A Most Versatile Relative Position Descriptor

by

Mohammad Naeem

A Thesis
presented to
The University of Guelph

In partial fulfilment of the requirements
for the degree of
Doctor of Philosophy
in
Computer Science

Guelph, Ontario, Canada
© Mohammad Naeem, May, 2016
ABSTRACT

A Most Versatile Relative Position Descriptor

Mohammad Naeem
University of Guelph, 2016

Adviser: Pascal Matsakis

In daily conversation, the relative positions of objects in space are described in terms of spatial relationships: topological (e.g., “is inside”), directional (e.g., “is above”) and distance (e.g., “is far from”) relationships. A relative position descriptor is a quantitative representation of the relative position of two spatial objects, and it is often used as a basis from which models of spatial relationships can be derived. Like colour, texture, and shape descriptors, it is a visual descriptor. Various relative position descriptors have been proposed, and they have found a variety of applications (e.g., graphical symbol retrieval, linguistic scene description, human-robot interaction, map-to-image conflation, land cover classification). In this thesis, we introduce a new relative position descriptor: the Φ-descriptor. It is a fast-to-compute, property-loaded tool that has many advantages over its competitors. Our approach borrows ideas from the Radon transform and Allen’s interval algebra. It is based on the concept of the F-histogram and on an original categorization of pairs of consecutive boundary points on a line. A spatial relationship is usually modeled either as a crisp relation, i.e., as a relation, in the standard mathematical sense, or as a fuzzy relation, which is a concept in fuzzy set theory. We show here that the Φ-descriptor can be used to develop crisp and fuzzy models of a large number of relationships. For example, the well-known RCC8 and DE-4IM topological relations are
definable in terms of the descriptor; the RCC8 relations can be fuzzified based on the
descriptor; the descriptor can be used to model directional relationships, including visual
surroundedness.

Keywords: Spatial relationships, relative position, relative position descriptor, visual
surroundedness.
Dedicated to my wife Qanitah Hussain, daughter Ayisha Muska, sons, Hassan Baryal and Zarak Yahya, brother Riaz Ahmad, and nephew, Dr. Nasar Ahmad for their patience and support in various ways to make me feel free and relieved of worries to continue working on the thesis.
Acknowledgements

My first and foremost acknowledgements go to Professor Pascal Matsakis who provided me precious guidance and advice in my work. Acknowledgements are also due to the member of my advisory committee, Herb Kunze, who gave me very helpful suggestions on my thesis and was available to review my work whenever I requested him. Thanks and acknowledgement also to Dr. Charlie Obimbo for his review of this work. Last, my thanks go to Professor William Gardener, Professor David Calvert, and Professor David Chiu of the School of Computer Science, University of Guelph, who were accessible and courteous to provide me assistance at various stages of this thesis writing.
Contents

Dedication IV

Acknowledgements V

1 Introduction ......................................................................................................................... 1
  1.1 Connection with spatial relationships ................................................................. 2
  1.2 Applications of relative position descriptors ..................................................... 3
  1.3 Current state of relative position description research ........................................ 3
    1.3.1 Some expectations from relative position descriptors ............................. 4
    1.3.2 Issues with the existing descriptors ......................................................... 5
    1.3.3 Contribution of the present work ............................................................. 6

2 Literature review .................................................................................................................... 9
  2.1 Models of spatial relationships ................................................................................... 10
    2.1.1 Crisp and fuzzy relations ........................................................................... 10
    2.1.2 Topological relations ................................................................................ 13
    2.1.3 Directional relations .................................................................................. 17
    2.1.4 Distance relations ...................................................................................... 19
    2.1.5 Surroundedness ........................................................................................ 20
  2.2 Models of relative position ......................................................................................... 21
    2.2.1 Categorization ............................................................................................. 22
    2.2.2 Main approaches ......................................................................................... 23
    2.2.3 Review of existing descriptors ...................................................................... 27
    2.2.4 Comparison of properties .......................................................................... 44
3 The $\Phi$-descriptor ........................................................................................................... 48

3.1 Background .................................................................................................................... 49
  3.1.1 The Radon transform .............................................................................................. 50
  3.1.2 The F-histogram .................................................................................................... 51
  3.1.3 The Allen relations ............................................................................................... 52

3.2 The $\Phi$-descriptor ....................................................................................................... 53
  3.2.1 Defining boundary points ..................................................................................... 54
  3.2.2 Categories of boundary points .............................................................................. 55
  3.2.3 Grouping boundary points into point-pairs ....................................................... 56
  3.2.4 Defining segment relations .................................................................................. 58
  3.2.5 Defining the F-histograms ................................................................................... 59
  3.2.6 The $\Phi$-descriptor ............................................................................................. 63

3.3 Properties ...................................................................................................................... 65
  3.3.1 Handling of both raster and vector objects ......................................................... 66
  3.3.2 Extraction of spatial relationships ....................................................................... 67
  3.3.3 Histogram comparison ....................................................................................... 68
  3.3.4 Affine properties ................................................................................................. 68
  3.3.5 Other properties .................................................................................................. 70

3.4 Summary ......................................................................................................................... 71

4 Topological relations ........................................................................................................ 73
  4.1 Qualitative models based on the $\Phi$-descriptor ....................................................... 73
    4.1.1 The RCC8 relations ......................................................................................... 73

### 4.1.2 The RCC23 relations and other relations

78

### 4.2 Quantitative models based on the Φ-Descriptor

80

#### 4.2.1 Expectations

80

#### 4.2.2 Meeting the expectations

83

#### 4.2.3 Fuzzifying disjointInteriors, subset and superset

85

#### 4.2.4 Fuzzifying follows

86

#### 4.2.5 Fuzzifying starts

90

#### 4.2.6 Fuzzifying encloses and divides

91

### 4.3 Experiments

92

#### 4.3.1 RCC8 relations

93

#### 4.3.2 Convex-hull based relationships

96

### 4.4 Summary

96

### 5 Directional spatial relationships

98

#### 5.1 Directional relationships

98

##### 5.1.1 The model M

99

##### 5.1.2 The emulation models K', F0', and F2'

104

#### 5.2 Experiments

110

#### 5.3 Summary

112

### 6 The surround relationship

113

#### 6.1 Models of surroundedness from the Φ-descriptor

113

##### 6.1.1 Qualitative models of surround relationship from the Φ-descriptor

114

##### 6.1.2 Quantitative models

116

#### 6.2 Experiments

121
6.3 Summary..................................................................................................................122
7 Computation of the Φ-descriptor ...............................................................................126
  7.1 An algorithm for computing the Φ-descriptor for raster objects .......................126
  7.2 Experiments ...........................................................................................................127
    7.2.1 Model descriptors for comparison .................................................................127
    7.2.2 The test objects ..............................................................................................129
    7.2.3 Results ...........................................................................................................129
  7.3 A note on computing the Φ-descriptor for vector data .......................................131
  7.4 Summary .............................................................................................................133
8 Summary and future work ..........................................................................................134
  8.1 Conclusions ........................................................................................................135
  8.2 Future work .........................................................................................................136
  8.3 Summary .............................................................................................................138
9 References ................................................................................................................139

Appendix A .....................................................................................................................154

Appendix B .....................................................................................................................159
List of tables

Table 2-1 Some topological relations definable in terms of \textit{conn}. ........................................ 15
Table 2-2 Some relations definable in terms of \textit{conn} and \textit{conv}. ..................................... 15
Table 2-3 The RCC8 vs. RCC23 relations. ...................................................................................... 15
Table 2-4 Classification of approaches to relative position description ...................................... 26
Table 2-5 Comparison of properties: ............................................................................................. 46
Table 3-1 Semantic inverse retrieval ............................................................................................... 71
Table 4-1 Some basic $\Phi$ relations. ............................................................................................... 75
Table 4-2 The $\Phi$ relations “closest” to the RCC8 relations. .......................................................... 75
Table 4-3 Some basic $\Phi$ relations (after fine-tuning). ................................................................. 77
Table 4-4 The RCC8 relations are $\Phi$ relations (after fine-tuning). ............................................... 77
Table 4-5 Simplifying the expressions in Table 4-4. ..................................................................... 78
Table 4-6 Other basic $\Phi$ relations ............................................................................................... 78
Table 4-7 The $\Phi$ relations “closest” to OUT, PIN, IN, OUT$_i$, PIN$_i$ and IN$_i$. ....................... 79
Table 4-8 The DE-4IM relations are $\Phi$ relations. ....................................................................... 79
Table 5-1 Truth values (in hundredths) produced by three fuzzy models. ....................... 106
Table 5-2 Emulating fuzzy models of directional relationships ................................................... 108
Table 6-1 Basic $\Phi$ relations to model surroundedness ................................................................. 115
Table 6-2 A model of surroundedness. .......................................................................................... 116
# List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Spatial relationship extraction from relative position descriptors</td>
<td>2</td>
</tr>
<tr>
<td>2-1</td>
<td>Object pairs and directions</td>
<td>11</td>
</tr>
<tr>
<td>2-2</td>
<td>Boundary intersections</td>
<td>12</td>
</tr>
<tr>
<td>2-3</td>
<td>Representing relations: convexity and connectedness</td>
<td>12</td>
</tr>
<tr>
<td>2-4</td>
<td>Representing relations: non-inclusion and inequality</td>
<td>12</td>
</tr>
<tr>
<td>2-5</td>
<td>The RCC8 conceptual neighbourhood graph</td>
<td>14</td>
</tr>
<tr>
<td>2-6</td>
<td>The RCC23 DC relations</td>
<td>16</td>
</tr>
<tr>
<td>2-7</td>
<td>The DE-4IM relations</td>
<td>17</td>
</tr>
<tr>
<td>2-8</td>
<td>The directional relations</td>
<td>18</td>
</tr>
<tr>
<td>2-9</td>
<td>Visual vs. topological surroundedness</td>
<td>20</td>
</tr>
<tr>
<td>2-10</td>
<td>Point and pixels-pair based approaches</td>
<td>24</td>
</tr>
<tr>
<td>2-11</td>
<td>Aligned section and segment based approaches</td>
<td>25</td>
</tr>
<tr>
<td>2-12</td>
<td>Calculation of the histogram of angles</td>
<td>27</td>
</tr>
<tr>
<td>2-13</td>
<td>Calculation of the force histogram</td>
<td>29</td>
</tr>
<tr>
<td>2-14</td>
<td>Calculation of the R-histogram</td>
<td>32</td>
</tr>
<tr>
<td>2-15</td>
<td>Calculation of the R*-histogram</td>
<td>33</td>
</tr>
<tr>
<td>2-16</td>
<td>Calculation of the visual area histogram</td>
<td>34</td>
</tr>
<tr>
<td>2-17</td>
<td>Calculation of the spread histogram</td>
<td>36</td>
</tr>
<tr>
<td>2-18</td>
<td>Calculation of the ratio-histogram</td>
<td>37</td>
</tr>
<tr>
<td>2-19</td>
<td>Calculation of the spatial template</td>
<td>39</td>
</tr>
<tr>
<td>2-20</td>
<td>Allen relations</td>
<td>41</td>
</tr>
<tr>
<td>2-21</td>
<td>Visual depiction of $Fr$ −histograms</td>
<td>42</td>
</tr>
</tbody>
</table>
Figure 2-22 Radial line model ................................................................. 43
Figure 3-1 Principle of the Radon transform ........................................ 50
Figure 3-2 Principle of the $F$-histogram .............................................. 52
Figure 3-3 Allen relations ................................................................. 53
Figure 3-4 Interpretation of the $\Phi$-descriptor ...................................... 54
Figure 3-5 Notation and terminology .................................................... 55
Figure 3-6 Boundary points and their categories ..................................... 56
Figure 3-7 The 36 point-pair categories ............................................ 57
Figure 3-8 Nonzero values for the functions $f_t, f_o, f_c, f_u, f_f, f_l, f_s$ ........ 59
Figure 3-9 Non-zero values for the functions $f_v, f_a, f_r$ .................... 60
Figure 3-10 Non-zero values for the functions $f_w, f_e, f_d$ ................. 60
Figure 3-11 The meanings of the $\Phi$-descriptor ................................... 64
Figure 3-12 Disconnected object A surrounds B .................................. 65
Figure 3-13 Histogram visualization of the functions ........................... 65
Figure 3-14 CPU time versus image size (N) and number of reference directions (K) 67
Figure 4-1 The link between RCC8, DE-4IM, and $\Phi$ relations ............. 76
Figure 4-2 Fine-tuning the definition of the $\Phi$-descriptor .................... 76
Figure 4-3 IN and PIN vs. $\Phi\text{IN}$ and $\Phi\text{PIN}$ .................................. 79
Figure 4-4 Examples of non-RCC23 non DE-4IM topological $\Phi$ relations .. 80
Figure 4-5 Fuzzifying EC ................................................................. 82
Figure 4-6 Fuzzifying TPP ............................................................... 83
Figure 4-7 Fuzzifying FFF0. Meeting the continuity condition ................ 84
Figure 4-8 Fuzzifying the basic $\Phi$ relations ....................................... 85
Figure 4-9 The fuzzy relation *follow* ................................................................. 86
Figure 4-10 Fuzzifying *follows* (the *trails* case) .............................................. 88
Figure 4-11 Fuzzifying *follows* (the ideal case) ................................................... 88
Figure 4-12 Fuzzifying *follows* (the *overlaps* case) .......................................... 89
Figure 4-13 Fuzzifying *follows* (the general case) ............................................... 89
Figure 4-14 Fuzzifying *follows* (impact of the region of interaction) .................. 89
Figure 4-15 Fuzzifying *encloses0, encloses1* and *encloses2* ............................ 92
Figure 4-16 Experimental results for the RCC8 relations: I .................................... 94
Figure 4-17 Experimental results for the RCC8 relations: II ................................... 95
Figure 4-18 Results of quantitative models for the convex-hull based relationships .... 97
Figure 5-1 fuzzification principle for θ and θ .................................................... 101
Figure 5-2 Definition of the devaluation factor ε ..................................................... 102
Figure 5-3 Configurations for comparing fuzzy models of directional relationships .... 105
Figure 5-4 Approximating a particular angle histogram value .................................. 108
Figure 5-5 Deriving heights and widths of regions from the Φ-descriptor .................. 108
Figure 5-6 Experimental results for directional relationships ................................. 110
Figure 6-1 A definition of surround ................................................................. 115
Figure 6-2 Distinction between visual and topological surroundedness .................. 115
Figure 6-3 Fuzzifying the basic Φ relations used to model surroundedness .......... 118
Figure 6-4 Surroundedness and disconnected objects ......................................... 119
Figure 6-5 A region of interaction as seen by the model *SURROUNDED1* ............ 121
Figure 6-6 Experimental results for Surroundedness: I ........................................ 123
Figure 6-7 Experimental results for surroundedness: II ....................................... 123
Figure 6-8 Experimental results for surroundedness: III ................................................. 124
Figure 6-9 Experimental results for surroundedness: IV ............................................... 125
Figure 6-10 Experimental results for surroundedness: V ............................................... 125
Figure 7-1 Relevant raster lines and pixels ................................................................. 127
Figure 7-2 Processing times for the Φ-descriptors and skeleton F-histogram .......... 131
Figure 7-3 Processing time for the force histogram and the Φ-descriptor .. .......... 131
Figure 7-4 Partitioning of the objects ........................................................................... 132
Figure 7-5 Tuples of trapezoids sorted in direction θ .................................................. 133
List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>QSR</td>
<td>Qualitative Spatial Reasoning</td>
</tr>
<tr>
<td>RCC</td>
<td>Region Connection Calculus</td>
</tr>
<tr>
<td>RCC8, RCC23</td>
<td>8 &amp; 23 relations Region Connection Calculi</td>
</tr>
<tr>
<td>conn</td>
<td>Connection (primitive)</td>
</tr>
<tr>
<td>conv</td>
<td>Convex hull</td>
</tr>
<tr>
<td>4IM</td>
<td>Four Intersection Model</td>
</tr>
<tr>
<td>DE-4IM</td>
<td>Dimensionally Extended Four Intersection Model</td>
</tr>
<tr>
<td>JEPD</td>
<td>Jointly Exhaustive Pair-wise Disjoint</td>
</tr>
<tr>
<td>DC</td>
<td>Disconnected</td>
</tr>
<tr>
<td>EC</td>
<td>Externally connected</td>
</tr>
<tr>
<td>PO</td>
<td>Partially overlaps</td>
</tr>
<tr>
<td>EQ</td>
<td>Equal</td>
</tr>
<tr>
<td>TPP</td>
<td>Tangential Proper Part of</td>
</tr>
<tr>
<td>NTPP</td>
<td>Non-tangential Proper Part of</td>
</tr>
<tr>
<td>TPPi</td>
<td>Tangential Proper Part of inverse</td>
</tr>
<tr>
<td>NTPPi</td>
<td>Non-tangential Proper Part of inverse</td>
</tr>
<tr>
<td>OUT (resp. OUTi)</td>
<td>Outside (resp. Outside inverse)</td>
</tr>
<tr>
<td>IN (resp. INi)</td>
<td>Inside (resp. Inside inverse)</td>
</tr>
<tr>
<td>PIN (resp. PINi)</td>
<td>Partially Inside (resp. Partially Inside inverse)</td>
</tr>
<tr>
<td>SURR (resp. SURRi)</td>
<td>Surrounds (resp. surrounded by)</td>
</tr>
<tr>
<td>Sim</td>
<td>Similarity transformations</td>
</tr>
<tr>
<td>AFF</td>
<td>Affine transformations</td>
</tr>
</tbody>
</table>
Chapter I

1 Introduction

Relative position is a visual feature. Like color, texture, and shape, it is a low-level visual feature [1]. It refers to the position of one object in space with respect to another. In daily life, knowledge about relative position of objects is conveyed through linguistic descriptions like “object A is to the left of object B, somewhat above and partially overlapping it”. Such qualitative statements use the concepts of direction, topology, and distance to describe relative position. Like other low-level visual features such as color, texture, and shape, information of relative position is important in various areas of image processing and computer vision. However, most practical image processing and computer vision tasks require quantitative representation of visual information instead of its qualitative description. Quantitative approaches have, therefore, been developed to the representation of relative position. Such quantitative models are called relative position descriptors. In simple terms, a relative position descriptor is a vector of numbers. Similar to color [2], texture [3], and shape descriptors [4, 5, 6] that describe color, texture, and shape content of objects, relative position descriptors provide information about the relative position of objects (See Figure 1-1). They can be used either independently or potentially combined with other image descriptors to assist tasks such as geospatial image indexing and retrieval [7], scene matching [8, 9], etc. Considerable attention has been paid to the modeling of relative position in image processing research and a number of relative position descriptors have been proposed. Some best-known relative position descriptors are angle histogram [10], force histogram [11], and R-histogram [12].
Spatial relationships can be extracted directly from the image space. Alternatively, they can be extracted from the relative position descriptor using some extraction method. The relative position descriptor in turn is generated from the image space.

1.1 Connection with spatial relationships

The concept of relative position is closely related to that of the spatial relationships. Spatial relationships also provide information about the arrangement of objects in space [13] [14] [15].

However, unlike the complex spatial organization that relative position denotes, spatial-relationships represent individual (and atomic) spatial relations between objects such as the “left” (e.g., “object A is to the left of object B”) or the “far” (“object A is far from object B”), etc. Due to this close connection, spatial relationships can be deduced from relative position information and relative position information can be conveyed in terms of spatial relationships.

Spatial relationships are commonly categorized into directional (“right”, “left”, “above”, “below”), topological (“touch”, “overlap”, “cover”, etc.), and distance (“far”, “near”)
relationships. They are important in many areas, e.g., image processing and computer vision, Geographic Information Systems (GISs), imaging databases, anatomical and medical investigation, and natural language processing [16] [17]. Although, a number of specialized qualitative and quantitative approaches have been developed to the modeling of spatial relationships, one way to model them is to derive their definition from *relative position descriptors* (See Figure 1-1). A key role of relative position descriptors is thus to serve as a basis for the modeling of *spatial relationships* [13] [14] [15]. This has been illustrated in Figure 1-1. As the figure shows, spatial relationships can be directly extracted from images. Alternatively, they can be deduced from relative position descriptors through the use of special extraction methods. The advantage of extraction from relative position description is that, descriptors capture information about a variety of spatial relationships and once the relative position descriptor is generated, interpretation of spatial relationship from the descriptor is fast.

**1.2 Applications of relative position descriptors**

Apart from the role of being a container for spatial relationships as discussed in the preceding section, relative position descriptors have applications in a number of other areas. Some of the applications of relative position descriptors are in human-robot interaction [18], semantic metadata generation for image digital libraries [19], suspected minefield risk estimation [20], melanocytic image analysis and recognition [21], geospatial information retrieval and indexing [22], scene matching [23], land cover classification [24], graphical symbol retrieval [25], shape matching [26], spatiotemporal reasoning [27], and map-to-image conflation [28].

**1.3 Current state of relative position description research**

Relative position modeling has attracted considerable research attention. A detailed survey of various techniques proposed for the modeling of the relative position is given in Chapter 2. Here a brief description of some issues in the existing approaches is given.

3
This is followed by a summary of the goals and main contributions of the work presented in this thesis.

1.3.1 Some expectations from relative position descriptors

As containers of spatial relationships, a key requirement in relative position descriptors is the possibility to extract maximum spatial relationships from a descriptor. At the minimum, relative position descriptors are expected to allow the extraction of primitive directional relationships and the surround relationship. The ability to allow the extraction of topological (set and non-set) and distance relationships (far, near) is considered an additional strength. Another important property is the low computational cost. This is a critical requirement when large images are involved. Usually, relative position descriptors are designed with 2D crisp objects in raster format in mind. Extension to 3D objects or to the handling of fuzzy objects or objects in vector format, therefore, is considered an important property of relative position descriptors.

Invariance to geometric transformations is another critical requirement in relative position modeling. The invariance problem arises when image-objects undergo changes in shape, size, etc. due to changes in imaging conditions (e.g., change in viewing angle or camera position). Geometric transformations provide a mean to model such changes. Usually, the similarity transformations such as rotation, translation, and scaling are used for invariance modeling. However, affine transformations (rotation, translation, scaling, and shearing) are considered more appropriate due to their ability to approximate more complex geometric transformations.

Two problems arise in relative position modeling when objects undergo affine transformations; the first problem called the direct problem in this work deals with obtaining the resulting descriptor from the original descriptor; the second problem called the inverse problem here is concerned with retrieving the transformation when the original and resulting descriptors are given. A related problem is the recovery problem, which is concerned with determining all object-pairs that can be associated with a given relative
position descriptor. These problems can be defined as follows. Let $A$ and $B$ be two spatial objects, and let $t$ be an affine transformation. Assume, $\varphi^{AB}$ describes the position of $A$ relative to $B$ and $\varphi^{t(A)t(B)}$ describes the relative position of $t(A)$ relative to $t(B)$. The three problems can be stated as,

1) the direct problem: knowing $\varphi^{AB}$ and $t$, find $\varphi^{t(A)t(B)}$
2) the inverse problem: knowing $\varphi^{AB}$ and $\varphi^{t(A)t(B)}$, find $t$
3) the recovery problem
   a) knowing $\varphi^{AB}$, find all the object pairs $(A',B')$ such that $\varphi^{A'B'} = \varphi^{AB}$
   b) knowing $\varphi^{AB}$ and object $B$, find object $A$ such that $\varphi^{A'B'} = \varphi^{AB}$

Solving these problems has practical utility from the stand-point of relative position computation and recognition. For example, solving the direct problem allows the computation of resulting relative position descriptor from the original descriptor without any knowledge of the objects. Solving the inverse problem can help match object pairs through their relative position descriptors. This later property is useful in task like image registration and restoration. Finally, solving the recovery problem can help retrieve the object-pair from their descriptor. This can be useful in image compression. Recovery of objects has been studied in the area of shape description as well and a number of shape descriptors have been proposed that allow the recovery of objects [29]. These problems are also important from the perspective of theoretical solvability and computational consideration of finding efficient algorithms for calculating the descriptor.

Finally, a crucial property is the possibility to normalize a relative position descriptor. This property is important because if a relative position descriptor can be normalized, it may be possible to make it invariant to affine transformations.

1.3.2 Issues with the existing descriptors

The existing descriptors don’t satisfy all the above properties (detailed survey in Section 2). For example, none of the existing relative position descriptors allows the extraction of
distance relationships. Additionally, only a few topological relationships can be extracted from the existing relative position descriptors and the extraction is not straightforward. Likewise, most descriptors consider only 2D crisp objects in raster format. Handling of 3D crisp raster objects or fuzzy objects (2D and 3D) is not easy or has not been studied in the case of many descriptors. Handling of vector objects has been addressed in the case of some descriptors only (e.g., force-histogram). Furthermore, computational time for many descriptors is high. Finally, although extensively studied in the case of other image descriptors such as color, texture, and shape descriptors, design of relative position descriptors that are invariance to affine transformations has received scant attention in the modeling of relative position. Of the three affine invariant problems, the direct problem has been solved for only the force-histogram. The inverse problem or the recovery problems have not been solved or considered for any existing relative position descriptor.

1.3.3 Contribution of the present work

The aim of this work was to design a relative position descriptor that would satisfy most of the above requirement. Specifically, low computational complexity, an easy normalization procedure, representation of maximum possible spatial relationships, and a known behavior to affine transformations were the intended features in the new descriptor. The study produced the following tangible results.

