RETINA Vision and Learning Group

# Modeling Spatial Relationships in Images

## Version 1.0

R. Gökberk Cinbiş
cinbis@ug.bilkent.edu.tr
Selim Aksoy
saksoy@cs.bilkent.edu.tr

Report Number: BU-CE-0702

RETINA Vision and Learning Group
Department of Computer Engineering
Bilkent University
Bilkent, 06800, Ankara, Turkey

Last Updated: January 19, 2007

Email: retina@cs.bilkent.edu.tr
URL: http://retina.cs.bilkent.edu.tr

# Contents



**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 flexible, intuitive 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 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.

# 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 to resolve these uncertainties. Furthermore, even when regions/patches can be classified correctly, two scenes with similar regions/patches can have different interpretations if the regions/patches 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. Consequently, development of context models has become a challenging problem in both statistical and structural pattern recognition. A structural way of modeling context in images is through quantification of spatial relationships. Typical relationships studied in the literature include topological (set relationships, adjacency), distance-based (near, far) and relative position-based relationships.

The methods used for computing these relationships depend on the way how objects/regions are modeled [7]. Examples include grid-based representations [5], centroids and minimum bounding rectangles [14]. Centroids and minimum bounding rectangles are useful when regions have circular or rectangular shapes but regions in natural scenes often do not follow these assumptions. When regions are represented as sets of points (pixels), adjacency of two regions can be measured as a fuzzy function of the distance between their closest points or using morphological dilations modeling connectivities [7]. Distance-based relationships can also be defined using fuzzy membership functions modeling symbolic classes such as near and far using the distance between boundary pixels. More complex representations of spatial relationships include spatial association networks [12], knowledge-based spatial models [10, 15], and attributed relational graphs [13]. However, these approaches require either manual delineation of regions by experts or partitioning of images into grids. Therefore, they are not generally applicable due to the infeasibility of manual annotation in large databases or because of the limited expressiveness of fixed sized grids that cannot capture large number of structures with different sizes.

In previous work [3, 4], we developed fuzzy models for pairwise spatial relationships based on overlaps between region boundaries (disjoined, bordering, invaded by, surrounded by), distances between region boundaries (near, far), and angles between region centroids (right, left, above, below). Then, we combined these pairwise relationships into higher order relationship models using fuzzy logic, and illustrated their use in image retrieval [4]. We also developed a Bayesian framework that learns image classes based on automatic selection of distinguishing (e.g., frequently occurring, rarely occurring) relations between regions [2]. Finally, we built attributed relational graph structures to model scenes by representing regions by the graph nodes and their spatial relationships by the edges between such nodes [1], and used relational matching techniques to find similarities between graphs representing different scenes. We demonstrated the effectiveness of these approaches in scenarios that cannot be expressed by traditional approaches but where the proposed models can capture both feature and spatial characteristics of scenes and model them according to their high-level semantic content.

## 1.1 Related work

In this report, we concentrate on relative position-based relationships: binary directional relationships and the ternary between relationship. Most of the existing methods for defining relative spatial positions rely on angle measurements between points of objects of interest where the angle corresponding to a pair of points is computed between the segment joining the points and the horizontal axis in the coordinate system [9].

A common approximation is to represent objects using their centroids and to use the angle corresponding to the centroids for defining the relative position of those objects. This approach is widely used because of its simplicity and computational feasibility. However, when objects are not compact, the results can be quite counterintuitive.

As an alternative, a histogram is constructed using the angles between all pairs of points from both objects. Then, the mean or the maximum angle computed from this histogram can be used to represent the relative position of these objects. When objects are large, this is an expensive method not only because angles between each pair of points of two objects should be computed, but also this process may be required for each object pair in the image. In addition, similar to the centroid-based approach, correctness of this method decreases as objects get spatially more extended relative to each other.

Matsakis and Wendling [11] proposed the histogram of forces as an alternative to the histogram of angles. This method computes the degree of satisfaction for a given angle using intersection of longitudinal sections of objects with lines having the desired direction.

The projection approach is different from the previous two in that it does not use any histogram. It is based on the projection of the reference object on the axis representing the direction of interest where the satisfaction of the relationship by a target object is proportional to the ratio of the number of pixels it has on the selected side of this projection.

The morphological approach is based on directional dilations where a fuzzy landscape for a reference object is created at a given angle and other objects are compared to this landscape to evaluate how well they match with the areas having high membership values. As the structuring element for dilation, the histogram of angles described above can be used after normalization by the maximum frequency [8]. Since only one landscape is calculated for each object for a particular direction of interest and all other objects are compared to this landscape, it is an efficient approach; however, in the original definition in [6], its computational burden increases exponentially as image gets larger since both structuring element size and landscape size increases.

Another relationship that is often used in daily life but has not been studied as thoroughly as the binary relationships is the between relationship (see [8] for an extensive review and a comparative study). The most common and intuitive approach is based on convex hulls. The landscape between two reference regions can be defined as the difference between the convex hull of these regions and the regions themselves. However, if a region is spatially extended relative to the other or if regions have concavities that are invisible from each other, this method is generally unsuccessful.

Another approach is using morphological dilations and separations, where a seed line is obtained by dilating two regions until they meet, and the between landscape is generated by dilating the seed with the same number of dilations. However, this method can only be applied to convex sets. In addition, as the distance between the regions increases, the landscape may become too extended.

A better approach is to use the watershed line (or skeleton by influence zones) to obtain the seed line and apply geodesic dilation until the convergence of the watershed line to generate the landscape. Although this method is also applicable for non-convex sets, when regions are compact, it produces the same landscape with the convex hull approach; thus, it has the same drawbacks.

Directional dilations are also useful for the between relationship. After obtaining an approximate relative angle between the reference regions, directional dilations are applied to both regions to extend them towards each other to generate the landscape. Angle histogram can be directly used to create the structuring element for dilation.

Bloch *et al.* [8] defined another approach called visibility using admissible segments, where the landscape is defined as the combination of the segments reaching the boundaries of both regions. A more flexible landscape is obtained by fuzzy visibility, where segments are defined to have intermediate points and the landscape is obtained using the angle between the line segments intersecting at the intermediate point. In addition, for the cases where one region is spatially extended, Bloch *et al.* [8] suggested to use the closest parts of the regions to compute the landscape and called this the myopic vision and pointed out that it is possible to combine the visibility and myopic vision approaches. However, these approaches can often be computationally costly.

## 1.2 Proposed approach

Intuitively, the influence of the shape of the object (e.g., concavities, extent) and the influence of the distance between 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 images and results.

In this report, we propose flexible, intuitive and efficient methods for modeling directional spatial relationships for object pairs and the between relation for three objects using fuzzy mathematical morphology. First, a fuzzy landscape is defined where each point is assigned a value that represents the degree of satisfaction for the point according to the reference object(s) and the type of relationship. Directional mathematical dilation with a fuzzy structuring element is used to compute this landscape. We provide flexible definitions of fuzzy structuring elements that are tunable along both radial and angular dimensions. Then, the definitions for the fuzzy landscape are extended to support sensitivity to visibility to handle image areas that are fully or partially enclosed by a reference object but are not visible from image points along the direction of interest. Given an object and a direction of interest that specifies a spatial relationship, the degree of satisfaction of this relation by a target object can be computed by integrating the landscape corresponding to this relation over the support of the target region.

Next, the definitions of the directional spatial relationships are combined to generate a landscape in which the degree of each image area being located between the reference objects is quantified. Our definition also handles the cases where one object is significantly spatially extended relative to the other by taking spatial proximity into consideration. Similarly, the satisfaction of this ternary relation relative to two reference objects by a target object is computed by integrating the corresponding landscape.

