

RELATIVE POSITION-BASED SPATIAL RELATIONSHIPS USING MATHEMATICAL MORPHOLOGY

R. Gökberk Cinbiş and Selim Aksoy

Department of Computer Engineering
Bilkent University
Bilkent, 06800, Ankara, Turkey

ABSTRACT

Spatial information is a crucial aspect of image understanding for modeling context as well as resolving the uncertainties caused by the ambiguities in low-level features. We describe intuitive, flexible and efficient methods for modeling pairwise directional spatial relationships and the ternary between relation using fuzzy mathematical morphology. First, a fuzzy landscape is constructed where each point is assigned a value that quantifies its relative position according to the reference object(s) and the type of the relationship. Then, the degree of satisfaction of this relation by a target object is computed by integrating the corresponding landscape over the support of the target region. Our models support sensitivity to visibility to handle areas that are partially enclosed by objects and are not visible from image points along the direction of interest. They can also cope with the cases where one object is significantly spatially extended relative to others. Experiments using synthetic and real images show that our models produce more intuitive results than other techniques.

Index Terms— Spatial relationships, mathematical morphology, fuzzy sets, relative position, between

1. INTRODUCTION

Traditional approaches to scene classification and retrieval have used global features for image representation. However, the object variability and background complexity in realistic data sets have increased the need for region-based analysis. More recently, local feature-based methods have received significant attention due to their invariance to translation, scale and rotation, and robustness to partial occlusion and clutter. However, the visual polysemy caused by similar local features (also called patches) occurring at semantically different parts of a scene leads to ambiguities if the classification methods do not exploit additional contextual information. Furthermore, even when regions/patches can be classified correctly, two scenes with similar regions/patches can have different interpretations if they have different arrangements. This especially becomes important and critical when the scenes contain complex structures like in medical or remote sensing images.

Contextual information has long been acknowledged for playing a very important role in both human and computer vision. A structural method for modeling context in images is through quantification of spatial relationships. Typical relationships studied in the literature include topological, distance-based, and relative position-based relationships. We have successfully used such relationships [1]

for image classification and retrieval in scenarios that cannot be expressed by traditional pixel- and region-based approaches.

In this paper, we concentrate on binary directional relationships and the ternary between relationship for modeling relative positions. Most of the existing methods for defining binary spatial positions rely on angle measurements between points of objects of interest [2]. The angle between object centroids or the histogram of angles between all pairs of points have been used for approximating relative positions. Alternatives include histogram of forces, projections, and morphological methods (see [2, 3] for reviews). The between relationship has not been studied as thoroughly as the binary relationships. Example models of between include convex hulls, combinations of line segments, and mathematical morphology (see [4] for an extensive review and a comparative study). However, most of these methods are computationally expensive, some give reasonable results only for compact objects, and many cannot handle the cases where one of the regions is spatially extended relative to the other or if regions have concavities that are invisible from each other.

Intuitively, the influence of the shape of the object (e.g., concavities, extent) and the influence of the distance between the objects are important points to be considered in the design of an algorithm. Mathematical morphology provides a strong basis for such studies. Furthermore, the ambiguities and subjectiveness inherent in the definitions of the relationships make fuzzy representation a promising approach for modeling the imprecision in both the images and the results.

In this paper, we propose intuitive, flexible and efficient methods for modeling pairwise directional relationships and the ternary between relation using fuzzy mathematical morphology. First, a fuzzy landscape is defined where each point is assigned a value that quantifies its relative position according to the reference object(s) and the type of the relationship. Directional mathematical dilation with fuzzy structuring elements is used to compute this landscape. Then, the degree of satisfaction of this relation by a target object is computed by integrating the corresponding landscape over the support of the target region.

Our main contributions in this paper are the flexible definitions for fuzzy structuring elements that are tunable along both radial and angular dimensions. Furthermore, the proposed methods support the notion of visibility to handle image areas that are fully or partially enclosed by a reference object and are not visible from image points along the direction of interest. Our definitions also handle the cases where one object is significantly spatially extended relative to the other by taking spatial proximity into consideration. The methods are illustrated and compared to other techniques using synthetic images and real satellite scenes.

