGHT) = 0.$$

If there are no effective forces, i.e., if $b_r(RIGHT)$ is equal to zero, then $a_r(RIGHT)$ is also equal to zero, and the value of $\alpha_r(RIGHT)$ is of no importance. Otherwise, $\alpha_r(RIGHT)$ is chosen such that $\left(\alpha_r(RIGHT), \int_{\theta_-}^{\theta_+} F_r^{AB}(\theta)d\theta\right)$ corresponds to the barycenter of the following weighted system, which depends on the threshold τ:

$$\left\{\left(0, \int_{\theta_-}^{\theta_+} \min\left\{\tau, F_r^{AB}(\theta)\right\} d\theta\right)\right\}$$
$$\cup \left\{\left(\theta, \max\left\{0, F_r^{AB}(\theta)-\tau\right\}\right)\right\}_{\theta\in[\theta_-,\theta_+]}.$$

The only element of the first set (left operand of the union) expresses that the optimal components of the histogram are assimilated to a unique force applied at point zero. The suboptimal components appear in the second set. The physical interpretation of $\alpha_r(RIGHT)$ is illustrated by Fig. 7(b). Hence, $\alpha_r(RIGHT)$ is equal to

$$\frac{\int_{\theta_-}^{\theta_+} \theta\left[\max\left\{0, F_r^{AB}(\theta)-\tau\right\}\right] d\theta}{\int_{\theta_-}^{\theta_+} F_r^{AB}(\theta)d\theta}.$$

The threshold τ can be any nonnegative real number. Let us now explain how to choose it. Let $\tau_{\min}$ and $\tau_{\max}$ be respectively the minimum and the maximum of the effective forces (on $[-\pi/2, \pi/2]$). The value of $\tau_{\min}$ is 0 if $\theta_- \ne -\pi/2$ or $\theta_+ \ne \pi/2$, or is $\min_{\theta\in[\theta_-,\theta_+]} F_r^{AB}(\theta)$ otherwise. The value of $\tau_{\max}$ is $\max_{\theta\in[\theta_-,\theta_+]} F_r^{AB}(\theta)$. If τ is chosen lower than or equal to $\tau_{\min}$, there are no optimal components. Then, the absolute value of $\alpha_r(RIGHT)$ is maximum, and $a_r(RIGHT)$ is minimum (i.e., no choice of τ can result in a lower value). If τ is greater than or equal to $\tau_{\max}$, there are no suboptimal components, $\alpha_r(RIGHT)$ is equal to zero, and $a_r(RIGHT)$ is maximum (i.e., equal to $b_r(RIGHT)$). Setting τ to the average—or a weighted average—of the effective forces constitutes a natural

TABLE I

THE EFFECT OF THE THRESHOLD τ ON THE ASSESSMENT OF THE DIRECTIONAL RELATIONSHIPS. (a) A WITH REGARD TO B. (b) C WITH REGARD TO D. (c) E WITH REGARD TO F. THE CONSIDERED PAIRS OF OBJECTS, (A, B), (C, D), AND (E, F), ARE THOSE REPRESENTED BY FIG. 3. THE DIRECTIONAL RELATIONS "TO THE RIGHT OF" AND "ABOVE" ARE ASSESSED USING TWO HISTOGRAMS (THE F_0-HISTOGRAM AND THE F_2-HISTOGRAM), AND FOUR THRESHOLDS: THE *MINIMUM* OF THE EFFECTIVE FORCES ($\tau_{\min}$), THE *MAXIMUM* ($\tau_{\max}$), THE *AVERAGE* ($\tau_{\overline{1}}$), AND THE *WEIGHTED* AVERAGE τ_s WHERE S DENOTES THE FUNCTION GRAPHED IN FIG. 10(a). NOTE THAT IN EACH CASE $a_0(LEFT)$, $a_2(LEFT)$, $a_0(BELOW)$, AND $a_2(BELOW)$ ARE EQUAL TO ZERO

τ	min	max	ave	wei
a_0(RIGHT)	1.00	1.00	1.00	1.00
a_0(ABOVE)	0.00	0.00	0.00	0.00
a_2(RIGHT)	1.00	1.00	1.00	1.00
a_2(ABOVE)	0.00	0.00	0.00	0.00

(a)

τ	min	max	ave	wei
a_0(RIGHT)	0.69	1.00	0.83	0.88
a_0(ABOVE)	0.31	0.58	0.39	0.38
a_2(RIGHT)	0.82	1.00	0.92	1.00
a_2(ABOVE)	0.18	0.37	0.23	0.21

(b)

τ	min	max	ave	wei
a_0(RIGHT)	0.23	1.00	0.39	0.23
a_0(ABOVE)	0.77	1.00	0.82	0.87
a_2(RIGHT)	0.23	1.00	0.39	0.23
a_2(ABOVE)	0.77	1.00	0.82	0.87

(c)

compromise. This is why we set τ to the ratio τ_s, defined as follows:

$$\tau_s = \frac{\int_{\theta_-}^{\theta_+} S(\theta) F_r^{AB}(\theta) d\theta}{\int_{-\pi/2}^{\pi/2} S(\theta) d\theta}.$$

The weighting function S, from $[-\pi/2, \pi/2]$ into $[0, 1]$, is even, continuous, decreasing on $[0, \pi/2]$, and takes the value 1 at zero. S characterizes the "directional sensitivity" that is given to the F_r-histograms. Fig. 10(a) shows the trapezoid function used in our experiments, while Fig. 10(b) shows how an analogy with the design features of a microphone, which picks up more or less the lateral sounds, can be established. The chosen function corresponds to a medium directional sensitivity, whereas $\overline{1}$—the function that associates 1 to any value—corresponds to a large sensitivity (a sound engineer would say "omnidirectional" instead of "large"). Note that $\tau_{\overline{1}}$ is the average of the effective forces. Table I illustrates how the choice of τ impacts the assessment of the directional relationships. The table refers to the objects represented in Fig. 3. Four thresholds are considered. The weighted average τ_s allows all the linguistic descriptions given in Fig. 3 to be accommodated. Moreover, the values obtained express the "clearest" opinions (the absolute differences $|a_0(RIGHT) - a_0(ABOVE)|$ and $|a_2(RIGHT) - a_2(ABOVE)|$ are systematically higher when using this threshold). Finally, applying τ_s to the F_2-histogram allows C to be assessed perfectly to the right of D. We will exploit these features in the next sections.

D. Examples

In this section, five families of directional spatial relations are considered: **K, M, W, F0,** and **F2**. The first, **K, M,** and **W,** are based on the construction of angle histograms: **K** is defined by the aggregation method [17], **M** by the compatibility method [15], and **W** by the neural approach [18]. The second, **F0** and **F2**, are based on the construction of F_0 and F_2-histograms, and the distinction between contradictory, compensatory and effective forces, as described above. Our aim here is to show through a few examples that, contrary to the existing families, **F0** and **F2** do have the properties we were looking for in Section III-A. Note that **K** and **M** are probably the most typical families that:

1) involve fuzzy relations and not "all-or-nothing" ones;
2) do not assimilate objects to very elementary entities such as a point (centroid) or a (bounding) rectangle;

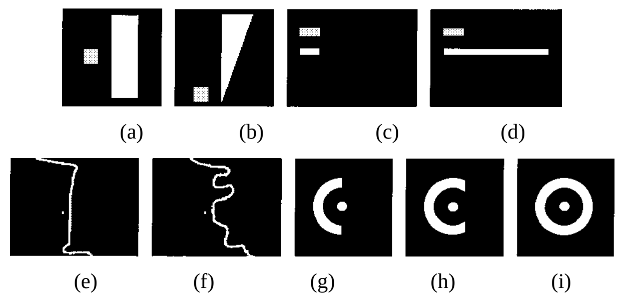


Fig. 11. Comparison between five families of directional relations. Images. Argument A is in white and referent B in gray.

3) fairly meet the basic axiomatic properties mentioned in Section III-B.

Therefore, like **F0** and **F2**, they can be hoped to constitute a satisfactory modeling of the directional relationships. **W** is used by the system for linguistic scene description introduced in [19]. We will come back to that system in Section V-B.