Our main contributions in this report are the flexible definitions for the directional structuring elements and efficient morphological formulation of the directional and between spatial relationships with support for visibility and extended objects. The rest of the report is organized as follows. Symbol conventions used in the report are listed in Section 2. Directional spatial relationships and between relationship are described in Sections 3 and 4, respectively. The computation of the degree of satisfaction of a relationship by a target object relative to the reference object(s) is discussed in Section 5. The proposed methods are illustrated and compared to other techniques using synthetic images and real satellite scenes in Section 6. Conclusions are given in Section 7. Appendices provide details of mathematical functions used in morphological processing.

## 2 Symbol convention

In this document, following symbol conventions are used:

$A$ Capital letters represent set of pixels in an object/region.

$A^c$ Represents complement of the set of pixels in an object/region.

$a$ Small letters represents a point in a region ($a \in A$).

$f(x)$ $f$ is the image function from the Euclidean 2-space into $[0,1]$. $f(x)$ is the value of this function at the point $x$. For crisp objects, $f(x) \in \{0,1\}$; for fuzzy objects, $f(x) \in [0,1]$.

$\overrightarrow{xy}$ A vector from point $x$ to point $y$.

$\alpha$ An angle value. In particular, it represents the direction of relationship between two regions.

$\beta$ A fuzzy landscape calculated using reference object(s) for a given type of relationship. Each point is assigned a value which represents the degree of satisfaction for the point according to the reference object(s) and the type of relationship. It is considered as an image having pixel values in $[0,1]$.

$\mu(A)$ Satisfaction degree of a relationship for the target object $A$ on a given landscape.

$area(A)$ Number of pixels in a crisp object $A$.

$\oplus$ Morphological dilation.

$\ominus$ Morphological erosion.

$\nu$ Structuring element. Similar to $f$, it is a function from the Euclidean 2-space into $[0,1]$. If it is a fuzzy structuring element, fuzzy mathematical morphology is used.

Definitions for crisp and fuzzy objects are available in the literature. However, only crisp objects are considered in this report.

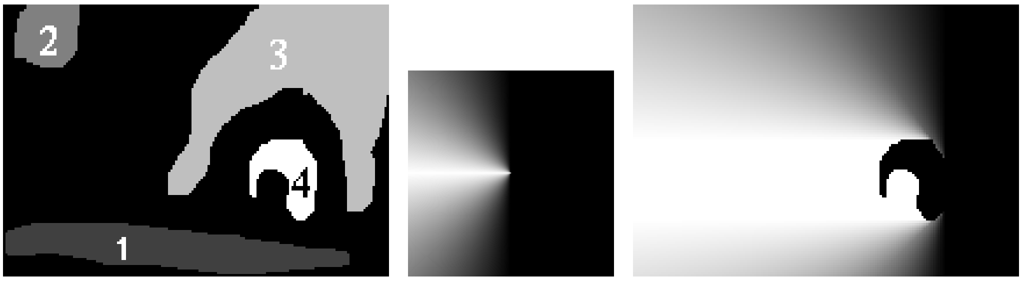
## 3 Directional spatial relationships

Directional relationships describe the spatial arrangement of two objects relative to each other. Although, it is a common approach to use right (east), left (west), above (north), and below (south) as the directions, for modeling purposes it is more convenient and generalizable to use angle-based definition of these relations where it is possible to calculate the degree of satisfaction of the relation for a given angle.

Given a reference object $B$ and a direction specified by the angle $\alpha$, our goal in this work is to generate a landscape in which the degree of satisfaction of the directional relationship at each image area relative to the reference object is quantified. Then, given a second object, its relation to the reference object can be measured using this landscape. The landscape will be denoted by $\beta_\alpha(B)$ in the rest of the report.

### 3.1 Morphological approach

The landscape $\beta_\alpha(B)$ around a reference object $B$ along the direction specified by the angle $\alpha$ 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 under examination where points in areas that satisfy the directional relation with a high degree have high membership values.




(a) Synthetic image (b) Structuring element $\nu_\alpha$ for $\alpha = \pi$ (c) Landscape of region 4 for $\alpha = \pi$

Figure 1: An example synthetic image and the directional landscape $\beta_\alpha$ for one of the regions using the structuring element $\nu_\alpha$ defined in (5).

This relationship can be defined in terms of the angle between the vector from a point in the reference object to a point in the image and the unit vector along the direction $\alpha$ measured with respect to the $x$-axis. 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 $\alpha$. Then, the value of the fuzzy landscape at an image point can be computed in terms of the smallest angle using a decreasing function $h : [0, \pi] \rightarrow [0, 1]$ as

$$\beta_\alpha(B)(x) = h\left(\min_{b \in B} \theta_\alpha(x, b)\right) \tag{1}$$

where $b$ is a point in $B$ and $x$ is a point in the image. $\theta_\alpha(x, b)$ is the angle between the vector $\overrightarrow{bx}$ and the unit vector $\vec{u}_\alpha = (\cos\alpha, \sin\alpha)^T$ along $\alpha$, and can be computed as

$$\theta_\alpha(x, b) = \begin{cases} \arccos\left(\frac{\overrightarrow{bx} \cdot \vec{u}_\alpha}{\|\overrightarrow{bx}\|}\right) & \text{if } x \neq b, \\ 0 & \text{if } x = b. \end{cases} \tag{2}$$

Bloch [6] used a linear function

$$h(\theta) = \max\left\{0, 1 - \frac{2\theta}{\pi}\right\} \tag{3}$$

for (1). It can be shown that this is equivalent to the morphological dilation of $B$,

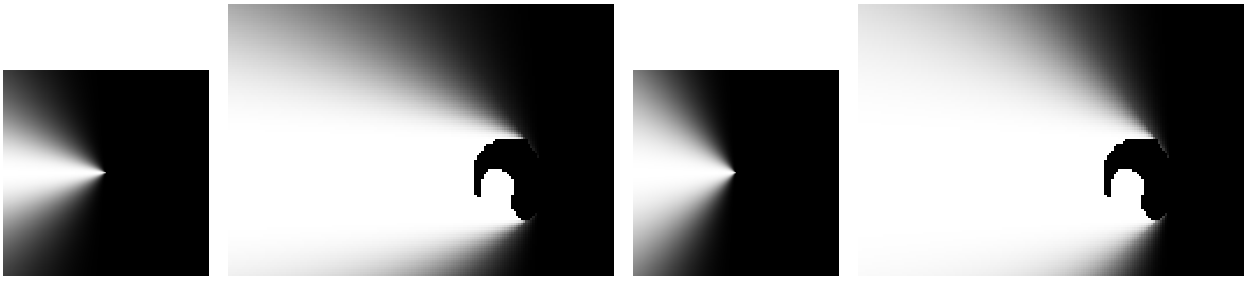
$$\beta_\alpha(B)(x) = (B \oplus \nu_\alpha)(x) \cap B^c, \tag{4}$$

using the fuzzy structuring element

$$\nu_\alpha(x) = \max\left\{0, 1 - \frac{2}{\pi}\theta_\alpha(x, o)\right\} \tag{5}$$

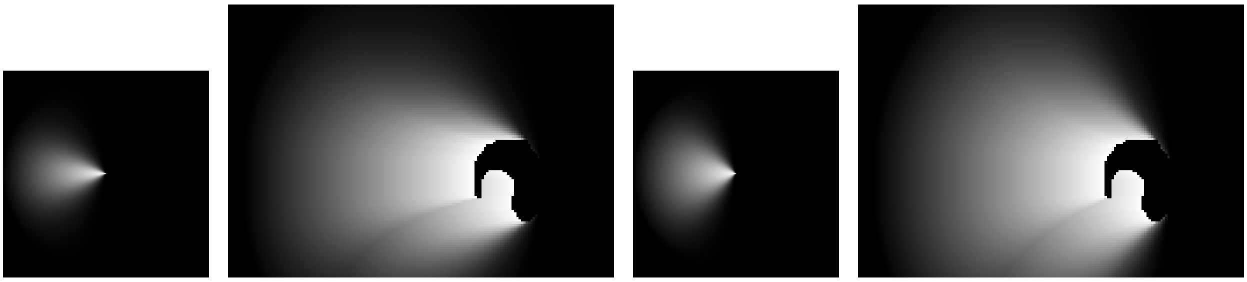
where $o$ is the origin (center) of the structuring element (see Appendix A for the definition of fuzzy morphological dilation) and $B$ is removed from the result of dilation in (4) ($c$ represents complement). An example synthetic image and fuzzy landscape examples using morphological dilation are given in Figure 1. In all figures in this report, white represents binary 1, black represents binary 0, and gray values represent the fuzziness in the range $[0, 1]$.