This work was supported in part by the TUBITAK CAREER Grant 104E074 and European Commission Sixth Framework Programme Marie Curie International Reintegration Grant MIRG-CT-2005-017504.

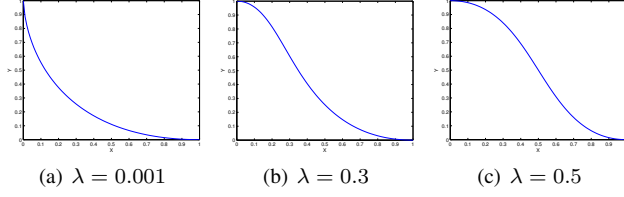


Fig. 1. Examples of $g_\lambda(x)$ with the shape of a cubic Bézier curve and a single parameter λ .

2. DIRECTIONAL SPATIAL RELATIONSHIPS

Directional relationships describe the spatial arrangement of two objects relative to each other. Although it is common to use right, left, above and below as the directions, it is more generalizable to use an angle-based definition of these relations. Given a reference object B and a direction specified by the angle α , the landscape $\beta_\alpha(B)$ around the reference object along the given direction can be defined as a fuzzy set such that the membership value of an image point corresponds to the degree of satisfaction of the spatial relation.

This relationship can be defined in terms of the angle $\theta_\alpha(x, b)$ between the vector from a point b in the reference object to a point x in the image and the unit vector along the direction α measured with respect to the horizontal axis. Bloch [5] suggested that the smallest such angle computed for a point in the image considering all points in the reference object corresponds to the visibility of the image point from the reference object in the direction α , and defined the landscape using a function linearly decreasing with θ as

$$\beta_\alpha(B)(x) = \max \left\{ 0, 1 - \frac{2}{\pi} \min_{b \in B} \theta_\alpha(x, b) \right\}. \quad (1)$$

It can be shown that this is equivalent to the morphological dilation of B ,

$$\beta_\alpha(B)(x) = (B \oplus \nu_\alpha)(x) \cap B^c, \quad (2)$$

using the fuzzy structuring element

$$\nu_\alpha(x) = \max \left\{ 0, 1 - \frac{2}{\pi} \theta_\alpha(x, o) \right\} \quad (3)$$

where o is its origin (center) and B is removed from the result of dilation in (2).

However, the linear function may not give realistic results for many cases (see Figure 2 and Section 4 for examples). We propose a more intuitive and flexible structuring element using a nonlinear function with the shape of a Bézier curve:

$$\nu_{\alpha, \lambda}(x) = g_\lambda \left(\frac{2}{\pi} \theta_\alpha(x, o) \right) \quad (4)$$

where λ determines the inflection point (see [3] for the derivation) and the nonlinear function enables different definitions of fuzziness for different cases. Figure 1 shows examples for different λ values.

The definition of the structuring element can be further extended to decrease the degree of a point's spatial relation to a reference object according to its distance to that object by introducing a new linear term

$$\nu_{\alpha, \lambda, \tau}(x) = g_\lambda \left(\frac{2}{\pi} \theta_\alpha(x, o) \right) \max \left\{ 0, 1 - \frac{\|\vec{ox}\|}{\tau} \right\} \quad (5)$$

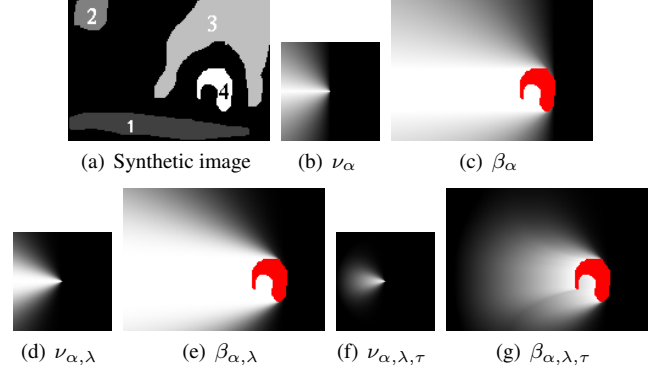


Fig. 2. A synthetic image and directional landscapes for region 4 using the parameters $\alpha = \pi$, $\lambda = 0.3$ and $\tau = 100$.