Fig. 11 presents nine pairs of objects. For each pair (A, B), the propositions "A is to the RIGHT of B", "A is ABOVE B", "A is to the LEFT of B" and "A is BELOW B" have been assessed. The degrees of truth produced by **K, M, W, F0,** and **F2** are displayed in Table II. First, we give some specific comments concerning the different configurations. If A is not perfectly below B in the case depicted by Fig. 11(c), when does this event occur? The fact is that according to **K, M,** and **W** a proposition such as "A is below B" is never totally true. **K** and **M** see the "house" of Fig. 11(e) (object A) rather south of the "river" (object B), or maybe north, but certainly not west. **W** sees the house rather west, but the values produced are very low. In Fig. 11(b), **F2** and **W** are alone in thinking that A is more to the right of B, even though they give a certain credit to the proposition "A is above B". In Fig. 11(d), as A becomes longer, **K** and **M** quickly affirm that A is essentially located to the right of B. **F0** eventually shares this point of view, but later on, and in a less definite way. **W** is uncertain, and gives the lowest degrees of truth. **F2** is the only family to maintain that A essentially remains below B. Now, according to **M**, object A of Fig. 11(g) is not much more to the left of object B than below or above it, and A is not much more to the left of B in image (g) than in (i). Finally, in Fig. 11(i), we ask if the ring is located to the left of the disc. **F0** and **F2** definitely say no. The families of directional relations cannot substitute for the spatial relation "surround." **W** behaves

TABLE II
COMPARISON BETWEEN FIVE FAMILIES OF DIRECTIONAL RELATIONS: RESULTS. THE DEGREES OF TRUTH ARE GIVEN IN HUNDREDTHS

	K	M	W	F0	F2	K	M	W	F0	F2	K	M	W	F0	F2	K	M	W	F0	F2
RIGHT	**71**	**78**	**83**	**100**	**100**	38	32	**66**	55	**86**	10	12	3	0	0	**60**	**71**	**45**	54	21
ABOVE	15	23	9	0	0	**67**	**68**	43	**73**	43	0	0	0	0	0	0	0	5	0	0
LEFT	0	0	0	0	0	0	0	1	0	0	10	12	8	0	0	2	12	1	0	0
BELOW	15	23	8	0	0	1	5	2	0	0	**80**	**88**	**86**	**100**	**100**	38	29	43	**76**	**99**
	(a)					(b)					(c)					(d)				

	K	M	W	F0	F2	K	M	W	F0	F2	K	M	W	F0	F2	K	M	W	F0	F2	K	M	W	F0	F2
RIGHT	3	7	1	0	0	1	3	1	0	0	1	1	0	0	0	4	10	1	0	0	**25**	**50**	11	0	0
ABOVE	33	80	3	3	0	17	45	1	0	0	25	48	9	0	0	29	**50**	10	0	0	**25**	**50**	5	0	0
LEFT	22	18	**21**	**75**	**99**	40	35	**31**	**87**	**99**	**51**	**52**	**68**	**96**	**95**	**38**	**50**	**55**	**48**	**44**	**25**	**50**	**25**	0	0
BELOW	**43**	**81**	5	7	2	**43**	**68**	13	20	5	25	48	13	0	0	29	**50**	11	0	0	**25**	**50**	6	0	0
	(e)					(f)					(g)					(h)					(i)				

strangely, and gives values that do not reflect the symmetry of the configuration.

Now, let us discuss the most distinguishing feature between **K, M,** and **W** on the one hand, and **F0** and **F2** on the other. Consider for instance the objects depicted in Fig. 11(h). Through a point of B (the disc), draw a vertical line. The right half-plane so defined contains some points of A (in white). For **K, M,** and **W**, it is enough to conclude that the proposition "A is to the right of B" cannot be totally false. The **F0** and **F2** families are much more exacting. According to them, an object cannot be simultaneously a bit to the left, and a bit to the right of another. Which family provides the "best" results? The answer obviously depends on context and the application considered. We just dealt here with what Gapp [30] called the *basic meanings* of spatial relations (the model proposed by Gapp to define the semantics of spatial relations distinguishes *context-specific conceptual knowledge* from the *basic meanings* of the relations). However, the new families **F0** and **F2** express opinions which are fully logical. As expected, and contrary to **K, M,** and **W**, they fit with the perception illustrated by Fig. 3, and do not assess an object to be in more than two primitive directions with respect to another.

IV. GENERATION OF LINGUISTIC DESCRIPTIONS

Now, we want to give a linguistic description of the relative position between any 2-D objects A and B. The description proposed in the present paper relies on the sole primitive directional relationships: "to the right of," "above," "to the left of," and "below." It is generated from F_0^{AB} and F_2^{AB}. Other histograms could have been considered. However, as observed in Sections II and III-A: 1) F_0-histograms coincide with angle histograms, which have been extensively used in the literature; 2) gravitational forces are a reality of our physical world; and 3) the F_0 and F_2-histograms have very different and interesting characteristics, which complement one another and allow for geometric interpretation. First, eight values are extracted from the analysis of F_0^{AB} and F_2^{AB}. These values, computed as in Section III, are: $a_r(RIGHT)$, $b_r(RIGHT)$, $a_r(ABOVE)$, $b_r(ABOVE)$, $a_r(LEFT)$, $b_r(LEFT)$, $a_r(BELOW)$, and $b_r(BELOW)$. They represent the "opinion" given by the considered family (family **F0** if r is 0, family **F2** if r is 2). Then, the two opinions—16 values—are combined. Four numeric and two symbolic features result from this combination. They feed a system of fuzzy rules that finally outputs the expected description.

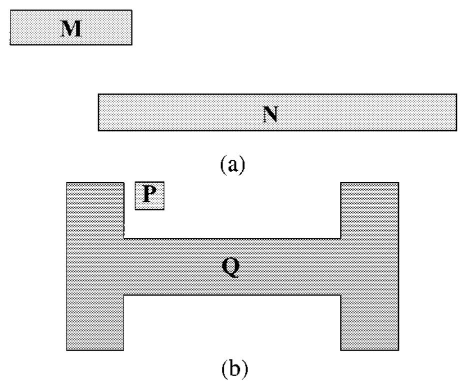


Fig. 12. Combination of **F0** and **F2**'s opinions. Objects. (a) The argument is M, and the referent is N. (b) The argument is P, and the referent is Q.

TABLE III
COMBINATION OF **F0** AND **F2**'S OPINIONS: RESULTS

δ	RIGHT	ABOVE	LEFT	BELOW	
$a_0(\delta)$	0	0.53	0.81	0	**δ1**=ABOVE
$b_0(\delta)$	0	1.00	0.93	0	**d1**=0.83
					m1=1.00
$a_2(\delta)$	0	0.83	0.49	0	
$b_2(\delta)$	0	1.00	0.89	0	**δ2**=LEFT
a(δ)	0 (case 3)	0.83 (case 3)	0.81 (case 3)	0 (case 3)	**d2**=0.81
					m2=0.89

(a)

δ	RIGHT	ABOVE	LEFT	BELOW	
$a_0(\delta)$	0	0.80	0.03	0	**δ1**=ABOVE
$b_0(\delta)$	0	0.85	0.04	0	**d1**=0.80
					m1=0.61
$a_2(\delta)$	0.50	0.55	0	0	
$b_2(\delta)$	0.51	0.61	0	0	**δ2**=LEFT
a(δ)	0 (case 1)	0.80 (case 2)	0.03 (case 2)	0 (case 3)	**d2**=0.03
					m2=0

(b)

		m_1			
		high	*medium-high*	*medium-low*	*low*
d_1	*high*	perfectly	$--_1$	$nearly_1$	*no primary direction*
	medium-high	$--_2$	$nearly_2$	$loosely_1$	
	medium-low	mostly	$loosely_2$	$loosely_3$	
	low	*no primary direction*			

(a)

		m_2		
		high	*medium*	*low*
d_2	*high*	somewhat	strongly	*no secondary direction*
	medium	a little	slightly	
	low	*no secondary direction*		

(b)

		$\min\{m_1,m_2\}$			
		high	*medium-high*	*medium-low*	*low*
d_2/d_1	*very high*	$--_3$			*no compound direction*
	NOT *very high*	*no comp. dir.*			
	high		$nearly_3$		
	NOT *high*		*no comp. dir.*		
	rather high			$loosely_4$	
	NOT *rather high*			*no comp. dir.*	

(c)

$\max\{m_1,m_2\}$		
high	*medium*	*low*
satisfactory	rather satisfactory	unsatisfactory

(d)

Fig. 13. Rule base. (a) Primary direction (9 rules + 2 meta-rules). (b) Secondary direction (4 rules + 2 meta-rules). (c) Compound direction (3 rules + 4 meta-rules). (d) Self-assessment (three rules).

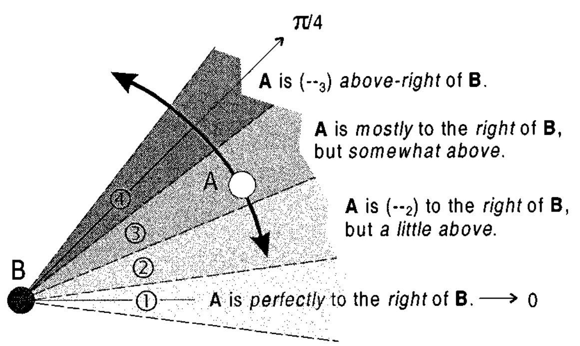
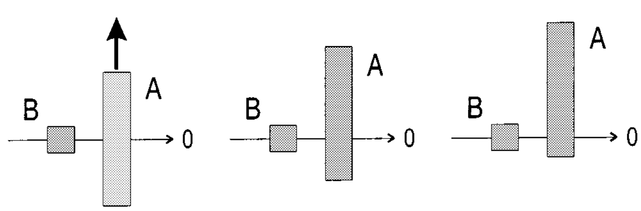


Fig. 14. Training configurations and terminology (I). Primary, secondary, and compound directions. In each case, the self-assessment of the description is: "The description is *satisfactory*."



(a) (b) (c)

Fig. 15. Training configurations and terminology (II). Secondary direction. Shifting. (a) A is *perfectly* to the *right* of B. (b) A is *perfectly* to the *right* of B, but *slightly* shifted *upward*. (c) A is *perfectly* to the *right* of B, but *strongly* shifted *upward*. In each case, the self-assessment of the description is: "The description is *satisfactory*."