However, the linear function in (3) and the corresponding structuring element in (5) may not be feasible for many cases (see Section 6 for examples). Instead of using linearly decreasing membership



(a) Structuring element for $\lambda = 0.3$ (b) Landscape for $\lambda = 0.3$ (c) Structuring element for $\lambda = 0.5$ (d) Landscape for $\lambda = 0.5$

Figure 2: Structuring element $\nu_{\alpha,\lambda}$ defined in (6) and directional landscape $\beta_{\alpha,\lambda}$ of region 4 for $\alpha = \pi$.



(a) Structuring element for $\lambda = 0.3$ (b) Landscape for $\lambda = 0.3$ (c) Structuring element for $\lambda = 0.5$ (d) Landscape for $\lambda = 0.5$

Figure 3: Structuring element $\nu_{\alpha,\lambda,\tau}$ defined in (7) and directional landscape $\beta_{\alpha,\lambda,\tau}$ of region 4 for $\alpha = \pi$ and $\tau = 100$.

values according to the angle, we propose a more intuitive and flexible structuring element using a nonlinear function with the shape of a Bézier curve:

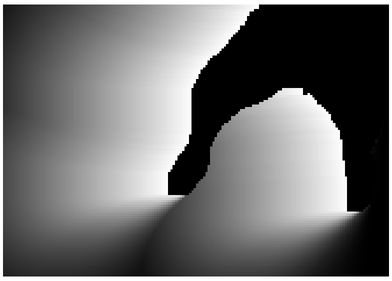
$$\nu_{\alpha,\lambda}(x) = g_\lambda\left(\frac{2}{\pi}\theta_\alpha(x,o)\right) \tag{6}$$

where $\lambda$ determines the inflection point of the curve (see Appendix B for more details) and the nonlinear function enables different definitions of fuzziness for different cases. Fuzzy landscape examples using this structuring element definition are given in Figure 2.

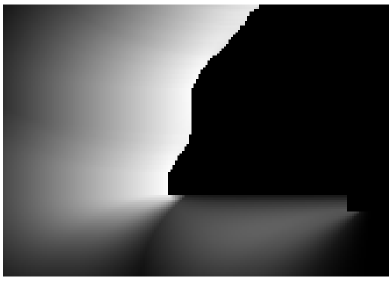
The definition of the structuring element can be further extended to decrease the degree of a point's spatial relation to a reference object according to its distance to that object by introducing a new term

$$\nu_{\alpha,\lambda,\tau}(x) = g_\lambda\left(\frac{2}{\pi}\theta_\alpha(x,o)\right)\ \max\left\{0, 1 - \frac{\|\overrightarrow{ox}\|}{\tau}\right\} \tag{7}$$

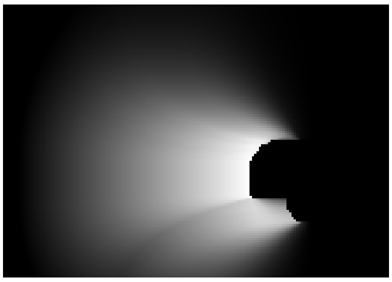
where $\|\overrightarrow{ox}\|$ is the Euclidean distance of point $x$ from the structuring element's center. In this definition, a point's spatial relation to the reference object decreases linearly with its distance to the object with $\tau$ corresponding to the distance where a point is no longer visible from the reference object. This definition also has a computational advantage because in the previous definitions the structuring element must be at least twice as large as the landscape of interest in the image space whereas in definition (7) a structuring element with size of at most $2\tau \times 2\tau$ is sufficient, leading to dramatical improvements in the efficiency of the algorithm. Fuzzy landscape examples using this structuring element definition are given in Figure 3.



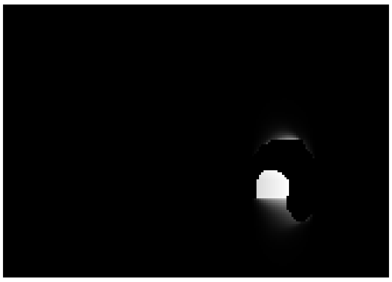
(a) Landscape of region 3 without visibility



(b) Landscape of region 3 with visibility for $\lambda' = 0.001$



(c) Landscape of region 4 with visibility for $\lambda' = 0.001$



(d) Difference between landscapes of region 4 with and without visibility

Figure 4: Directional landscape $\beta_{\alpha,\lambda,\lambda',\tau}$ for $\alpha = \pi$, $\lambda = 0.3$ and $\tau = 100$ with and without the visibility extension. (a) uses the structuring element definition in (7) without visibility, (b) and (c) use the definition in (8) with visibility, (d) illustrates the difference between landscapes with and without visibility.

### 3.2 Visibility

In directional dilation of (4), 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 shown in Figures 3 and 4. To overcome this problem, we propose the following definition

$$\beta_{\alpha,\lambda,\lambda',\tau}(B)(x) = (B \oplus \nu_{\alpha,\lambda,\tau})(x) \cap (B \oplus \nu_{\alpha+\pi,\lambda'})(x)^c \tag{8}$$

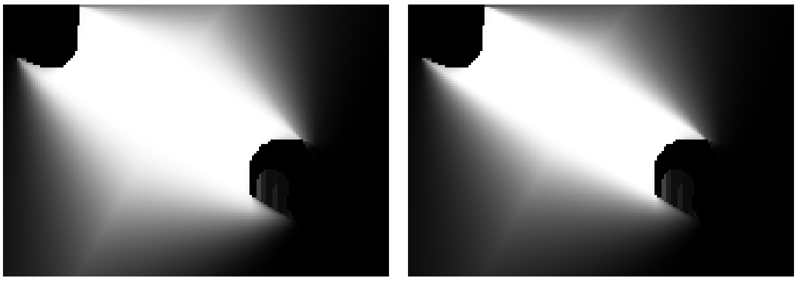
where the first dilation uses the structuring element defined in (7) and the second dilation uses the structuring element defined in (6). We compute fuzzy intersection using multiplication as the $t$-norm operator and compute fuzzy complement by subtracting the original values from 1. The proposed definition of visibility is illustrated in Figure 4.

## 4 Between relationship

Between relationship is a ternary relationship defined by two reference objects and a target object. Given two reference objects $B$ and $C$, our goal in this work is to generate a landscape in which the degree of each image area being located between the reference objects is quantified. Then, given a third object, its relation to the reference objects can be determined using this landscape. The landscape will be denoted by $\beta_{\between}(B, C)$ in the rest of the report.