1. **A first comprehensive survey of relative position descriptors:** A first comprehensive survey of the existing work on relative position descriptors was carried out that produced the following results.
   a. **A first classification of the current approaches** to the modeling of relative position of objects was proposed. It is given in Section 2.2.2 and Table 2-4 (Chapter 2).
   b. **A set of properties for a relative position descriptor** to possess to satisfy user's expectation and meet requirements of practical image processing applications was determined. For detail, see Section 2.2.4, Table 2-5 (Chapter 2).
2. **A new relative position descriptor: The $\Phi$ – Descriptor.** A new relative position descriptor, i.e. the $\Phi$-descriptor (phi-descriptor) is proposed. The $\Phi$-descriptor is based on the original idea of associating spatial relation categories with the directionally aligned boundary-points of the two objects and modeling the descriptors in terms of the categories. The new approach combines ideas from the modeling approaches used in Radon transform, F-histogram, and Allen relations with the innovative idea of boundary-point categories to define a versatile relative position descriptor (See Chapter III). The $\Phi$-descriptor has many nice features. It satisfies many of the properties described in the preceding section. For example, it allows the extraction of a wide range of spatial relationships. This includes the directional and distance relationships, the set and non-set topological relationships (RCC8, RCC23), and the surround relationship. Moreover, more complex relationships, e. g., between and among, can also be potentially interpreted from the descriptor. Additionally, the descriptor has a known behavior to affine transformations; therefore, the direct problem has already a solution in the case of the $\Phi$-descriptor. The inverse problem can be solved for the descriptor up to translation under certain conditions. Furthermore, the $\Phi$-descriptor can be normalized and the normalized descriptor is invariant to geometric transformations. Additionally, objects crisp/fuzzy, raster/vector, or 2D/3D can be handled. A most important property of the new descriptor is its linear time complexity (for raster objects) that makes its calculation fast.

3. **Models of spatial relationships based on the $\Phi$-descriptor:** Methods are also proposed for extracting spatial relationships from the $\Phi$-descriptor. These include methods for extracting both the qualitative and quantitative topological relationships defined by the well-known Region Connection Calculus (RCC) framework. The models have been proposed for the core RCC8 model as well as its RCC23 extension. Extraction of qualitative topological spatial relations defined in the 4-Intersection Model (4IM) has also been considered. Additionally, methods
for extracting (1) directional, (2) surround, and (3) distance spatial relations (see Appendix A) have also been proposed in the present work.

The rest of this work has been organized as follows. Chapter II gives a detailed review of the existing descriptors and their properties. Chapter III describes the theory and properties of the new descriptor. Chapter IV proposes some methods for extracting topological relationships (crisp/fuzzy RCC8/RCC23 and 4IM) from the descriptor and chapter V and VI propose models of directional and surround relationships respectively based on the Φ-descriptor. Chapter VII describes the algorithms for calculating the Φ-descriptor for raster images and Chapter VIII gives conclusions and directions for future work.
Chapter II

2 Literature review

Spatial propositions, like above, inside, near, denote spatial relationships between spatial regions. A relative position descriptor is a quantitative representation of the relative position of two spatial regions, and a basis from which models of spatial relationships can be developed. This chapter gives a review of relative position modeling. It also includes a review of the models of spatial relationship proposed in literature. Models for spatial relationships are discussed in Section 2.1 and those of relative position descriptors in Section 2.2.

First, however, a specification of what a spatial region means here is given. In its general sense, a spatial region may be a subset of the Euclidean plane, or a subset of the Euclidean space of dimension 3. It may be a 3-D region (e.g., ball), a 2-D region (e.g., disk), a 1-D region (e.g., arc), a 0-D region (e.g., point), or any combination of these. It may be vague, i.e., with imprecise or indeterminate boundaries. Vague regions induce uncertainty on spatial relationships, which has an influence on spatial reasoning. A good amount of effort has been devoted to the issue of spatial vagueness. Many approaches are based on fuzzy set theory [30, 31, 32, 33]; the handling of vague regions represented as fuzzy sets can often be reduced to that of crisp sets (their $\alpha$-cuts) using some general scheme [34, 35]. Approaches based on rough set theory [36], stochastic approaches [37, 38], the broad boundary, egg-yolk and super-valuationist approaches [39, 40, 41] are other examples of approaches for dealing with spatial vagueness. The spatial regions considered in this thesis are nonempty bounded regular closed sets of the Euclidean plane. Here, they are called objects.
2.1 Models of spatial relationships

Many categorizations of spatial relationships have been proposed in the literature. A common categorization of spatial relationships is the classification into topological, directional, and distance relationships. Examples of topological spatial relationships are “touch”, and “overlap”, that of directional relationships are “right”, “left”, and of distance relationships, “far” and “near”. Another classification categorizes spatial relationships into binary and non-binary spatial relationships. Examples of binary spatial relationship are “left” and “right” and those of non-binary relationships, “between” and “among” [42]. However, binary relationships are far more common. Spatial relationships are also categorized into crisp or fuzzy relations (See Section 2.1.1). In the case of crisp relationships, relationship values are mapped to true/false values whereas in the case of fuzzy relationships, relationship values are represented on a continuum of truth-degrees.

The work presented in this thesis focusses on binary spatial relationships (fuzzy/crisp) only. The binary relationships that are explored include topological (set and non-set), directional, and distance relationships including spatial relationships associated with “surroundedness” (some of which are directional and some topological). Sections 2.1.1 through 2.1.5 presents an overview of these various categorizations of spatial relationships.

2.1.1 Crisp and fuzzy relations

A crisp relation—or relation for short—is a binary relation on the set of objects. It can be seen as a function from the set of all object pairs to the set \{true, false\}, or to the set \{0, 1\}, where 0 stands for false and 1 for true. Likewise, the logical negation can be seen as a function from \{false, true\} to \{false, true\}, or from \{0, 1\} to \{0, 1\}, and the logical conjunction as a function from \{false, true\}^2 to \{false, true\}, or from \{0, 1\}^2 to \{0, 1\}. When \{0, 1\} is replaced with [0, 1], the relation becomes a fuzzy relation, the negation a fuzzy negation and, the conjunction a fuzzy conjunction. The elements of [0, 1] are called truth values.
Crisp relations are commonly used as qualitative models of spatial relationships whereas fuzzy relations are used as quantitative models. Fuzzy relations may be preferred to crisp ones depending on whether the spatial regions are vague or not, and whether the spatial relationships are inherently vague (e.g., directional) or not (e.g., topological). For example, it is often difficult or undesirable to differentiate between situations where two objects are very close to each other but nonintersecting, and situations where their boundaries intersect. This is illustrated in figures 2.1-2.4. In the figures, a pair \((A, B)\) of objects is composed of the argument object, \(A\), and the reference object, \(B\). In all the figures, \(A - B\) appears in light gray, \(B - A\) in dark gray and \(A \cap B\) in medium gray. The direction 0 is from left to right, the direction \(\pi/2\) from bottom to top, etc. (Figure 2-1). A white dot indicates that the object boundaries intersect in an isolated point and a black dot that they intersect in an arc (Figure 2-2). A relation may be represented by an object pair: convex objects are then preferred to concave ones, and connected objects are preferred to disconnected ones (Figure 2-3); non-inclusion is preferred to inclusion, and inequality is preferred to equality (Figure 2-4).

Figure 2-1 Object pairs and directions. (a) The argument object \(A\) is in light gray and the referent object \(B\) is in dark gray (b) The cardinal directions. The direction 0 corresponds to “right”, direction \(\pi/2\) to “above”, \(\pi\) to left, and \(3\pi/2\) to “below”.

Figure 2-2 Object pairs and directions. (a) The argument object \(A\) is in light gray and the referent object \(B\) is in dark gray (b) The cardinal directions. The direction 0 corresponds to “right”, direction \(\pi/2\) to “above”, \(\pi\) to left, and \(3\pi/2\) to “below”.
Figure 2-2 Boundary intersections. (a) The boundaries of $A$ and $B$ intersect. (b) They intersect in an isolated point. (c) They intersect in an arc.

Figure 2-3 Representing relations: convexity and connectedness. Choose (a), instead of (b) or (c), to indicate that $A$ and $B$ are disjoint. Choose (b), instead of (c), to indicate that the convex hull of $A$ includes $B$.

Figure 2-4 Representing relations: non-inclusion and inequality. Choose (a), instead of (b) or (c), to indicate that $A$ and $B$ intersect. Choose (b), instead of (c), to indicate that $A$ is a subset of $B$. 
2.1.2 Topological relations

The *modeling of topological relationships* has been extensively studied in AI (Artificial Intelligence), mostly from a qualitative point of view. The two best-known qualitative frameworks that have been proposed for representing and reasoning with topological relationships are the RCC8 (Region Connection Calculus) [43] and the 4IM (Intersection Model) [44]. The RCC8 uses the notion of connection between objects, while the 4IM is based on the intersections between the interiors and the boundaries of objects. Most other qualitative frameworks are extensions of the RCC8 and 4IM, e.g., the RCC23 and the DE-4IM. Comprehensive reviews can be found in [45, 46, 47]. Due to their importance in approximate reasoning—a type of reasoning in which a possible imprecise conclusion is deduced from a collection of imprecise premises [48], quantitative models of topological relationships have also received good attention in literature. Most of the work in this area has been on quantitative models of RCC and IM relations. Following sections include an overview of RCC and 4IM frameworks.

2.1.2.1 The RCC8 and RCC23 relations

Let \( \text{conn} \) be a binary relation on the set of objects such that for any \( A \) and \( B \), \( \text{conn}(A,A) \) and \( \text{conn}(A,B) \rightarrow \text{conn}(B,A) \). In other words, \( \text{conn} \) is reflexive and symmetric. \( \text{conn}(A,B) \) can be interpreted as \( A \cap B \neq \emptyset \), hence the name of the relation (\( \text{conn} \) as in *connection*). It gives rise, through first-order logic, to various topological relations. These relations are given in Table 2-1. Now, consider the set \( \text{RCC8} = \{ \text{DC}, \text{EC}, \text{PO}, \text{EQ}, \text{TPP}, \text{NTPP}, \text{TPPi}, \text{NTPPi} \} \), where \( \text{TPPi} \) and \( \text{NTPPi} \) (with \( i \) as in *inverse*) are defined by \( \text{TPPi}(A,B) = \text{TPP}(B,A) \) and \( \text{NTPPi}(A,B) = \text{NTPP}(B,A) \). Its elements, the RCC8 relations, are JEPD (Jointly Exhaustive and Pairwise Disjoint): they are jointly exhaustive because for any objects \( A \) and \( B \), there can be found an RCC8 relation \( \text{rel} \) such that \( \text{rel}(A,B) \); they are pairwise disjoint because for any objects \( A \) and \( B \) and for any distinct RCC8 relations \( \text{rel1} \) and \( \text{rel2} \), we have \( \neg \text{rel1}(A,B) \) or \( \neg \text{rel2}(A,B) \). The RCC8 relations can be represented by a graph: two relations are neighbours in the graph if and only if continuous changes in the shape or position of an object lead to a direct change from one relation to the other.
The graph is given in Figure 2-5. The JEPD property makes the RCC8 a relational algebra that is closed under composition and that allows efficient reasoning about topological relations through a computational mechanism such as the path-consistency algorithm [49].

The RCC23 [50] is an extension of the RCC8. Let conv(A) be the convex hull of the object A, and let IN, PIN and OUT be the binary relations on the set of objects defined as in Table 2-2. The RCC23 replaces DC and EC with 17 new relations. See Table 2-3 and Figure 2-6. Note that OUTi(A,B) = OUT(B,A), PINi(A,B) = PIN(B,A) and INi(A,B) = IN(B,A). The RCC23 relations are JEPD. The new relations allow more distinctions to be made between configurations. They are not, however, topological in a strict mathematical sense. For example, Figure 2-6 represents eight different RCC23 relations but the same topological relation, since the object pairs are homeomorphic.

Fuzzy extensions of the RCC8 have been proposed in [51, 52, 53, 54]: the relation conn is redefined as a fuzzy relation, the logical operators are replaced with fuzzy logical operators, and the universal and existential quantifiers are modeled as infima and suprema of truth values.

Figure 2-5 The RCC8 conceptual neighbourhood graph. In each case, the argument A is the lighter object and the referent B is the darker object.
Table 2-1 Some topological relations definable in terms of \( \text{conn} \).

<table>
<thead>
<tr>
<th>name</th>
<th>Definition</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>( \neg \text{conn}(A, B) )</td>
<td>is not connecting with</td>
</tr>
<tr>
<td>P</td>
<td>( \forall C [\text{conn}(C, A) \rightarrow \text{conn}(C, B)] )</td>
<td>is a part of</td>
</tr>
<tr>
<td>PP</td>
<td>( P(A, B) \land \neg P(B, A) )</td>
<td>is a proper part of</td>
</tr>
<tr>
<td>EQ</td>
<td>( P(A, B) \land P(B, A) )</td>
<td>is equal to</td>
</tr>
<tr>
<td>O</td>
<td>( \exists C, [P(C, A) \land P(C, B)] )</td>
<td>overlaps</td>
</tr>
<tr>
<td>DR</td>
<td>( \neg O(A, B) )</td>
<td>is discrete from</td>
</tr>
<tr>
<td>PO</td>
<td>( O(A, B) \land \neg P(A, B) \land \neg P(B, A) )</td>
<td>partially overlaps</td>
</tr>
<tr>
<td>EC</td>
<td>( \text{conn}(A, B) \land \neg O(A, B) )</td>
<td>is externally connecting with</td>
</tr>
<tr>
<td>TPP</td>
<td>( PP(A, B) \land \exists C[EC(C, A) \land EC(C, B)] )</td>
<td>is a tangential proper part of</td>
</tr>
<tr>
<td>NTPP</td>
<td>( PP(A, B) \land \neg TPP(A, B) )</td>
<td>is a non-tangential proper part of</td>
</tr>
</tbody>
</table>

Table 2-2 Some relations definable in terms of \( \text{conn} \) and \( \text{conv} \).

<table>
<thead>
<tr>
<th>name</th>
<th>Definition</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUT</td>
<td>( DR(A, \text{conv}(B)) )</td>
<td>is outside</td>
</tr>
<tr>
<td>PIN</td>
<td>( DR(A, B) \land PO(A, \text{conv}(B)) )</td>
<td>is partially inside</td>
</tr>
<tr>
<td>IN</td>
<td>( DR(A, B) \land P(A, \text{conv}(B)) )</td>
<td>is inside</td>
</tr>
</tbody>
</table>

Table 2-3 The RCC8 vs. RCC23 relations.

<table>
<thead>
<tr>
<th>RCC8</th>
<th>RCC23</th>
<th>RCC8</th>
<th>RCC23</th>
<th>RCC8</th>
<th>RCC23</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>( OUT \land OUTi \land DC )</td>
<td>EC</td>
<td>( OUT \land OUTi \land EC )</td>
<td>PO</td>
<td>PO</td>
</tr>
<tr>
<td></td>
<td>( OUT \land PINi \land DC )</td>
<td></td>
<td>( OUT \land PINi \land EC )</td>
<td>EQ</td>
<td>EQ</td>
</tr>
<tr>
<td></td>
<td>( OUT \land INi \land DC )</td>
<td></td>
<td>( OUT \land INi \land EC )</td>
<td>TPP</td>
<td>TPP</td>
</tr>
<tr>
<td></td>
<td>( PIN \land OUTi \land DC )</td>
<td></td>
<td>( PIN \land OUTi \land EC )</td>
<td>NTPP</td>
<td>NTPP</td>
</tr>
<tr>
<td></td>
<td>( PIN \land PINi \land DC )</td>
<td></td>
<td>( PIN \land PINi \land EC )</td>
<td>TPP</td>
<td>TPP</td>
</tr>
<tr>
<td></td>
<td>( PIN \land INi \land DC )</td>
<td></td>
<td>( PIN \land INi \land EC )</td>
<td>NTPP</td>
<td>NTPP</td>
</tr>
<tr>
<td></td>
<td>( IN \land OUTi \land DC )</td>
<td></td>
<td>( IN \land OUTi \land EC )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( IN \land PINi \land DC )</td>
<td></td>
<td>( IN \land PINi \land EC )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( IN \land INi \land DC )</td>
<td></td>
<td>( IN \land INi \land EC )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2-6 The RCC23 DC relations. The RCC23 replaces the RCC8 relation DC with eight pairwise disjoint relations.

### 2.1.2.2 The 4IM and DE-4IM relations

The 4IM (4-Intersection Model) is a theoretical framework that provides a definition of topological relations between two simply-connected regions based on the intersection primitive (\(\cap\)) and the point-set topological notions of interior and boundary. The model is described as follows. Let A and B be two simply-connected regions. Let \(A^I\) and \(B^I\) denote the interior (I) of A and B and \(A^B\) and \(B^B\) the boundaries (B) of A and B, then the intersection of \(A^I\), \(B^I\), \(A^B\), and \(B^B\) gives rise to 16 possible relations represented by the matrix \([A^I \cap B^I, A^I \cap B^B, A^B \cap B^I, A^B \cap B^B]\). Of these relations, eight are valid relations and are named as \(\text{disjoint, meet, overlap, covered by, inside, covers, contains, and equal}\). These relations correspond to the RCC8 relations of DC, EC, PO, TPP, NTPP, TPpi, NTPPi, and EQ respectively (see Figure 2-7) [44].

The DE-4IM (Dimensionally-extended 4IM) is an extension of the 4IM model that also takes into account the dimensions of the intersection (\(\cap\)) of the entities A and B. The dimension of the empty intersection set \(\emptyset\) (e.g., \(A^I \cap B^I = \emptyset\)) is denoted by \(F\) or -1 and that of the non-empty set \(\neg\emptyset\) by the maximum number of dimensions of the intersection.
Thus 0 denotes a point, 1 a line, and 2, an area. The model is, therefore, defined over the domain \{0,1,2,F\}. The DE-4IM is useful in modeling relations between points, lines, and regions. Each relation can then be denoted by a string ssss where \(s \in \{0,1,2,F\}\). For example, 2110 represents the relation where \(A \cap B\) has dimension 2, \(A \cap B^B\) and \(A^B \cap B\) have dimension 1, and \(A^B \cap B^B\) has dimension 0. The different DE-4IM relations have been shown in Figure 2-7.

Other extensions of the intersections models have also been proposed. The 9IM model also considers the exterior (E) besides the interiors (I) and boundaries (B) of objects. Thus a large number of relations can be represented. The DE-9IM is a dimensionally extended counter-part of 9IM where the dimensions \{0,1,2,F\} are also considered [55, 56, 57].

2.1.3 Directional relations

Qualitative directional relations have been extensively studied in the area of AI (artificial intelligence). Some of the models of the directional relations proposed are the cone-shaped direction relations [58], projection-based direction relations [59], 2D-String [60], and 2D-CString [61]. A detailed survey can be found in [47]. In the present work, only quantitative directional relations are considered.
A directional relation between two objects (reference object and argument object) \( A \) (argument object) and \( B \) (reference object) is a binary relation that is represented by the angle \( \theta \in [0,2\pi] \) between \( A \) and \( B \). The cardinal directions “right”, “left”, “above”, and “below” correspond to \( \theta = 0, \theta = \pi/2, \theta = \pi, \) and \( \theta = 3\pi/2 \). (See Figure 2-8)

Models of directional relations based on the direct and indirect method (extraction of spatial relationship from a relative position descriptor, see Figure 1-1) have been proposed in the literature. An example of the direct method is the method of fuzzy landscape introduced in [62]. A number of methods for extracting directional relations from relative position descriptors have also been proposed. The compatibility method [10, 63] treats the normalized descriptor (e.g., the angle histogram) as an unlabelled fuzzy set and computes its compatibility with the given fuzzy directional relation (e.g., “RIGHT”). The center of gravity of the compatibility set is then regarded as the acceptability degree of the proposition (e.g., “A is to RIGHT of B”). Similar to the compatibility method, the descriptor is normalized (e.g., angle histogram) in the aggregation method [34, 64, 65] and treated as an unlabelled fuzzy set. The normalized histogram values are then used to compute the weighted average of the degrees of alignment between the directions in the histogram and the given direction. The weighted average represents the acceptability degree of the proposition being assessed. The method of effective forces assesses the validity-degree of the proposition from the degree of alignment between the average direction of the histogram and the given direction [66, 67]. Besides, a method based on
machine learning has been proposed in [19]. It uses fuzzy k-NN classifier to extract directional relations from R-histogram.

2.1.4 Distance relations

Distance is a scalar property associated with the position of two objects. It possibly also depends on other attributes such as shape. The concept of distance has been used in many contexts in image processing. Two common notions of distance used in image processing literature are distance as a measure of similarity between objects as in shape matching and distance (between objects) as a feature of space as in spatial modeling and qualitative spatial reasoning (QSR) [68]. This later conception of distance is referred to as spatial distance. The work in this thesis is restricted to this second notion of distance. Spatial distance has been studied from two angles in spatial modeling; qualitative modeling of distance; quantitative modeling of distance. Qualitative models focus on definitions of distance that map metric distance information (expressed as ranges of distance or geometric intervals) to qualitative expressions or semantic labels (e.g., far, near, medium far, etc.) [69] [70]. In the qualitative approach, factors such as frame of reference in which the objects are situated, isotropy or anisotropy of the space, shape and size of objects, etc. are taken into account when modeling the distance. The quantitative approaches mostly rely on fuzzy techniques for modeling spatial distances. Many approaches for fuzzy modeling of spatial distance have been proposed. One approach is to generalize distance defined as a crisp concept into fuzzy distance [71] [72]. Other approaches deduce fuzzy distance from set or other relationships [73] [74]. The approach where the fuzzy linguistic variables of “far” and “near” are associated with the quantitative measure of distance has been studied in such contexts as spatial ontology for image interpretation [75, 76], spatial reasoning [77], and scene description [78]. Specifically, such variables are called as distance spatial relationships.
2.1.5 Surroundedness

Qualitatively, B is surrounded by A if any path from object B to the region outside A must intersect A; otherwise B is not surrounded by A. However, it is more common to use a quantitative definition of surroundedness. Rosenfeld and Klette [79] have provided two quantitative definitions of surroundedness; topological surroundedness and visual surroundedness. The difference between the two definitions is illustrated in Figure 2-9. In Figure 2-9(a)(b), there are paths leading out of B through A that don’t intersect A and, therefore, B is not surrounded by A in the topological sense. However, in Figure 2-9(c), A surrounds B. In the visual sense, however, A surrounds B in the cases of Figure 2-9(b)(c) because B can see A in all directions. In literature, it is more common to use visual surroundedness due to its simplicity. There is also a link between surroundedness and the RCC23 (See Figure 2-6). For example, an object A can be considered to be surrounded (to some degree) by an object B iff $IN(A, B)$ or $PIN(A, B)$. Likewise, it may be concluded that A surrounds B iff $INi(A, B)$ or $PINi(A, B)$.

Approaches to the modeling of “surround” can be broadly divided into two groups. In the traditional approach, the surround relationship between A and B is inferred from the existence of cardinal directional relations (i.e., right, left, above, below) between objects [80, 10]. Thus if object A is in each of the cardinal directions of B, there is evidence for the relationship “A surrounds B”. An implicit assumption in the approach is that the objects don’t intersect. This is a significant limitation in that until this constraint is imposed on the objects, there is no way to tell whether the relationship is that of the “surround” or “part

![Figure 2-9 Visual vs. topological surroundedness. Visual surroundedness: (a) A does not surround B; (b)(c) A surrounds B. Topological surroundedness: (a)(b) A does not surround B; (c) A surrounds B.](image-url)
The more specialized approach to modeling the surround relies on Rosenfeld’s visual surrounded-ness” [79]. The main principle in this approach consists in finding a measure of the angle that one or more open intervals in object A subtend at object B and feeding it to a mapping function to obtain the degree of surround. Using the notion of visual surroundedness, models of surround have been proposed using both the direct and indirect approaches. The direct method proposed in [81] uses mathematical morphology to define a fuzzy region (the so called fuzzy landscape) that surrounds the reference object. The region is then used to derive the surround degree. A feature of the method is that influence of the distance to the reference object on the surround relationship is taken into account. However, only connected and non-intersecting objects can be handled in this method. The indirect methods in [82, 63, 34] extract the “surround” from the histogram of angles. The objects considered are connected and non-intersecting. Likewise, models of surround based on force histogram [83] [84] and spread histogram [21] have been proposed. However, the angle and force histograms are not used as generic entities when modeling the surround but are instead customized for the surround relationship. Thus in way the spread histogram attempts to tackle the surround relationship.

2.2 Models of relative position

Relative position description has received considerable attention in recent years. The main attention has been focussed on finding effective approaches to the modeling of relative position. Consequently, many approaches have been proposed. One key difference between different approaches is the way objects are represented. Some methods approximate objects by bounding boxes while other methods use characteristic points such as centroids. Methods also vary in as whether the descriptor is derived from the boundaries of objects or their interior regions. Another distinction between the different approaches is the way in which objects are handled. Some methods reduce the handling of objects to the handling of simpler components such as 1D sections when evaluating relative position. Other methods model relative position at a more elementary level, e. g., at the level of points and pixels. One important consideration is the choice
between crisp or fuzzy representation of objects. The advantage of crisp representation is fast computation whereas that of the fuzzy representation is efficient processing of objects with vague boundaries. Extraction of spatial relationships from the descriptor is also an important consideration in relative position modeling. Although relative position descriptors allow the extraction of crisp spatial relationships, extraction of fuzzy spatial relationships from descriptors has received relatively more attention. Efficient algorithms and 3D extensions are other areas of research attention. Finally, invariance to geometric transformations (e.g., similarity or affine) is an important consideration in relative position modeling.