A. Input Variables

Let Δ be the set of the four 2-D primitive directions: $\Delta = \{RIGHT, ABOVE, LEFT, BELOW\}$. Consider an element δ of Δ. A degree of truth $a(\delta)$ has to be attached to the proposition "A is in direction δ of B." We work on the principle that $a_0(\delta)$, the value proposed by **F0**, is never too optimistic, but is often too cautious. We attribute the previous drawback to the fact that **F0** only has a global view of the situation, and we correct it considering **F2**, which focuses on the closest parts of the objects. However, just because of this characteristic, **F2**'s opinion may be excessive: sometimes excessively pessimistic, and sometimes excessively optimistic. We will use the examples presented in Fig. 12 and Table III to illustrate **F0** and **F2**'s behavior. There are actually three cases.

1) $a_2(\delta) > b_0(\delta)$ (which is equivalent to: $a_2(\delta) > b_0(\delta) \geq a_0(\delta)$)

According to **F0**, the value $b_0(\delta)$ is the maximum degree of truth that can reasonably be attached to the proposition "A is in direction δ of B". Therefore, **F2** conflicts with **F0**. The value $a_0(\delta)$ may be too cautious, but $a_2(\delta)$ seems excessively optimistic. We choose a compromise solution and set: $a(\delta) = b_0(\delta)$. For instance, as shown in Fig. 12(b) and Table III-B, **F2** assesses P to be somewhat to the right of Q: $a_2(RIGHT) = 0.50$. The reason is that P is actually to the right of the closest part of Q. **F0**, which examines the configuration from a global point of view, finds **F2**'s opinion inordinate. It considers that, reasonably, "P is to the right of Q" cannot be but completely false: $b_0(RIGHT) = 0$. We support **F0**'s analysis and set $a(RIGHT)$ to zero.

2) $a_0(\delta) > b_2(\delta)$ (which implies that: $a_2(\delta) < b_0(\delta)$).

According to **F2**, the value $b_2(\delta)$ is the maximum degree of truth that can reasonably be attached to the proposition "A is in direction δ of B." Therefore, **F0** conflicts with **F2**. We ignore the excessive pessimism of **F2** and set: $a(\delta) = a_0(\delta)$. Let us take again the example with P and Q. **F0** estimates at 0.80 the degree of truth of "P is above Q," whereas **F2** considers that it should not exceed 0.61. This is a severe opinion, due to the fact that the top-left part of Q, very close to P, eclipses the rest of the reference object in **F2**'s analysis.

3) $a_2(\delta) \leq b_0(\delta)$ and $a_0(\delta) \leq b_2(\delta)$.

There is no conflict. We set: $a(\delta) = \max\{a_0(\delta), a_2(\delta)\}$. For instance, as shown in Fig. 12(a) and Table III-A, **F2** gives great credit to the proposition "M is above N." $a_2(ABOVE)$

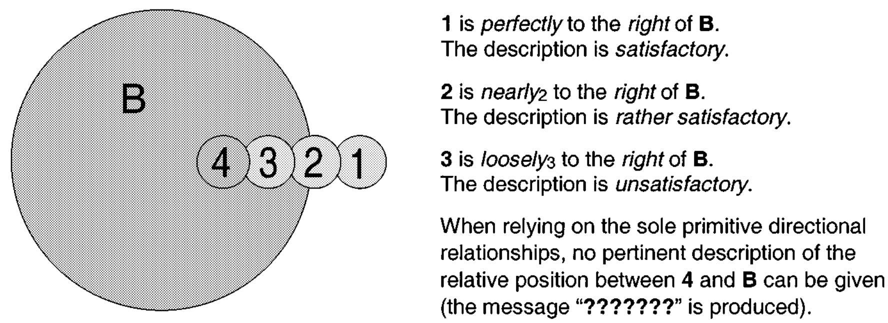


Fig. 16. Training configurations and terminology (III). Self-assessment.

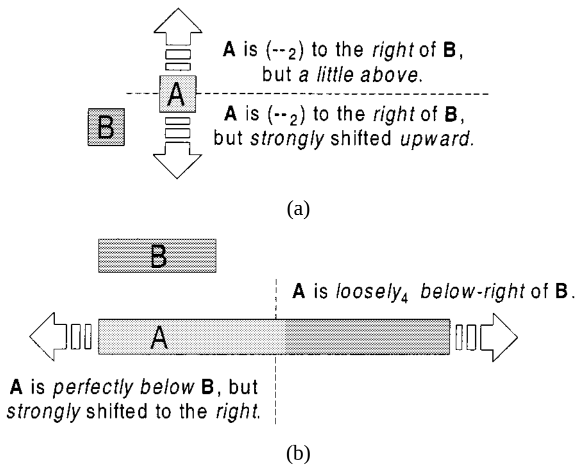


Fig. 17. Connection between: (a) "but *above*" and "but shifted *upward*," and (b) "*below* but shifted to the *right*" and "*below-right*." In each case, the self-assessment of the description is: "The description is *satisfactory*."

equals 0.83. **F0** is more cautious ($a_0(ABOVE)$ equals 0.53), but admits that **F2**'s opinion is defensible ($b_0(ABOVE)$ equals 1). Part of M is actually right above some part of N. On the other hand, though no part of M is perfectly to the left of N, many points of M are mostly to the left of many points of N, and **F0**'s opinion ($a_0(LEFT)$ equals 0.81) is also defensible ($b_2(LEFT)$ equals 0.89). We finally set $a(ABOVE)$ to $a_2(ABOVE)$, and $a(LEFT)$ to $a_0(LEFT)$.

It is easy to see that in the three cases:

$$a(\delta) = \max\{a_0(\delta), \min\{a_2(\delta), b_0(\delta)\}\}.$$

Moreover, in the first and second cases (conflict):

$$a(\delta) \geq \min\{b_0(\delta), b_2(\delta)\}.$$

And in the third one (no conflict):

$$a(\delta) \leq \min\{b_0(\delta), b_2(\delta)\}.$$

The value $\min\{b_0(\delta), b_2(\delta)\}$ measures to what extent both sources of information agree on the fact that A can be considered in direction δ of B. Finally, six parameters are extracted from the analysis of the histograms F_0^{AB} and F_2^{AB}, and used in order to generate the linguistic description of the relative position between A and B. These values δ_1, d_1, m_1, δ_2, d_2, and m_2 are defined as follows:

$$\delta_1 = \arg\max_{\delta\in\Delta} a(\delta), \quad \delta_2 = \arg\max_{\delta\in\Delta-\{\delta_1\}} a(\delta),$$
$$d_1 = a(\delta_1), \quad d_2 = a(\delta_2),$$
$$m_1 = \min\{b_0(\delta_1), b_2(\delta_1)\}, \quad m_2 = \min\{b_0(\delta_2), b_2(\delta_2)\}.$$

Here, δ_1 is the primary direction, and δ_2 the secondary direction. The degree of truth $a(\delta)$ attached to the proposition "A is in direction δ of B" is maximum when δ is δ_1. See the two examples presented in Fig. 12 and Table III.

B. The Rule Base

A system of 27 fuzzy rules and meta-rules, displayed in Fig. 13, handles a set of 16 adverbs, and allows precise linguistic descriptions to be produced (such as "A is *perfectly* to the right of B," "A is *mostly* to the right of B, but *somewhat* above"). The symbol -- that appears in Fig. 13 represents the "void" adverb. For instance, "A is -- to the right of B" should be read "A is to the right of B," and can be considered equivalent to "A is *almost perfectly* to the right of B." Finding appropriate terms to distinguish—in a natural way—between closely related configurations is sometimes difficult. Moreover, a given word may be suited for linguistic descriptions of completely different configurations. This is why, in the rule base, some adverbs appear more than once. The subscripts will be used only in this section, for tracking purposes. Note that the adverbs are stored in a dictionary of terms, and can be tailored to individual users. Two dictionaries are currently available: one in English, and one in French. The linguistic values, such as *high*, *medium*, *low*, will be discussed subsequently.

The description of the relative position between two objects A and B will generally be composed of three parts. The first part is the main part of the description (e.g., "A is to the *right* of B"). It involves the primary direction δ_1. The second part supplements the description (e.g., "but a little *above*"). It involves the secondary direction δ_2. The third part indicates to what extent the four primitive directional relationships are suited to describing the relative position of the objects (e.g., "The description is *satisfactory*"). In other words, it indicates to what extent it is necessary to turn or not to other spatial relations (e.g., "surrounds").