### 4.1 Morphological approach

Similar to the directional spatial relationships described in Section 3.1, the landscape $\beta_{\between}(B, C)$ between two reference objects $B$ and $C$ can be defined as a fuzzy set such that image points with a high degree of the spatial relation have high membership values. This landscape can be computed as the intersection



(a) Landscape for $\lambda = 0.3$ and $\lambda' = 0.001$ (b) Landscape for $\lambda = 0.15$ and $\lambda' = 0.001$

Figure 5: Between landscape $\beta_{\between}$ of regions 2 and 4 using the definition in (11). The relative angle for these regions is $\theta_{\between} = -30.04°$.

of the directional dilations of the reference regions along the directions $\alpha = \theta_{\between}$ and $\alpha = \theta_{\between} + \pi$ where $\theta_{\between}$ is the relative position of the reference objects. This relative position can be calculated using the maximum or mean value in the histogram of angles between all pairs of points of the reference objects [8]. Using the horizontal axis as the axis of reference, the histogram of angles for the objects $B$ and $C$ can be computed as

$$h_{B,C}(\theta) = |\{(b,c)|b \in B, c \in C, \angle(\overrightarrow{bc}, \vec{u}_{\alpha=0}) = \theta\}| \tag{9}$$

and normalized as

$$H_{B,C}(\theta) = \frac{h_{B,C}(\theta)}{\max_{\theta'} h_{B,C}(\theta')}. \tag{10}$$

Then, using $\theta_{\between}$ as the relative position obtained from this histogram (as the maximum or mean value), the landscape between the reference objects $B$ and $C$ is computed as

$$\beta_{\between}(B,C)(x) = \beta_{\alpha=\theta_{\between},\lambda,\lambda'}(B)(x) \cap \beta_{\alpha=\theta_{\between}+\pi,\lambda,\lambda'}(C)(x) \tag{11}$$

where the directional landscape $\beta_{\alpha,\lambda,\lambda'}$ is computed as

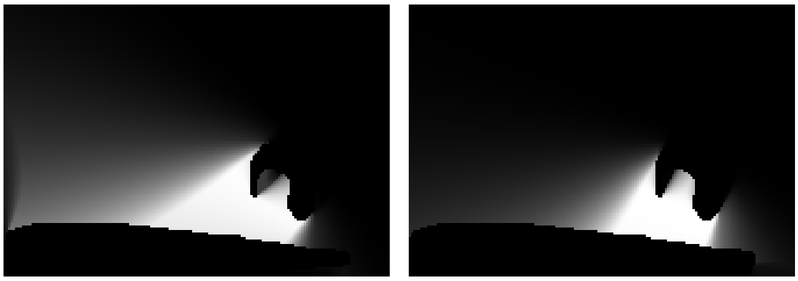
$$\beta_{\alpha,\lambda,\lambda'}(B)(x) = (B \oplus \nu_{\alpha,\lambda})(x) \cap (B \oplus \nu_{\alpha+\pi,\lambda'})(x)^c \tag{12}$$

using the structuring element definition in (6). Since the landscape should include only the areas that are visible from both reference objects, the notion of visibility defined in Section 3.2 is used in the computation. Fuzzy landscape examples for the between relationship using this definition are given in Figure 5.

## 4.2 Myopic vision

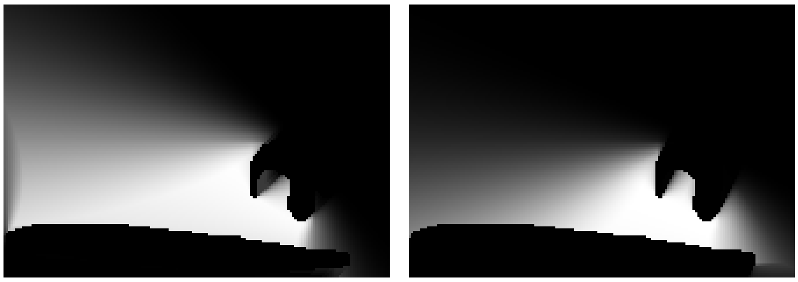
Although histogram of angles generally provides a good approximation to the relative position of two objects, it fails in the cases where one object is significantly spatially extended relative to the other [8] (see Figure 6 for examples). We propose to solve this problem by taking into account only the part of the spatially extended region close to the other region. (Bloch *et al.* [8] called this the "myopic vision" and suggested to use the distance map to find close parts of regions, approximate these parts using line segments, and apply conditional dilation to compute the landscape, but did not specify the details of the method.)

Spatial proximity for handling spatially 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 $\max\{0, 1 - \|\overrightarrow{bc}\|/\tau_{\text{myopic}}\}$ (instead of a constant weight of 1 in (9)) where $\overrightarrow{bc}$ is the Euclidean distance between the points $b$ and $c$, and $\tau_{\text{myopic}}$ 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 6 using regions 1 and 4 where region 1 is spatially extended relative to region 4.



(a) Landscape for $\lambda = 0.15$ and $\lambda' = 0.001$ without myopic vision

(b) Landscape for $\lambda = 0.15$ and $\lambda' = 0.001$ with myopic vision



(c) Landscape for $\lambda = 0.5$ and $\lambda' = 0.001$ without myopic vision

(d) Landscape for $\lambda = 0.5$ and $\lambda' = 0.001$ with myopic vision

Figure 6: Between landscape $\beta_{\between}$ of regions 1 and 4 with and without myopic vision. $\tau_{\text{myopic}}$ is taken as the half of the width of the image. The relative angles are 42.28° and 63.40° for figures without and with myopic vision, respectively. For larger values of $\lambda$, error in landscape without myopic vision becomes more significant.

# 5 Degree of satisfaction of a relationship

After calculating the landscape $\beta$ for a spatial relation as in Sections 3 or 4, the degree of satisfaction of this relation by a target object $A$ can be computed as

$$\mu(A) = \frac{1}{area(A)} \sum_{a \in A} \beta(a). \tag{13}$$

However, this definition would not be suitable in cases where the target region has a large spatial extent relative to the reference region(s) [8]. In such cases, although intuitively the target region would satisfy the relationship, because of the normalization using its area, resulting $\mu$ might be very small.

An alternative definition is proposed by Bloch *et al.* [8] for the between relationship where the target region has a large spatial extent as follows: $core(\beta_{\between})$ is the area of the landscape where membership values are 1, and $supp(\beta_{\between})$ is its whole support. If the reference regions $B$ and $C$ are not connected to each other, then $supp(\beta_{\between}) \setminus core(\beta_{\between})$ has two connected components, which are denoted by $R_1$ and $R_2$. Then, the satisfaction degree can be defined as the intersection of the target region $A$ with both $R_1$ and $R_2$, such that

$$\mu'(A) = \min \left\{ \sup_{x \in R_1} (A \cap \beta_{\between}^c)(x), \sup_{x \in R_2} (A \cap \beta_{\between}^c)(x) \right\}. \tag{14}$$

However, if $A \cap core(\beta_{\between})^c = \emptyset$, then the relationship degree should be zero. Moreover, the following special cases should be handled separately:

- If neither $R_1$ nor $R_2$ exist, the previous $\mu(A)$ definition in (13) should be used.

- If only one of $R_1$ or $R_2$ exists, the degree of satisfaction should be set to 0 or to the normalized intersection value restricted to the points of $A$ having nonzero intersection.

A second alternative suggested by Bloch *et al.* [8] for the between relationship is to use the maximum $\beta_{\between}$ membership value at points of $A$ along the direction orthogonal to the relative orientation as the satisfaction degree.

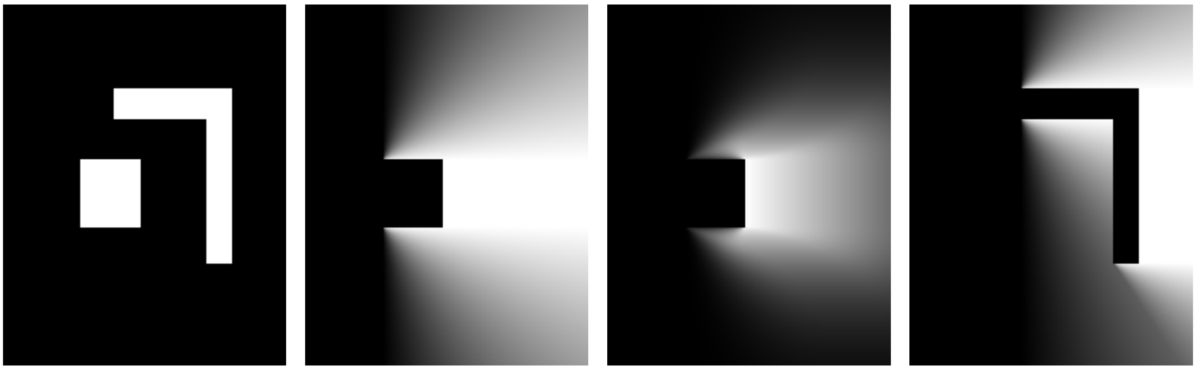
## 6 Illustrative examples

In this section, comparisons with other major methods for binary directional and ternary between relationships are presented with numerical and graphical results using both synthetic and real images.