In the following sections, a review of different approaches to the modeling of relative position in image processing is given. The review involves a classification of various approaches to relative position modeling and a comparison of their properties besides a review of individual descriptors. Other aspects such as applications have also been highlighted where applicable. Sections 2.2.1 and 2.2.2 propose a classification of the existing approaches to relative position modeling, Section 2.2.3 gives an overview of individual descriptors. Section 2.2.4 provides a comparison of the properties of different descriptors.

2.2.1 Categorization

Existing approaches to relative position description involves the histogram construction of some measure of relative position. The histogram is calculated from the angle relationship between the elements of the objects (reference and argument). Sometimes, other relationships between elements such as Allen relations in addition to angle are considered. The elements chosen are either pixels or larger blocks of the objects such as longitudinal sections, segments, or square sub-regions. Moreover, the entire or part of the object-content may be used to calculate the relative position. Last, relative position may be calculated from pixels on the boundaries of the objects or both the boundary pixels and pixels in the interior regions. Based on these distinctions, this work proposes the classification in Section 2.2.2 of the approaches to relative position modeling.
Definitions of the various terms used is as follows. An object is a nonempty, regular closed subset of the Cartesian plane. A pixel is a unit square whose sides are parallel to the $x$- and $y$-axis and whose center has integer coordinates. A raster object is the union of a finite number of pixels. A vector object is an object whose boundary is the union of a finite number of line segments. Consider two objects $A$ and $B$. The position of $A$ relative to $B$ is usually represented by a histogram $H^{AB}$, or by a real function improperly called a histogram, or by a tuple of such functions. A working assumption is that the objects may be too close to each other to be approximated by their centroids or minimum bounding boxes.

2.2.2 Main approaches

Based on whether the relative position is modeled at a higher level (coarse-grained) or a more elementary level (fine-grained) of the object’s content, the methods are categorized as: pixel-pair based methods; point-pair based methods, method based on aligned-sections or cores, and method based on aligned-segments. In the pixels-pair approach [10, 63, 12, 21], relationships at the level of pixels are considered to calculate the histogram (Figure 2-10). This involves pairing a pixel of the object $B$ (reference object) with a (directionally-aligned) pixel of object $A$ (argument object) and deriving a measure of the relative position of the pixels.

The pixels considered may belong to boundaries of the objects [12] or both boundaries and interior regions [10]. In some cases, blocks of pixels are considered instead of pixel-pairs [85]. This method is used with only raster objects in mind. The point-pair based approach [11] [86] has been proposed for the relative position modeling of both vector and raster objects in mind. In the method, the histogram value is calculated from relative position between points in a point-pair that involves one point from the reference object and the other from the argument object. The points are defined by a beam of parallel directed lines that intersect the objects in the given direction (see Figure 2-11(a)(e)).
Figure 2-10 Point and pixels-pair based approaches. (a) Pixels $a_i \in A$, $b_i \in B$ are related pair-wise. (b) Histogram values are calculated from some measure of the relative position values for point-pairs given by some function, $\delta$; $\text{ope}$ is an operator for aggregating $\delta$. (c) The histogram.

The core-pair and segment-pair based approaches use a beam of directed lines as in the point-pair approach to define pairs of 1D aligned sections or cores of the two objects such that each pair has one section from the reference object and one section from the argument object (See Figure 2-11 (d)). A core or section may be further partitioned into two or more segments depending on the structure of the object (e.g., an object with holes). See Figure 2-11 (a), (b), and (c). Then there are two possibilities; relative position may be defined at the level of segments; or it may be modeled at the level of sections or core. The former gives rise to the segment-pair based approach whereas the later to the core-pair or section-pair based approach [11] [26]. Example of the former are the $f$-histogram and that of the core-based approach the $\mathcal{F}$-histograms. R-histogram (Sections 2.2.3.4) and Ratio-histogram (Section 2.2.3.7) are specific cases of the $\mathcal{F}$-histograms.

Another useful categorization may be boundary-based methods if the points/pixels in the pairs considered belong to the boundaries of the object or region-based methods if they can also belong to the interior regions. Example of the later approach is the histogram of angles (Section 2.2.3.1) and the former, the visual area histogram (Section 2.2.3.5) and R-histogram (Section 2.2.3.3). A summary of the approaches is given in Table 2-4.
Figure 2-11 Aligned section and segment based approaches. (a) Longitudinal sections \( A_\theta(l_i), B_\theta(l_i), i = 1 \ldots 6 \) of \( A, B \) defined by lines \( l_i, i = 1 \ldots 6 \). (b) Values of some relationship \( r \) between \( A_\theta(l_i) \) and \( B_\theta(l_i) \) aggregated using some aggregation operator \( ope \) to obtain histogram values. (c) The resulting histogram. (d) Histogram calculation based on relationship between segments. \( A_\theta(l_2) = I_1 \cup I_2, \ B_\theta(l_2) = J_1 \cup J_2 \). (d) Histogram calculation based on relationship between points. (e) Histogram calculation from relationship between points/pixels.

The descriptors can also be categorised as whether they are generic or specific. Generic approaches provide theoretical frameworks for modeling relative position of objects without giving any quantitative definition of relative position. Example is the \( \mathcal{F} \) —histogram which allows the relative position to be modeled at the level of aligned point-pair, segment-pair, or core-pair. The specific methods on the other hand use particular definitions of relative position, which model relative position at specific granularity level using a particular quantitative representation. Examples of the specific methods are angle...
and force histograms. Angle histogram involves counting directionally aligned pixel-pairs of the given objects whereas force histogram uses the aggregation of force-values between the points of objects to calculate relative position.

Table 2-4 Classification of approaches to relative position description

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Pixel/point/segment/core</th>
<th>Boundary, region</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>generic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\phi)-histogram</td>
<td>point-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>f-histogram</td>
<td>segment-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>F-histogram</td>
<td>core-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td><strong>specific</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>angle histogram</td>
<td>pixel-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>force-histogram ((\phi)-histogram)</td>
<td>point-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>Allen-histogram</td>
<td>core-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>R-histogram</td>
<td>pixel-pair-based</td>
<td>boundary-based</td>
</tr>
<tr>
<td>R*-histogram</td>
<td>pixel-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>spread-histogram</td>
<td>pixel-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>visual area histogram</td>
<td>pixel-pair-based</td>
<td>boundary-based</td>
</tr>
<tr>
<td>radial line model</td>
<td>core-pair-based</td>
<td>region-based</td>
</tr>
<tr>
<td>Ratio histogram</td>
<td>core-pair-based</td>
<td>region-based</td>
</tr>
</tbody>
</table>
2.2.3 Review of existing descriptors

2.2.3.1 Angle histogram

The histogram of angles [10, 63] was the first relative position descriptor proposed. The idea was earlier introduced in [87]. The angle histogram is a pixel-pair-based, region-based, and non-generic (specific) relative position descriptor. The main principle in the angle histogram consists in representing the relative position of two objects $A$ and $B$ by the frequency distribution of the angles in the collection $\theta = \{\theta_{ij} | \theta_{ij} = \angle (a_i, b_j), a_i \in A, b_j \in B, i = 1..n, j = 1..m \}$. The descriptor $H_{\theta}(A, B)$ is thus the histogram of the number of point-pairs $f_{\theta}$ that can be associated with each angle $\theta \in \theta$. Figure 2-12 describes the principle. The histogram is given by,

$$H_{\theta}(A, B) = \{(\theta, f_{\theta})\}$$  \hspace{1cm} (2-1)

A variant of angle histogram called the quad-tree histogram was proposed in [85, 88]. In the method of quad-tree histogram, the object regions are partitioned into a number of

![Diagram of Angle Histogram](image)

Figure 2-12 Calculation of the histogram of angles

(a) Calculation of the frequencies of angles: $f_{\theta_k}$ is the frequency of the angle $\theta_k$, $f_{\theta_l}$ is the frequency of angle $\theta_l$, and so on. (b) The histogram of angles denoted a $H_{\theta}(A, B)$ is constructed from the angle frequencies.
square sub-regions. A weighted angle histogram is calculated from the angles $\alpha_{ij}$ between the centroids $\{g_i^A, g_j^B\}$ of the pairs of the sub-regions as follows.

$$QH_\theta(A, B) = \begin{cases} +\text{Area}_i^A \times \text{Area}_j^B & \text{if } \alpha_{ij} \in \left[\theta - \frac{\tau}{2}, \theta + \frac{\tau}{2}\right] \\ 0 & \text{otherwise} \end{cases}$$ (2-2)

where the $\text{Area}_i^A, i = 1..n$, and $\text{Area}_j^B, j = 1..m$, are the areas of the sub-regions of objects $A$ and $B$, and $\tau$ is the step quantization of the interval $[-\pi, \pi]$. Angle histogram is invariant to rotation and change of coordinates system. It can handle both crisp and fuzzy objects. As its computation is based on the whole object region, it is sensitive to changes in shape, size, and orientation of objects. Models of primitive directional spatial relationships (right, left, above, and below) and surround relationship based on angle histogram have been proposed. These methods are based on fuzzy methods. Two such methods are the aggregation and compatibility methods [89]. Angle histogram has a high computation cost and can represent only limited spatial relationships (Table 2-5). In particular, topological and distance relationships cannot be extracted from the angle histogram. Furthermore, the relationships “between” and “among” cannot be easily modeled and vector data cannot be handled. Although its extension to 3D images is possible, no 3D variant of angle histogram has been proposed so far. The Quad-tree variant of angle histogram suffers from same limitations as the classical angle histogram. For example, for smaller square blocks, the computation time degrades to that of the angle histogram. Moreover, the overhead of partitioning the object into sub-regions adds to the computation time.

2.2.3.2 Histogram of forces

Histogram of forces is a member of the family of F-Histograms [11, 90]. It is the type $\varphi – \text{Histogram}$, i.e., the histogram is calculated from the relation between points along the oriented sections of the two objects. The principle is shown in Figure 2-13. The histogram is a function $F_{r}^{AB}: [0, \pi) \rightarrow \mathbb{R}$ defined by,
Figure 2-13 Calculation of the force histogram. Point $p \in A_{\theta}(v)$ whereas $A_{\theta}(v) = I_1 \cup I_2$. Likewise point $q \in B_{\theta}(v)$ and $B_{\theta}(v) = J_1 \cup J_2$. The angle $\theta$ represents the direction of the line $v$.

$$F_{r}^{AB}(\theta) = \int_{-\infty}^{+\infty} F_{r}(\theta, A_{\theta}(v), B_{\theta}(v)) dv = \int_{-\infty}^{+\infty} \left( \int_{p \in A_{\theta}(v)} \int_{q \in B_{\theta}(v)} \varphi_{r}(\theta, p, q) dp dq \right) dv \quad (2-3)$$

where $r, v \in \mathbb{R}$ and points $p$ and $q$ can be seen as particles of unit mass that attract each other; $q$ exerts an elementary force on $p$ whose direction is $\tilde{\theta}$ and whose magnitude is given by,

$$\varphi_{r}(\theta, p, q) = \begin{cases} \frac{1}{|p\vec{q}|^r} & \text{if } p \neq q, \text{ and } \vec{p}\vec{q} \cdot \tilde{\theta} = 0 \\ 0 & \text{otherwise} \end{cases} \quad (2-4)$$

The reason for calling $F_{r}^{AB}(\theta)$ is due to the fact that for $r = 2$, $F_{r}^{AB}(\theta)$ can be seen as the scalar resultant of elementary gravitational forces that the points of $B$ exert on points of $A$ in direction $\theta$. Thus the closer parts of the objects receive more importance. When $r = 1$, all parts are given equal importance and the histogram can be considered as a histogram of constant forces. When $r = 0$, the force histogram behaves like an angle histogram thus generalizing and superseding the angle histogram. An equivalent definition of force histogram based on the idea of spatial correlation is given in [91]. The spatial correlation $\Psi: \vec{P} \rightarrow \mathbb{R}$ for objects $A$ and $B$ in vector plane $\vec{P}$ is defined by,
WC AB (v̂) = \int_{u̅ \in \mathbb{P}} A(\omega + u̅)B(\omega + u̅ + v̂)d\bar{u} \hfill (2.5)

Where \(\bar{u}\) and \(\bar{v}\) are two vectors and \(\omega\) is the origin of the coordinate system. The force histogram is then given by the equation (2.6).

\[ F_{r}^{AB}(\theta) = \int_{0}^{+\infty} \left[ \Psi^{AB}(k\bar{\theta})/k^{r} \right]dk \hfill (2.6) \]

Force histogram has many good properties. It can be normalized. When normalized, it is invariant to rotation, scaling, and translation. It allows the handling of complex objects, e.g., convex, concave, objects with holes, and objects with multiple disconnected components. It also makes the handling of vector objects possible [92]. Furthermore, 3D extensions of F-histogram for both raster and vector data have been proposed [93] [94]. Extraction of primitive directional relationships and that of surround, between, and among from force histogram has been explored. Reducing the handling of objects to the handling of their 1D sections reduces computational time for force histogram relative to angle histogram. The computation time for the classical form of the force histogram (equations (2-3) & (2-4)) is \(O(KN\sqrt{N})\), where \(K\) denotes the number of directions in which forces are considered and \(N\) is the number of pixels in the image. For the equivalent definition (eq., (2-5) & (2-6), the complexity is \(O(N \log N)\). A limitation of force histogram is that not all topological or distance spatial relationships can be easily extracted from the descriptor. Nevertheless, several applications of force histogram have been suggested. These include use in fuzzy scene matching [23] [8], robot navigation through linguistic spatial description derived from range sensor data [95], geospatial information retrieval and indexing [96], and recognition of graphical symbols in technical line drawing [97]. A specific method, i.e., the method of effective forces, has been proposed in [98] for extracting spatial relationships from the histogram of force. Solutions to the direct and inverse problems for force histogram have also been explored [99].
2.2.3.3  R-Histogram

R-histogram [12] is a modification of angle histogram for handling the topological spatial relationship of overlap (Figure 2-14). The difference with the angle histogram is that only pixels on the boundaries of the objects are considered. In addition to angles, $R$-histogram also uses information about the distance between the points and the condition of the occurrence of the points of one object inside the other. In this way a labelled distance measure $LD(p, q)$ is associated with each point-pair that indicates whether a point in a pair is in the overlapped or non-overlapped regions of the objects. The tuple $LD(p, q)$ has two elements; the Euclidean distance $d(p, q)$ between the points $p \in A$ and $q \in B$, where $A$ is the reference object and $B$ is the argument object; the label $l(p, q)$ indicating whether $p \in A$ is inside $B$ or $q \in B$ is inside $A$. This is illustrated in 2-14(a), where $LD(p_1, q_1) = (d_1, 0)$. The distance $d_1$ between $p_1, q_1$ is labelled 0 because neither $p_1 \in A$ is inside $B$ nor $q_1 \in B$ is inside $A$. Likewise, $LD(p_2, q_2) = (d_2, 1)$ is labelled 1 because $q_2 \in B$ is inside $A$ but $p_2 \in A$ is not inside $B$. Based on the inclusion of a point of one object in the other, the distances have 4 possible labels: $L = \{0, 1, 2, 3\}$. The histogram bin is defined as below.

$$H(I, J, L) = \begin{cases} 
(H(I, J, L) + 1 & \text{if } \theta = \angle(p, q) \in A_I \text{ and } d(p, q) \in D_J \text{ and } l(p, q) \in L \\
H(I, J, L) & \text{otherwise}
\end{cases}$$  \hspace{1cm} (2-7)

where $A_I$ is the range of angles, $D_J$ the range of distance values, and $L = \{0, 1, 2, 3\}$ is the set of labels associated with the values of distances that the bin $H(I, J, L)$ spans. The final R-histogram is the normalized histogram $RH(B, A)$ given by the equation (2.8). In the equation, $n_A$ denotes the number of angles considered and $n_D$ the range of distances.

$$RH(B, A) = \frac{H(I, J, L)}{\sum_{I'}^{n_A} \sum_{J'}^{n_D} \sum_{L'}^{3} H(I', J', L')}$$  \hspace{1cm} (2-8)
Figure 2-14 Calculation of the R-histogram. (a) $d_1$ is labelled 0 because neither $p_1$ is inside $B$ nor $q_1$ is inside $A$. (b) R-histogram is a combination of 4 sub-histograms corresponding to the 4 labels mapped to separate quadrants.

R-histogram is invariant to similarity transformations when normalized. Certain topological relations, e.g., “overlaps” and “contains”, can be extracted from R-histogram but extraction of some other topological relationships such as touch and “cover”, or “distance” relationships “far” and “near” is difficult. Vector images cannot be handled in the R-histogram and extension to 3D or fuzzy objects is possible but has not been studied. The R-histogram has a time complexity of $O(N)$ but can have the complexity of $O(N^2)$ in the worst case. The application of R-Histogram has been studied in similarity-based image retrieval and shape matching.

2.2.3.4 $R^*$-Histograms

In [86] Wang et al. proposed the $R^*$-Histogram. $R^*$-Histogram is a generalization of R-Histogram in which all pixels of the objects are considered including the boundary pixels. The pixels are selected along a set $L_\theta$ of parallel lines intersecting both the objects (Figure 2-15). A histogram $H_{l_\theta}(I,J)$ of labelled distances is generated for the set of pixels on a line.
Figure 2-15 Calculation of the R*-histogram

$l_\theta$ according to equation (2.7). Then $H_\theta(I,J)$ for pixel pairs on all lines in $L_\theta$ is calculated as $H_\theta(I,J) = \sum_{l_\theta \in L_\theta} H_{l_\theta}(I,J)$. A combined histogram $H(I,J,K)$ is built for all angles $\theta \in \Theta$ where $\Theta$ is a set of angles uniformly spaced in $[-\pi, \pi]$ as $H(I,J,K) = H_\theta(I,J)$ such that $\theta$ is the $kth$ angle in $\Theta$. The final histogram is the normalized histogram given by equation (2.8).

The R*-Histogram is invariant to translation and rotation. It can be invariant to scale if the distances are divided by the maximum distance. Due to the use of parallel lines to approximate objects, complex objects and vector images can be handled by the R*-histogram. The method can potentially be extended to 3D objects but that possibility has not been explored so far. The computational time is $O(KN \sqrt{N})$ (with $K$ the number of directions and $N$, the number of pixels) but can be $O(KN \log N)$ with an FFT based algorithm proposed by the authors. The use of R*-Histogram has been investigated in QBE based image retrieval system.

2.2.3.5 Visual area histogram

Visual area histogram [65] describes relative position under the deictic reference frame instead of the extrinsic reference frame used in angle and force histograms. In the deictic approach, the evaluation of the relative position of the argument object is based on the reference object’s own view instead of the view of an external observer.
Figure 2-16 Calculation of the visual area histogram. (a) $V_{AB}$ is the visual area of A with respect to B with corner points as $p_1, p_n, q_m$, and $q_1$. $p_i \in L_A = \{p_1, p_2, ..., p_n\}, q_j \in L_B = \{q_1, q_2, ..., q_m\}, \theta_{ij} = \angle(p_i, q_j)$ (b) Visual area histogram (VAH). $f_\theta$ is calculated from angle frequency and distance between $p_i$ and $q_j$.

Such a view is comprised of the region between the reference and the argument objects as shown by $V_{AB}$ in Figure 2-16 (a). It is thus a boundary based pixel-pair approach. The histogram is constructed by first finding the point sets, $L_A = \{p_1, p_2, ..., p_n\}$ and $L_B = \{q_1, q_2, ..., q_m\}$ of the edges of the reference and argument objects respectively. Then frequencies $f_{\theta_{ij}}$ of angles $\Theta = \{\theta_{ij}| \theta_{ij} = \angle(p_i, q_j), i = 1, ..., n, j = 1, ..., m\}$ in for all $n \times m$ point-pairs are found as in equation (2.9). Distance information is taken into account. The quantity $d$ in equation (2.9) represents the distance $d_{ij}(p_i, q_j)$ and $d_{\text{min}}$ represents the minimum distance, $d_{ij}, i = 1..n, j = 1..m$. The quantity $r$ is a real. The visual area histogram (VAH$_{AB}$) is given by the normalized histogram in equation (2.10).

\[
f_\theta = f_\theta + \frac{1}{\left(\frac{d}{d_{\text{min}}}\right)^r}, \quad \angle(p_i, q_j) = \theta, 1 \leq f_\theta \leq n \times m \tag{2-9}
\]

\[
f_\theta = f_\theta / \sum f_\theta \tag{2-10}
\]

Because only points on the adjacent boundaries are considered, the computational time for the visual area histogram is reasonable. However, it can be high if the objects have multiple disconnected components. Another limitation of the visual area histogram is that,
extension to objects with multiple disconnected components or objects with complex topologies is difficult. Furthermore, the capability of the visual area histogram to represent distance or topological relationships is limited because of the presence of an asymptote in the frequency calculation (frequency in equation (2.9) cannot be calculated when \( d = 0 \)). As can be seen, the histogram can be normalized and, therefore, can be made invariant to geometric transformations. Last, the visual area histogram can be extended to 3D. This, however, has not been investigated.

2.2.3.6 Spread histogram

The method in [21] proposes the idea of the coefficient of spread to model topological relationships such as inside, outside, encompass, and partially encompass. Given the reference object \( A \) and the argument object \( B \), the coefficient of spread \( \beta_i \) for any point \( p_i \in B \) is defined as the maximum difference \( \alpha_{m+1} - \alpha_m \) between consecutive angles \( \alpha_m, \alpha_{m+1} \) in the sorted list of angles \( (\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n) \) that the vectors \( r_{ij} \) joining the pair-points \( (p_i, q_j), p_i \in B, q_j \in A, i = 1..n, j = 1..m \) make with the \( x - axis \) (Figure 2-17). More precisely \( \beta_i = \max(\alpha_2 - \alpha_1, \alpha_3 - \alpha_2, \alpha_4 - \alpha_3, ..., \alpha_n - \alpha_{n-1}, 2\pi - \alpha_n + \alpha_1) \). In this way, if \( p_i \) is inside \( A \) and \( P_i \) is not on the border of \( A \) or \( p_i \) is encompassed by \( A \), then \( (n \to \infty) \Rightarrow (\beta \to 0) \) otherwise \( (n \to \infty) \Rightarrow (\beta \gg 0) \). The un-normalized spread histogram (S-Histogram) of the coefficient of spread \( \beta \) denoted as \( S'(A,B,\beta) \) is calculated as below.

\[
\forall i \in \{1,2,3,...,m-1\}, \quad S'(A,B,\beta) = \begin{cases} S'(A,B,\beta) + 1 & \text{if } \beta_i = \beta \\ S'(A,B,\beta) & \text{otherwise} \end{cases} \quad (2-11)
\]

The normalized spread histogram is obtained as follows.

\[
S(A,B,\beta) = \frac{S'(A,B,\beta)}{\sum_{\beta \in [0,2\pi]} S'(A,B,\beta)} \quad (2-12)
\]
Figure 2-17 Calculation of the spread histogram. (a) Coefficient of spread for some point \( p \in B, \beta = \max(\alpha_2 - \alpha_1, \alpha_3 - \alpha_2, ..., \alpha_{i+1} - \alpha_i, ..., \alpha_n - \alpha_{n-1}, 2\pi - \alpha_n + \alpha_1) \) (b) \( \beta \) represents inside-ness or outside-ness; smaller values indicating inside-ness; larger values outside-ness.

The method can be used along-side histogram of angles to evaluate special relations such as INSIDE, OUTSIDE, SURROUND, and PARTIALLY SURROUND. For example, the relationship OUTSIDE is given by,

\[
Outside(A,B) = \sum_{\beta \in [\pi,2\pi)} S(A,B,\beta) / \sum_{\beta \in [0,2\pi)} S(A,B,\beta)
\]  

(2-13)

As mentioned earlier, spread histogram can model the topological relationships of inside, outside, encompass, and partially encompass but the cost of computation can be high if used independently due to the overhead of sorting the list of angles involved. The complexity is \( O(nm + mk) \) where \( n \) and \( m \) are number of points in \( B \) and \( A \) respectively and \( k \) is the number of buckets in bucket sorting used to calculate the \( \beta \) values.

Moreover, the method cannot differentiate between INSIDE and IS ENCOMPASSED BY (SURROUNDED BY) easily and directional relationships cannot be represented until used alongside angle histogram. Also, distance relationships cannot be extracted from the spread histogram. The histogram can be normalized and the normalized histogram is rotation and scale invariant.
2.2.3.7 Ratio histogram

The method proposed in [26] introduces the notion of Ratio-histogram. The basic principle of Ratio Histogram is similar to that of \(F - Histograms\) and \(R^* - Histogram\). 2D objects are approximated by 1D longitudinal sections (Figure 2-18).

A longitudinal section can have multiple disjoint segments. However, only one direction parallel to the line joining the centroid of the object is considered instead of multiple directions as in \(R^*\) or \(F - Histograms\). The histogram is defined by,

\[
v_n = v_{\text{min}} + n \times \frac{v_{\text{max}} - v_{\text{min}}}{N}, \quad n = 0, 1, 2, ..., N
\]  

(2-14)

where \(L(E^A(v_n))\) and \(L(E^B(v_n))\) are the lengths of the longitudinal sections \(E^A(v_n)\) and \(E^B(v_n)\) of the argument and reference objects respectively along an oriented line \(l(v_n)\) with the intercept \(v_n\). The quantities \(v_{\text{min}} = \max(v_{\text{min}}^A, v_{\text{min}}^B)\) and \(v_{\text{max}} = \min(v_{\text{max}}^A, v_{\text{max}}^B)\) are the minimum and maximum intercepts of valid lines \(l(v)\), i.e., the

\[
R_{AB}(n) = \frac{L(E^A(v_n))}{L(E^B(v_n))}, \quad n = 0, 1, 2, ..., N
\]

(2-15)

Figure 2-18 Calculation of the ratio-histogram. Calculation of ratio histogram: lines \(l(v_0) \ldots l(v_n)\) parallel to the line joining the centroid \(C_A, C_B\) are considered. \(v_0 \ldots v_N\) are the intercepts of the lines. \(v_{\text{min}} = v_0, v_{\text{max}} = v_N\).
lines intersecting both the objects. \( v_{\text{min}}^A, v_{\text{max}}^A \) and \( v_{\text{min}}^B, v_{\text{max}}^B \) are the minimum and maximum intercepts of the non-empty intersections of the valid lines \( l(v) \) with \( A \) and \( B \).