where $\|\vec{ox}\|$ is the Euclidean distance of point x from the structuring element's center and τ is a threshold corresponding to the distance where a point is no longer visible from the reference object. This definition also has a computational advantage because in (3) and (4) the structuring element must be at least twice as large as the landscape of interest in the image space (landscape computation has quadratic complexity with respect to image size) whereas in (5) a structuring element with size of at most $2\tau \times 2\tau$ is sufficient (linear complexity), leading to dramatical improvements in the efficiency of the algorithm. As can be seen in Figure 2, definition ν_α given in (3) as proposed in [5] leads to a landscape with a large spread and unintuitive transitions when the angle departs from α whereas $\nu_{\alpha, \lambda}$ given in (4) and $\nu_{\alpha, \lambda, \tau}$ given in (5) provide more intuitive landscapes with more compact support (see [3] for more examples).

In directional dilation of (2), the areas that are fully or partially enclosed by the reference object but are not visible from image points along the direction of interest may have high values as in the cavity of region 4 in Figure 2. To overcome this problem, we propose the following definition

$$\beta_{\alpha, \lambda, \lambda', \tau}(B)(x) = (B \oplus \nu_{\alpha, \lambda, \tau})(x) \cap (B \oplus \nu_{\alpha + \pi, \lambda'})(x)^c \quad (6)$$

where the first dilation uses the structuring element defined in (5) and the second dilation uses the structuring element defined in (4). We compute fuzzy intersection using multiplication as the t -norm operator and compute fuzzy complement by subtracting the original values from 1. This definition of visibility is illustrated in Figure 3.

3. BETWEEN RELATIONSHIP

Given two reference objects B and C , the landscape $\beta_{\delta}(B, C)$ between them can also be defined as a fuzzy set. This landscape can be computed as the intersection of the directional dilations of the reference regions along the directions $\alpha = \theta_{\delta}$ and $\alpha = \theta_{\delta} + \pi$ where θ_{δ} is the relative position of the reference objects. This relative position can be calculated using the maximum or average value in the histogram of angles between all pairs of points of the reference objects [4]. Then, the landscape is computed as

$$\beta_{\delta}(B, C)(x) = \beta_{\alpha = \theta_{\delta}, \lambda, \lambda'}(B)(x) \cap \beta_{\alpha = \theta_{\delta} + \pi, \lambda, \lambda'}(C)(x) \quad (7)$$

where the directional landscape $\beta_{\alpha, \lambda, \lambda'}$ is computed as in (6) without considering τ (using (4) instead of (5)). Since the landscape should include only the areas that are visible from both reference objects, the notion of visibility described in Section 2 is used.

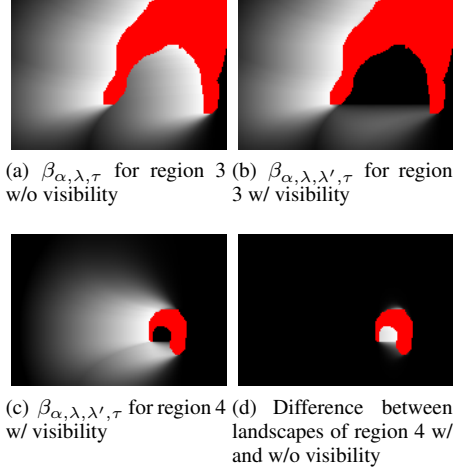


Fig. 3. Directional landscapes with and without visibility using the parameters $\alpha = \pi$, $\lambda = 0.3$, $\lambda' = 0.001$ and $\tau = 100$.

Although histogram of angles generally provides a good approximation to the relative position of two objects, it fails in the cases where one object is significantly spatially extended relative to the other [4] (see Figure 4 for examples). We propose to solve this problem by taking into account only the part of the spatially extended region close to the other region. (Bloch *et al.* [4] called this the “myopic vision” and suggested to use line segments to approximate close parts of the regions but did not specify the details of the method.)

Spatial proximity for handling extended regions is incorporated into our morphological approach using a weighted histogram of angles where the contribution of the angle between each point pair in the histogram is weighted by the term $\max\{0, 1 - \|\vec{bc}\|/\tau_{\text{myopic}}\}$ (instead of a constant weight of 1 in [4]) where \vec{bc} is the Euclidean distance between the points b and c , and τ_{myopic} is the threshold for the maximum distance between two points for allowing them to contribute to the histogram. The proposed definition of myopic vision is illustrated in Figure 4.