The first part of the description depends on the input variables δ_1, d_1 and m_1. It is generated by the set of rules shown in Fig. 13(a). Let us assume for instance that the primary direction δ_1 is RIGHT, and that the two objects A and B can be assimilated to points. In this case, the configuration is not ambiguous, there is no possibility of conflict between the two sources of information F_0^{AB} and F_2^{AB}, and m_1 is *high* (actually equal to 1). However, the object A can be "more or less" to the right of B. The primary direction rule base offers three adverbs as candi-

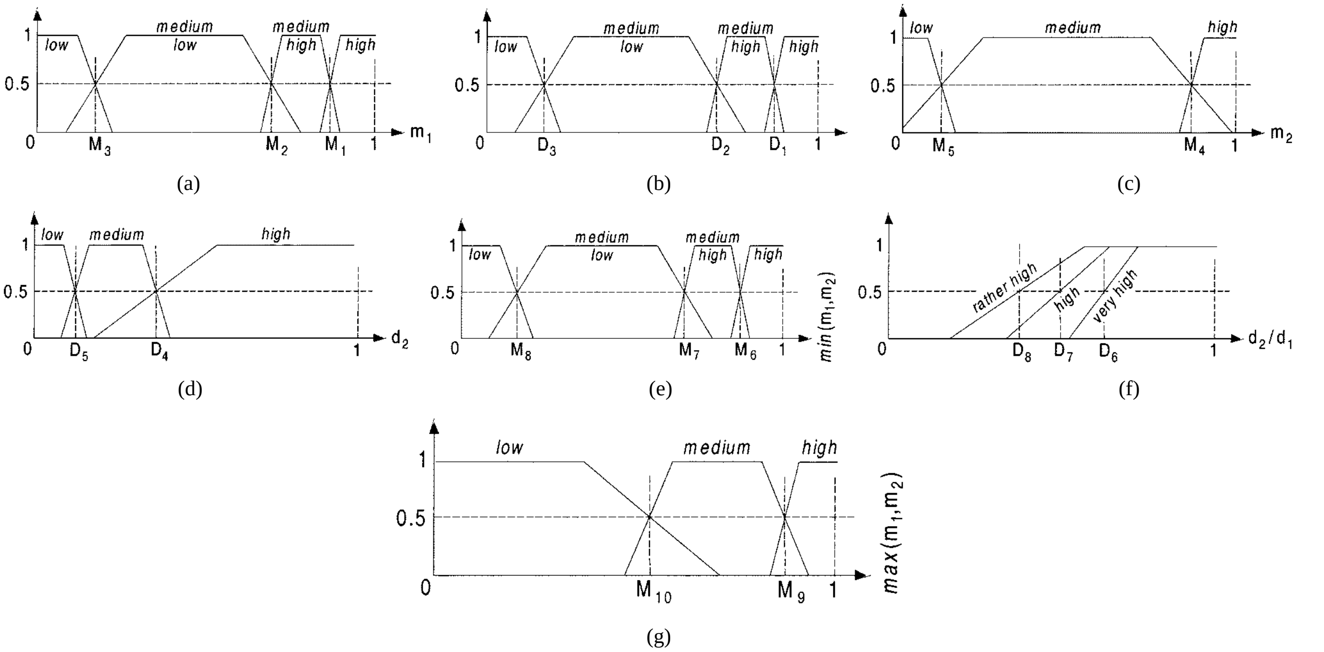


Fig. 18. Linguistic values used in the rule base. (a) Linguistic values for m_1. In our experiments: $M_1 = 0.96$, $M_2 = 0.80$, $M_3 = 0.20$. (b) Linguistic values for d_1. In our experiments: $D_1 = 0.92$, $D_2 = 0.75$, and $D_3 = 0.20$. (c) Linguistic values for m_2. In our experiments: $M_4 = 0.93$, $M_5 = 0$. (d) Linguistic values for d_2. In our experiments: $D_4 = 0.25$ and $D_5 = 0.08$. (e) Linguistic values for $\min\{m_1, m_2\}$. In our experiments: $M_6 = 0.95$, $M_7 = 0.80$, and $M_8 = 0.20$. (f) Linguistic values for d_2/d_1. In our experiments: $D_6 = 0.71$, $D_7 = 0.56$, and $D_8 = 0.42$. (g) Linguistic values for $\max\{m_1, m_2\}$. In our experiments: $M_9 = 0.90$ and $M_{10} = 0.50$.

dates for a hedge in the main part of the linguistic description. The adverbs *perfectly*, --$_2$ (void) and *mostly* have been chosen here. They are presented in sectors 1, 2, and 3 in Fig. 14. The selection among the three words is made according to d_1, the degree of truth of the proposition "A is to the right of B."

It is clear that the 2-D objects A and B cannot always be assimilated to points. The configuration may be ambiguous, and m_1 may not be *high*. Depending on the amount of ambiguity, *perfectly* degenerates into --$_1$ or $nearly_1$, --$_2$ into $nearly_2$ or $loosely_1$, etc. Note that if m_1 or d_1 are *low* (very serious conflict, very ambiguous configuration), the primary direction is meaningless. Then, no pertinent linguistic description relying on the sole primitive directional relationships can be given, and the system produces the message "???????" This usually happens when A and B intersect, or one surrounds the other.

Unless it is equal to the series of question marks, the main description is likely to be supplemented using the set of rules shown in Fig. 13(b). Suppose, for instance, that the primary and secondary directions, δ_1 and δ_2, are, respectively, RIGHT and ABOVE. The supplementary description depends on δ_2, d_2, and m_2. For *high* values of m_2, it can be "but (the object A is) *a little* above (the object B)" or "but *somewhat* above." The two possibilities are shown in sectors 2 and 3 in Fig. 14. The choice between the competing adverbs is made according to d_2, the degree of truth of "A is above B." For *medium* values of m_2, the possibilities turn into "but (the object A is) *slightly* shifted upward (relative to B)" and "but *strongly* shifted upward." The signification of these expressions—which could not be obtained without the contribution of the histogram of gravitational forces—is illustrated by Fig. 15. The connection between the two kinds of expressions ("but… above," "but… shifted upward") is illustrated by Fig. 17(a). Note that if m_2 or d_2 are *low*, the secondary direction is meaningless, and the main description is not supplemented [see sector 1 in Fig. 14 and Fig. 15(a)].

Unless the message "???????" is to be generated, the two first parts of the description may be combined, using one of the four compound directions: ABOVE-RIGHT, ABOVE-LEFT, BELOW-LEFT and BELOW-RIGHT. This happens according to the rules shown in Fig. 13(c). Look at sectors 2, 3, and 4 in Fig. 14. The two sources of information agree that A can be considered both to the right of and above B (i.e., $\min\{m_1, m_2\}$ is *high*). However, only in sector 4 the degree of truth of "A is above B" can actually be compared to the degree of truth of "A is to the right of B" (i.e., only in sector 4 the ratio d_2/d_1 is *very high*). Now, look at Fig. 19(a). Whatever the argument, d_2/d_1 is *very high*. However, most configurations are ambiguous, and only argument 5 makes $\min\{m_1, m_2\}$ *high*. The connections between 1-piece descriptions (such as "A is above-right of B") and two-piece descriptions are illustrated in sectors 3 and 4 in Fig . 14 and Fig. 17(b).

Finally, if a pertinent linguistic description relying on the sole primitive directional relations can be given (i.e., except for "???????"), then the description assesses itself using the last set of rules shown in Fig. 13(d). The description is *satisfactory* when there is at least one direction that wins both sources of information over (i.e., when $\max\{m_1, m_2\}$ is *high*). This is illustrated by Fig. 16.

C. Linguistic Values

Each linguistic value used in the rule base (like *high*, *medium*, *low*) corresponds to a fuzzy set whose membership function λ is represented by a trapezoid. The trapezoid is symmetrical (unless truncated), and chosen so that the core is half the support.

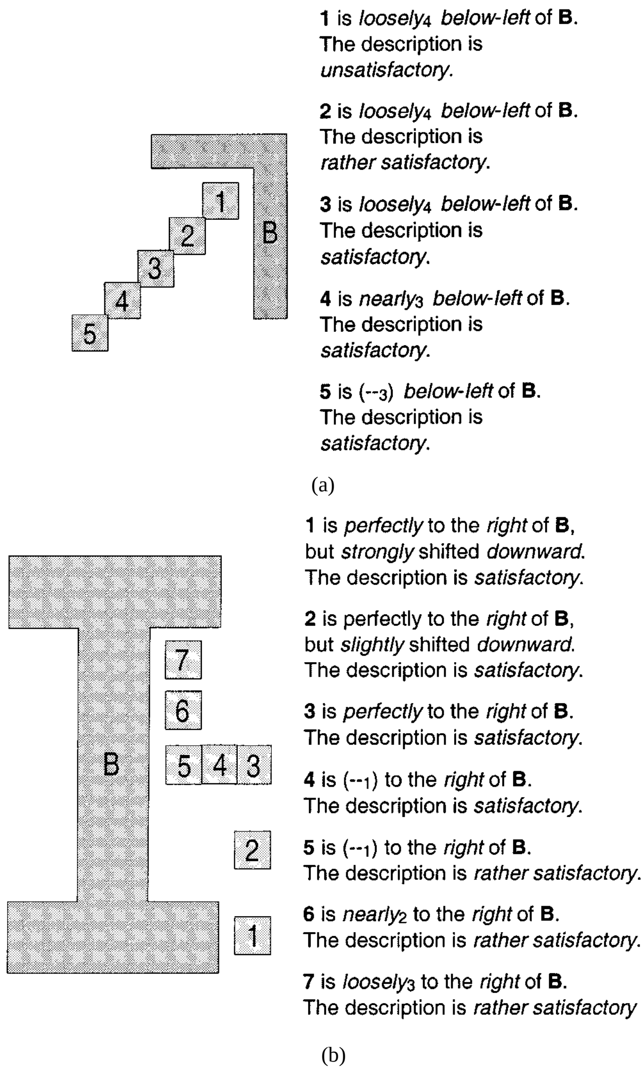


Fig. 19. Two series of tests on synthetic data.