The Ratio histogram is invariant to affine transformations. The computation time is low because only one direction is considered and only lengths of segments are used. However, the power to describe the relative position fully may be limited because of the use of only one direction to calculate the histogram. Extraction of spatial relationships from Ratio-Histogram is possible but has not been studied. Apparently, extension to 3D or vector objects or handling of objects with complex topologies may be easier with Ratio histogram. The use of Ratio histogram has been studied in similarity-based image retrieval and shape-matching.

2.2.3.8 Spatial template

Spatial templates were introduced by Logan et al. [100] for modeling the perceptual processes that underlie the linguistic spatial categories of \textit{above}, \textit{below}, \textit{under}, \textit{over}, \textit{left}, and \textit{right}. In spatial template, the reference-argument object configuration is represented by a \( 7 \times 7 \) grid in which the reference or \textit{land-mark object} occupies the central grid cell (cell 4,4) and the argument or \textit{trajector object} any of the remaining cells. The argument object is a point-object whereas the reference object is a two-dimensional object. A computational model is used to compute the relative position of the trajector object with respect to the landmark object for one of the spatial relationships. The resulting descriptor is visualized as a three-dimensional histogram in which the \( x \)-axis represents the columns of the grid, \( y \)-axis the rows, and the \( z \)-axis the weight that can be associated with the validity of the given relationship between the objects. Example of the spatial template for the “above” is shown in Figure 2-19(e). The template represents the degree of the “above” spatial relationship between the trajector and landmark objects for various positions of the trajector with respect to the landmark across the grid. The relationship degree is categorized as good, acceptable, or bad as shown in the figure.
Figure 2-19 Calculation of the spatial template. (a) Vectors rooted at each point of the landmark object pointing upwards towards the trajector object. (b) The attentionally weighted vectors. The vector closer to the trajector receive more attention. (c) Direction of the attentionally weighted vector-sum. (d) Orientation of the vector-sum relative to vertical upright. (e) Spatial template for the relation “above”.

Regier et al. [101] extended the work by proposing four computational models to calculate the spatial template; the bouncing-box (BB), the proximal and center-of-mass (PC) model, the hybrid model, and the attentional vector-sum model (AVM). Through experiment, they determined that the predictions by the vector-sum model better match human perception compared to prediction by other models. So here, only the Attentional Vector-sum Model (AVM) will be discussed.

The attentional vector-sum model is based on two observations; that the perception of spatial relationships in the visual field requires attention; that the overall direction (for the entire object) is the vector-sum of a set of constituent directions (of the elements or cells of the object). Based on these observations, the AVM model for the relationship “above” is defined by,
AVS model: above = \( g(\sum a_i\vec{c}) \times \text{height} \)  \((2-16)\)

In the formula, the \( a_i \) is the attention being paid to cell \( i \) of the landmark object, \( \vec{c} \) is the direction rooted at cell \( i \), and the sum taken over the cell-population is the resulting direction. Moreover, \( g(\cdot) \) is a function that gives the alignment of the resulting direction with the vertical upright direction (in the case of the spatial relationship above) and height is function that gives the height of the trajector object relative to the top of the landmark object. The method considers the trajector object (i.e., the argument object) as a point object but can be extended to 2D objects. It can be used for complex as well as 3D and fuzzy objects. It can also be potentially extended to vector objects. However, behaviour to geometric transformation is unknown.

### 2.2.3.9 Allen F-Histograms

Allen relations [102] are a set of 13 mutually exclusive jointly exhaustive relations used for reasoning about temporal phenomena. The relations are: before, meet, overlap, start, finish, during, equal, during-by, finished-by, started-by, overlapped-by, meet-by, and after, given by the set \( \mathcal{A} = \{<, m, o, s, f, d, eq, di, fi, si, oi, mi, >\} \). Allen relations are illustrated on a conceptual graph as in Figure 2-20. Matsakis et al. and Salamat et al. proposed Allen F-Histograms, which can be used either for modeling topological and directional relationships from the perspective of Allen relations or Allen relations from the perspective of directional relationships [103] [104] [105] [106]. Allen F-Histograms are in reality F-Histograms (Section 2.2.1) defined for spatiotemporal domain. Detail can be found in the cited papers. Here only a brief overview will be presented.

Given objects \( A \) and \( B \), and the direction \( \theta \), Allen F-Histogram attaches the weight \( F^A_B(\theta) \) to the proposition \( P^A_B(\theta) \equiv \"ArB in direction \theta\" \) where \( r \) is an Allen relation belonging to the set \( \mathcal{A} = \{<, m, o, s, f, d, eq, di, fi, si, oi, mi, >\} \). \( F^A_B(\theta) \) is given by,
Figure 2-20 Allen relations. The conceptual neighbourhood graph depicts the Allen relations as continuous deformation of preceding relations representing semantic closeness between the relations.

\[ F_{r}^{AB}(\theta) = \int_{-\infty}^{\infty} F_{r}(\theta, A_{\theta}(v), B_{\theta}(v)) dv \]  \hspace{1cm} (2-17)

Function \( F_{r}(\theta, A_{\theta}(v), B_{\theta}(v)) \) in turn attaches weight to the proposition “\( A_{\theta}(v)rB_{\theta}(v) \)” such that \( A_{\theta}(v) \) and \( B_{\theta}(v) \) are the longitudinal sections of \( A \) and \( B \) in direction, \( \theta \). \( F_{r}^{AB}(\theta) \) is thus evaluated from the weights attached to the given Allen proposition for the aligned sections. As Allen relations are sensitive to slight changes in the data, e.g., missing pixel from the middle or at the end of a segment or extraneous pixels, fuzzy techniques are used to handle the resulting imprecision in the evaluation of the weights. The value \( \sum_{r \in \mathcal{R}} F_{r}^{AB}(\theta) \) indicates as to what degree are the objects involved in some spatial relationships along the direction \( \theta \). If this information is considered unimportant, the \( F_{r} \)-histograms can be normalized as follows.

\[ \forall \theta \in \mathbb{R}, \overline{F_{r}^{AB}(\theta)} = \frac{F_{r}^{AB}(\theta)}{\sum_{p \in \mathcal{R}} F_{p}^{AB}(\theta)} \]  \hspace{1cm} (2-18)
For a given direction, $\theta$, the normalized $F_r$-histograms define a fuzzy 13-partition of the set of all object pairs, and each class corresponds to an Allen histogram. Figure 2-21 gives an example of the visualization of $F_r$—histograms stacked upon each other where each histogram is represented with a gray color.

The computational complexity of Allen F-histograms is $O(nN\sqrt{N})$ where $n$ is the number of direction and $N$, the number of pixels. Salamat et al. [107] proposed polygonal object approximation to reduce the number of longitudinal sections to be processed and so the computational complexity of Allen-F histograms. With the approach, the computational complexity is reported to reduce to $O(nM \log M)$ where $M$ is the number of vertices of the polygons.

Allen F-histogram can be normalized so invariance to geometric transformation like scaling, rotation, and translation is possible. Behaviour to affine transformation has not been investigated. Topological and directional spatial relationships can be easily extracted from F-histogram. The main application of Allen F-histograms is in spatiotemporal such as motion estimation where knowledge of the location of an object in space and the time of presence at a location is important.

Figure 2-21 Visual depiction of $F_r$—histograms. (a) Allen relations and attached gray-scale values. (b) A pair of objects. (c) Corresponding $F_r$-histogram.
2.2.3.10 Radial line model

Consider two objects $A$ and $B$, and a reference point $p$ determined by the minimum bounding rectangles of the objects, as suggested in Figure 2-22. Partition $(-\pi, \pi]$ into $n$ intervals $\theta_1$, $\theta_2$, etc. (the direction bins). The unbounded sector extending from $p$ and defined by $\theta_i$ intersects $A$ (respectively $B$) in some region $A_i$ (respectively $B_i$). The histogram value $H^A(i)$ (respectively $H^B(i)$) is the area of $A_i$ (resp. $B_i$) over the area of $A$ (resp. $B$). The position of $A$ relative to $B$ is represented by the pair $(H^A, H^B)$. See [25].

Radial Line Model. $H^A(i)$ is the total area of the two darker regions in $A$, divided by the area of $A$. Another option is to define $H^A(\theta)$ as the total length of the two black segments.

The Radial Line Model (RLM) targets directional and set relationships. However, extraction methods and models of such relationships based on the RLM have not been investigated. Besides, the RLM does not always allow us to determine whether two objects overlap, or whether one includes the other. The behaviour of the RLM under similarity transformations is unknown, and similarity invariance cannot be obtained. These two properties would hold, however, if the model was defined as follows: choose $p$ as the centroid of $A \cup B$; the half-line that extends from $p$ in direction $\theta$ intersects $A$ (resp. $B$) in a union (possibly empty) of pairwise disjoint segments (Figure 2-22); set the histogram value $H^A(\theta)$ (resp. $H^B(\theta)$) to the total length of these segments, and represent

![Figure 2-22 Radial line model](image-url)
the position of A relative to B by the pair \((H_A, H_B)\). The RLM has been used for graphical symbol retrieval.

2.2.4 Comparison of properties

From the review in the preceding sections, a set of properties can be established that a good relative position descriptor should aim to satisfy. These properties form a basis for comparing the different relative position descriptors. A brief description of the properties was given in Chapter I. A more detailed explanation is included here. A requirement in relative position descriptors is the satisfaction of the semantic inverse property, which states that given the descriptor for object pair \((A, B)\), it should be possible to derive the descriptor for the pair \((B, A)\). Most existing descriptors satisfy the semantic-inverse property. Easy extraction of spatial relationships from the descriptor is considered an important property of relative position descriptors. Most existing descriptors allow extraction of directional spatial relationships. Specific topological relationships can also be extracted from some descriptors. However, the extraction of distance relationships has not been addressed in any of the existing descriptors so far. Another crucial requirement is the efficient computation of the descriptor. As can be seen from Table 2-4, pixel-pair region based approaches such as angle histogram are computationally more expensive than boundary-based or section/segment based approaches.

Handling of different types of objects is also a useful property in relative position description. Such objects include 3D, fuzzy, or vector objects, overlapped objects, concave and convex objects, objects with holes, and objects with disconnected components. Most existing descriptors consider only 2D crisp raster images. Theoretically, however, all descriptors can handle 3D or fuzzy objects. Practically, extension to 3D objects has been proposed only for the histogram of forces. Likewise, vector objects have been considered only in the case of histogram of forces. Objects with holes or objects with disconnected components can be handled in section/segment based methods and overlapped objects potentially in R-histogram and histogram of forces. Invariance to geometric transformations such as similarity transformations or affine
transformations is a critical property desired in relative position descriptors. Solving three problems are useful in determining the invariant properties of a descriptor; the **normalization problem**, i.e., can the descriptor be normalized; the **direct problem**, i.e., can the descriptor for the transformed objects be derived given the transformation and the descriptor for the original objects; the **inverse problem**, i.e., can the transformation be recovered given the descriptors before and after the transformation.

The first question is important because if the descriptor can be normalized, it can possibly be made invariant to geometric transformations. Solving the second problem allows the calculation of the resulting descriptor from the original descriptor directly.

Solving the inverse problem is useful in applications like image registration and image warping. Almost all existing descriptors can be normalized. However, the direct and inverse problems have been studied only for the histogram of forces. Last, two potentially useful but unexplored properties are the recovery of the objects from the descriptor and shape description potential of a descriptor. Table 2-5 gives a comparison of the existing descriptors in terms of the above properties.

### 2.3 Summary

A review of spatial relationships and relative position descriptors was presented in this chapter. The RCC8, RCC23, and DE-4IM frameworks were discussed and other spatial relationships such as directional, distance, and surroundedness were described. A detailed review of relative position modeling together with a classification and comparison of the properties of the different descriptors was given. Spatial relationships have a natural link with relative position description. Relative position is described qualitatively in terms of spatial relationships and spatial relationships can be inferred from relative position descriptors. As can be seen from the review, although a good amount of research work exists on relative position descriptors, there is a room for work in some areas.
Table 2-5 Comparison of properties: An empty cell means that, to the author’s knowledge, the property has not been investigated, and that, as far as we can tell, there is no straightforward evidence towards the property.

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Handling of objects</th>
<th>Target relationships</th>
<th>sim(ilarities) and AFF(inities)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2D crisp raster</td>
<td>Vector (2D crisp)</td>
<td>Fuzzy (2D crisp raster)</td>
</tr>
<tr>
<td>angle histogram</td>
<td>$O(N \log N)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>force histogram m</td>
<td>$O(N \log N)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Allen histograms</td>
<td>$O(nN^2)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R-histogram</td>
<td>$O(N^2)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R*-histogram</td>
<td>$O(nN \log N)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>spread histogram</td>
<td>$O(N^2)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>visual area histogram</td>
<td>$O(N^2)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>radial line model</td>
<td>$O(N)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ratio histogram</td>
<td>$O(N)$</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Where $N$ is the number of pixels in the image and $n$, the number of directions $\theta$ considered in the case of certain histogram; sim denotes similarity transformations; AFF, the affine transformations.**
Specifically, only a limited range of spatial relationships can be extracted from existing descriptors and the extraction may not be easy. Most of the descriptors target the modeling of directional relations only. Extraction of topological relationships such RCCn relations or the surround relations have not been considered in most relative position descriptors. Other areas where there is a scope for work is relative position descriptors with low computation time and invariance to geometric transformations. Extension to 3D objects or handling of vector objects are also some potential topics for work on relative position description.
Chapter III

3 The $\Phi$-descriptor

The last chapter gave a detailed review of the existing relative position descriptors. A comparison of the properties of the descriptors was also presented (Table 2-5). A set of desirable properties for evaluating relative position descriptors was identified that included low computational time, the ability to handle both vector and raster objects (in 2D as well as 3D), efficient representation of a wider range of spatial relationships (distance, directional, topological, etc.), easy normalization, and invariance to geometric transformations. As can be seen from the Table 2-5, none of the existing relative position descriptors satisfies the full range of these properties. For example, the force-histogram can tackle both raster and vector objects but doesn’t allow the extraction of topological or distance spatial relationships. The angle histogram has a higher computational time besides having all the limitations of the force histogram and can only handle raster objects. Other descriptors such as R, R*, Spread or Visual Area histograms suffer from the same or more serious limitations. Additionally, for most of the descriptors, the realization of affine properties such as a simple affine formulation or the extraction of affine transformation is not straightforward.

In this chapter a new relative position descriptor, i.e. the $\Phi$—descriptor, is introduced. The intended aim of the new descriptor is the satisfaction of most of the above properties. Specifically, the goal is the representation of a wider range of spatial relationships and the preservation of affine invariant properties. The descriptor is based on concepts from F-histogram, Radon transform, and Allen relations. In the following sections, a detailed explanation of the theory and properties of the $\Phi$—descriptor is given. Section 3.1 gives a brief background of Radon transform, F-histogram, and Allen relations. Section 3.2 describes the main principle and theoretical framework of the new descriptor in detail. Section 3.3 discusses the key properties and capabilities of the descriptor. Section 3.4 gives the summary.
### 3.1 Background

The Radon transform, F-histogram, and Allen relations have many features that are suitable for modeling spatial relations. Both the Radon transform and F-histogram use a beam of parallel lines (along a given direction) to approximate objects in two dimensions. Such an approach has two advantages; first the handling of multi-dimensional (2D) objects is reduced to the handling of simpler one-dimensional (1D) entities; second, the directional nature of the beam naturally serves as a tool for modeling directional spatial relations. Additionally, with the selection of suitable functionales (e.g., length or areas) as the quantitative representation of spatial information, the computational complexity of the descriptor can be reduced to a lower order (linear or logarithmic, etc.). This last feature can also make it easier to design relative position descriptors with known affine properties. Last the Radon transform has a well-established theory and many nice properties. Some of these properties are linearity, shifting, scaling, rotation, convolution, etc. These properties can potentially be inherited by the new descriptor. The $\Phi$-descriptor exploits the computational approach in the Radon transform to calculate relative position descriptions at the level of 1D segments and sections. Consequently, quantities in the descriptors appear as area measures.

The F-histogram considers multiple objects (two objects) instead of the single object configuration used in the Radon transform. It is thus a natural extension of the technique of the parallel beam geometry to an object-pair. Besides, F-histogram allows modeling of the relative position of 2D objects at the level of 1D sections (one from each object) as defined by the directed lines. This match well the framework of Allen relations for modeling binary relations between temporal phenomena represented as 1D entities. A number of studies have exploited this feature to propose methods for extracting combined topological and directional spatial relations from images [103, 104, 105]. In the $\Phi$-descriptor, the framework of F-histogram has been used to describe relative position of objects. At the level of 1D sections, relationships resembling but distinct from Allen
relations have been used to model relative position. In the following three sections, a brief review of the Radon transform, F-histogram, and Allen relations is given.

3.1.1 The Radon transform

Radon transform is an integral transform that consists of the integral of a function over straight lines. It is widely used in computed axial tomography for producing cross-sectional scans of an object. The scans generate projection data that can be used to reconstruct an image of the object’s interior. Other applications of Radon transform are electron microscopy, barcode scanning, reflection seismology, and solution of hyperbolic partial differential equations [108]. Radon transform has been defined in many ways [109]. Here the definition of the Radon transform in two-dimensions will be given [110]. The geometric principle of the normal Radon transform is illustrated in Figure 3.1.

There is a beam of directed parallel lines intersecting an object defined in a two-dimensional Euclidean space. Let \( f(x,y) : \mathbb{R}^2 \to \mathbb{R} \) be some function associated with the object. The Radon transform \( g(\rho, \theta) : \mathbb{R}^2 \to \mathbb{R} \) of \( f(x,y) \) is then defined by,

\[
g(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \delta(\rho - x\cos\theta - y\sin\theta) \, dx \, dy
\]  

(3-1)

Figure 3-1 Principle of the Radon transform
where \( \delta(t) \) is the Dirac Delta function ( \( \delta(t) = 0 \) when \( t \neq 0 \) and \( \delta(t) = \infty \) otherwise). The inclusion of the Dirac Delta function in the definition causes the integration of \( f(x,y) \) along the line,

\[
\rho - x \cos \theta - y \sin \theta = 0
\]

The radial coordinate \( \rho \) denotes the length of the normal \( p \perp L \) from the origin to the line \( L \) and \( \theta \) is the angle the normal makes with the \( x \)-axis. The integral of each line is called the Radon sample and the collection of all the Radon samples gives rise to the Radon transform of the function. The Radon transform produces the projections of \( f(x,y) \) across the image at different orientations \( \theta \) and offset \( \rho \). The offset is taken relative to the parallel line that passes through the image center [110].

3.1.2 The \( F \)-histogram

The principle of \( F \)-histogram is illustrated in Figure 3-2. It is defined as follows. Let \( S \) be the Euclidean space with the arbitrary point \( \omega \) as the origin. Let \( \theta \) be a direction vector (unit vector), \( \theta(p) \) a line in direction \( \theta \) that passes through the point \( p \in S \), and \( \theta^\perp(p) \) the subspace that is orthogonal to \( \theta \) and passing through the point \( p \). Let \( A \) and \( B \) be nonempty bounded subsets of \( S \). Then \( A \cap \theta(p) \) called the core of \( A \) is a closed set with a finite number of connected components. Likewise, \( B \cap \theta(p) \) is the core of \( B \). The \( F \)-histogram associated with the pair \((A, B)\) is the function \( F^{AB}(\theta): \mathbb{R} \rightarrow \mathbb{R} \) defined by:

\[
F^{AB}(\theta) = \int_{p \in \theta^\perp(\omega)} F(\theta, A \cap \theta(p), B \cap \theta(p)) \, dp
\]

Where the integrand \( F: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) (on the right-side of the equation) is a function that takes an input of the form \((\theta, S_1, S_2)\). The intended purpose of \( F \)-histogram is to represent, in some way, the position of object \( A \) (the \textit{argument object}) with respect to object \( B \) (the \textit{reference object}).
Figure 3-2 Principle of the $F$-histogram. The core of A or longitudinal section of A, i.e., $A \cap \theta(p)$ has two components whereas the core of B or longitudinal section of B $B \cap \theta(p)$ has only one component. The $F$-histogram $F^{AB}(\theta)$ calculates the relative position of A with respect to B from the relative position of $A \cap \theta(p)$ with respect to $B \cap \theta(p)$ for all $p \in \theta^\perp(\omega)$.

The idea and assumptions behind the concept of the $F$-histogram [11] [90] are that acceptable representations of relative positions can be obtained by reducing the handling of multidimensional objects to the handling of 1D entities. The force histogram, the $R^*$-histogram, the ratio histogram and the Allen histograms mentioned in Section 1 are based on this concept.

### 3.1.3 The Allen relations

The $F$-histogram is a natural container of directional information. There is, therefore, a reason to use a fuzzy approach and select the function $F$ in a way that the real value $F(\theta, A \cap \theta(p), B \cap \theta(p))$ represents the degree to which a given topological relationship holds between $A \cap \theta(p)$ and $B \cap \theta(p)$. Assume there are $n$ possible topological relationships that hold between two cores, then $n$ histograms can be derived (one per relationship) that should provide maximum quantitative information on the directional and topological relationships between $A$ and $B$. Unfortunately, the number of binary relations that can be defined in the algebra generated by unions of segments on a directed line is infinitely large [111]. For example, when there are only a segment and the union of two disjoint segments involved, the topological relationships are already 40 [44].
The 13 Allen relations between two aligned segments (Allen, 1983).

<table>
<thead>
<tr>
<th>Relation</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precedes</td>
<td>P</td>
</tr>
<tr>
<td>Preceded by</td>
<td>Pi</td>
</tr>
<tr>
<td>Meets</td>
<td>M</td>
</tr>
<tr>
<td>Met by</td>
<td>Mi</td>
</tr>
<tr>
<td>Overlaps</td>
<td>O</td>
</tr>
<tr>
<td>Overlapped by</td>
<td>Oi</td>
</tr>
<tr>
<td>Starts</td>
<td>S</td>
</tr>
<tr>
<td>Started by</td>
<td>Si</td>
</tr>
<tr>
<td>During</td>
<td>D</td>
</tr>
<tr>
<td>Contains</td>
<td>Di</td>
</tr>
<tr>
<td>Contains</td>
<td>Fi</td>
</tr>
<tr>
<td>Equals</td>
<td>E</td>
</tr>
<tr>
<td>F = Finish by</td>
<td>F</td>
</tr>
</tbody>
</table>

In each case, the argument is the light gray segment and the referent is the dark gray segment.

It then is more realistic to rely on the well-known 13 Allen relations and avoid a combinatorial explosion of relations between two segments [102]. See Figure 3-3. For every segment (i.e., connected component) $I$ of $A \cap \theta(p)$ and for every segment $J$ of $B \cap \theta(p)$, the value $F(\theta,I,J)$ then represents the degree to which $I$ and $J$ have a given Allen relation (fuzzified). Finally, the value $F(\theta,A \cap \theta(p),B \cap \theta(p))$ can be considered as an aggregate of all the $F(\theta,I,J)$ values.

### 3.2 The Φ-descriptor

The Φ-descriptor denoted as $\Phi^{AB}$ is a $n$-tuple of $F$-histogram together with other quantitative measures. Each $F$-histogram corresponds to a region of the interaction of object A and B in some direction. More specifically, an $F$-histogram represents the area (2D) or volume (3D) of the region of interaction. The Figure 3-4 illustrates the physical interpretation of the Φ-descriptor. The regions colored in dark, gray, medium-gray, or dotted-gray represent the regions of interaction of A and B (in some direction $\theta$). An interaction is a spatial relation between the cores of A and B, i.e., between $A \cap \theta(p)$ and $A \cap \theta(p)$ respectively.
Figure 3-4 Interpretation of the $\phi$-descriptor. A $\phi$-descriptor is a collection of $F$-histograms. Each $F$-histogram corresponds to the area of a region in which objects $A$ and $B$ interact in some way along the direction $\theta$. Examples of the regions are the regions with dark-gray, medium-gray, gray, and dotted-gray color.

In the following sections, the detailed definition of the interactions and the regions they give rise to is given. The $\phi$-descriptor is formally defined in Section 3.2.6.

### 3.2.1 Defining boundary points

Let $A$ be the argument object and $B$ be the reference object. Further, let $A$ and $B$ be defined as finite bounded sets in 2D Euclidean space $S$, and $L$ be a set of parallel lines intersecting $A$ and $B$ in direction $\theta$. Let $L \in L$ be a line. This is illustrated in Figure 3-5(a).