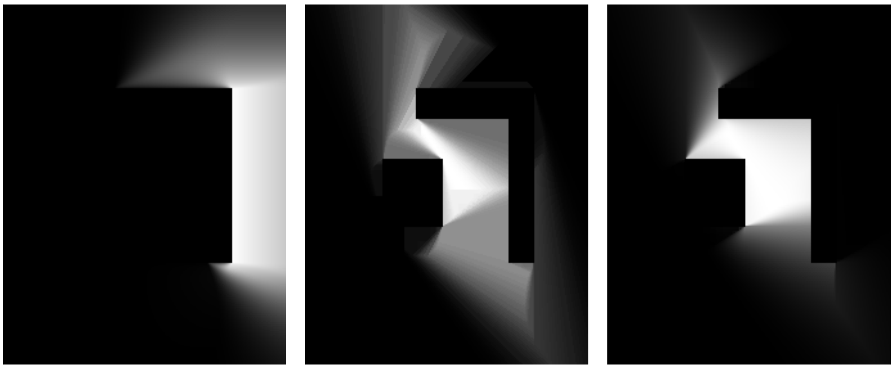
Figure 7 shows directional and between landscapes of regions using Bloch's directional landscape definition in [6] ((5) in this report), Bloch *et al.*'s between landscape definition (17) in [8], and our definitions (8) and (11) on a synthetic image that was also used in [6, 9, 7, 8]. Figures 7(b) and 7(c) illustrate the differences between the fuzzy landscapes obtained using Bloch's and our structuring element definitions. The latter is sensitive to the distance to the object according to the constant $\tau$ and the landscape's fuzziness is more centralized along the main direction of interest by the help of the constant $\lambda$. Figures 7(d) and 7(e) present the importance of the support for visibility in our definition for directional relationships. Although both landscapes for the direction "right" have similar distributions to the right and above of the reference region, the first one also has nonzero values on the left of the region, which contradicts the intuition. Figures 7(f) and 7(g) shows the differences in the definition of between. The first landscape, which is generated according to the definition in [8], is spatially too extended in the upper and lower parts of the image. It also includes non-smooth transitions that contradict the intuition. On the other hand, the second landscape is more compact and fully covers the expected between area.

In Tables 1, 2 and 3, experimental statistics using the synthetic image in Figure 1(a) are given. For finding objects' satisfaction of the specified relationships, definition (13), that returns a value in the interval $[0, 1]$ is used. For landscapes calculated using our definitions, constants are set as: $\lambda = 0.3$, $\lambda' = 0.001$, $\tau = 150$. As the $t$-norm operator, minimum is used in all definitions, except for visibility in directional relationships where multiplication is used as suggested in Section 3.2.

- Table 1 presents directional relationship satisfaction degrees of all object pairs in the directions left, right, above and below, where $\alpha$ value corresponds to $\pi$, 0, $\pi/2$ and $-\pi/2$, respectively. Results for three different methods are given: centroid-based method using the definitions in [9] and the cosine fuzzy membership functions defined in [4]; Bloch's morphological directional landscape definition in [6] ((5) in this report); and the proposed morphological definition (8). It is worth noting that, in all rows, as regions in relation get further away from each other, relationship degree decreases in our definition. This is the result of the proposed metric information in directional relationships, which depends on the value of $\tau$. However, it should also be noted that our flexible definition allows to create adequate structuring elements according to the context. Some rows of the table are also worth mentioning. For reference region 1 and target region 4, our method decides that 4 is mostly above 1. (According to our definition, above relation is the highest with 0.79, which is followed by the right relation with 0.40.) 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 along a wide angular range 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. Specifically, for reference region 3 and target region 4, centroid-based method and our definition perform similarly by deciding that region 4 is below region 3. However, according to Bloch's structuring element definition, region 4 is also at the left (1.00), right (0.99) and above (0.34) of region 3. Although above relation can be considered as an error, right and left relations may be acceptable based on application requirements. In fact, by not using visibility in our definition (i.e., using



(a) Synthetic image (b) Directional landscape of $A$ for $\alpha = 0$ using Bloch's definition (5) (c) Directional landscape of $A$ for $\alpha = 0$ using our definition (8) (d) Directional landscape of $B$ for $\alpha = 0$ using Bloch's definition (5)



(e) Directional landscape of $B$ for $\alpha = 0$ using our definition (8) (f) Between landscape calculated using Bloch *et al.*'s definition (17) in [8] (g) Between landscape using our definition (11)

Figure 7: Examples of directional and between landscapes using different definitions on a synthetic image that was also used in [6, 9, 7, 8]. The square object is denoted by $A$ and the other object by $B$. The between landscape in (f) is calculated using dilation by a structuring element derived from the histogram of angles as defined in (17) in [8]. For landscapes calculated using our definitions given in text, constants are set as follows: $\lambda = 0.3$, $\lambda' = 0.001$, $\tau = 200$.

$\beta_{\alpha,\lambda,\tau}$ instead of $\beta_{\alpha,\lambda,\lambda',\tau}$), it is possible to obtain the same behavior. Rest of the cases give similar results for all methods.

- Table 2 presents the relative angles (in degrees) for all object pairs using three methods: centroid-based as described in Section 1; mean angle obtained from the histogram of angles defined in (10); and mean angle obtained from the histogram of angles computed using myopic vision as described in Section 4.2. In relative angle values computed using myopic vision, Inf represents that objects under consideration are too distant to be related. The Inf threshold directly depends on the constant $\tau_{\text{myopic}}$ used in histogram calculation. This behavior of myopic vision is an advantage of the proposed method because it also identifies the reference object pairs where the between relationship calculation is meaningless (and computationally expensive). For all cases, our myopic vision definition gives more intuitive results. For example, for relative degrees between objects (regions) 1-2, 1-3 and 1-4, myopic vision returns a degree closer to the expected direction north (90°).

- Table 3 presents the between relationship satisfaction degrees for all reference and target object triplets using two methods: Bloch *et al.*'s between landscape definition (17) in [8] and the proposed definition (11) where the histogram of angles with myopic vision is used to calculate the relative angles between the reference objects. In some rows of the table, there are significant differences between the results of the two methods. For reference objects 1 and 2, results are similar for target object 3; however, for target object 4, our definition can be considered as performing better intuitively (see Figures 8(a) and 8(b) for the corresponding landscapes). We can intuitively say that object 4 is between 1 and 3 more than it is between 1 and 2. Our definition gives 0.95 for the former, whereas Bloch *et al.*'s method returns 0.77. Visibility in the between relationship is also important. We can see that object 4 is not between 2 and 3, and 2 is not between 3 and 4. (For example, a person at 2 cannot see a person at 4 in any way.) However, Bloch *et al.*'s definition erroneously gives 0.41 for the former case and 0.27 for the latter but our definition results in almost 0 for both cases (see Figures 8(c)–8(f) for the corresponding landscapes). In all of the above cases, our results are much closer to expectations than the results of the method proposed in [8]. The case for reference objects 2 and 4 illustrates the effect of $\tau_{\text{myopic}}$ in myopic vision. Since these objects are reported to be too far from each other in Table 2, the degree of the between relationship for any target object is 0 without any further computation. For target object 3, Bloch *et al.*'s method gives 0.16. Our definition can be adjusted to give a similar result by increasing $\tau_{\text{myopic}}$. Finally, one particular case we would like to note is the reference objects 2 and 3 where our method gives 0.23 for the satisfaction degree for target object 1. This is because the myopic vision effect in the histogram of angles reports object 3 to be more below object 2 than the histogram of angles without myopic vision does, which causes the creation of the between landscape closer to object 1 (see Figures 8(c) and 8(d) for the landscapes). Overall, the results obtained by the definitions proposed in this report are much closer to expectations than the results of the method proposed in [8].