Finally, after calculating the landscape β for a spatial relation as in Sections 2 or 3, the degree of satisfaction of this relation by a target object A can be computed as

$$\mu(A) = \frac{1}{\text{area}(A)} \sum_{a \in A} \beta(a). \quad (8)$$

4. ILLUSTRATIVE EXAMPLES

In Tables 1, 2 and 3, experimental statistics using the synthetic image in Figure 2(a) are given (see [3] for more details). For the landscapes calculated using our definitions, the constants are set as: $\lambda = 0.3$, $\lambda' = 0.001$, $\tau = 150$. Table 1 presents the directional relationship satisfaction degrees of several object pairs in the directions left, right, above and below, where α value corresponds to π , 0, $\pi/2$ and $-\pi/2$, respectively. For reference region 1 and target region 4, our method decides that 4 is mostly above 1. This decision is consistent with intuition. However, the centroid-based method says that 4 is more to the right than above, and Bloch’s definition erroneously gives 0.41 for left because of its large spread in the landscape. Bloch’s definition also gives conflicting results for the reference-target relations 1-2, 1-3 and 3-4 because of the same problem. The rest of the cases give similar results for all methods.

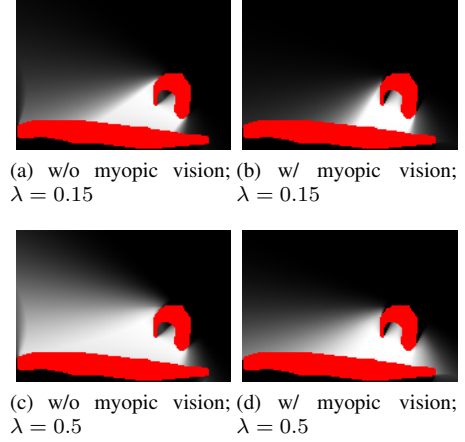


Fig. 4. Between landscape examples for regions 1 and 4 where 1 is spatially extended relative to 4. τ_{myopic} is set to half of image width and $\lambda' = 0.001$. The relative angles are 42.28° and 63.40° for figures without and with myopic vision, respectively. For larger values of λ , the error in landscape without myopic vision becomes more significant.

Table 2 presents the relative angles for several object pairs. For our myopic vision definition, Inf represents that objects under consideration are too distant to be related (determined by τ_{myopic}). This behavior is an advantage of the proposed method because it also identifies the reference object pairs where the between relationship is meaningless. For all cases, our myopic vision definition gives more intuitive results.

Table 3 presents the between relationship satisfaction degrees. We can intuitively say that object 4 is between 1 and 3 more than it is between 1 and 2. We can also see that object 4 is not between 2 and 3, and 2 is not between 3 and 4. Our results are much closer to these expectations than the results of the method proposed in [4]. Both methods perform similarly for the rest of the cases.

Figure 5 shows a LANDSAT scene of British Columbia in Canada and its segmentation using the method in [1]. Figure 6 illustrates the scenario for searching for bridges where a bridge is defined as a region classified as asphalt or concrete and is between two water regions. Figure 7 illustrates the scenario of finding the fields to the north (above) of a river (water). The directional landscape without visibility in Figure 7(a) erroneously covers some areas that are to the south of the river. Introducing visibility using the structuring element in (5) with $\alpha = \pi/2$, $\lambda = 0.5$, $\tau = 150$ for the first dilation in (6) and $\alpha = -\pi/2$, $\lambda = 0.001$, $\tau = 100$ for the second dilation in (6) produces the landscape in Figure 7(b) where areas with water regions closer to them from below than above have high membership values for the “field above water” relationship. Finally, restricting the size of the structuring element to 200×200 in the second dilation in (6) gives the landscape in Figure 7(c) where areas with a water region closer than 200 pixels (corresponding to 6 km) from above are ignored in the relationship.