Note that for the sake of clarity, the figure only shows $L$. Then the intersection of $L$ with $A$ and $B$ consists of a finite set of directed components and each component is a line segment. The line $L$ has components $\overline{p_1p_2}, \overline{p_2p_3}$ and so on. Let $p$ and $q$ be the end-points of one such segments. If $p \neq q$ and $\overline{pq}/|\overline{pq}| = \theta$, where $\overline{pq}$ denotes the vector from $p$ to $q$ and $|\overline{pq}|$ denotes the length of this vector, then $p$ is an $A$ – entry (or $A$ – exit, or $B$ – entry, or $B$ – exit) on $L$ in direction $\theta$ and $q$ is an $A$ – exit (or $A$ – entry, or $B$ – entry, or $B$ – exit).

For example, in the case of the line segment $\overline{p_1p_2}$, $p_1$ is $B$ – entry and $p_2$ is $A$ – entry. The point $p_2$ is the successor point to $p_1$. Likewise, for the segment $\overline{p_2p_3}$, $p_2$ is $A$ – entry and $p_3$ is $B$ – exit. Moreover $p_3$ is the successor to $p_2$. 
Figure 3-5 Notation and terminology. (a) The line $L$ intersects the objects $A$ and $B$ in direction $\theta$. $p_1..p_6$ are entry and exit points into/from $A$ and/or $B$ on $L$ in direction $\theta$. Point $p_2$ is the successor of $p_1$, $p_4$ the successor of $p_3$, and so on. $p_1$ is B-entry, $p_2$ is A-entry, $p_3$ is B-exit in $A$, $p_4$ is A-exit, and so on. (b) $p$ is B-exit in $A$ whereas $q$ is both A-exit and B-entry.

3.2.2 Categories of boundary points

Let $\{p_1, p_2, ..., p_n\}$ be the set of all the end-points of the segments of $L \cap (A \cup B)$ in direction $\theta$. Let $\frac{|\overrightarrow{p_ip_{i+1}}|}{|\overrightarrow{p_{i}p_{i+1}}|} = \bar{\theta}$ for any $i$. Let $(p, q)$ be the end-points of a line-segment such that $p$ corresponds to point $p_i$ and $q$ corresponds to $p_{i+1}$. For clarity, let $p_{i+1}$ be called the successor of $p_i$. Then there can be one or more of the following patterns incident on an end-point; A-entry, B-entry, A-exit, B-exit, in A, in B, not in A, and not in B.

Based on the entry-exit pattern incident on $p_i$ (or $p_{i+1}$) and the occurrence of $p_i$ (or $p_{i+1}$) in A or B, twelve categories of the end-points of line-segments are defined. These categories are shown in Figure 3-6. Example of the category "A - entry not in B" point is the point $p_5$ (one of the two end-points of $\overrightarrow{p_4p_5}$) in Figure 3-5(a). It is in the "A - entry not in B" category because it has A - entry pattern incident on it and it is not inside B. On the other hand $p_2$ (an end-point of $\overrightarrow{p_2p_3}$) belongs to the category "A - entry in B" because the pattern A - entry is incident on it and it lies in (the interior of) B. Likewise, $p_3$ (the other end-point of $\overrightarrow{p_2p_3}$) belongs to "B - exit in A", $p_4$ is of type "A - exit not in B", $p_5$ again is "A - entry not in B" and so on.
Figure 3-6 Boundary points and their categories. The light-gray segment represents a part of A and the dark-gray segment a part of B. In “A-entry not in B”, the end-point of the line-segment is an A-entry that is not in B. In A-entry B-entry, the point is both A-entry and B-entry. The other categories have similar meanings.

3.2.3 Grouping boundary points into point-pairs

In the next step, end-points are grouped into pairs and categories are assigned to the pairs according to the exit-entry patterns the points exhibit. This results into 36 possible categories of point-pairs. These categories are shown in Figure 3-7 (numbered 1 through 36). The 36 pairs are further organized into 9 groups of 4 categories each. Each group is labelled according to the exit-entry pattern it represents. For example, all the points in the grouping labelled A-A belong to the pattern “A-entry, A-exit” or “A-exit, A-entry”. This means for a given point pair \((p, q)\) such that \(q\) is the successor and \(p\) the predecessor, \(p\) is of type \(A - entry \ in \ B\) or \(A - entry \ not \ in \ B\) whereas \(q\) is of type \(A - exit \ in \ B\) or \(A - exit \ not \ in \ B\). Or conversely \(p\) can be \(A - exit \ in \ B\) or \(A - exit \ not \ in \ B\) and \(q\) can be \(A - entry \ in \ B\) or \(A - entry \ not \ in \ B\). For example, the point-pair \((p_4, p_5)\) that contains the end-points of the segment \(p_4p_5\) is A-A type because \(p_4\) is \(A - exit \ not \ in \ B\) and \(p_5\) is \(A - entry \ not \ in \ B\). Similarly, the pair \((p_5, p_6)\) is also A-A type because \(p_5\) is \(A - entry \ not \ in \ B\) and \(p_6\) is \(A - exit \ not \ in \ B\). On the contrary consider the pair \((p_2, p_3)\) which is A-B because \(p_2\) is \(A - entry \ in \ B\) and \(p_3\) is \(B - exit \ in \ A\).
Figure 3-7 The 36 point-pair categories
Next descriptions are assigned to the point-pairs. These descriptions represent relationship types between the segments of A and B. For example, the pair \((p_2, p_3)\) corresponding to the segment \(p_2p_3 \in A\) and \(p_2p_3 \in B\) can be assigned the description “overlap” because the exit-entry pattern it represents defines the overlap relation (or interaction type) between the segments. As can be seen, there are ten distinct descriptions assigned (e.g., start, overlap, trails, etc.). Each description represents a relationship category.

3.2.4 Defining segment relations

In the next step the point-pair categories (Figure 3-7) are divided into two groups. In the first group are included the point-pairs that give rise to some useful relationship between the segments of A and B. For example, in the pair of points numbered 3 (belonging to the group A-A), the configuration “A-entry, in B” followed by “A-exit, in B” results in the “overlaps” relationship between the segments. Likewise, in the AB-AB group, the sequence “A-entry, B-entry” followed by “A-exit, B-exit” (numbered 33) gives rise to the start relations or interaction between the segments of A and B. These relationships are assigned names that represent linguistic verb expressions such as “overlaps”, “starts”, etc. Furthermore, redundant or directionally inverse (e.g., lead, follow) relations are ignored. The number of useful relationships is thus reduced to seven. The final set of relationships includes “overlaps”, “leads”, “trails”, “follows”, “covers”, “uncovers”, and “starts” (Figure 3-8). In the second group are included the point-pairs that don’t represent any useful spatial relationship between the respective segments. For example, in the group A-A, the point sequence “A-entry, not in B” followed by “A-exit, not in B” (numbered 1) only defines a segment of A and doesn’t represent any meaningful relationship between segments of A and B. There are three such configurations described by the noun expressions “argument”, “referent”, and “void” (Figure 3-9). To these two sets of categories is added a final set of three more patterns; that is “encloses”, “divides”, and “width” (See Figure 3-10).
3.2.5 Defining the F-histograms

Let $A$ and $B$ be two objects and $L$ a line intersecting $A$ and $B$ in direction $\theta$. Let $S = \{p_1, p_2, ..., p_n\}$ be the set of all $A$ – entries, $A$ – exits, $B$ – entries, and $B$ – exits defined on $L$ in direction $\theta$.

![Figure 3-8 Nonzero values for the functions $f_t, f_o, f_c, f_u, f_p, f_i$, and $f_s$.](image-url)
As explained in the preceding section, the grouping of the end-points of the segments of \( L \) gives rise to different spatial relationships (or interaction types) between the 1D segments of A and B that the (directed) line \( L \) defines. These relations have been shown in the Figure 3-8, Figure 3-9, and Figure 3-10. Next with each relationship type is associated a real function \( f \) that takes the triplet \((\theta, p, q)\) as an input and maps it to a real value. Consequently, the functions \( f_a, f_c, f_i, f_o, f_r, f_s, f_v, f_u \) are defined that are associated with the relations argument, covers, follows, leads, overlaps, referent, starts, trails, void, and uncovers respectively. Each function \( f \) is calculated as follows. A function \( f(\theta, p, q) \) is 0 if the \((p, q)\) does not represent the pattern \((p_i, p_{i+1})\) associated with the relation that the function \( f \) corresponds to. Otherwise, \( f(\theta, p, q) \) is calculated as \( f(\theta, p, q) = |pq| \) (and \( f(-\theta, p, q) = 0 \)) if the category is not its own directional inverse. If the directional inverse of \( f(\theta, p, q) \) does not exist and the category is its own directional inverse, then \( f(\theta, p, q) = |pq|/2 \) (and \( f(-\theta, p, q) = |pq|/2 \)).
Three other functions denoted as $f_e, f_d, f_w$ are also defined. These functions are associated with the encloses, divides, and width interaction types. The function $f_e(\theta, p, q) = 0$ is calculated as follows. $f_e(\theta, p, q) = B \cap [p_i, p_{i+1}]$ if $(p, q)$ is the pair $(p_i, p_j)$ such that $j > i$ and $p_i$ is $A$–exit and $p_j$ is $A$–entry and no point $p_k, i < k < j$ is $A$–exit or $A$–entry; $f_e(\theta, p, q) = 0$ otherwise. Likewise, the value $f_d(\theta, p, q) = 0$ until $(p, q)$ is the point $(p_i, p_{i+1})$ such that $p_i$ is $B$–exit and $p_{i+1}$ is $B$–entry, and no other point between them is either a $B$-exit or a $B$-entry in which case $f_d(\theta, p, q) = B \cap [p_i, p_{i+1}]$. Finally, the function $f_w(\theta, p, q) = 0$ unless $(p, q)$ is the point-pair $(p_1, p_n)$ in which case $f_w(\theta, p, q) = [p_1, p_n]$.

Based on the functions defined above, thirteen new functions, $F_a, F_c, F_d, F_e, F_f, F_l, F_o, F_r, F_s, F_t, F_u, F_v$, and $F_i$ are defined as follows. A function $F(\theta, A \cap L, B \cap L) = 0$ if the set of all $A$–entry, $A$–exit, and $B$–entries, $B$–exits is empty; otherwise,

$$F_a(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_a(\theta, p_i, p_{i+1})$$  \hspace{1cm} (3.4)

$$F_c(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_c(\theta, p_i, p_{i+1})$$ \hspace{1cm} (3.5)

$$F_f(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_f(\theta, p_i, p_{i+1})$$ \hspace{1cm} (3.6)

$$F_l(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_l(\theta, p_i, p_{i+1})$$ \hspace{1cm} (3.7)

$$F_o(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_o(\theta, p_i, p_{i+1})$$ \hspace{1cm} (3.8)

$$F_r(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_r(\theta, p_i, p_{i+1})$$ \hspace{1cm} (3.9)
\[ F_s(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_s(\theta, p_i, p_{i+1}) \quad (3-10) \]

\[ F_t(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_t(\theta, p_i, p_{i+1}) \quad (3-11) \]

\[ F_u(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_u(\theta, p_i, p_{i+1}) \quad (3-12) \]

\[ F_v(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} f_v(\theta, p_i, p_{i+1}) \quad (3-13) \]

\[ F_e(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f_e(\theta, p_i, p_j) \quad (3-14) \]

\[ F_d(\theta, A \cap L, B \cap L) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f_d(\theta, p_i, p_j) \quad (3-15) \]

\[ F_i(\theta, A \cap L, B \cap L) = f_w(\theta, p_1, p_n) \quad (3-16) \]

Next, a histogram of each of the above functions is defined. For example, the histogram of \( F_v(\theta, A \cap L, B \cap L) \) is defined as,

\[ F_v^{AB}(\theta, A, B) = \int_{p \in \theta^\perp(\omega)} F_v(\theta, A \cap L, B \cap L) dp \quad (3-17) \]

where \( \theta^\perp(\omega) \) is the direction perpendicular to \( \theta \) and passing through the origin \( \omega \) (see Figure 3.2). In this way, 13 F-histograms are defined that are denoted by \( F_v^{AB}, F_a^{AB}, F_t^{AB}, F_e^{AB}, F_d^{AB}, F_o^{AB}, F_s^{AB}, F_u^{AB}, F_c^{AB}, F_f^{AB}, F_l^{AB}, F_i^{AB} \) and \( F_i^{AB} \). Each histogram corresponds to the area (or volume in 3D case) of a part of the region in which \( A \) and \( B \) are aligned along some direction \( \theta \). The histogram \( F_i^{AB} \) gives the area of the total region of the \( \theta \)-alignment.
of objects A and B (called here the region of interaction of objects A and B in direction \( \theta \)). These histograms are, therefore, called *area F-histograms*.

### 3.2.6 The \( \Phi \)-descriptor

#### 3.2.6.1 Length histograms

A variety of length histograms can be derived from the histograms of areas. Width histograms are one type of length histograms. A width histogram represents the average width of the area corresponding to an area F-histogram. In the definition of the \( \Phi \)-descriptor that follows, only the width histogram \( W_{i}^{AB}(\theta) \) corresponding to the area-histogram \( F_{i}^{AB}(\theta) \) is used. It is calculated as follows. Let \( \delta: \mathbb{R} \rightarrow \mathbb{R} \) be a function such that \( \delta(x) = 0 \) if \( x = 0 \) and \( \delta(x) = 1 \) for \( x \neq 0 \), then the length histogram \( W_{i}^{AB}(\theta) \) is given by,

\[
W_{i}^{AB}(\theta) = \frac{\int_{p \in \theta \perp \omega} \delta\left(F_{i}(\theta, A \cap \theta(p), B \cap \theta(p))\right)dp}{\int_{p \in \theta \perp \omega} \delta\left(F_{i}(\theta, A \cap \theta(p), B \cap \theta(p))\right)dp}
\]  

(3-18)

\( W_{i}^{AB} \) represents the average width of the region of interaction of A and B along direction \( \theta \). Actually, it is the average non-zero value of \( f_{w} \) in direction \( \theta \). As can be seen from equation (3.17), \( W^{AB} \) is undefined for a \( \theta \) when \( F_{i}(\theta, A \cap L, B \cap L) = 0 \). Width histograms, e.g., \( W_{t}^{AB} \), \( W_{o}^{AB} \), \( W_{s}^{AB} \), corresponding to the area histograms \( F_{t}^{AB} \), \( F_{o}^{AB} \), \( F_{s}^{AB} \), can be derived in the same manner. Like \( W_{i}^{AB} \), each of these histograms represents the average non-zero value of the \( f \) (i.e., the \( f_{t} \), \( f_{o} \), \( f_{s} \), etc.) function and is undefined when the corresponding \( F \) (i.e., \( F_{t}^{AB} \), \( F_{o}^{AB} \), \( F_{s}^{AB} \), etc.) is zero. Similar to the width histogram, *height histograms* \( H_{t}^{AB}(\theta) \), \( H_{o}^{AB}(\theta) \), \( H_{s}^{AB}(\theta) \), etc. can be defined. These histograms can be derived from area and width histograms. For example, the height histogram \( H_{t}^{AB}(\theta) \) can be defined as \( H_{t}^{AB}(\theta) = F_{t}^{AB}(\theta)/W_{t}^{AB}(\theta) \), \( H_{o}^{AB}(\theta) \) as \( H_{o}^{AB}(\theta) = F_{o}^{AB}(\theta)/W_{o}^{AB}(\theta) \), \( H_{s}^{AB}(\theta) \) as \( H_{s}^{AB}(\theta) = F_{s}^{AB}(\theta)/W_{s}^{AB}(\theta) \), and so on. In case an \( F^{AB}(\theta) = 0 \), \( W^{AB}(\theta) \) and \( H^{AB}(\theta) \) are set to zero. Finally, other length histograms such *minimum and maximum width*
histograms can be used for the Φ-descriptor that can be calculated from minimum and maximum non-zero f values in direction θ.

### 3.2.6.2 Defining the ϕ-descriptor

The Φ-descriptor is defined as a tuple Φ_{AB} of area F—histograms (in the 2-dimensional case) or volume F—histograms (in the 3-dimensional case) and the length histograms. It can be used to describe the relative position of the argument object (object A) with respect to the reference (object B). A definition of the Φ-descriptor is given in equation (3.19) although other definitions including additional length histograms are also possible. In the equation, |A| and |B| represent the areas (in the case of 2D objects) or volumes (in the case of 3D objects) of objects A and B respectively. The physical interpretation of the descriptor has been illustrated through Figure 3-11.

$$\Phi_{AB} = (F^A_d, F^A_c, F^A_e, F^A_f, F^A_l, F^A_o, F^A_r, F^A_s, F^A_t, F^A_u, F^A_v, F^A_i, W^A_i, |A|, |B|)$$

(3-19)

Figure 3-11 The meanings of the Φ-descriptor. Each histogram corresponds to a part of the region in which objects A and B are aligned along direction θ (and −θ) in (a), (b), and (c), such regions are shown by dotted, lined, and solid regions in dark, gray, and medium-gray colors. The regions are labelled d, e, etc. The region labelled d corresponds to the F-histogram $F^A_d$, the one labelled e to $F^A_e$, and so on.
3.3 Properties

Consider the pair of objects in Figure 3-12. In the figure, a disconnected object A surrounds a simple connected object B. The histogram representation of the relative position of A and B is given in Figure 3-13. The descriptor was generated using 360 reference directions. Only the histograms $F_{e}^{AB}$, $F_{a}^{AB}$, $F_{r}^{AB}$, and $F_{t}^{AB}$ corresponding to the interactions “encloses”, “argument”, “referent”, and “trails” are shown here.

Figure 3-12 Disconnected object A surrounds B.

Figure 3-13 Histogram visualization of the functions $F_{e}^{AB}$, $F_{r}^{AB}$, $F_{a}^{AB}$, $F_{t}^{AB}$. $F_{e}^{AB}$ corresponds to the “encloses” category, $F_{r}^{AB}$ to the “referent” category, $F_{a}^{AB}$ to the argument, and $F_{t}^{AB}$ to trails.
The $\Phi$-descriptor has many useful properties. In the following sections, a brief description of some of its properties is given. Detailed proof of these properties will be presented in the future work.

### 3.3.1 Handling of both raster and vector objects

Both vector and raster objects can be handled in the $\Phi$-descriptor framework. Moreover, the handling is simple. For any direction $\theta$, handling of raster objects involves partitioning the objects into parallel raster lines along direction $\theta$ and computing the histograms $F^{AB}(\theta)$ and $F^{AB}(-\theta)$ by processing the raster lines one at a time. For vector objects, the histograms can be calculated from the areas of the polygons defined by the boundaries of the objects and the direction $\theta$. The objects in both cases can have arbitrary shapes. The $\Phi$-descriptor has a low computational complexity (See Chapter VII). The efficiency comes from the fact that updating a histogram value only requires associating a pixel with anyone histogram according to the definition (of the histograms) given in the previous sections. Thus as each pixel in a raster line is examined, all the histogram values can be updated simultaneously. This reduces the task of computing the histograms to counting the pixels between the exit and entry points in different raster lines and updating the values of different histograms on the fly. If $K$ is the number of directions and $N$ is the number of pixels in both objects, the computational time of the descriptor for raster objects has then the linear complexity $O(KN)$. As can be seen, the time depends on the number of directions for fixed size objects. It may be possible to determine an adequate number of directions for the performance to be optimum. However, in most practical cases, only a limited number of directions is required to calculate the descriptor without any significant loss of information [90]. The computational time (CPU time) as a function of image size ($N$) and number of reference directions ($K$) is shown in Figure 3.14. The linear nature of the time growth is evident from the graph.

For the vector objects, the computation of histograms $F^{AB}(\theta)$ and $F^{AB}(-\theta)$ will depend on the total number of vertices of the polygons into which the objects are partitioned.
If $\eta$ is the number of object vertices, then the computational time of the $\Phi$-descriptor is $O(K\eta^3)$. However, this is the worst-case performance, which improves to $O(K\eta^2)$ as number of vertices decreases to $\eta$ (objects intersecting in $\eta$ points is typical in practice).

### 3.3.2 Extraction of spatial relationships

The $\Phi$-descriptor encapsulates rich spatial relations information. It thus allows the modeling of a wide variety of spatial relationships. Both qualitative and qualitative models of spatial relationships can be derived from the $\Phi$-descriptor. Examples of relationships that can be extracted include the distance relationships, surround relationships, and directional relationships. Furthermore, set relationships representing the intersection of objects or inclusion of one object into another and topological non-set relationships describing the intersection between the boundaries or interior of objects can be easily interpreted from the $\Phi$-descriptor.

Some models of topological, directional, surround, and distance relationships based on the $\Phi$-descriptor have been proposed in the following chapters. Chapter IV presents some simple methods for extracting fuzzy and crisp $RCC_n(RCC_8$ and $RCC_23)$ and DE-4IM relations from the $\Phi$-descriptor. Chapter V explores the extraction of directional spatial relationships from the descriptor whereas Chapter VI that of the surround relationship.
3.3.3 Histogram comparison

The similarity of the relative position of two objects can be assessed by comparing their Φ descriptors. A simple way to accomplish this will be to compare the corresponding components \( h_1(\theta) \) and \( h_2(\theta) \) (representing, e.g., \( F_t^{AB} \) and \( F_t^{A'B'} \), etc.) of the descriptors \( \Phi^{AB} \) and \( \Phi^{A'B'} \) associated with the object pairs \( (A,B) \) and \( (A',B') \) respectively. Many similarity measures can be considered. An appropriate measure introduced by [112] is given by,

\[
\frac{\sum_\theta \min(h_1(\theta), h_2(\theta))}{\sum_\theta \max(h_1(\theta), h_2(\theta))}
\]  

(3-20)

The same measure can be used to calculate the similarity between the areas \( a_1 \) and \( a_2 \) of two objects. The formulation would be as follows.

\[
\frac{\min(a_1, a_2)}{\max(a_1, a_2)}
\]  

(3-21)

Finally, the similarity between \( \Phi^{AB} \) and \( \Phi^{A'B'} \) can be defined as the minimum similarity between the corresponding histograms and object areas.

3.3.4 Affine properties

Invariance to geometric transformations is one of the key requirements in image descriptors. It is valuable in computer vision tasks like object and pattern-recognition. For this reason, a great deal of attention has been paid in literature to the design of image descriptors (e.g., shape, texture, and color) that are invariant to geometric transformations. Usually, invariance to similarity transformation is considered important. However, invariance to affine transformations is considered more desirable from the point of view of handling complex transformations. The Φ-descriptor due to its particular design has good affine properties. These properties are described below.
3.3.4.1  Behavior to affine transformations

As the histograms values in the $\Phi$-descriptor are computed from areas (of regions) and areas are scaled by the absolute value of the determinant of the linear part of an affine transformation, the $\Phi$-descriptor has a pre-determined affine behavior. This can be stated as follows. Let $aff$ be an invertible affine transformation and $\Phi^{AB}$ be the relative position descriptor of objects A and B. Then each element of $\Phi^{AB}$ is scaled by the absolute value of the determinant of the linear part of $aff$. For example, for $F^{AB}_e$, the following holds.

$$F^{aff(A)aff(B)}_e = |m|F^{AB}_e$$  \hspace{1cm} (3-22)

Where $m$ is the value of the determinant of the linear part of $aff$. The same holds for $F^{AB}_d$, $F^{AB}_c$, $F^{AB}_f$, and so on. Thus if $aff$ is known, the above property allows calculation of the resulting descriptor (i.e., $\Phi^{aff(A)aff(B)}$) directly from the original descriptor $\Phi^{AB}$ (and vice versa) instead of calculation from the objects.

3.3.4.2  Normalization

The $\Phi$-descriptor can be easily normalized. A simple procedure for normalizing the $\Phi$—descriptor descriptor is as follows. Let $(A, B)$ be a pair of well-behaved objects. Then there exists a linear transformation $lin$ such that,

$$\overline{\Phi^{AB}} = \Phi^{lin(A)lin(B)}$$  \hspace{1cm} (3-23)

Moreover, for the following condition holds for any pair of objects that is well-behaved and any affine transformation $aff$ that is invertible.

$$\Phi^{aff(A)aff(B)} = \overline{\Phi^{AB}}$$  \hspace{1cm} (3-24)

In other words, the normalized $\Phi$-descriptor is affine invariant. The basic principle underlying the $\Phi$-descriptor normalization is to extract from $\Phi^{AB}$ a vector basis intrinsic to the pair $(A, B)$. The uniqueness of $lin$ is the consequence of the uniqueness of the
transformation used to change the vector-basis of one pair into that of another. The procedure for normalization uses the width histogram $\overline{F_i^{AB}}$ and the fact that the $\Phi$-descriptor has a known behaviour to affine transformation (see equation 3.21).

### 3.3.4.3 Inference of affine transformation

If there are two pairs of well-behaved objects $(A, B)$ and $(A', B')$ and an invertible affine transformation $\text{aff}$ such that,

$$A' = \text{aff}(A) \quad \text{and} \quad B' = \text{aff}(B)$$  \hfill (3-25)

then $\text{aff}$ can be obtained from $\Phi^{AB}$ and $\Phi^{A'B'}$ (up to a translation) with the normalization procedure described in the previous section. Further, let $t$ be an invertible transformation (not necessarily affine) such that,

$$A' = t(A) \quad \text{and} \quad B' = t(B)$$  \hfill (3-26)

then the linear transformation that best approximates $t$ (up to translation) can be found and the quality of approximation assessed e.g., through a similarity measure described in Section 3.3.3.