Therefore, λ is completely determined by one or two parameters: the values of x such that $\lambda(x) = 0.5$. All the membership functions are depicted in Fig. 18. The different parameters, D_1 to D_8 and M_1 to M_{10}, have been set considering typical configurations such as those in Figs. 14–16. Most of the values D_1 to D_8 result from precise computations, the others have been determined empirically. We refer to this tuning process as the training of the system. D_1 for instance [Fig. 18(b)] has been deduced from the sequence of configurations represented by Fig. 14 and involving point-like objects. In each case, m_1 is *high* (equal to 1). The parameter D_1, related to d_1's linguistic values *high* and *medium-high*, therefore determines when the void adverb --$_2$ is preferred to *perfectly* [Fig. 13(a)]. It has been naturally set to $\mu(\pi/24)$, where μ is the membership function graphed in Fig. 9, and $\pi/24$ the angle common to sectors 1 and 2 in Fig. 14. In other words, D_1 is the degree of truth of the proposition "point A is to the right of point B," when A is in fact exactly in direction $\pi/24$ of B. Most of the values D_1 to D_8 have been deduced from such geometric observations, considering nonambiguous configurations. D_4 for example, which is related to d_2's linguistic values *high* and *medium* [Fig. 18(d)], determines when *somewhat* is preferred to *a little* [Fig. 13(b)] and has been set to $\mu(3\pi/24 - \pi/2)$. It is the degree of truth of "point A is above point B," when A is actually in direction $3\pi/24$ of B (the direction common to sectors 2 and 3). In a similar way, D_6 [Fig. 13(c) and Fig. 18(f)] has been set to $\mu(5\pi/24 - \pi/2)/\mu(5\pi/24)$. The numerator is the degree of truth of the proposition "point A is above point B," when A is actually in direction $5\pi/24$ of B, and the denominator is the degree of truth of "A is to the right of B."

In the description of the relative position between point-like objects, each adverb is selected among a few other candidates, and the selection rules are determined by D_1, D_2, etc., through the linguistic values. When the objects cannot be assimilated to points, the set of candidates depends on the amount of ambiguity. For a "low" ambiguity (Fig. 13(a), m_1 is *high*), the adverb in the main description is picked from $\{perfectly, \text{--}_2, mostly\}$. For a "medium" ambiguity (m_1 is *medium-high*), it is picked from $\{\text{--}_1,\ nearly_2,\ losely_2\}$ instead. The parameter M_1 (Fig. 18(a)) thus determines when the second set of candidates is preferred to the first one. The decision, of course, is rather subjective. That is why, contrary to D_1, D_2, etc., the parameters M_1 to M_{10} have all been assessed empirically, according to our own intuition. For example, M_9 and M_{10} [Fig. 18(g)] have been chosen considering the sequence of configurations represented by Fig. 16.

It is clear that the linguistic values can be tailored—like the dictionary of terms—to individual users. In particular, a coarser language can be easily obtained by choosing a smaller set of adverbs, fewer linguistic values, and fewer rules (e.g., the set of rules shown in Fig. 13(b) can be ignored if no supplementary description is desired). A fuzzy rule base is a natural mechanism to allow users to remove, add and test new adverbs, rules, and membership distributions.

V. Experimental Results

A. Synthetic Data

Many results on synthetic data have already been presented in Section IV (see Figs. 14–17). Fig. 19 shows two more series of configurations, which have not been used in the training stage for the determination of the linguistic values. Other series can be found in [29]. In fact, all these data are part of an animation that we built to evaluate our system. Structured round 35 key configurations, it is made up of more than two thousand images, and lasts about three and a half minutes. Six short movies supplement the electronic version of the paper and cover a large part of the animation: the first movie, MATSA01.AVI, can be related to Fig. 17(a); MATSA02.AVI is a variant that involves intersecting objects; MATSA03.AVI is linked to Fig. 3, Fig. 11(c) and (d), and Fig. 17(b); MATSA04.AVI shows the series of configurations presented by Fig. 19(b), and MATSA05.AVI shows the series presented by Fig. 19(a); the last movie, MATSA06.AVI, is related to Fig. 4, Fig. 11(g)–(i), and Fig. 16. Note that the six video clips use a pseudo-polar representation of the histograms, more expressive than the Cartesian one (Fig. 20). An example of a frame is shown and described in Fig. 21.

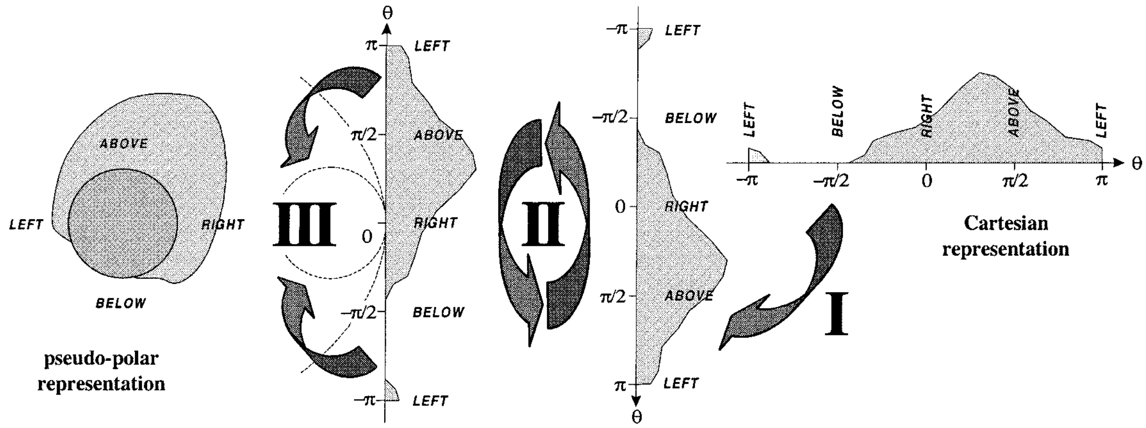


Fig. 20. From Cartesian to pseudo-polar histogram representation: I. Rotate. II. Turn. III. Wrap.

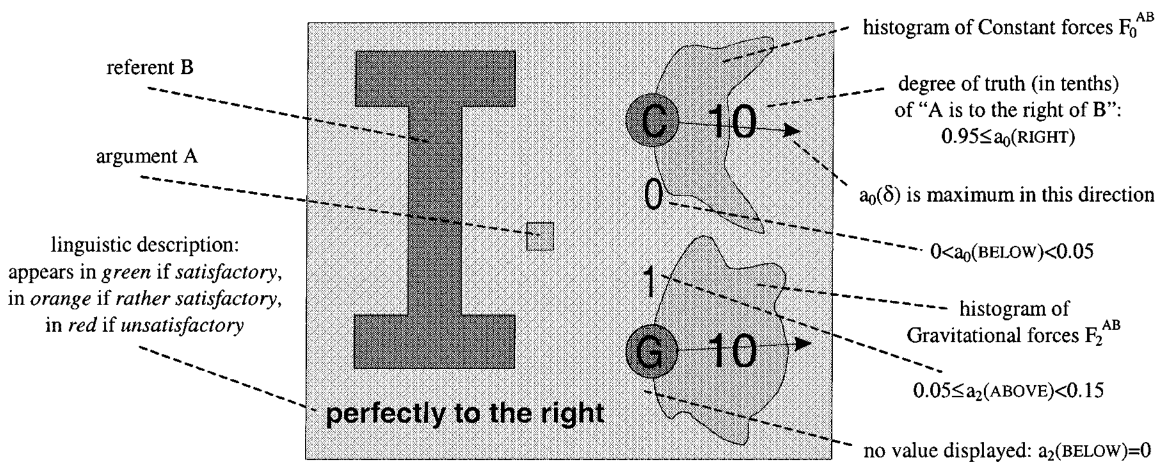


Fig. 21. Sample frame to understand the movies that supplement the electronic version of the paper.

B. Real Data

In our experiments on real data, we used two images provided by the Naval Air Warfare Center (Fig. 22). These images are LADAR (Laser Radar) range images [31] of the power-plant at China Lake, CA. They were processed by applying first a median filter, and then the pseudo-intensity filter $1/\sqrt{1+G_x^2+G_y^2}$, where G_x and G_y are the Sobel gradient magnitudes in a 3×3 window. Finally, the filtered images were segmented and labeled manually. Wang and Keller used the same real data to test a fuzzy rule-based approach for linguistic scene description [19]. Below, their system is referred to as the **WK** system, and the system introduced in the present paper is referred to as the **MK** system. Figs. 23 and 24 show the 24 pairs of objects (or groups of objects) that have been considered in our experiments. For each pair, the results from both **WK** and **MK** are displayed in Tables IV and V.