Finally, Figure 9 shows a LANDSAT scene of British Columbia in Canada and its segmentation using the method in [2]. We present two example scenarios:

- Figure 10 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. The between landscape for two water regions is successfully constructed so that the integration of this landscape over the region classified as asphalt/concrete results in the detection of the bridge with a high confidence. Note that, a bridge cannot be detected by conventional methods that look only at pixels or individual regions.

- Figure 11 illustrates the scenario of finding the fields to the north (above) of a river (water). The directional landscape without visibility in Figure 11(a) erroneously covers some areas that

Table 1: Satisfaction degrees of directional relationships for all object pairs in the synthetic image in Figure 1(a). Ids of reference objects and target objects are given in the first two columns, respectively.

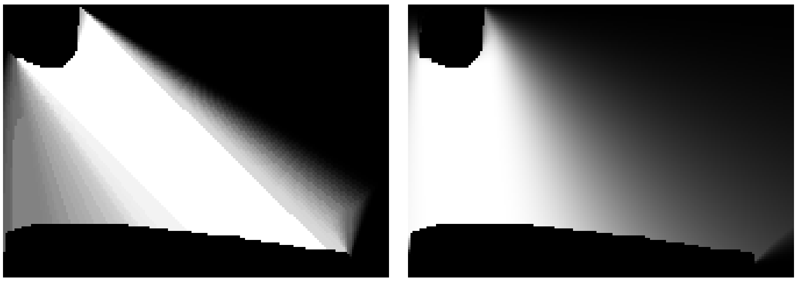
| | | Centroid-based | | | | Bloch's definition (5) | | | | Our definition (8) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ref. | Target | left | right | above | below | left | right | above | below | left | right | above | below |
| 1 | 2 | 0.24 | 0.00 | 0.76 | 0.00 | 0.60 | 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 |
| 2 | 1 | 0.00 | 0.24 | 0.00 | 0.76 | 0.02 | 0.39 | 0.00 | 0.78 | 0.00 | 0.04 | 0.00 | 0.26 |
| 2 | 3 | 0.00 | 0.96 | 0.00 | 0.04 | 0.00 | 0.92 | 0.03 | 0.23 | 0.00 | 0.30 | 0.00 | 0.03 |
| 2 | 4 | 0.00 | 0.75 | 0.00 | 0.25 | 0.00 | 0.74 | 0.00 | 0.43 | 0.00 | 0.18 | 0.00 | 0.04 |
| 3 | 1 | 0.38 | 0.00 | 0.00 | 0.62 | 0.82 | 0.24 | 0.00 | 0.83 | 0.39 | 0.08 | 0.00 | 0.49 |
| 3 | 2 | 0.96 | 0.00 | 0.04 | 0.00 | 1.00 | 0.00 | 0.59 | 0.09 | 0.50 | 0.00 | 0.16 | 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 |
| 4 | 1 | 0.72 | 0.00 | 0.00 | 0.28 | 0.75 | 0.07 | 0.00 | 0.59 | 0.40 | 0.02 | 0.00 | 0.34 |
| 4 | 2 | 0.75 | 0.00 | 0.25 | 0.00 | 0.75 | 0.00 | 0.43 | 0.00 | 0.19 | 0.00 | 0.04 | 0.00 |
| 4 | 3 | 0.00 | 0.01 | 0.99 | 0.00 | 0.25 | 0.36 | 0.79 | 0.04 | 0.14 | 0.19 | 0.52 | 0.01 |

Table 2: Relative angles (in degrees) between all object pairs in the synthetic image in Figure 1(a). Ids of the object pairs for which the angle is computed are given in the first two columns.

| Obj.1 | Obj.2 | Centroid | Hist. of angles | Hist. of angles with myopic vision |
|---|---|---|---|---|
| 1 | 2 | 119.25 | 115.98 | 93.98 |
| 1 | 3 | 51.70 | 56.70 | 74.93 |
| 1 | 4 | 31.88 | 42.29 | 63.41 |
| 2 | 3 | -10.99 | -12.03 | -21.97 |
| 2 | 4 | -29.92 | -30.04 | Inf |
| 3 | 4 | -96.01 | -80.13 | -73.13 |

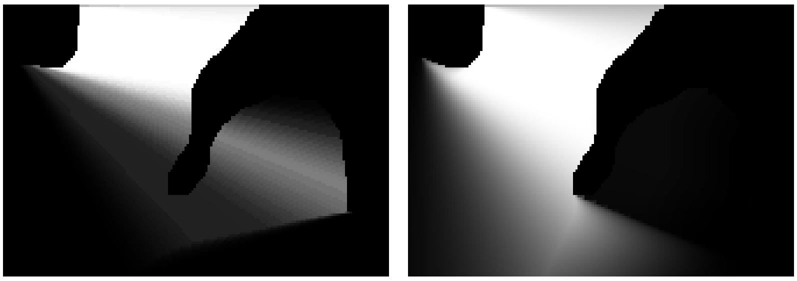
Table 3: Satisfaction degrees of between relationship for object triplets in the synthetic image in Figure 1(a). Ids of reference objects are given in the first two columns and the target object is given in the third column.

| Ref.1 | Ref.2 | Target | Bloch *et al.*'s definition (17) in [8] | Our definition (11) |
|---|---|---|---|---|
| 1 | 2 | 3 | 0.12 | 0.10 |
| 1 | 2 | 4 | 0.52 | 0.22 |
| 1 | 3 | 2 | 0.09 | 0.05 |
| 1 | 3 | 4 | 0.77 | 0.95 |
| 1 | 4 | 2 | 0.00 | 0.00 |
| 1 | 4 | 3 | 0.06 | 0.02 |
| 2 | 3 | 1 | 0.01 | 0.23 |
| 2 | 3 | 4 | 0.41 | 0.02 |
| 2 | 4 | 1 | 0.00 | 0.00 |
| 2 | 4 | 3 | 0.16 | 0.00 |
| 3 | 4 | 1 | 0.09 | 0.01 |
| 3 | 4 | 2 | 0.27 | 0.00 |



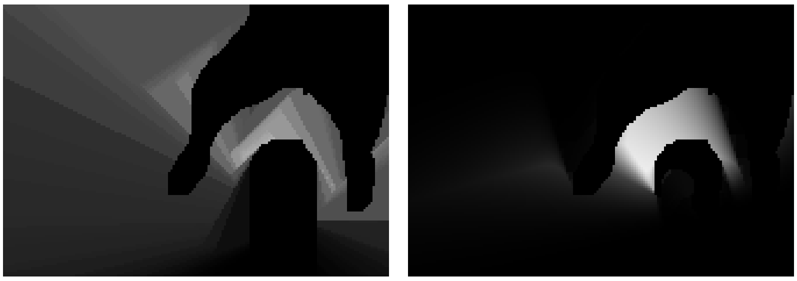
(a) Between landscape for objects 1 and 2 calculated using Bloch *et al.*'s definition (17) in [8]

(b) Between landscape for objects 1 and 2 using our definition (11)



(c) Between landscape for objects 2 and 3 calculated using Bloch *et al.*'s definition (17) in [8]

(d) Between landscape for objects 2 and 3 using our definition (11)



(e) Between landscape for objects 3 and 4 calculated using Bloch *et al.*'s definition (17) in [8]

(f) Between landscape for objects 3 and 4 using our definition (11)

Figure 8: Examples of between landscapes used in Table 3.