### 3.3.5 Other properties

The $\Phi$-descriptor allows retrieval of semantic inverse, i.e., the position of $B$ relative to $A$ ($\Phi^{BA}$) can be derived from the position of $A$ relative to $B$ ($\Phi^{AB}$). Then Table 3-1 describes the rules for semantic inverse retrieval. Moreover, histograms $F_v^{AB}(\theta), F_a^{AB}(\theta), F_r^{AB}(\theta), F_e^{AB}(\theta), F_d^{AB}(\theta)$, and $F_i^{AB}(\theta)$ are even such that $F_v^{AB}(\theta) = F_v^{BA}(-\theta), F_a^{AB}(\theta) = F_a^{BA}(-\theta)$, and so on. The same applies to the corresponding width histograms $W_v^{AB}, W_a^{AB}, W_r^{AB}, W_e^{AB}, W_d^{AB}$, and $W_i^{AB}$ and height histograms $H_v^{AB}, H_a^{AB}, H_r^{AB}, H_e^{AB}, H_d^{AB}$, and $H_i^{AB}$.

A number of other measures can be derived from the descriptor. Let $F_{\delta}^{AB}(\theta)$ denote $F_{\delta}^{AB}(-\theta), F_{\epsilon}^{AB}(\theta)$ denote $F_{\epsilon}^{AB}(-\theta), F_{\gamma}^{AB}(\theta)$ denote $F_{\gamma}^{AB}(-\theta)$, and so on. Further, let $F_{|a|}^{AB}$
denote \( F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(-\theta) \) and so on. The quantities \( F_{[\theta]}^{AB}(\theta), F_{[\theta]}^{AB}(\theta), F_{[\theta]}^{AB}(\theta), \) and \( F_{[\theta]}^{AB}(\theta) \) as defined by equations (3.27) - (3.29) can be obtained from the \( \Phi \)-descriptor. The quantity \( F_{[\theta]}^{AB}(\theta) \) represents the sub-region (of the region of interaction in direction \( \theta \)) occupied by both objects A and B whereas \( F_{[\theta]}^{AB}(\theta) \) denotes the sub-region occupied by object A (See Figure 3.11). Corresponding to \( F_{[\theta]}^{AB}(\theta) \) and \( F_{[\theta]}^{AB}(\theta) \), the measures \( W_{[\theta]}^{AB}(\theta), W_{[\theta]}^{AB}(\theta) \) and \( H_{[\theta]}^{AB}(\theta) \) can be defined that give the widths and heights of these regions. Likewise, the area measures \( F_{[\theta]}^{AB}(\theta), F_{[\theta]}^{AB}(\theta), \) and \( F_{[\theta]}^{AB}(\theta) \) are associated with the sub-regions of the region of interaction in direction \( \theta \) that are neither occupied by A nor B, occupied by only A, occupied by only B, and occupied by A or B respectively. The corresponding width and height measures are \( W_{[\theta]}^{AB}(\theta), W_{[\theta]}^{AB}(\theta), \) etc. and \( H_{[\theta]}^{AB}(\theta), H_{[\theta]}^{AB}(\theta) \) etc.

\[
F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(-\theta) \\
(3-27)
\]

\[
F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) \\
(3-28)
\]

\[
H_{[\theta]}^{AB} = (H_{[\theta]}^{AB}(-\theta) + H_{[\theta]}^{AB}(-\theta)) \\
(3-29)
\]

\[
F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(\theta) + F_{[\theta]}^{AB}(-\theta) \\
+ F_{[\theta]}^{AB}(-\theta) + F_{[\theta]}^{AB}(-\theta) + F_{[\theta]}^{AB}(-\theta) + F_{[\theta]}^{AB}(-\theta) \\
(3-30)
\]

Examples of other measures that can be defined are \( H_{[\theta]}^{AB}(\theta) \), etc.

<table>
<thead>
<tr>
<th>Table 3.1 Semantic inverse retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(-\theta) )</td>
</tr>
<tr>
<td>( F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) )</td>
</tr>
<tr>
<td>( F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) )</td>
</tr>
<tr>
<td>( F_{[\theta]}^{AB}(\theta) = F_{[\theta]}^{AB}(\theta) )</td>
</tr>
</tbody>
</table>

3.4 Summary

A new relative position descriptor was introduced in this chapter. The descriptor is based on the original idea of categorizing pairs of consecutive boundary points of objects on a directed line to define spatially related regions of the objects. With each region is
associated an F-histogram that either represents its area (if 2D) or volume (if 3D) and the collection of histograms for all the regions gives rise to the relative position descriptor of the objects. The approach exploits modeling concepts from Radon transform and Allen relations. It has many advantages as described in Section 3.3. Calculation of the descriptor is fast (linear time) and it is possible to extract a broader range of spatial relationships from the descriptor. Some relationship types that can be extracted from the Φ-descriptor are the topological relationships, directional relationships, and distance relationships. Other useful features of the descriptor are easy normalization, a known behavior to affine transformations, and the possibility to extract the affine transformation from the descriptor. In the following chapters, models of topological (Chapter IV), directional (Chapter V), and surround relationships (Chapter VI) based on the Φ-descriptor would be proposed. Future work on the descriptor has been discussed in Chapter VIII.
Chapter IV

4 Topological relations

The well-known DE-4IM and RCC8 are sets of topological relations. The RCC23—an extension of the RCC8—is a set of topological and convex hull based relations. See sections 2.1.1 and 2.1.2. In this chapter, it will be shown that the DE-4IM, RCC8 and RCC23 relations are definable in terms of the $\Phi$-descriptor. It will also be shown in later chapters that the descriptor can be used to model many other topological and convex hull based relationships and the relationship of surroundedness. Qualitative models of these relations are discussed in Section 4.1 and quantitative models in Section 4.2. Validation experiments are presented in Section 4.3 and conclusions are given in Section 4.4.

4.1 Qualitative models based on the $\Phi$-descriptor

The RCC8, RCC23 and DE-4IM relations are definable in terms of the $\Phi$-descriptor, i.e., they are $\Phi$ relations (See sections 4.1.1 and 4.1.2). Moreover, many other topological and convex hull based relationships can be qualitatively modeled using the $\Phi$-descriptor (See Chapter VI).

4.1.1 The RCC8 relations

A large number of relations can be defined from the $\Phi$-descriptor. Initially, a small set of basic relations may be directly defined from the descriptor and then other relations may be built from that set using logical operators. See Table 4-1 and Table 4-2. (Note that follows $(A, B)$ is true if and only if leads $(A, B)$ is true; subset $(A, B)$ is true iff $A$ is a subset of $B$; superset $(A, B)$ is true iff $A$ is a superset of $B$; disjointInteriors $(A, B)$ is true iff the interiors of $A$ and $B$ do not intersect.) Unfortunately, the RCC8 relations are not $\Phi$ relations. The $\Phi$ relations that are the "closest" to the RCC8 relations are those shown in Table 4-2. See Figure 4-1 for an understanding of the difference. The issue is that the $\Phi$-
descriptor cannot distinguish the case where the object boundaries are disjoint from the case where they intersect in a set of isolated points.

This issue can be resolved by fine-tuning the definition of the $\Phi$-descriptor and the definition of an object. First, let there be a very special element, $0^+$, to be added to the set of nonnegative real numbers. The intention here is not to discuss number systems that include infinitesimals, like the system of dual numbers [113], or the system of hyperreal numbers [114]. Nonetheless, $0^+$ may be seen as an infinitesimal quantity such that $0 < 0^+ < x$ for any positive real number $x$. Now, consider two cores $I$ and $J$ aligned in direction $\theta$. Assume $F_f(\theta, I, J)$ is 0; assume one of the segments $[p, q]$ of $I$, with $p \neq q$ and $\bar{pq}/|\bar{pq}| = \tilde{\theta}$, is such that $[q, q]$ is one of the segments of $J$; then $F_f(\theta, I, J)$ is changed to $0^+$. Likewise, assume $F_l(\theta, I, J)$ is 0; assume one of the segments $[p, q]$ of $J$, with $p \neq q$ and $\bar{pq}/|\bar{pq}| = \tilde{\theta}$, is such that $[q, q]$ is one of the segments of $I$; then $F_l(\theta, I, J)$ is changed to $0^+$. Next, consider one of the 10 groups of point pair categories and the function $F$ attached to it. The changes, if any, made to the function $F$ do not affect the area histogram value $F^{AB}(\theta)$ unless $F^{AB}(\theta)$ is found to be 0 and there is an element $p$ of $\theta^*(\omega)$ such that $F(\theta, A \cap \theta(p), B \cap \theta(p)) \geq 0^+$; in that case, $F^{AB}(\theta)$ is changed to $0^+$. Finally, an object is redefined as being a set $S$ of points such that: $S$ is not empty; $S$ is bounded; $S$ is a regular closed set; for any line $L$, the set $S \cap L$ is a core; the boundary of $S$ is not self-intersecting; for any point $p$ of $S$, there exists a nondegenerate segment in $S$ ending at $p$.

The changes above affect the basic $\Phi$ relations follows, leads and starts. See Figure 4-2. Let there be a new set of basic $\Phi$ relations, as in Table 4-3. This set can be used to retrieve the RCC8 relations, as shown in Table 4-4. In other words, the RCC8 relations are now $\Phi$ relations. Note that most right-hand sides in Table 4-4 can be simplified. See Table 4-5. For example, if $A$ is a subset of $B$ then the interiors of $A$ and $B$ cannot be disjoint. Therefore, $EQ = subset^+ \land superset^+$. On the other hand, the equalities in Table 4-4 make it clear that the 8 relations are JEPD.
Table 4-1 Some basic $\Phi$ relations.

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>follows</td>
<td>$\exists \theta, F^{AB}_f(\theta) \neq 0$</td>
</tr>
<tr>
<td>leads</td>
<td>$\exists \theta, F^{AB}_l(\theta) \neq 0$</td>
</tr>
<tr>
<td>starts</td>
<td>$\exists \theta, F^{AB}_s(\theta) \neq 0$</td>
</tr>
<tr>
<td>subset</td>
<td>$\forall \theta, (F^{AB}_c(\theta) = 0 \land F^{AB}_l(\theta) = 0 \land F^{AB}_a(\theta) = 0)$</td>
</tr>
<tr>
<td>superset</td>
<td>$\forall \theta, (F^{AB}_d(\theta) = 0 \land F^{AB}_f(\theta) = 0 \land F^{AB}_r(\theta) = 0)$</td>
</tr>
<tr>
<td>disjointInteriors</td>
<td>$\forall \theta, (F^{AB}_a(\theta) = 0 \land F^{AB}_s(\theta) = 0)$</td>
</tr>
</tbody>
</table>

Note that $\text{follows}(A,B)$ is true if and only if $\text{leads}(A,B)$ is true; $\text{subset}(A,B)$ is true iff $A$ is a subset of $B$; $\text{superset}(A,B)$ is true iff $A$ is a superset of $B$; $\text{disjointInteriors}(A,B)$ is true iff the interiors of $A$ and $B$ do not intersect.

Table 4-2 The $\Phi$ relations “closest” to the RCC8 relations.

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi_{DC}$</td>
<td>$\text{disjointInteriors} \land \neg \text{follows}$</td>
</tr>
<tr>
<td>$\Phi_{EC}$</td>
<td>$\text{disjointInteriors} \land \text{follows}$</td>
</tr>
<tr>
<td>$\Phi_{PO}$</td>
<td>$\neg \text{disjointInteriors} \land \neg \text{subset} \land \neg \text{superset}$</td>
</tr>
<tr>
<td>$\Phi_{EQ}$</td>
<td>$\neg \text{disjointInteriors} \land \text{subset} \land \text{superset}$</td>
</tr>
<tr>
<td>$\Phi_{NTPP}$</td>
<td>$\neg \text{disjointInteriors} \land \text{subset} \land \neg \text{superset} \land \neg \text{starts}$</td>
</tr>
<tr>
<td>$\Phi_{TPP}$</td>
<td>$\neg \text{disjointInteriors} \land \text{subset} \land \neg \text{superset} \land \text{starts}$</td>
</tr>
<tr>
<td>$\Phi_{NTPPi}$</td>
<td>$\neg \text{disjointInteriors} \land \neg \text{subset} \land \text{superset} \land \neg \text{starts}$</td>
</tr>
<tr>
<td>$\Phi_{TPPi}$</td>
<td>$\neg \text{disjointInteriors} \land \neg \text{subset} \land \text{superset} \land \text{starts}$</td>
</tr>
</tbody>
</table>
Figure 4-1 The link between RCC8, DE-4IM, and Φ relations. The RCC8 relations (Section 2.1.2.1), the DE-4IM relations (Section 2.1.2.2), and some Φ relations. Object A is in light-gray; object B is in dark-gray.

Figure 4-2 Fine-tuning the definition of the Φ-descriptor and the definition of an object. Before the changes: (a)(b)(c)(d) → starts; (e)(f)(g) → follows. After the changes: (a)(b) → starts if such objects were considered, but they are not; (c)(d) starts; (e)(f)(g) follows.
Table 4-3 Some basic $\Phi$ relations (after fine-tuning).

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>follows</td>
<td>$\exists \theta, F_{f}^{AB}(\theta) &gt; 0$</td>
</tr>
<tr>
<td>leads</td>
<td>$\exists \theta, F_{l}^{AB}(\theta) &gt; 0$</td>
</tr>
<tr>
<td>starts</td>
<td>$\exists \theta, F_{s}^{AB}(\theta) &gt; 0$</td>
</tr>
<tr>
<td>follows$^+$</td>
<td>$\exists \theta, F_{f}^{AB}(\theta) &gt; 0^+$</td>
</tr>
<tr>
<td>leads$^+$</td>
<td>$\exists \theta, F_{l}^{AB}(\theta) &gt; 0^+$</td>
</tr>
<tr>
<td>starts$^+$</td>
<td>$\exists \theta, F_{s}^{AB}(\theta) &gt; 0^+$</td>
</tr>
<tr>
<td>subset$^+$</td>
<td>$\forall \theta, (F_{c}^{AB}(\theta) \leq 0^+ \land F_{f}^{AB}(\theta) \leq 0^+ \land F_{a}^{AB}(\theta) \leq 0^+)$</td>
</tr>
<tr>
<td>superset$^+$</td>
<td>$\forall \theta, (F_{u}^{AB}(\theta) \leq 0^+ \land F_{f}^{AB}(\theta) \leq 0^+ \land F_{r}^{AB}(\theta) \leq 0^+)$</td>
</tr>
<tr>
<td>disjointInterior$^+$</td>
<td>$\forall \theta, (F_{o}^{AB}(\theta) \leq 0^+ \land F_{s}^{AB}(\theta) \leq 0^+)$</td>
</tr>
</tbody>
</table>

Note that $\text{follows}(A, B)$ may be true while $\text{leads}(A, B)$ is false and vice versa; $\text{follows}^+(A, B)$ is true if and only if $\text{leads}^+(A, B)$ is true; $\text{subset}^+(A, B)$ is true iff $A$ is a subset of $B$; $\text{superset}^+(A, B)$ is true iff $A$ is a superset of $B$; $\text{disjointInterior}^+(A, B)$ is true iff the interiors of $A$ and $B$ do not intersect.

Table 4-4 The RCC8 relations are $\Phi$ relations (after fine-tuning).

<table>
<thead>
<tr>
<th>DC = disjointInterior$^+$ $\land \neg$ follows</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC = disjointInterior$^+$ $\land$ follows</td>
</tr>
<tr>
<td>PO = $\neg$disjointInteriors$^+$ $\land \neg$subset$^+$ $\land \neg$superset$^+$</td>
</tr>
<tr>
<td>EQ = $\neg$disjointInteriors$^+$ $\land$ subset$^+$ $\land$ superset$^+$</td>
</tr>
<tr>
<td>NTPP = $\neg$disjointInteriors$^+$ $\land$ subset$^+$ $\land \neg$superset$^+$ $\land$ starts</td>
</tr>
<tr>
<td>TPP = $\neg$disjointInteriors$^+$ $\land$ subset$^+$ $\land \neg$superset$^+$ $\land$ starts</td>
</tr>
<tr>
<td>NTPPi = $\neg$disjointInteriors$^+$ $\land$ subset$^+$ $\land$ superset$^+$ $\land$ starts</td>
</tr>
<tr>
<td>TPPi = $\neg$disjointInteriors$^+$ $\land$ subset$^+$ $\land$ superset$^+$ $\land$ starts</td>
</tr>
</tbody>
</table>
Simplifying the expressions in Table 4-4.

| DC      | disjointInteriors⁺ ∧ ¬follows |
| EC      | disjointInteriors⁺ ∧ follows |
| PO      | ¬disjoint Interiors⁺ ∧ ¬subset⁺ ∧ ¬superset⁺ |
| EQ      | subset⁺ ∧ superset⁺ |
| NTPP    | subset⁺ ∧ ¬starts |
| TPP     | subset⁺ ∧ superset⁺ ∧ starts |
| NTPPi   | superset⁺ ∧ ¬starts |
| TPPi    | ¬subset⁺ ∧ superset⁺ ∧ starts |

4.1.2 The RCC23 relations and other relations

Consider the basic Φ relations in Table 4-3 and Table 4-6. They can be used to define the Φ relations in Table 4-7. If the objects are not restricted to connected objects (i.e., an object may be made of multiple pieces) then ΦOUT, ΦPIN, etc., differ from but are comparable to OUT, PIN, etc. See Figure 4-3. If the objects are restricted to connected objects then ΦOUT = OUT, ΦPIN = PIN, etc., and, therefore, the RCC23 relations are Φ relations. The DE-4IM relations are Φ relations as well. See Table 4-8. Likewise, examples of topological relations other than RCC23 and DE-4IM abound. See Fig. 4.4.

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>encloses0</td>
<td>( \exists \theta, F_e^{AB}(\theta) = 0 )</td>
</tr>
<tr>
<td>encloses1</td>
<td>( \exists \theta, (F_e^{AB}(\theta) \neq 0 \land F_e^{AB}(\theta) \neq</td>
</tr>
<tr>
<td>encloses2</td>
<td>( \exists \theta, F_e^{AB}(\theta) =</td>
</tr>
<tr>
<td>divides0</td>
<td>( \exists \theta, F_d^{AB}(\theta) = 0 )</td>
</tr>
<tr>
<td>divides1</td>
<td>( \exists \theta, (F_d^{AB}(\theta) \neq 0 \land F_d^{AB}(\theta) \neq</td>
</tr>
<tr>
<td>divides2</td>
<td>( \exists \theta, F_d^{AB}(\theta) =</td>
</tr>
</tbody>
</table>
Table 4-7 The Φ relations “closest” to OUT, PIN, IN, OUTi, PINi and INi.

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΦOUT</td>
<td>disjointInteriors⁺ ∧ ¬ divides1 ∧ ¬ divides2</td>
</tr>
<tr>
<td>ΦPIN</td>
<td>disjointInteriors⁺ ∧ divides1 ∧ ¬ divides2</td>
</tr>
<tr>
<td>ΦIN</td>
<td>disjointInterior⁺ ∧ divides2</td>
</tr>
<tr>
<td>ΦOUTi</td>
<td>disjointInteriors⁺ ∧ ¬encloses1 ∧ ¬ encloses2</td>
</tr>
<tr>
<td>ΦPINi</td>
<td>disjointInterior⁺ ∧ encloses1 ∧ ¬ encloses2</td>
</tr>
<tr>
<td>ΦINI</td>
<td>disjointInteriors⁺ ∧ encloses2</td>
</tr>
</tbody>
</table>

Figure 4-3 IN and PIN vs. ΦIN and ΦPIN. (a) IN and ΦIN. (b) IN and ΦPIN. (c) PIN and ΦPIN

Table 4-8 The DE-4IM relations are Φ relations.

\[ FFFF = disjointInteriors⁺ \land \neg \text{follows} \]
\[ FFF0 = disjointInteriors⁺ \land \text{follows} \land \neg \text{follows} + \]
\[ FFF1 = disjointInteriors⁺ \land \text{follows}⁺ \]
\[ 2110 = \neg disjointInteriors⁺ \land \neg \text{subset}⁺ \land \neg \text{superset}⁺ \land \neg \text{starts}⁺ \]
\[ 2111 = \neg disjointInteriors⁺ \land \neg \text{subset}⁺ \land \neg \text{superset}⁺ \land \text{starts}⁺ \]
\[ 2FF1 = \text{subset}⁺ \land \text{superset}⁺ \]
\[ 2F1F = \text{subset}⁺ \land \neg \text{starts} \]
\[ 2F10 = \text{subset}⁺ \land \text{starts} \land \neg \text{starts}⁺ \]
\[ 2F11 = \text{subset}⁺ \land \neg \text{superset}⁺ \land \text{starts}⁺ \]
\[ 21FF = \text{superset}⁺ \land \neg \text{starts} \]
\[ 21F0 = \text{superset}⁺ \land \text{starts} \land \neg \text{starts}⁺ \]
\[ 21F1 = \neg \text{subset}⁺ \land \text{superset}⁺ \land \text{starts}⁺ \]

It should be noted that the right-hand sides have been simplified using various properties of the basic Φ relations (follows⁺ → follows, starts⁺ → starts, subset⁺ → ¬disjointedInteriors⁺, etc.).
Examples of non-RCC23 non-DE-4IM topological $\Phi$ relations. (a) $PO \land \neg follows^+ \land \neg starts^+$, (b) $PO \land \neg follows^+ \land starts^+$, (c) $PO \land follows^+ \land \neg starts^+$, (d) $PO \land follows^+ \land starts^+$. Note that the RCC23 cannot distinguish between (a)(b)(c)(d) while the DE-4IM can only distinguish (a) from (b)(c)(d).

4.2 Quantitative models based on the $\Phi$-Descriptor

A fuzzification process, denoted by $\sim$, involves mapping crisp relations to fuzzy relations. For example, $\overline{RCC8}$ is the set of fuzzified RCC8 relations and $\overline{DC}$ is the RCC8 relation DC after fuzzification. There is, of course, an infinite number of ways to fuzzify the RCC relations. The aim here is twofold: it is to design a fuzzification process based on the $\Phi$-descriptor and such that the fuzzified relations fit with a natural perception of the world; just as importantly, it is to show that the $\Phi$-descriptor lends itself, with great flexibility, to the quantitative modeling of spatial relationships. Section 4.2.1 lists the expectations for the fuzzified relations. Section 4.2.2 shows that the task of fuzzifying $\Phi$ relations reduces to that of fuzzifying the basic $\Phi$ relations. Sections 4.2.3 to 4.2.6 cover the fuzzification of the basic $\Phi$ relations introduced in Section 4.1.

4.2.1 Expectations

Let $REL$ be a set of relations. As specified by the conditions below, the fuzzified relations are expected to satisfy some properties. In particular, it is expected that some of the properties satisfied by the crisp relations are transferred to the fuzzified relations. It could be argued whether the perception illustrated by these conditions is the most intuitive, but it is one possible perception, and it will be used for the discussion here.
**Boundary conditions**

Let $\text{rel}$ be a relation in $\text{REL}$. If there exist object pairs $(A, B)$ such that $\text{rel}(A, B) = 0$, then there exist pairs such that $\tilde{\text{rel}}(A, B) = 0$. If there exist pairs such that $\text{rel}(A, B) = 1$, then there exist pairs such that $\tilde{\text{rel}}(A, B) = 1$. Moreover, for any objects $A$ and $B$:

\[
\tilde{\text{rel}}(A, B) = 0 \rightarrow \text{rel}(A, B) = 0
\]

\[
\tilde{\text{rel}}(A, B) = 1 \rightarrow \text{rel}(A, B) = 1
\]

**Reflexivity and symmetry conditions**

Let $\text{rel}$ be a relation in $\text{REL}$. If $\text{rel}$ is reflexive, then $\tilde{\text{rel}}$ is reflexive. For example, $\text{EQ}$ is expected to be reflexive. If $\text{rel}$ is symmetric, then $\tilde{\text{rel}}$ is symmetric. For example, $\text{DC}$ is expected to be symmetric.

**Condition on inverse relations**

Let $\text{rel1}$ and $\text{rel2}$ be relations in $\text{REL}$.

If $\text{rel1}(B, A) = \text{rel2}(A, B)$) for any pair $(A, B)$, then $\tilde{\text{rel1}}(B, A) = \tilde{\text{rel2}}(A, B)$ for any pair $(A, B)$.

**JEPD condition**

By definition (Section 2.1.2), the relations in $\text{REL}$ are JEPD iff for any object pair $(A, B)$ there is one and only one relation $\text{rel}$ in $\text{REL}$ such that $\text{rel}(A, B) = 1$. Equivalently, they are JEPD iff for any $(A, B)$ we have:

\[
\sum_{\text{rel} \in \text{REL}} \text{rel}(A, B) = 1
\]  

(4-1)

In other words, the relations define an 8-partition of the set of all object pairs (each class of the partition corresponds to a relation in $\text{REL}$). Equation (4.1) provides a natural way
to extend the JEPD concept to fuzzy relations. We say that the fuzzy relations in $\tilde{REL}$ are JEPD iff for any $(A, B)$ we have:

$$\sum_{\text{rel} \in \text{REL}} \tilde{r}(A, B) = 1 \quad (4-2)$$

In other words, the fuzzy relations define a fuzzy 8-partition of the set of all object pairs. The JEPD condition is the following: if the relations in $REL$ are JEPD, then the fuzzy relations in $\tilde{REL}$ are JEPD.

**Continuity condition**

Continuous changes in the shape or position of an object must translate into continuous changes in the truth values. In particular, $\tilde{EC}$ and $\tilde{TPP}$ are expected to behave according to Figure 4-5 and Figure 4-6. For example, Figure 4-5(a), when the smaller square A is far enough from the bigger square B, the truth value $\tilde{EC}(A, B)$ is 0; it increases when B gets closer to B, until the two boundaries intersect, at which point it is 1; when A gets further in B, the value decreases, and eventually becomes 0.