Before examining the results, let us describe briefly **WK**. The system accepts 373 input variables. Eleven inputs are related to some geometric features (e.g., areas of the objects, distance between them) and 181 come from the angle histogram defined as in [15]. The remaining 181 are from a second histogram that requires the computation, for each point P of the reference object, of the angle made by the two tangents from P to the argument object. The first 192 inputs feed four neural networks fused with the Choquet fuzzy integral, and trained on aggregate responses from a panel of people. The results are the degrees of truth of "A is to the right of B," "A is above B," "A is to the left of B," and "A is below B." The 181 remaining inputs feed a multilayer perceptron that produces the degree of truth of "A surrounds B" (again, trained by the human panel responses). Finally, the five outputs of the neural networks are used as the inputs of a fuzzy rule base containing 242 rules. The **WK** system is able to generate ten different linguistic descriptions. Eight are related to the primitive and compound directions: "A is to the right of B," "A is above-right of B," "A is above B," "A is above-left of B," "A is to the left of B," "A is below-left of B," "A is below B," and "A is below-right of B." The two others are "A surrounds B" and "A is among B." No self-assessment is provided.

MK has not been given the ability to recognize the last two relationships. However, its vocabulary is much richer. Examination of Tables IV and V shows that this richness is generally very well employed. The outputs of the two systems **WK** and **MK** can sometimes be found equivalent (in Table IV, compare the results about the objects 2, 3, 7, and 10; in Table V, compare those about 4, 7, and 9). Nevertheless, **MK** is often much more precise than **WK**. Consider for instance the object 4 of Fig. 23: the descriptions agree for the most part, but **MK** notes that the relationship is not a perfect above-left, and uses the adverb "loosely" to indicate a bias in one direction. Con-

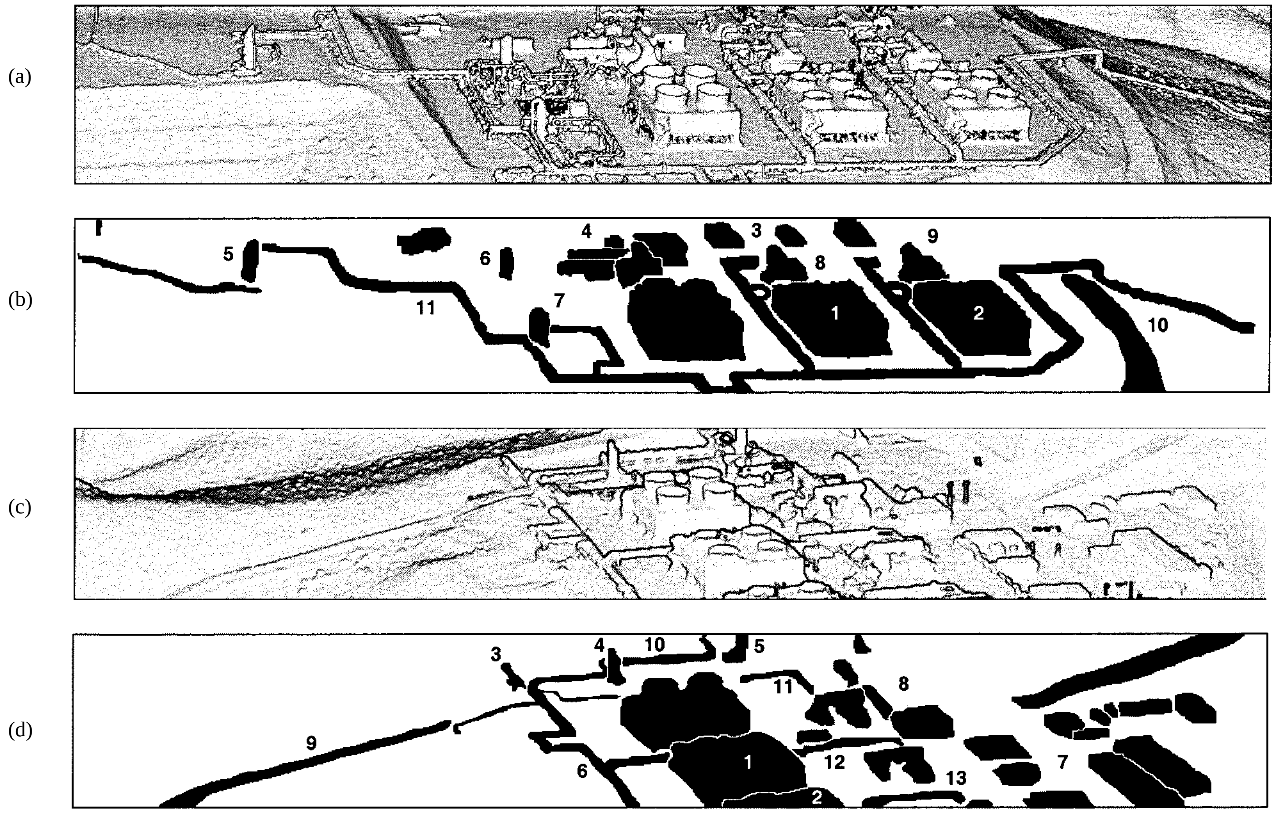


Fig. 22. Real data. (a) LADAR range image NAWC 20675, after filtering. (b) LADAR range image NAWC 20675, after filtering, hand-segmentation and labeling. (c) LADAR range image NAWC 20695, after filtering. (d) LADAR range image NAWC 20695, after filtering, hand-segmentation and labeling.

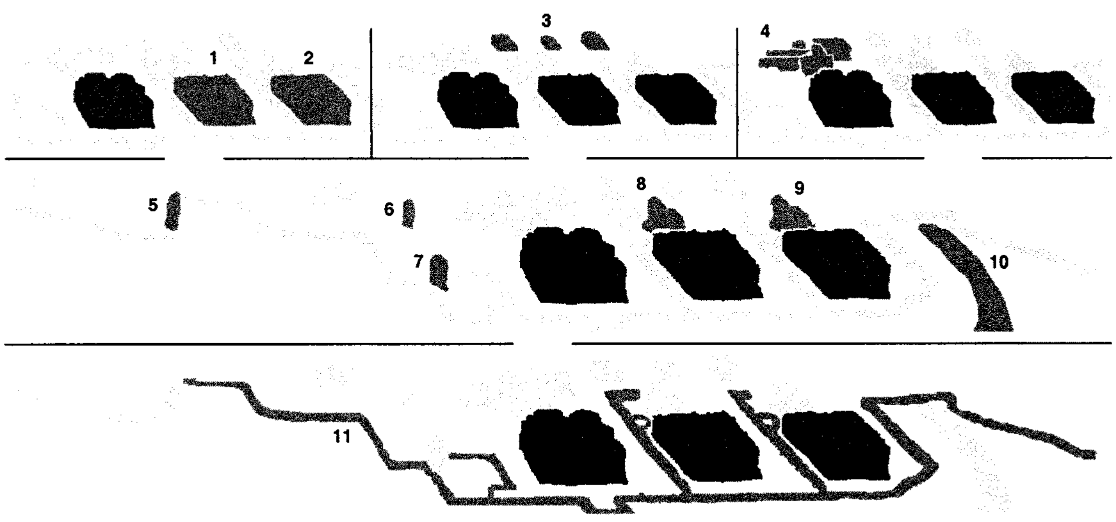


Fig. 23. First series of tests on real data. Configurations. For each image, the reference object is in black, and the argument(s) in dark gray. The light gray objects are ignored.

sider now object 8: the descriptions agree on the primary direction, but **MK** points out that the storehouse is slightly shifted to the left. These expanded vocabulary and increased descriptive ability give to **MK** a higher resolving power than **WK**. For instance, our system distinguishes between the two configurations with objects 5 and 6 of Fig. 23. Concerning object 5, **MK** ignores the "above" relationship, insignificant in distance compared to the "left" relation. Concerning object 6, it expresses that the tower is more to the left of than above the stackbuildings. **WK** does not make any distinction and sees both arguments above-left. Many other examples can be found in Fig. 24 and Table V (consider the objects 1 and 2, 3 and 4, and 10 and 11). We also point out that, contrary to **WK**, the **MK** system guarantees the two following properties: 1) the linguistic descriptions are not sensitive to scale changes and 2) the semantic inverse [10] principle is respected. In Fig. 15(b) for instance, **MK** finds object A perfectly to the *right* of B, but slightly shifted *upward*. Therefore, we can be sure that B will be found perfectly to the