5. CONCLUSIONS

We presented new, flexible and efficient definitions for modeling binary directional relationships and the ternary between relationship using fuzzy mathematical morphology techniques. Our definitions support the notion of visibility for handling areas that are partially

Table 1. Satisfaction degrees of the directional relationships for object pairs in the synthetic image in Figure 2(a).

Ref.	Target	Centroid-based [2]				Bloch's definition (3)				Our definition (6)			
		left	right	above	below	left	right	above	below	left	right	above	below
1	2	0.24	0.00	0.76	0.00	0.60	0.13	1.00	0.00	0.05	0.01	0.46	0.00
1	3	0.00	0.38	0.62	0.00	0.19	0.70	0.98	0.00	0.03	0.14	0.53	0.00
1	4	0.00	0.72	0.28	0.00	0.41	0.87	1.00	0.00	0.05	0.40	0.79	0.00
3	4	0.01	0.00	0.00	0.99	1.00	0.99	0.34	1.00	0.05	0.01	0.00	0.72

Table 2. Relative angles (in degrees) between object pairs in the synthetic image in Figure 2(a).

Obj.1	Obj.2	Centroid	Hist. of angles	Hist. of angles with myopic vision
1	2	119.25	115.98	93.98
1	4	31.88	42.29	63.41
2	3	-10.99	-12.03	-21.97
2	4	-29.92	-30.04	Inf

Table 3. Satisfaction degrees of the between relationship for object triplets in the synthetic image in Figure 2(a).

Ref.1	Ref.2	Target	Bloch <i>et al.</i> 's defn. (17) in [4]	Our defn. (7)
1	2	3	0.12	0.10
1	2	4	0.52	0.22
1	3	4	0.77	0.95
2	3	4	0.41	0.02
3	4	2	0.27	0.00

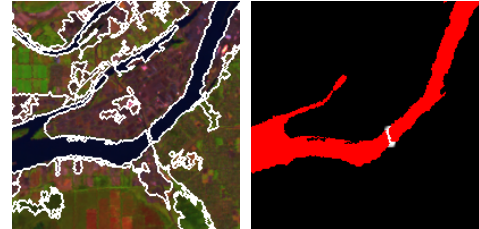
enclosed by objects and are not visible from image points along the direction of interest. They also cover the cases where one object is significantly spatially extended relative to the other. Numerical and visual examples showed that our models often produce more intuitive results than the state-of-the-art techniques. Future work includes using these models for image classification and retrieval.

6. REFERENCES

- [1] S. Aksoy, K. Koperski, C. Tusk, G. Marchisio, and J. C. Tilton, "Learning Bayesian classifiers for scene classification with a visual grammar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 581–589, March 2005.
- [2] I. Bloch and A. Ralescu, "Directional relative position between objects in image processing: A comparison between fuzzy approaches," *Pattern Recognition*, vol. 36, no. 7, pp. 1563–1582, July 2003.
- [3] R. G. Cinbis and S. Aksoy, "Modeling spatial relationships in images," Tech. Rep. BU-CE-0702, Department of Computer Engineering, Bilkent University, Ankara, Turkey, January 2007, <http://retina.cs.bilkent.edu.tr/papers/BU-CE-0702.pdf>.
- [4] I. Bloch, O. Colliot, and R. M. Cesar, "On the ternary spatial relation "between"," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 36, no. 2, pp. 312–327, April 2006.
- [5] I. Bloch, "Fuzzy relative position between objects in image processing: A morphological approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 7, pp. 657–664, July 1999.

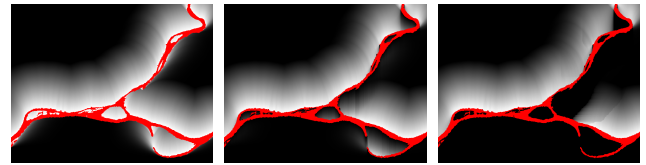


Fig. 5. LANDSAT scene of British Columbia in Canada.



(a) Zoomed sub-image (b) Between landscape of two water regions using st. el. of size 10×10

Fig. 6. Searching for bridges in the sub-image marked with a red rectangle in Figure 5 (see text for details).



(a) Without visibility (b) With visibility (c) With visibility using restricted st. el.

Fig. 7. Searching for fields to the north of a river in the sub-image marked with a yellow rectangle in Figure 5 (see text for details).