![Figure 4-5 Fuzzifying EC](image)

Figure 4-5 Fuzzifying EC. Continuous changes in the shape or position of an object implies continuous changes in the truth value. The upright arrows indicate that the truth value increases (or, at least, that it does not decrease), and the downright arrows that it decreases (or, at least, that it does not increase).
4.2.2 Meeting the expectations

Consider the three configurations in Figure 4-7. The value $|A|$ is the same in (a)(b)(c). The same applies to $|B|, F^{AB}_{t0}(0), F^{AB}_o(0), F^{AB}_{c0}(0)$, etc., and $W^{AB}_{t0}(0), W^{AB}_o(0), W^{AB}_{c0}(0)$, etc. (except that $F^{AB}_{f0}(0)$, is 0 in (a)/(b) and is $0^+$ in (c)). This would not hold, of course, if another direction than 0 was considered. Nonetheless, the continuity condition as illustrated by Figure 4-7 seems hard to meet (unless the trails minimum width histogram is considered so as to record the distance between the two objects; see Section 3.2). The option here is to get around the issue and ignore the distinction between FFF0 and FFF1—or, to be more specific, between 0 and $0^+$. Consider a $\Phi$ relation to be fuzzified. Let it first be expressed as a conjunction of basic $\Phi$ relations (possibly negated) in which each basic relation appears exactly once. For example:

$$\Phi_{PIN} \land \Phi_{IN} \land DC = (\text{disjointInteirors}^+ \land \text{divides}1 \land \neg \text{divides}2)$$

$$\land (\text{disjointInteirors}^+ \land \text{encloses}2)$$

$$\land (\text{disjointInteirors}^+ \land \neg \text{follows})$$

$$= \text{disjointInteirors}^+ \land \text{divides}1 \land \neg \text{divides}2 \land \text{encloses}2 \land \neg \text{follows}$$

Next, replace $\text{follows} \land \neg \text{follows}^+ \text{ with } \text{follows}^+$, and replace $\text{starts} \land \neg \text{starts}^+ \text{ with } \text{starts}$. Then, replace $\text{follows}^+ \text{ with } \text{follows}, \text{starts}^+ \text{ with } \text{starts}, \text{disjointInteirors}^+ \text{ with } \text{disjointInteirors}, \text{subset}^+ \text{ with } \text{subset}, \text{and superset}^+ \text{ with } \text{superset}$. For example,
\( \neg \text{follows}^+ \land \text{disjointInteriors}^+ \land \text{follows} \) becomes \( \text{follows} \land \text{disjointInteriors}^+ \) which becomes \( \text{follows} \land \text{disjointInteriors} \). The next step is to tackle the JEPD condition. If the \( \Phi \) relations to be fuzzified are JEPD, then it is required that the expressions clearly reflect the JEPD property. For example, RCC8 relations as in Table 4-4 may be used instead of the relations in Table 4-5. To fuzzify the \( \Phi \) relations, logical operators \( \neg \) and \( \land \) and the basic \( \Phi \) relations are fuzzified. Then let \( x \mapsto 1 - x \) for \( \neg \) (standard fuzzy negation) and \( (x, y) \mapsto xy \) for \( \land \) (product \( t \)-norm) JEPD property to be transferred from the crisp relations to the fuzzy ones. For example, consider the JEPD relations

\[
\text{rel}1 = \text{disjointInteriors} \land \text{follows}, \\
\text{rel}2 = \text{disjointInteriors} \land \neg \text{follows}, \\
\text{rel}3 = \neg \text{disjointInteriors},
\]

and the object pair \((A, B)\). Let \( x = \text{disjointInteriors} \,(A, B) \) and \( y = \text{follows} \,(A, B) \). Then the result is,

\[
\text{rel}1(A, B) + \text{rel}2(A, B) + \text{rel}3(A, B) = xy + x(1 - y) + (1 - x) = 1
\]

In other words, the fuzzy relations \( \text{rel}1, \text{rel}2, \) and \( \text{rel}3 \) are JEPD. In the end, fuzzifying \( \Phi \) relations comes down to fuzzifying the basic \( \Phi \) relations.

![Diagram](image)

**Figure 4-7** Fuzzifying FFF0. Meeting the continuity condition.
4.2.3 Fuzzifying disjointInteriors, subset and superset

There is a straightforward fuzzification of these basic $\Phi$ relations:

\[
disjointInteriors = \max \left( 0, 1 - 2 \times \frac{|A \cap B|}{\min(|A|, |B|)} \right)
\]

(4-3)

\[
\text{subset} = \max \left( 0, 2 \times \frac{|A \cap B|}{|A|} - 1 \right)
\]

(4-4)

\[
\text{superset} = \max \left( 0, 2 \times \frac{|A \cap B|}{|B|} - 1 \right)
\]

(4-5)

See Figure 4-8. Other thresholds of course may be chosen for the triangular functions, or trapezoidal functions instead of triangular ones, but the main principle is not changed. Also $|A|$, $|B|$, and $|A \cap B|$ are either elements of or can be derived from the $\Phi$-descriptor.

Figure 4-8 Fuzzifying the basic $\Phi$ relations. (a) disjointInteriors (b) subset (c) superset.
4.2.4 Fuzzifying follows

Since the crisp relation follows is defined from the homonymous area histogram, one may think of defining the fuzzy relation \( \text{follows} \) from that same histogram. This is not, however, a viable idea. For example, assume,

\[
\text{follows}(A, B) = \frac{\max_{\theta} F_{\text{follows}}^{AB}(\theta)}{|B|}
\]  
(4-6)

The truth value is 1 in the case of Figure 4-9(a)—which is as expected—but it is 0 in the case of Figure 4-9(b), and the continuity condition is violated. An alternative is to define \( \text{follows} \) from the \( \text{trails} \) area histogram: consider two objects with disjoint interiors; if one object \( \text{trails} \), the other, then it does not \( \text{follow} \); and if it does not \( \text{trail} \), then it necessarily \( \text{follows} \). It can be shown, however, that the idea is not viable either.

The solution presented here involves \( \text{follows} \), \( \text{trails} \) and \( \text{overlaps} \) length histograms. First, different elementary truth values are assigned to the proposition “\( A \text{ follows} B \) in direction \( \theta \).” Each refers to one of the three basic \( \text{follows} \) configurations shown in Figure 4-5(a)(b)(c):

\[
\varphi_{\text{follows}}^{AB}(\theta) = \max \left( 0, 1 - 2 \times \frac{W_{\text{trails}}^{AB}(\theta)}{H_{\text{trails}}^{AB}(\theta)} \right)
\]  
(4-7)

\[
\varphi_{\text{follows}}^{AB}(\theta) = 1 \]  
(4-8)

Figure 4-9 The fuzzy relation \( \text{follow} \) cannot be defined from the homonymous area histogram. The continuity condition would be violated.
\[
\varphi_{o}^{AB}(\theta) = \max \left( 0, 1 - 2 \times \frac{W_{o}^{AB}(\theta)}{H_{o}^{AB}(\theta)} \right) \quad (4-9)
\]

See Figure 4-10, Figure 4-11 and Figure 4-12. In the general case, there is a mixture of these three basic configurations, as in Figure 4-13. A weighted average of the elementary truth values is therefore calculated:

\[
\varphi^{AB}(\theta) = \text{weight}_{t}^{AB}(\theta) \times \varphi_{t}^{AB}(\theta) + \text{weight}_{f}^{AB}(\theta) \times \varphi_{f}^{AB}(\theta) + \text{weight}_{o}^{AB}(\theta) \times \varphi_{o}^{AB}(\theta) \quad (4-10)
\]

where

\[
\text{weight}_{t}^{AB}(\theta) = \frac{H_{t}^{AB}(\theta)}{H_{t}^{AB}(\theta) + H_{f}^{AB}(\theta) + H_{o}^{AB}(\theta)},
\]

\[
\text{weight}_{f}^{AB}(\theta) = \frac{H_{f}^{AB}(\theta)}{H_{t}^{AB}(\theta) + H_{f}^{AB}(\theta) + H_{o}^{AB}(\theta)},
\]

\[
\text{weight}_{o}^{AB}(\theta) = \frac{H_{o}^{AB}(\theta)}{H_{t}^{AB}(\theta) + H_{f}^{AB}(\theta) + H_{o}^{AB}(\theta)}
\]

and

\[
H_{t}^{AB}(\theta) = H_{t}^{AB}(\theta) + H_{f}^{AB}(\theta) + H_{o}^{AB}(\theta)
\]

It should be noted that \( \varphi^{AB}(\theta) \) ignores the expected impact of the region of interaction in direction \( \theta \). For example, in Figure 4-14(a)(b)(c)(d), we have \( \varphi^{AB}(\theta) = 1 \). The final truth value assigned to the proposition “A follows B in direction \( \theta \)” depends on the proportion of A or B in the region of interaction. It is:

\[
\text{weight}^{AB}(\theta) \times \varphi^{AB}(\theta)
\]

where

\[
\text{weight}^{AB}(\theta) = \max \left( \frac{F_{|\text{class}|}^{AB}(\theta)}{|A|}, \frac{F_{|\text{ufros}|}^{AB}(\theta)}{|B|} \right)
\]
It should be noted that $F_{\text{clus}}(\theta)$ is the area of a subregion of the region of interaction in direction $\theta$: the sub-region occupied by $A$. Likewise, $F_{\text{ufros}}(\theta)$ is the area of the sub-region occupied by $B$. In the end, the truth value assigned to the proposition “$A$ follows $B$” is:

$$\text{follows} = \max_{\theta} \left( \text{weight}^{AB}(\theta) \times \varphi^{AB}(\theta) \right)$$

(4-11)

Note that in the equations above, a fraction is assumed to be 0 if its denominator is 0.

Figure 4-10  Fuzzifying follows (the trails case). (a) The truth value of the proposition “$A$ follows $B$ in direction 0” is expected to be close to 1. (b) The truth value of “$A$ follows $B$ in direction 0” is expected to be much lower than 1. (c) An answer to these expectations.

Figure 4-11  Fuzzifying follows (the ideal case). We expect the truth value of the proposition “$A$ follows $B$ in direction 0” to be 1.
Figure 4-12 Fuzzifying follows (the overlaps case). (a) The truth value of the proposition “A follows B in direction 0” is expected to be close to 1. (b) The truth value of the proposition is expected to be much lower than 1. (c) An answer to these expectations.

Figure 4-13 Fuzzifying follows (the general case). The general case is a mixture of the trails, follows and overlaps cases. Note that the truth value of the proposition “A follows B in direction 0” is the same in (a) as in (b).

Figure 4-14 Fuzzifying follows (impact of the region of interaction). In (a), the whole object A follows a part of B in direction 0; in (b), a part of A follows the whole object B; in both cases, the truth value of the proposition “A follows B in direction 0” is expected to be 1. In (c) and (d), only a part of A and a part of B are involved; the truth value is expected to be much lower than 1.
### 4.2.5 Fuzzifying starts

The relation *starts* is fuzzified the same way as follows. Different elementary truth values are assigned to the proposition “*A starts B* in direction $\theta$”. Each refers to one of the three basic *starts* configurations as shown in Figure 4-6(a1)(a3)(a4):

$$
\sigma_c^{AB}(\theta) = \max\left(0, 1 - 2 \times \frac{W_c^{AB}(\theta)}{H_c^{AB}(\theta)}\right) \quad (4-12)
$$

$$
\sigma_s^{AB}(\theta) = 1 \quad (4-13)
$$

$$
\sigma_u^{AB}(\theta) = \max\left(0, 1 - 2 \times \frac{W_u^{AB}(\theta)}{H_u^{AB}(\theta)}\right) \quad (4-14)
$$

where

$$
W_c^{AB} = W_c^{AB}(-\theta) \text{ and } H_c^{AB}(\theta) = H_c^{AB}(-\theta)
$$

$$
W_u^{AB} = W_u^{AB}(-\theta) \text{ and } H_u^{AB}(\theta) = H_u^{AB}(-\theta)
$$

Notice that in the case of Figure 4-6(a2), $F_c^{AB}(\theta) = 0$ and $F_c^{AB}(-\theta) \neq 0$; in the case of Figure 4-6(a3) we have $F_s^{AB}(-\theta) = 0$ and $F_s^{AB}(\theta) \neq 0$; in the case of Figure 4-6(a4), we have $F_u^{AB}(\theta) = 0$ and $F_u^{AB}(-\theta) \neq 0$. This is why the direction $-\theta$ is considered for *covers* and *uncovers* while $\theta$ is considered for *starts*. In the general case, we have a mixture of the three basic configurations, and a weighted average of the elementary truth values is calculated:

$$
\sigma^{AB}(\theta) = \text{weight}_c^{AB}(\theta) \times \sigma_c^{AB}(\theta) + \text{weight}_s^{AB}(\theta) \times \sigma_s^{AB}(\theta) + \text{weight}_u^{AB}(\theta) \times \sigma_u^{AB}(\theta) \quad (4-15)
$$

where,

$$
\text{weight}_c^{AB}(\theta) = \frac{H_c^{AB}(\theta)}{H_{c\cap s\cap u}(\theta)}
$$

90
weight_s^{AB}(\theta) = \frac{H_s^{AB}(\theta)}{H_{csu}^{AB}(\theta)},

weight_u^{AB}(\theta) = \frac{H_u^{AB}(\theta)}{H_{csu}^{AB}(\theta)}

and

H_{csu}^{AB}(\theta) = H_c^{AB}(\theta) + H_s^{AB}(\theta) + H_u^{AB}(\theta)

It should be noted (see Table 4-4 and Table 4-8) that starts is always used along with subset or superset. In these cases, the region of interaction in direction \( \theta \) necessarily includes one of the objects. Therefore, it is not necessary to consider a global weight as done for follows. In the end, the truth value assigned to the proposition “\( A \) starts \( B \)” is:

\[
\text{start}(A, B) = \max_{\theta} \left( \sigma^{AB}(\theta) \right)
\] (4-16)

4.2.6 Fuzzifying encloses and divides

There is a straightforward fuzzification of the basic \( \Phi \) relations defined by Table 4-6. The fuzzification of encloses0, encloses1 and encloses2 is based on the following observations: if there is a direction \( \theta_0 \) such that \( F_e^{AB}(\theta_0) \) is close to 0 then \( \min_{\theta} F_e^{AB}(\theta) \) is close to 0 (and vice versa); if there is a direction \( \theta_0 \) such that \( F_e^{AB}(\theta_0) \) is far from 0 and far from \( |B| \) then \( F_e^{AB}(\theta_0) \) and \( |B| - F_e^{AB}(\theta_0) \) are far from 0, i.e., \( \min_{\theta} F_e^{AB}(\theta_0), |B| - F_e^{AB}(\theta_0) \) is far from 0, i.e., \( \max_{\theta} \min_{\theta} F_e^{AB}(\theta_0), |B| - F_e^{AB}(\theta_0) \) is large (and the largest possible value is \( |B|/2 \)); if there is a direction \( \theta_0 \) such that \( F_e^{AB}(\theta_0) \) is close to \( |B| \) then \( \max_{\theta} F_e^{AB}(\theta_0) \) is close to \( |B| \). See Figure 4-15 and equations (4.17) to (4.19). The fuzzification of divides0, divides1 and divides1 is based on similar observations. See equations (4.20) to (4.22). Again, other thresholds for the triangular functions, or
trapezoidal functions instead of triangular ones could be chosen, but the principle is the same.

\[ \text{encloses0}(A,B) = \max \left( 0, 1 - 2 \times \frac{\min_\theta F_e^{AB}(\theta)}{|B|} \right) \]  
\[ \text{encloses1}(A,B) = \frac{2}{|B|} \max \min (F_e^{AB}(\theta), |B| - F_e^{AB}(\theta)) \]  
\[ \text{encloses2}(A,B) = \max \left( 0, 2 \times \frac{\min_\theta F_e^{AB}(\theta)}{|B|} - 1 \right) \]  
\[ \text{divides0}(A,B) = \max \left( 0, 1 - 2 \times \frac{\min_\theta F_d^{AB}(\theta)}{|A|} \right) \]  
\[ \text{divides1}(A,B) = \frac{2}{|A|} \max \min (F_d^{AB}(\theta), |A| - F_d^{AB}(\theta)) \]  
\[ \text{divides2}(A,B) = \max \left( 0, 2 \times \frac{\max_\theta F_d^{AB}(\theta)}{|A|} - 1 \right) \]

4.3 Experiments

The models were tested on synthetic images in grayscale. The Φ-descriptors of the images were produced with \(2\pi/5\) to \(\pi\) reference directions. Results were generated both for the qualitative and quantitative models. Valid results were obtained for the qualitative models. As the interpretation of results for qualitative models is straightforward, in the following sections, only results for the quantitative models will be presented.
For the quantitative RCC8 and RCC23 models, images were grouped into sets of object-configurations. The configurations in each group represent the transformation of a JEPD relationship into a conceptually neighbouring relationship through a process of continuous deformation. For example, in Figure 4-16(a1)-(a4), the DC relationship between objects A and B is transformed into the EC relationship by continuously deforming the original configuration. Section 4.3.1 presents results for the RCC8 relations whereas sections Section 4.3.2 includes results for the RCC23 relations.

4.3.1 RCC8 relations

Figure 4-16 and Figure 4-17 give the test images and the results for the RCC8 relationships that hold for the objects. Each set is labelled by a conceptually neighbouring JEPD relationship pair. For example, the set of images labelled “DC to EC” represents the transformation from $\overline{DC}$ to $\overline{EC}$ (see Figure 4-16(a1)-(a4) and Figure 4-16(a5)-(a8)). Note that only non-zero relationship degrees have been given for each configuration in the figures whereas zero relationships degrees have been suppressed.

As shown by the results, the models correctly infer the RCC8 relations from the descriptor and satisfy the conditions imposed on them in Section 4.2.1. For instance, the models for $\overline{EC}$ and $\overline{DC}$ satisfy the continuity and JEPD conditions for both the sets of configurations (i.e., the configuration in Figure 4-16 (a1)-(a4) and that in Figure 4-16(a5)-(a8)) that show the transformation of the relationship between the objects from DC to EC. It can be seen that there is a continuous change in the truth degrees of $\overline{EC}$ and $\overline{DC}$ as the position of the objects relative to each other changes. Likewise the JEPD property is also preserved (See equation (4.1)). In the $\overline{EC}$ to $\overline{PO}$ transformation, the continuity condition is met as well as the JEPD property is satisfied. For example, compare the configurations in 4.16(b1)-(b4) with the fuzzification principle given in Figure 4.5. Likewise, the results for other configuration also meet the expectations in Section 4.2.1.
Figure 4-16 Experimental results for the RCC8 relations: I. Truth degrees of RCC8 relations for the DC to EC, EC to PO, and PO to EQ deformations. Only the non-truth degrees have been shown.
Figure 4-17 Experimental results for the RCC8 relations: II. Truth degrees for RCC8 relations when the deformations are PO to TPP, TPP to EQ, EQ to NTPP, and TPP to NTPP. Only relations with non-zero truth degrees have been shown.
4.3.2 Convex-hull based relationships

Results for the convex-hull based relationships are given in Figure 4-18. All cases are the $\bar{D}C$ cases. The transformations include $\bar{P}IN \, \bar{P}IN_i \, \bar{D}C$ to $\bar{O}UT \, \bar{O}UT_i \, \bar{D}C$, $\bar{P}IN \, \bar{P}IN_i \, \bar{D}C$ to $\bar{P}IN \, \bar{O}UT_i \, \bar{D}C$ to $\bar{O}UT \, \bar{O}UT_i \, \bar{D}C$, and $\bar{I}N \, \bar{O}UT_i \, \bar{D}C$ to $\bar{I}N \, \bar{P}IN_i \, \bar{D}C$ transition. As shown by the result, the continuity condition is met as well as the JEPD property is preserved, e.g., see PIN PINi DC to OUT OUTi DC.

4.4 Summary

Methods for extracting topological relations from images using $\Phi$—descriptor were explored in this chapter. Models of qualitative topological spatial relationships based on the RCC8, RCC23, and the DE-4IM frameworks were proposed and it was shown that all these relations are in fact $\Phi$ relations. Additionally, some fuzzification techniques for the elementary $\Phi$ relations were suggested and quantitative models of RCC8, RCC23, and surround relationships that satisfy the properties of continuity, the JEPD condition, etc. were derived from the elementary relations. Experimental results with synthetic images demonstrated the accuracy of the models and the capability of the $\Phi$-descriptor to provide a basis for modeling a variety of topological and convex-hull based relations. In the present study, extraction of topological relations for simple convex objects were considered. In future, extraction of such relations in the case of more complex objects, e.g., disconnected objects when the relationship is surround, will be investigated. Furthermore, relationships combining topological and directional relations will be studied. The content presented in this chapter shows the versatility of the $\Phi$—descriptor. For example, the change in the truth-degrees of the $\bar{O}UT$, $\bar{P}IN$, $\bar{O}UT_i$, and $\bar{P}IN_i$ relationships reflect the change in the spatial configuration of the two objects. Moreover, the JEPD condition is met.
Figure 4-18 Results of quantitative models for the convex-hull based relationships
Chapter V

5 Directional spatial relationships

Directional spatial relations are an important category of spatial relationships. Examples of directional relationships are the cardinal relations “right”, “left”, “above”, and “below” (or correspondingly, “west”, “east”, “north”, and “south”). A good number of qualitative models for directional spatial relationships have been proposed in the area of AI [47] (see Section 2.1.3). However, qualitative models don’t lead to satisfactory results even in situations of moderate complexity because directional relationships are inherently vague and the crisp true-false evaluation are inadequate to deal with the fuzziness in the definition of directional relationships [115, 116]. Thus, quantitative approaches based on fuzzy techniques have been recommended to model directional relationships [117, 42, 79]. The discussion in this chapter focuses on the fuzzy models of directional spatial relationships. Two sets of methods for modeling the directional relationships from the \( \Phi \)-descriptor are presented; a model based on the calculation of average angle; and the emulation models. The first model (called the M model here) infers (the degree of) a directional spatial relationship from the average angle between the referent and argument objects. The emulation models (denoted as \( K' \), \( F0' \), and \( F2' \) models here) provide the \( \Phi \)-descriptor based emulation of the aggregation [87, 10] and effective force [66] methods for extracting directional relations from angle [10] and force histograms [11]. See sections 2.1.3, 2.2.3.1, and 2.2.3.2. The models are described in Section 5.1. Section 5.2 presents experimental results and Section 5.3 gives the summary.

5.1 Directional relationships

Many efficient methods for the extraction of the directional spatial relationships from relative position descriptors have been proposed in the literature [118, 116]. Aggregation method [87, 10] and the method of effective forces [66] are two such methods. Common properties sought in these methods are computational efficiency and quality of results in
terms of the satisfaction of user perception and expectations. Additionally, preservation of the property of symmetry (i.e., object B is the same degree to the left of object A as the degree to which object A is to the right of object B), sensitivity to distance between objects and object shapes, and satisfaction of the semantic inverse property (i.e., if object A is to the right of object B, object B is to the left of object B) are considered useful. Likewise, the behavior to boundary cases (where the relationship cannot be, e.g., just “right” but both “right and above”) is also an important consideration in these extraction methods. A goal in the models of directional relationships presented in this chapter is the satisfaction of these conditions. Each model basically associates a validity degree with the spatial proposition of the form, “object A is in direction $\alpha$ of object B” and works by taking the direction $\alpha$ and the $\Phi$-descriptor as inputs and returning the acceptability degree of the proposition “object A is in direction $\alpha$ of object B”. For the purpose of illustration, here the direction $\alpha = 0$ ($\alpha$ is the angle) or “A is to the right of B” is considered. To find the truth degree of the proposition for other directions, the same computation can be repeated. It should be noted that the directional relations ("right", "left", "above", etc.) here should be interpreted from the standpoint of an extrinsic reference frame [119], i.e., from the viewpoint of an external observer. Other reference frames such as deictic (or egocentric) or intrinsic [65] are also possible but are not considered here.

5.1.1 The model M

Generally, the perception of “A is to the right of B” (angle $\alpha$) is formed by locating the object B in a field subtending $\pi$ degrees from angle $\alpha - \pi/2$ to $\alpha + \pi/2$ [42]. This field is here called $\alpha - \text{viewfield}$ or the right view field of B. The model of directional relationship presented in this section, therefore, is based on the calculation and fuzzification of two angles; the average angle $\hat{\theta}$; the around angle $\check{\theta}$. The average angle $\hat{\theta}$ gives the average direction of the location of object A (argument object) in the $\alpha - \text{viewfield}$ of B (reference object). The around angle $\check{\theta}$ represents the angle by which A extends around B. The fuzzification principle has been illustrated in Figure 5-1(a)(b)(c) and Figure 5-1(e)(f)(g). Explanation is as below.
In Figure 5-1(a), where object A lies entirely to the right of object B (i.e., in direction \( \alpha = 0 \)), the acceptability degree of “A is to the right of B” can be considered maximum. In Figure 5-1(b), as only a part of object A is located in the direction \( \alpha \) of B, the acceptability degree can be considered lower. In Figure 5-1(c) no parts of A lie to the right of B but A is still somewhat to the right of A. It may be argued that the relationship in the case of the configuration in (c) more belongs to the category “above and to the right” instead of the category “to the right”. Thus, the truth degree of “A is to the right of B” is the lowest in this case.

Generally, if the average direction of A is more aligned with the given direction \( \alpha \), the truth degree of “A is to the direction \( \alpha \) of B” is higher but as the average direction deviates from \( \alpha \), the truth degree gets lower until A is outside the viewfield when it is the lowest. A similar principle is used to take the around factor into account when directional relationships between objects are assessed. When the extent of object A is wholly located within the \( \alpha - \text{viewfield} \) of B, the acceptability degree of A being within the viewfield of B is the highest but as A extends beyond the viewfield, the truth degree becomes lower.