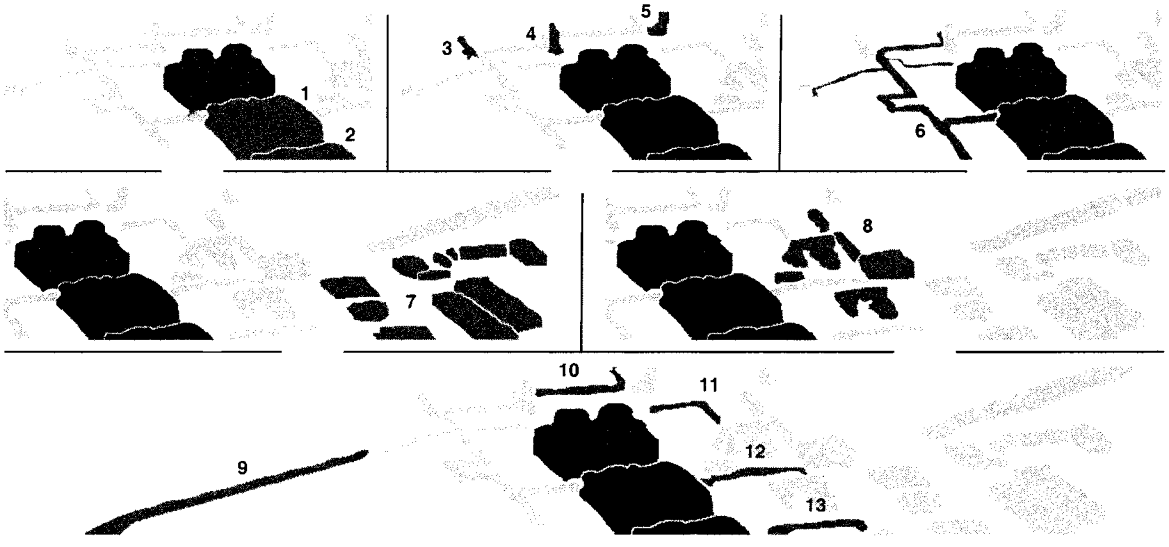


Fig. 24. Second series of tests on real data. Configurations. For each image, the reference object is in black, and the argument(s) in dark gray. The light gray objects are ignored.

TABLE IV
FIRST SERIES OF TESTS ON REAL DATA: RESULTS. **WK** IS THE SYSTEM INTRODUCED IN [19], AND **MK** THE SYSTEM DEFINED IN THE PRESENT PAPER

(a)	**WK**:	The stackbuilding **1** is *to the right of* the reference stackbuilding.
	MK:	The stackbuilding **1** is *perfectly to the right of* the reference stackbuilding, *but slightly shifted upward.* The description is *satisfactory*.
	WK:	The stackbuilding **2** is *to the right of* the reference stackbuilding.
	MK:	The stackbuilding **2** is *perfectly to the right of* the reference stackbuilding. The description is *satisfactory*.
(b)	**WK**:	The group **3** of buildings is *above* the stackbuildings.
	MK:	The group **3** of buildings is *perfectly above* the stackbuildings. The description is *satisfactory*.
(c)	**WK**:	The group **4** of storehouses is *above-left of* the stackbuildings.
	MK:	The group **4** of storehouses is *loosely above-left of* the stackbuildings. The description is *satisfactory*.
(d)	**WK**:	The tower **5** is *above-left of* the stackbuildings.
	MK:	The tower **5** is *perfectly to the left of* the stackbuildings. The description is *satisfactory*.
	WK:	The tower **6** is *above-left* of the stackbuildings.
	MK:	The tower **6** is *to the left of* the stackbuildings, *but a little above*. The description is *satisfactory*.
	WK:	The tower **7** is *to the left of* the stackbuildings.
	MK:	The tower **7** is *perfectly to the left of* the stackbuildings. The description is *satisfactory*.
	WK:	The storehouse **8** is *above* the stackbuildings.
	MK:	The storehouse **8** is *perfectly above* the stackbuildings, *but slightly shifted to the left*. The description is *satisfactory*.
	WK:	The storehouse **9** is *above-right of* the stackbuildings.
	MK:	The storehouse **9** is *perfectly above* the stackbuildings, *but slightly shifted to the right*. The description is *satisfactory*.
	WK:	The road **10** is *to the right of* the stackbuildings.
	MK:	The road **10** is *perfectly to the right of* the stackbuildings. The description is *satisfactory*.
(e)	**WK**:	The pipe **11** surrounds the stackbuildings.
	MK:	???????

left of A, but slightly shifted *downward*. This explains why the term "shifted" has been preferred to "extends." One could say that "A is perfectly to the right of B, but *extends* upward". However, "B is perfectly to the left of A, but *extends* downward" is obviously incorrect. Remember that **MK** does not use any information (area, compactness, etc.) about an individual object.

Although the **MK** system globally performs very well, some results are not totally satisfactory. Dealing with a language as rich as **MK**'s is tricky (it is always easier to be right when vague and imprecise). For instance, **MK** affirms that the stackbuilding 1 of Fig. 23 is perfectly to the right of the referent, but it also describes—and this piece of information is questionable—a secondary direction, "upward." The relation exists because of the top left corner of the argument, which is nearly above the bottom right corner of the referent. This minute detail is caught due to the F_2-histogram's local view. The least satisfactory description

TABLE V
SECOND SERIES OF TESTS ON REAL DATA: RESULTS. **WK** IS THE SYSTEM INTRODUCED IN [19], AND **MK** THE SYSTEM DEFINED IN THE PRESENT PAPER

(a)	**WK**:	The stackbuilding **1** is *below* the reference stackbuilding.
	MK:	The stackbuilding **1** is *loosely below-right of* the reference stackbuilding. The description is *satisfactory*.
	WK:	The stackbuilding **2** is *below* the reference stackbuilding.
	MK:	The stackbuilding **2** is *nearly below-right of* the reference stackbuilding. The description is *satisfactory*.
(b)	**WK**:	The building **3** is *above-left* of the stackbuildings.
	MK:	The building **3** is *to the left of* the stackbuildings, *but a little above*. The description is *satisfactory*.
	WK:	The tower **4** is *above-left* of the stackbuildings.
	MK:	The tower **4** is *above-left* of the stackbuildings. The description is *satisfactory*.
	WK:	The tower **5** is *above* the stackbuildings.
	MK:	The tower **5** is *perfectly above* the stackbuildings, *but strongly shifted to the right*. The description is *satisfactory*.
(c)	**WK**:	The pipe **6** is *to the left of* the stackbuildings.
	MK:	The pipe **6** is *loosely to the left of* the stackbuildings, *but slightly shifted downward*. The description is *rather satisfactory*.
(d)	**WK**:	The group **7** of buildings is *to the right of* the stackbuildings.
	MK:	The group **7** of buildings is *perfectly to the right of* the stackbuildings. The description is *satisfactory*.
(e)	**WK**:	The group **8** of storehouses is *to the right of* the stackbuildings.
	MK:	The group **8** of storehouses is *loosely above-right of* the stackbuildings. The description is *satisfactory*.
(f)	**WK**:	The road **9** is *to the left of* the stackbuildings.
	MK:	The road **9** is *perfectly to the left of* the stackbuildings. The description is *satisfactory*.
	WK:	The pipe **10** is *above* the stackbuildings.
	MK:	The pipe **10** is *perfectly above* the stackbuildings, *but slightly shifted to the left*. The description is *satisfactory*.
	WK:	The pipe **11** is *above* the stackbuildings.
	MK:	The pipe **11** is *loosely above-right of* the stackbuildings. The description is *satisfactory*.
	WK:	The pipe **12** is *to the right of* the stackbuildings.
	MK:	The pipe **12** is *to the right of* the stackbuildings, *but strongly shifted upward*. The description is *rather satisfactory*.
	WK:	The pipe **13** is *to the right of* the stackbuildings.
	MK:	The pipe **13** is *perfectly to the right of* the stackbuildings, *but slightly shifted downward*. The description is *satisfactory*.

might be the one that concerns the object 6 of Fig. 24. A piece of the pipe extends between the uppermost and middle stackbuildings. At the end of the extension, the pipe has a strong "downward" relationship with the uppermost building (and a weak "upward" relationship with the middle one). As a result, the **MK** system assesses the argument to be slightly shifted downward relative to the referent. A similar phenomenon can be observed with the object 12 of the same figure. However, note that in both cases, **MK** itself considers the description *rather* satisfactory. Finally, note that **MK** produces the message "???????" to describe the relative position between the pipe 11 and the stackbuildings of Fig. 23. The output is appropriate, since none of the directional relationships are relevant. The fact is confirmed by **WK**, which has been given the ability to recognize "surrounds."

VI. CONCLUSION

In this paper, we have examined the issues involved in utilizing consistent spatial relationship information to produce linguistic descriptions of natural scenes. The methodology is based on histograms of forces that capture essential elements of relative position with well defined properties. By imposing physical considerations on the histograms, we have introduced new families of fuzzy directional relations. Our system interfaces these families with a fuzzy rule base and handles a rich language to describe the spatial organization of scene regions, as demonstrated by the many examples shown. The system is even able to provide a degree of self-assessment concerning the linguistic descriptions. Good intuitive results are displayed for most cases. The descriptions in general agree with those produced by the Keller and Wang system for straightforward objects (hence, they are compatible with the human panel responses). However, in many cases, they provide better information. Moreover, the fuzzy rule base in our approach is much smaller, and the computation of extra geometric features is not needed. In the future, such features could be used to improve and refine even more the descriptions. Relations like "surrounds" and "is surrounded by," "includes," and "contains," could also be incorporated. The ability to easily adapt the fuzzy rules, their adjectives, and the fuzzy sets defining the adjective meanings, constitute another possible area for enhancement. We are currently working on using our approach to retrieve images and object images given a linguistic description (by a human or by the system, say, from a different aspect angle).