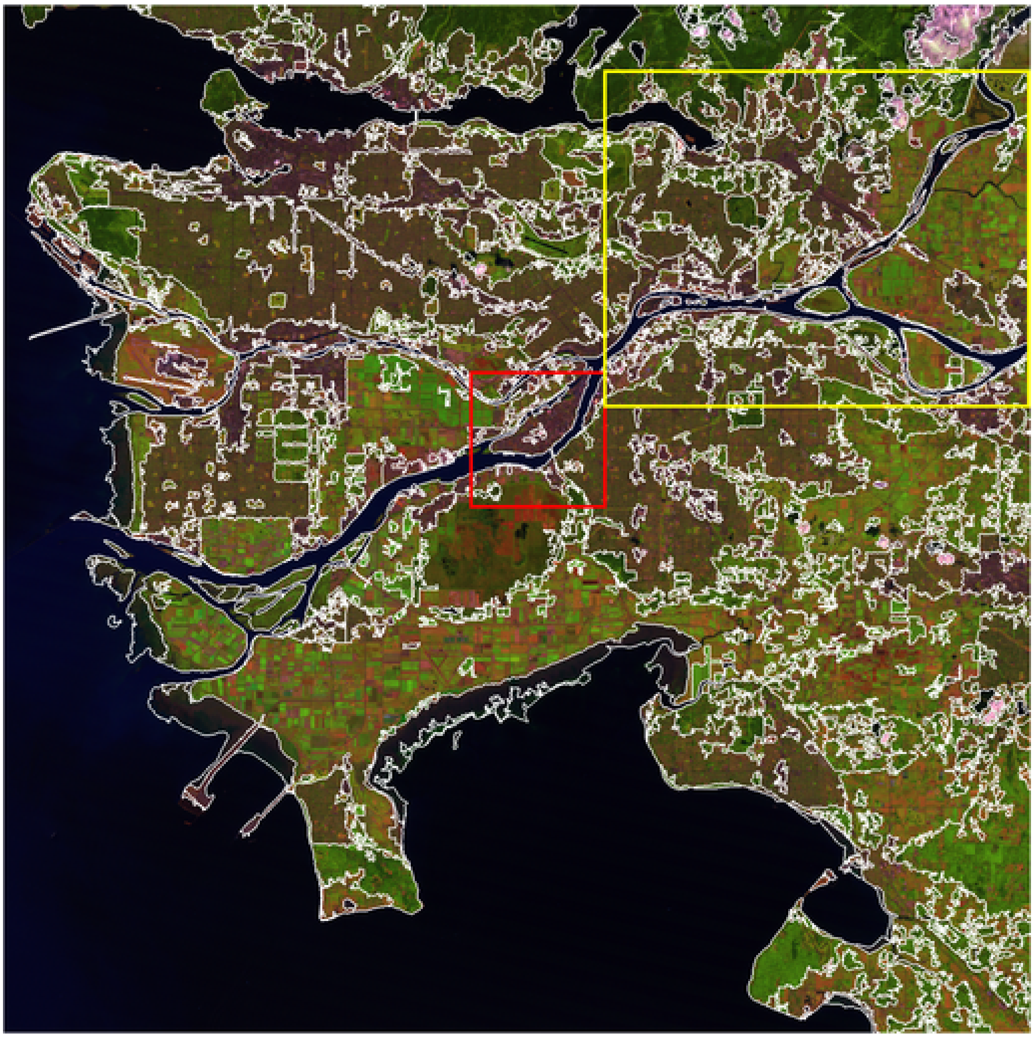
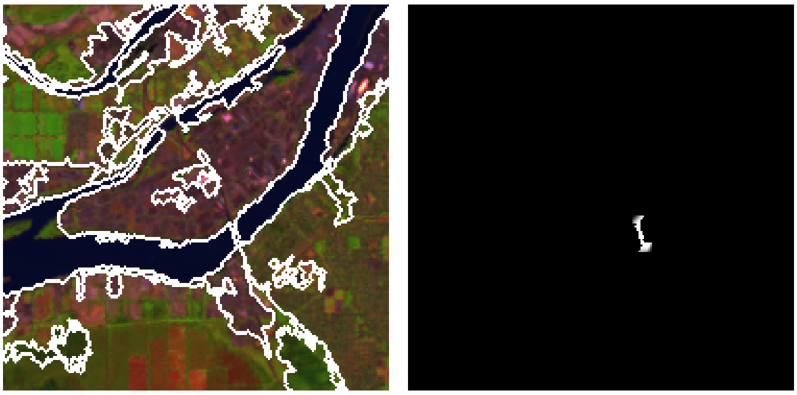


Figure 9: LANDSAT scene of British Columbia in Canada with its segmentation overlaid.

are to the south of the river. Introducing visibility using the structuring element in (7) with $\alpha = \pi/2$, $\lambda = 0.5$, $\tau = 150$ for the first dilation in (8) and $\alpha = -\pi/2$, $\lambda = 0.001$, $\tau = 100$ for the second dilation in (8) produces the landscape in Figure 11(b) where areas with water regions closer to them from below than above have high membership values for the "field above water" relationship. As an alternative, restricting the size of the structuring element to $100 \times 100$ in the second dilation in (8) gives the landscape in Figure 11(c) where areas with a water region closer than 100 pixels from above are ignored in the relationship. Figure 11(d) shows the results of restricting the size of the structuring element to $200 \times 200$.

## 7 Conclusions

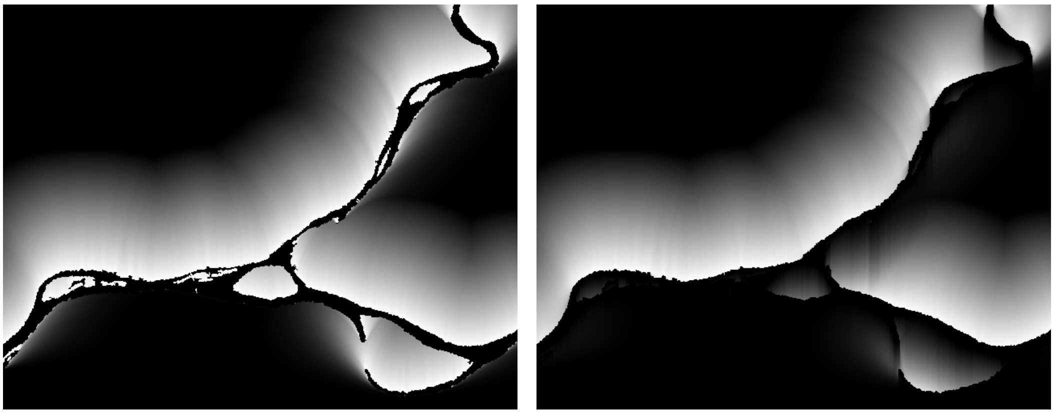
We presented new, flexible and efficient definitions for modeling binary directional relationships and the ternary between relationship using fuzzy mathematical morphology techniques. Our definitions support the notion of visibility for handling areas that are partially enclosed by objects and are not 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.