ACKNOWLEDGMENT

P. Matsakis, J. Keller, J. Marjamaa, and O. Sjahputera wish to thank the personnel of the Naval Air Warfare Center for their assistance with the LADAR imagery

REFERENCES

[1] R. Brooks, "Symbolic reasoning among 3-D models and 2-D images," *Artif. Intell.*, vol. 17, pp. 285–348, 1981.

[2] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[3] D. M. Russell, "Where do I look now?," in *Proceedings, Pattern Recognition, and Image Processing*, 1979, pp. 175–183.

[4] K. Andress and A. Kak, "Evidence Accumulation and Flow of Control in a Hierarchical Spatial Reasoning System," in *Advances in Spatial Reasoning*, S. Chen, Ed. Norwood, NJ: Ablex, 1990, vol. 1, pp. 133–179.

[5] E. Walker, M. Herman, and T. Kanade, "A framework for representing and reasoning about three-dimensional objects in vision," *Artif. Intell.*, vol. 9, no. 2, pp. 47–58, 1988.

[6] H. Kang and E. Walker, "Characterizing and controlling approximation in hierarchical perceptual grouping," *Fuzzy Sets Syst.*, vol. 65, pp. 187–224, 1994.

[7] R. Antony, "A Hybrid Spatial/Object-Oriented DBMS to Support Automated Spatial, Hierarchical, and Temporal Reasoning," in *Advances in Spatial Reasoning*, S. Chen, Ed. Norwood, NJ: Ablex, 1990, vol. 1, pp. 63–132.

[8] S. Medasani and R. Krishnapuram, "A fuzzy approach to content-based image retrieval," in *Proc. FUZZ-IEEE'99*, Seoul, Korea, pp. 1251–1260.

[9] P. Winston, "Learning Structural Descriptions from Examples," in *The Psychology of Computer Vision*, P. Winston, Ed. New York: McGraw-Hill, 1975.

[10] J. Freeman, "The modeling of spatial relations," *Comput. Graph. Image Process.*, vol. 4, pp. 156–171, 1975.

[11] S. Dutta, "Approximate spatial reasoning: Integrating qualitative and quantitative constraints," *Int. J. Approx. Reason.*, vol. 5, pp. 307–331, 1991.

[12] S. Y. Lee and F. J. Hsu, "Spatial reasoning and similarity retrieval of images using 2-D C-string knowledge representation," *Pattern Recognit.*, vol. 25, no. 3, pp. 305–318, 1992.

[13] T. S. Levitt and D. T. Lawton, "Qualitative navigation for mobile robots," *Artif. Intell.*, vol. 44, no. 3, pp. 305–360, 1990.

[14] J. Sharma and D. M. Flewelling, "Inferences from combined knowledge about topology and directions," in *Proc. LNCS 951*, 1995, pp. 279–291.

[15] K. Miyajima and A. Ralescu, "Spatial organization in 2-D segmented images: Representation and recognition of primitive spatial relations," *Fuzzy Sets Syst.*, vol. 65, no. 2/3, pp. 225–236, 1994.

[16] J. M. Keller and X. Wang, "Comparison of spatial relation definitions in computer vision," in *Proc. ISUMA-NAFIPS'95*, College Park, MD, 1995, pp. 679–684.

[17] R. Krishnapuram, J. M. Keller, and Y. Ma, "Quantitative analysis of properties and spatial relations of fuzzy image regions," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 3, pp. 222–233, 1993.

[18] X. Wang and J. M. Keller, "Human-based spatial relationship generalization through neural/fuzzy approaches," *Fuzzy Sets Syst.*, vol. 101, no. 1, pp. 5–20, 1999.

[19] J. M. Keller and X. Wang, "A fuzzy rule-based approach for scene description involving spatial relationships," *Comput. Vis. Image Understand.*, vol. 80, pp. 21–41, 2000.

[20] P. Matsakis, "Relations spatiales structurelles et interprétation d'images," Ph. D. dissertation, Inst. Recherche Inform. Toulouse, France, 1998.

[21] P. Matsakis and L. Wendling, "Représentation de la position relative d'objets 2-D au moyen d'un histogramme de forces," *Trait. Signal*, vol. 15, no. 1, pp. 25–38, 1998.

[22] ——, "A new way to represent the relative position between areal objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 634–643, July 1999.

[23] I. Bloch, "Fuzzy relative position between objects in image processing: New definition and properties based on a morphological approach," *Int. J. Uncertainty Fuzz. Knowl.-Based Syst.*, vol. 7, no. 2, pp. 99–133, 1999.

[24] P. D. Gader, "Fuzzy spatial relations based on fuzzy morphology," in *Proc. IEEE Int. Conf. Fuzzy Systems*, vol. 2, Barcelona, Spain, 1997, pp. 1179–1183.

[25] L. T. Kòczy, "On the description of relative position of fuzzy patterns," *Pattern Recognit. Lett.*, vol. 8, pp. 21–28, 1988.

[26] P. Matsakis, J. Keller, and L. Wendling, "F-Histograms and fuzzy directional spatial relations," in *Proc. LFA'99 (French-Speaking Conf. Fuzzy Logic and Its Applications)*, vol. 1, Valenciennes, France, pp. 207–213.

[27] K. P. Gapp, "Angle, distance, shape, and their relationship to projective relations," in Proc. 17th Conf. Cognitive Science Soc., 1995.

[28] G. Retz-Schmidt, "Various views on spatial prepositions," *Artif. Intell.*, vol. 9, no. 2, pp. 95–105, 1988.

[29] J. M. Keller and P. Matsakis, "Aspects of high level computer vision using fuzzy sets," in *Proc. 8th IEEE Int. Conf. Fuzzy Systems*, Seoul, Korea, 1999, pp. 847–852.

[30] K. P. Gapp, "Basic meanings of spatial relations: Computation and evaluation in 3-D space," in *AAAI'94*, Seattle, WA, pp. 1393–1398.

[31] M. Suk and S. M. Bhandarkar, *Three-Dimensional Object Recognition from Range Images*, T. L. Kunii, Ed. New York: Springer-Verlag, 1992.

Pascal Matsakis received the B.Sc. degree both in mathematics and computer science in 1992, and the Ph.D. degree in computer science from Paul Sabatier University, Toulouse, France in 1998.

He worked at the Toulouse Institute of Research in Computer Science, in the Group of Image Processing and Understanding, until August 1999. He is currently an Assistant Professor of Computer Engineering and Computer Science at the University of Missouri, Columbia. His research interests include computer vision, pattern recognition, satellite image analysis, human-machine interaction, fuzzy set theory, and fuzzy logic.

James M. Keller (M'79–SM'92–F'00) received the Ph.D. degree in mathematics from the University of Missouri, Columbia, in 1978.

He has had faculty appointments in the Bioengineering/Advanced Automation Program, the Electrical and Computer Engineering Department, and the Computer Engineering and Computer Science Department at the University of Missouri, Columbia, where he currently holds the rank of Professor. He is also the R. L. Tatum Research Professor in the College of Engineering. His research interests include computer vision, pattern recognition, fuzzy set theory and fuzzy logic, fractal geometry, and neural networks. He has coauthored over 175 technical publications.

Professor Keller is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) for whom he has presented live and video tutorials on fuzzy logic in computer vision, is a national lecturer for the Association for Computing Machinery (ACM), is an IEEE Neural Networks Council Distinguished Lecturer, and is a past President of the North American Fuzzy Information Processing Society (NAFIPS). He is the Editor-in-Chief of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, is an Associate Editor of the *International Journal of Approximate Reasoning*, and is on the editorial board of *Pattern Analysis and Applications*, *Fuzzy Sets and Systems*, and the *Journal of Intelligent and Fuzzy Systems*. He was the conference chair of the 1991 NAFIPS workshop, program co-chair of the 1996 NAFIPS meeting, program co-chair of the 1997 IEEE International Conference on Neural Networks, and the program chair of the 1998 IEEE International Conference on Fuzzy Systems.

Laurent Wendling received the Ph.D. degree in computer science from University Paul Sabatier, Toulouse, France, in 1997.

From 1993 to 1999, he worked at the Toulouse Institute of Research in Computer Science, in the field of pattern recognition using fuzzy segmentation and graph matching. He is currently an Assistant Professor at the University Henri-Poincare and a member of the ISA research team at the LORIA Laboratory in Nancy. His present research interests are in graphical document analysis and indexing strategy of images.

Jonathon Marjamaa received the B.A. degree in computer science from St. Louis University, St. Louis, MO, in 1999. He is currently pursuing the M.Sc. degree in the Computer Engineering and Computer Science Department at the University of Missouri, Columbia. His main research interests are in computer graphics.

Ozy Sjahputera was born on December 14, 1970, in Sabang, Aceh, Indonesia. He received the B.Sc. degree in computer engineering in 1994, the B.Sc. degree in electrical engineering in 1994, and the M.Sc. degree in electrical engineering in 1996, from the University of Missouri, Columbia, where he is currently pursuing the Ph.D. degree. His research interests are in high level computer vision.

He worked as a member of the research team at the Laboratory of Aerodynamics, Gas Dynamics, and Vibration, of the Agency of Assessment and Application of Technology of the Republic of Indonesia, Serpong, Jakarta, Indonesia, from 1996 to 1998.