(a) Zoomed sub-image

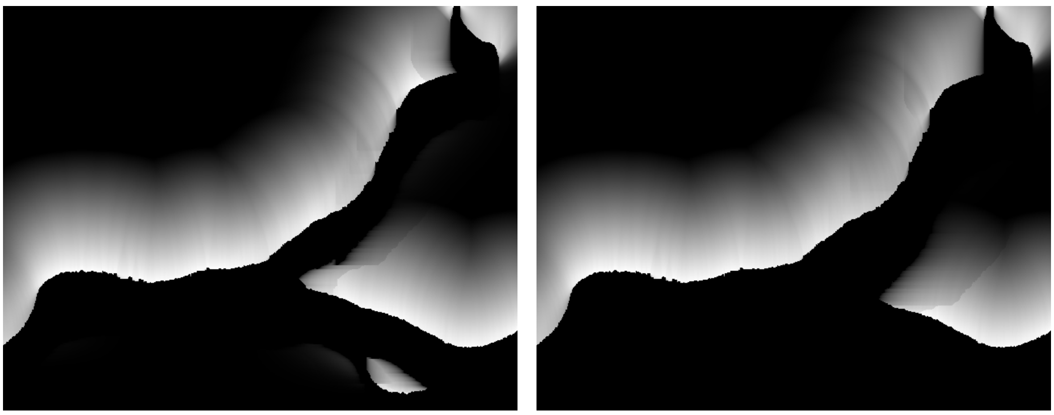
(b) Between landscape of two water regions using structuring element of size $10 \times 10$

Figure 10: Searching for bridges in the sub-image marked with a red rectangle in Figure 9 (see text for details).



(a) Without visibility

(b) With visibility using structuring element in (7)



(c) With visibility using structuring element restricted to size $100 \times 100$

(d) With visibility using structuring element restricted to size $200 \times 200$

Figure 11: Searching for fields to the north of a river in the sub-image marked with a yellow rectangle in Figure 9 (see text for details).

## References

[1] S. Aksoy. Modeling of remote sensing image content using attributed relational graphs. In *Proceedings of IAPR International Workshop on Structural and Syntactic Pattern Recognition*, pages 475–483, Hong Kong, August 17–19, 2006. Lecture Notes in Computer Science, vol. 4109.

[2] 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*, 43(3):581–589, March 2005.

[3] S. Aksoy, G. Marchisio, K. Koperski, and C. Tusk. Probabilistic retrieval with a visual grammar. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 1041–1043, Toronto, Canada, June 24–28, 2002.

[4] S. Aksoy, C. Tusk, K. Koperski, and G. Marchisio. Scene modeling and image mining with a visual grammar. In C. H. Chen, editor, *Frontiers of Remote Sensing Information Processing*, pages 35–62. World Scientific, 2003.

[5] S. Berretti, A. Del Bimbo, and E. Vicario. Modelling spatial relationships between colour clusters. *Pattern Analysis & Applications*, 4(2/3):83–92, 2001.

[6] I. Bloch. Fuzzy relative position between objects in image processing: A morphological approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):657–664, July 1999.

[7] I. Bloch. Fuzzy spatial relationships for image processing and interpretation: A review. *Image and Vision Computing*, 23(2):89–110, February 2005.

[8] I. Bloch, O. Colliot, and R. M. Cesar. On the ternary spatial relation "between". *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, 36(2):312–327, April 2006.

[9] I. Bloch and A. Ralescu. Directional relative position between objects in image processing: A comparison between fuzzy approaches. *Pattern Recognition*, 36(7):1563–1582, July 2003.

[10] W. W. Chu, C.-C. Hsu, A. F. Cardenas, and R. K. Taira. Knowledge-based image retrieval with spatial and temporal constructs. *IEEE Transactions on Knowledge and Data Engineering*, 10(6):872–888, November/December 1998.

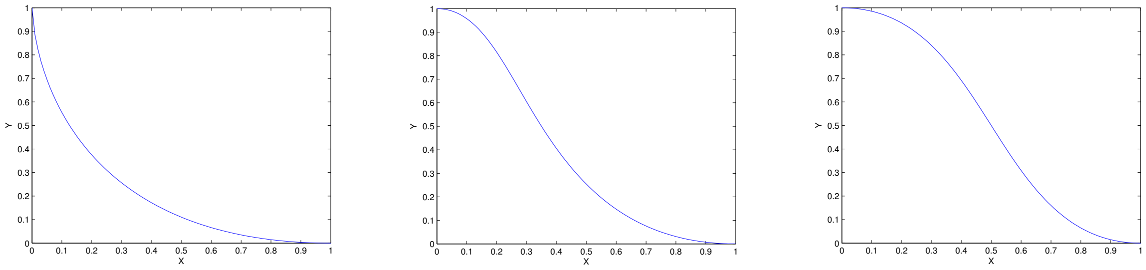
[11] P. Matsakis and L. Wendling. A new way to represent the relative position between areal objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):634–643, July 1999.

[12] P. J. Neal, L. G. Shapiro, and C. Rosse. The digital anatomist structural abstraction: A scheme for the spatial description of anatomical entities. In *Proceedings of American Medical Informatics Association Annual Symposium*, Lake Buena Vista, FL, November 1998.

[13] E. G. M. Petrakis, C. Faloutsos, and K.-I. Lin. Imagemap: An image indexing method based on spatial similarity. *IEEE Transactions on Knowledge and Data Engineering*, 14(5):979–987, September/October 2002.

[14] J. R. Smith and S.-F. Chang. VisualSEEk: A fully automated content-based image query system. In *Proceedings of ACM International Conference on Multimedia*, pages 87–98, Boston, MA, November 1996.

[15] H. L. Tang, R. Hanka, and H. H. S. Ip. Histological image retrieval based on semantic content analysis. *IEEE Transactions on Information Technology in Biomedicine*, 7(1):26–36, March 2003.



(a) $\lambda = 0.001$ (b) $\lambda = 0.3$ (c) $\lambda = 0.5$

Figure 12: One-dimensional functions $g_\lambda(x)$ with the shape of a cubic Bézier curve and a single parameter $\lambda$ (see text for details).

# A Fuzzy mathematical morphology

Dilation of an object $A$ with a fuzzy structuring element $\nu$ is defined as [6]

$$(A \oplus \nu)(x) = \max_y \{t[f(y), \nu(x-y)]\} \tag{15}$$

where $f$ is the function representing region $A$, $\nu$ is the structuring element, $t$ is a $t$-norm operator for fuzzy intersection, and $y$ is taken over all points in the image.

This is equal to

$$(A \oplus \nu)(x) = \max_y \{t[f(x-y), \nu(y)]\} \tag{16}$$

when it is assumed that the pixels outside the boundaries of the structuring element are zero.

It should be noted that fuzzy dilation is similar to dilation using a non-flat structuring element.

# B Bézier curves

Bézier curve is a parametric curve defined using a number of reference points. Four points $a_0, a_1, a_2, a_3$ on a plane define a cubic Bézier curve where the curve starts at $a_0$ going toward $a_1$ and arrives at $a_3$ coming from the direction of $a_2$. The parametric form of the curve is

$$b(t) = (1-t)^3 a_0 + 3t(1-t)^2 a_1 + 3t^2(1-t)a_2 + t^3 a_3 \tag{17}$$

where $t$ is the parameter having values in $[0, 1]$.

To construct a one-dimensional function that has the shape of a Bézier curve and maps each $x \in [0, 1]$ to a $y \in [0, 1]$, we set the reference points as

$$a_0 = (0, 1),\ a_1 = (\lambda, 1),\ a_2 = (\lambda, 0),\ a_3 = (1, 0) \tag{18}$$

where $\lambda \in (0, 1)$ so that the cubic curve has only one parameter. Then, equation (17) reduces to

$$b_x(t) = 3t(1-t)^2\lambda + 3t^2(1-t)\lambda + t^3 \tag{19}$$
$$b_y(t) = (1-t)^3 + 3t(1-t)^2 + 3t^2(1-t) \tag{20}$$

and for any $x \in [0, 1]$, $b_x(t)$ can be solved for $t$, and the corresponding $y \in [0, 1]$ can be computed using $b_y(t)$.

In this report, this function/mapping is donated as $g_\lambda(x)$. This function has an inflection point at $x = \lambda$. Examples of $g_\lambda(x)$ for different $\lambda$ values are shown in Figure 12.