A question may be asked as why it is important to consider the around factor in assessing the validity degree of “A is to the right of B”? The case in Figure 5-1(g) suggests the answer. An assessment of the configuration in Figure 5-1(g) may be that all the directional relationships, “above”, below”, “left”, “right”, etc., hold for the objects in the configuration. However, such an assessment may be unsatisfactory because people generally don’t combine more than two propositions when communicating visual information (in language) [120, 121]. The purpose in considering the around condition, thus, is to identify two dominant directional relationships between the objects and to distinguish between the “surround” and the directional cases.

The truth degree of “A is to the right of B” is derived from the combination of the above two factors, i.e., the directional and around factors as follows. Let \( \text{right}(A, B) \) be the
acceptability degree of the proposition “A is to the right of B”, then \( right(A, B) \) is defined by

\[
right(A, B) = \mu_{A-right-B}(\hat{\theta}) \times \mu_{A-around-B}(\tilde{\theta})
\]

(5-1)

where \( \mu_{A-right-B} : [-\pi/2, \pi/2] \rightarrow [0,1] \) is a membership function that assigns a truth-degree to “A is to the right of B” in \( A - right - B \). See Figure 5-1(d). As can be seen, \( \mu_{A-right-B} \) is a trapezoidal function that takes the average angle \( \hat{\theta} \in [-\pi/2, \pi/2] \) as an argument and map it to a truth-degree in [0,1]. The set \( A - right - B \) is the fuzzy set that represents the “right” relationship. The function \( \mu_{A-around-B} : [0,2\pi] \rightarrow [0,1] \) tackles the around condition and maps the around angle \( \tilde{\theta} \) to a truth degree. See Figure 5-1(h). Both these functions are defined in the following sections.

![Figure 5-1 fuzzification principle for \( \hat{\theta} \) and \( \tilde{\theta} \).](image)
5.1.1.1 Definition of $\mu_{A-right-B}(\theta)$

The overall direction of A with respect to B is represented by the average angle $\bar{\theta}$. When objects are closer to each other in direction “right”, the closer parts of A and B have more influence on the acceptability level of the directional spatial proposition, “A is to the right of B”. To take this influence into account, an adjusted angle $\hat{\theta}$ is then derived from $\bar{\theta}$. Finally, $\mu_{A-right-B}$ is calculated from the adjusted average angle $\hat{\theta}$. The average angle $\bar{\theta}$ is defined by,

$$\bar{\theta} = \frac{\int_{-\pi/2}^{\pi/2} \theta \times F_{\text{lafr}}^{AB}(\theta) \times \epsilon d\theta}{\int_{-\pi/2}^{\pi/2} F_{\text{lafr}}^{AB}(\theta) \times \epsilon \, d\theta}$$  \hspace{1cm} (5-2)$$

where

$$F_{\text{lafr}}^{AB}(\theta) = F_{\text{lafr}}^{AB}(\theta) + F_{\text{lafr}}^{AB}(-\theta)$$ and

$$F_{\text{lafr}}^{AB}(\theta) = F_l^{AB}(\theta) + F_a^{AB}(\theta) + F_f^{AB}(\theta) + F_r^{AB}(\theta)$$

$$F_{\text{lafr}}^{AB}(-\theta) = F_l^{AB}(-\theta) + F_a^{AB}(-\theta) + F_f^{AB}(-\theta) + F_r^{AB}(-\theta).$$

The quantity $F_{\text{lafr}}^{AB}(\theta)$ represents the part of the region of interaction of A and B in direction $\theta$ that is occupied by A or B (here non-intersecting objects are considered). It serves as a weight for the direction $\theta$.

![Figure 5-2 Definition of the devaluation factor $\epsilon$. The weight for a direction $\theta$ is devaluated if $\alpha + \pi/4 \leq \theta \leq \pi/2$ or $-\pi/2 \leq \theta \leq \alpha - \pi/4$](image)
The symbol $\varepsilon$ is a devaluation factor. It represents the devaluation of the weight $F_{|lafr|}^{AB}(\theta)$ of the angle $\theta$ according as whether $\theta$ is closer to $\alpha = 0$, i.e., “A is to the right of B”, or $\alpha = \pi/2$ (resp. $\alpha = -\pi/2$), i.e., “A is above B” or (resp. “A is below B”). See Figure 5-2. The choice to devaluate the directions that drift away from the target direction is inspired by the cone model of directional relations [122]. In order to calculate the adjusted average angle $\hat{\theta}$, an average $\overline{F}_{|lafr|}^{AB}$ is also calculated as follows.

$$\overline{F}_{|lafr|}^{AB} = \frac{\int_{-\pi/2}^{\pi/2} F_{|lafr|}^{AB}(\theta) d\theta}{\int_{-\pi/2}^{\pi/2} d\theta}$$ (5-3)

Next, the adjusted average angle $\hat{\theta}$ is defined as,

$$\hat{\theta} = \frac{\int_{-\pi/2}^{\pi/2} \left( \theta \overline{F}_{|lafr|}^{AB} + \theta \lambda(\theta) F_{|lafr|}^{AB}(\theta) \right) \times \varepsilon d\theta}{\int_{-\pi/2}^{\pi/2} \left( \overline{F}_{|lafr|}^{AB} + \lambda(\theta) F_{|lafr|}^{AB}(\theta) \right) \times \varepsilon d\theta}$$ (5-4)

where $\lambda [0, 1]$ is a factor that represents the effect of the closer parts of A to B on the acceptability level of the proposition. It can be defined in many ways, e. g., using an inverse square law such as the gravitational law. Here, it is defined by:

$$\lambda(\theta) = \frac{F_{|lafr|}^{AB}(\theta)}{F_{|lafr|}^{AB}(\theta) + F_{|lafr|}^{AB}(\theta)}$$ (5-5)

where $F_{|lafr|}^{AB}(-\theta)$ is an element of $\Phi$-descriptor that represents the area of the region between A and B. As can be seen, parts of the objects A and B are given importance according to their distance from each other with closer parts receiving more importance and farther parts receiving less importance. Thus $\hat{\theta}$ is adjusted according to distance such that when the entire objects move far from each other, the adjusted average angle $\hat{\theta}$ approaches the original average angle $\bar{\theta}$. Finally, the membership degree $\mu_{A\text{--right--B}}$ is calculated from $\hat{\theta}$ as follows.
\[ \mu_{A-right-B}(\theta) = \frac{\pi}{2} - \max(k, |\theta|) \]

As can be seen, the core of \( \mu_{A-right-B}(\theta) \) is defined over \([-k, k]\) where \( k \) is a constant such that \( 0 \leq k \leq \frac{\pi}{2} \). For pessimistic evaluation of the proposition, \( k \) can be set to a small value and for optimistic evaluation, it can be set to a large value.

### 5.1.1.2 Definition of \( \mu_{A-around-B}(\theta) \)

The angle \( \tilde{\theta} \) is the angle by which object A extends around B. If \( \tilde{\theta} > |\alpha + \pi/2| - |\alpha - \pi/2| \), i.e., parts of object A lie outside the given view-field, the proposition “A is to the right of B” is weakened by a corresponding amount; otherwise it receives full weight. \( \tilde{\theta} \) is calculated as follows.

\[ \tilde{\theta} = \frac{1}{2\pi \times |B|} \int_0^{2\pi} F_{e}^{AB}(\theta) d\theta \]

Finally, \( \mu_{A-around-B}(\theta) \) is defined as,

\[ \mu_{A-around-B}(\theta) = \frac{2\pi - \max(\tilde{\theta}, \pi)}{\pi} \]

### 5.1.2 The emulation models \( K', F0', \) and \( F2' \)

This section is based on a comparative study presented in [98]. In that study, three models of directional relationships are considered: K, F0 and F2. They are typical fuzzy models in the sense that they do not approximate the objects by, e.g., their centroids, or minimum bounding rectangles, and they fairly meet some widely accepted properties [11]. In particular, they are not sensitive to scale changes, there is no preferred direction, and the semantic inverse principle is respected (e.g., object A is to the left of object B as much as B is to the right of A).

As in [98], the configurations in Figure 5.3 are used to illustrate the main characteristics of the models. For each pair (A, B), the propositions “object A is to the right of object B”
“A is above B,” “A is to the left of B” and “A is below B” are assessed. The truth values $\text{RIGHT}(A,B)$, $\text{ABOVE}(A,B)$, $\text{LEFT}(A,B)$ and $\text{BELOW}(A,B)$ produced by K, F0 and F2 are presented in Table 5-1 (in hundredths). With the model K, most values are greater than 0; all values are less than 1 (correspondingly, 100); the sum of the four truth values is 1. Consider for instance the objects in Figure 5-3(c). If a vertical line is drawn through the middle of object B, the half-planes so defined contain points of A. As a result, A is somewhat to the right as well as to the left of B — and cannot, therefore, be perfectly below it. With F0 and F2, opposite directions usually exclude each other; an object may be perfectly in some direction of another; the sum of the truth values may be 1, but it may also be greater than 1, or less than 1, and even 0 (which indicates that other spatial relationships should be used to describe the relative position of the objects). In Figure 5-3(b), the model F2 alone asserts that A is more to the right of B, even though it gives a certain weight to the proposition “A is above B.” In Figure 5-3(c)(d), as A becomes longer, K quickly affirms that A is essentially located to the right of B. The model F0 eventually shares this point of view, but later on, and in a less definite way. F2 is the only model to maintain that A essentially remains below B.

![Figure 5-3 Configurations for comparing fuzzy models of directional relationships. In each case, the argument A is in light grey and the referent B in dark grey.](image)
Table 5-1 Truth values (in hundredths) produced by three fuzzy models of directional relationships. The configurations are those in Figure 5-3.

<table>
<thead>
<tr>
<th></th>
<th>K F0</th>
<th>F2</th>
<th>K F0</th>
<th>F2</th>
<th>K F0</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RIGHT</strong></td>
<td>71</td>
<td>100</td>
<td>38</td>
<td>55</td>
<td>86</td>
<td>10</td>
</tr>
<tr>
<td><strong>ABOVE</strong></td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>LEFT</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td><strong>BELOW</strong></td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>80</td>
</tr>
</tbody>
</table>

(a) (b) (c) (d)

<table>
<thead>
<tr>
<th></th>
<th>K F0</th>
<th>F2</th>
<th>K F0</th>
<th>F2</th>
<th>K F0</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RIGHT</strong></td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>ABOVE</strong></td>
<td>33</td>
<td>3</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td><strong>LEFT</strong></td>
<td>22</td>
<td>75</td>
<td>99</td>
<td>40</td>
<td>87</td>
<td>99</td>
</tr>
<tr>
<td><strong>BELOW</strong></td>
<td>43</td>
<td>7</td>
<td>43</td>
<td>20</td>
<td>5</td>
<td>25</td>
</tr>
</tbody>
</table>

(e) (f) (g) (h) (i)

5.1.2.1 Emulating K

The truth values RIGHT(A,B), ABOVE(A,B), LEFT(A,B) and BELOW(A,B) produced by the model K are derived, using the aggregation method [34][10], from the angle histogram (see Section 2.2.3.1) associated with (A, B). It is shown here that the angle histogram — and hence the model K — can be emulated by the Φ-descriptor.

Consider two disjoint sets of pixels, A and B. For any pixels p of A and q of B, let \( \angle(p,q) \) be the angle between the horizontal axis (from left to right) and the directed line that passes through the center of p and then of q. This angle belongs to \((-\pi, \pi]\). Partition\((-\pi, \pi]\) into n intervals of equal lengths, \(\theta_1, \theta_2, \text{etc.} \) (the direction bins). The angle histogram associated with (A,B) maps each \( i \) in \( 1..n \) to the number of pixel pairs \((p,q)\)\( \in A \times B \) such that \( \angle(p,q) \in \theta_i \).

Now, assume \( A \) is composed of \( W_A \times h \) pixels and \( B \) of \( W_B \times h \) pixels, as in Figure 5-4. Consider the direction bin \( \theta_i = (-\delta/2, \delta/2] \), where \( \delta = 2\pi/n \). For a given pixel p of \( A \), what is the number of pixels q of B such that \( \angle(p,q) \in (-\delta/2, \delta/2] \)? A reasonable estimate
seems to be \( W_B \times k \) if \( k \leq h \) and \( W_B \times h \) otherwise, where \( k = 2W \tan(\delta/2) \). In other words, the angle histogram maps \( i \) to a number that is roughly equal to \( W_A \times h \times W_B \times \min(k, h) \) i.e., \( W_A W_B h \min(2W \tan \pi/n, h) \).

Next, consider the two objects \( A \) and \( B \) in Figure 5-4. Each one is the union of three components. Let \( \theta \) be the horizontal direction, from left to right. Then,

\[
h = H^A_{t\theta} + H^A_{t\theta}, \quad h' = H^A_{t\theta} + H^A_{t\theta}, \quad w = W^A_{\theta}, \quad w' = W^A_{\theta}, \quad W_A = \frac{F^A_{\theta} + F^A_{\theta}}{H^A_{t\theta} + H^A_{t\theta}} \]

\[
W_B = \frac{F^A_{\theta} + F^A_{\theta}}{H^A_{t\theta} + H^A_{t\theta}} \tag{5-9}
\]

Note that \( F^A_{\theta} + F^A_{\theta} \) is the area of the part of \( A \) that interacts with \( B \) in direction \( \theta \). Likewise, \( F^A_{\theta} + F^A_{\theta} \) is the area of the part of \( B \) that interacts with \( A \).

The angle histogram associated with a pair \((A, B)\) of disjoint objects is emulated by considering \( n \) evenly distributed directions, \( \theta_1, \theta_2 \) etc., and mapping each one of these directions, \( \theta \), to \( W_A W_B h \min(2W \tan \pi/n, h) \), where \( h, W, W_A, \) and \( W_B \) are calculated as in equation (5.9). Using the aggregation method [34, 10] to derive truth values from emulated angle histograms defines a new fuzzy model \( K' \), which emulates \( K \) through the \( \Phi \)-descriptor. Overall, the truth values produced by \( K \) and \( K' \) are remarkably close, which can be observed by comparing Table 5-1 and Table 5-2. In particular, the cardinal direction that gives the greatest truth value is the same in each case, and the ranking of directions is the same in each case but one.
Figure 5-4 Approximating a particular angle histogram value.

Figure 5-5 Deriving heights and widths of regions from the $\Phi$-descriptor.

Table 5-2 Emulating fuzzy models of directional relationships through the $\Phi$-descriptor. The values can be compared with those in Table 5-1.
5.1.2.2 Emulating F0 and F2

The truth values RIGHT\((A, B)\), ABOVE\((A, B)\), LEFT\((A, B)\), and BELOW\((A, B)\) produced by the model F0 (resp. F2) are derived, using the effective-force method [98], from the constant (resp. gravitational) force histogram [11] (see Section 2.2.3.2) associated with \((A, B)\). It is shown here that these histograms — and hence the models F0 and F2 — can be emulated by the \(\Phi\)-descriptor.

Consider two disjoint objects A and B. Assume that every point \(q\) of B exerts on every point \(q'\) of A an elementary force \(pq/|pq|^{r+1}\), where \(r\) is a real number and \(pq\) is the vector from \(p\) to \(q\) with length \(|pq|\). This force is in the direction \(pq/|pq|\), and its magnitude is \(1/|pq|^r\). If \(r\) is 0 then the force is constant, i.e., its magnitude does not depend on \(p\) or \(q\). If \(r\) is 2 then the force is gravitational. The force histogram associated with \((A, B)\) maps each direction \(\theta\) to the scalar resultant of all the elementary forces in direction \(\theta\). For example, when the two objects are as in Figure 5-4 and when \(\theta\) is the left-to-right horizontal direction, then the scalar resultant is \(W_A W_B h\) if \(r\) is 0, and it is

\[
h \log \left( \frac{W_A + W}{W(W_A + W + W_B)} \right)
\]

if \(r\) is 2. It should be noted that these two expressions are the result of a triple integration [91].

The constant (resp. gravitational) force histogram associated with a pair \((A, B)\) of disjoint objects is emulated by considering \(n\) evenly distributed directions, \(\theta_1, \theta_2, \ldots\), and mapping each one of these directions, \(\theta\), to the first (resp. second) expression, where \(h\), \(W\), \(W_A\), and \(W_B\) are as in (5.9). Using the effective-force method [66] to derive truth values from emulated constant (resp. gravitational) force histograms defines a new fuzzy model \(F0'\) (resp. \(F2'\)), which emulates \(F0\) (resp. \(F2\)) through the \(\Phi\)-descriptor. Here again, the truth values produced by \(F0\) and \(F0'\) (resp. \(F2\) and \(F2'\)) are remarkably close: See Table 5-1 and Table 5-2.
5.2 Experiments

A comparison of the models (i.e., M, K', F0', and F2') for images in Figure 5.3 is given in Figure 5-6. The descriptors (i.e., the Φ−descriptor) were generated using 72 reference directions (more reference directions gave similar results). As can be seen, the models produce comparable results for most of the configurations. For example, for configuration in Figure 5-6(a), all the models agree that A is to the right of B. The models M, F0', and F2', however, assert it with more confidence by reporting the highest truth degree of 1.00 for the proposition.

![Figure 5-6 Experimental results for directional relationships.](image-url)
Likewise, for configurations in Figure 5-6(c)(d)(f), all the models agree. The point of departure occurs in Figure 5-6 (e) where there is a disagreement between the models as whether A is more to the right of B or above B. In this case, whereas M and F2’ assert that A is more to the right of B, K’ and F0’ strongly assert that A is above B (truth degrees assigned to “A is above B” by K’ and F0’ are 0.62 and 0.72). The model M is more cautious and assigns a higher acceptability degree (i.e., 0.68) to the proposition “A is to the right of B” but also give sufficient weight (i.e., 0.41) to the possibility of “A is above B”. F2’ strongly asserts that A is to the right of B. It appears that K’ and F0’ favour the direction in which the more massive part of the object is located whereas F2’ gives more weight to the case when portions of the argument object are located in the target direction. This is more evident in the case of Figure 5-6(g) where object A is elongated (extends to the right) relative to B. For this case, F0’ and F2’ report the dominant relation between A and B to be the “below” relationship whereas K’ reports it to be the “right” relationship. M on the other hand takes a moderate position and whereas it identifies the “BELOW” relationship to be the dominant relationship between A and B nevertheless it also gives some credibility to the proposition “A is to the right of B”.

The results for the remaining configurations can be interpreted in a similar way. For Figure 5-6(h), M says that A is predominantly to the left of B but also is to a degree below and above it. F2’ makes a very strong assertion that A is to the left of B but negligibly below B. K’ reports A to be predominantly below B but also to a degree in other directions. The surround and partially surround cases in (j)(k)(l) are interpreted comparably by M, F0’, and F2’ but differently by K.

Which model is better is difficult to decide and depends on the context in which the models are used. If the requirement is to find the dominant relationship between objects A and B, then F2’ may be better but if the aim is to have the full picture, then the remaining models may be useful.
5.3 Summary

Two sets of models for directional relationships based on the $\Phi$-descriptor were proposed in this chapter; an original model $M$ and the emulation models, $K'$, $F0'$, and $F2'$ of angle and force histogram based models of directional relationships reported in literature. Experiments were carried out on synthetic images used in previous research. The results confirm that directional relationships can be interpreted from the $\Phi$-descriptor with good precision. The extraction is easier and the results meet expectations. Moreover, the descriptor allows to capture important properties such as the semantic inverse property in the models. Overall, it can be concluded that the $\Phi$-descriptor can be used to emulate existing models and build models of directional spatial relations.
Chapter VI

6 The surround relationship

Surroundedness plays an important role in scene analysis and description. Two notions of surroundedness have been used in the literature; visual surroundedness and topological surroundedness [79] (see Section 2.1.5). Topological surroundedness is defined in terms of topology [123]; visual surroundedness is based on the concept of visual direction. Methods have been proposed to interpret surroundedness directly from images using the notion of visual surroundedness [81]. Its extraction from relative position descriptors has, however, been considered in the case of spread histogram only [21, 124]. The problem with the modeling from spread histogram is that directional relationships cannot be captured and the cost of computation is high. The work in this chapter focuses on the modeling of surround relationship from the \( \Phi \)-descriptor using the notion of visual surroundedness. The aim is to demonstrate that it is possible to derive qualitative and quantitative models of visual surroundedness from the \( \Phi \)-descriptor in a variety of ways and that, the extraction is straightforward. The chapter is organized as follows. Section 6.1 discusses models of the “surround” relationships. Section 6.2 includes experimental results and Section 6.3 gives the summary.

6.1 Models of surroundedness from the \( \Phi \)-descriptor

The following sections introduce qualitative and quantitative models of the “surround” relationship. The qualitative definition is derived from the \( \Phi \)-descriptor based models of RCC23 convex hull relations of PIN and IN (sections 2.1.1 and 2.1.5). Two sets of quantitative models are proposed. The first model uses the fuzzification of the basic \( \Phi \) relations to calculate the degree of surroundedness between objects. The other models, i.e., the SURROUNDED0 and SURROUNDED1, derive models of surroundedness from the area functions of the \( \Phi \)-descriptor. Section 6.1.1 presents the qualitative models of the surround relationship and Section 6.1.2, the quantitative models.
6.1.1 Qualitative models of surround relationship from the \( \Phi \)-decriptor

Consider \textit{DISJOINTINTERIORS} and the basic \( \Phi \) relations in Table 6-1. They define a set of seven JEPD \( \Phi \) relations: \text{OUT0}, \text{OUT1}, \text{PIN0}, \text{PIN1}, \text{IN0}, \text{IN1} and \text{IN2}, as in Table 6-2. Four other relations are defined from that set: \( \text{OUT} = \text{OUT0} \lor \text{OUT1} \), \( \text{PIN} = \text{PIN0} \lor \text{PIN1} \), \( \text{IN} = \text{IN0} \lor \text{IN1} \lor \text{IN2} \) and \( \text{SURR} = \text{PIN} \lor \text{IN} \), where \text{PIN} stands for Partly IN and \text{SURR} for SURRounded. It can be noted that \( \neg \text{divides0} \land \neg \text{divides1} \land \neg \text{divides2} \) and \( \text{divides0} \land \neg \text{divides1} \land \text{divides2} \) are always false (the latter is a contradiction because the area function \( F_{AB}^d \) is continuous). \( \neg \text{divides1} \land \neg \text{divides2} \) (see OUT1) is, therefore, equivalent to \( \text{divides0} \land \neg \text{divides1} \land \neg \text{divides2} \). Likewise, \( \neg \text{divides1} \land \text{divides2} \) (see IN2) is equivalent to \( \neg \text{divides0} \land \neg \text{divides1} \land \text{divides2} \).

Now, assume the objects are restricted to connected objects (i.e., an object cannot be made of multiple pieces). Then \( \text{SURR}(A,B) \), i.e., “A is surrounded by B”, if the interiors of the objects do not intersect and if the convex hull of \( B \) intersects \( A \); Likewise, the condition, \( \text{PIN}(A,B) \) holds if the convex hull of \( B \) intersects but does not include \( A \); and \( \text{IN}(A,B) \) if the convex hull of \( B \) includes \( A \). Moreover, if \( \text{PIN0}(A,B) \) or \( \text{IN0}(A,B) \) then \( A \) can escape linearly and indefinitely in some direction without colliding with \( B \) (there is an opening large enough for \( A \) to pass through \( B \)); if \( \text{PIN1}(A,B) \) or \( \text{IN1}(A,B) \) then only a part of \( A \) can be moved that way (there is an opening, but not large enough for \( A \) to go through); and if \( \text{IN2}(A,B) \) then no part of \( A \) can be moved that way (there is no opening—or at least none visible from \( A \)). See Figure 6-1. Except for the rightmost one, all of the configurations in Figure 6-1 are homeomorphic to each other and, therefore, topologically equivalent. Again, surroundedness must be understood as visual surroundedness [79], not as topological surroundedness (Figure 6-2). Also, the model presented here is related to the \text{RCC23}, which is an extension of the \text{RCC8} composed of 23 JEPD relations. Indeed, if the objects are restricted to connected objects then OUT, PIN and IN coincide with the relations \text{OUTSIDE}, \text{P-INSIDE} and \text{INSIDE} that give rise to the \text{RCC23} [50]. As a corollary, the \text{RCC23} relations are \( \Phi \) relations.
Finally, the inverse relations \( \text{SURRI}_i, \text{PIN}_i, \text{PIN0}_i, \) etc., can also be easily defined: this can be done by simply replacing \( F_d^{AB} \) with \( F_e^{AB} \); \( |A| \) with \( |B| \); \( \text{divides}_0, \text{divides}_1, \) and \( \text{divides}_2 \) with \( \text{encloses}_0, \text{encloses}_1, \) and \( \text{encloses}_2 \); \( \text{SURR} \) with \( \text{SURR}_i \), \( \text{OUT} \) with \( \text{OUT}_i \), \( \text{OUT0} \) with \( \text{OUT0}_i \), etc. in Table 6-1 and Table 6-2.

![Diagram](image)

**Figure 6-1** A definition of surround. The relation \( \text{SURR} \) (resp. \( \text{SURR}_i \)) can be defined in terms of five pairwise disjoint relations: \( \text{PIN0}, \text{PIN1}, \text{IN0}, \text{IN1}, \text{IN2} \) (resp. \( \text{PIN0}_i, \text{PIN1}_i, \text{IN0}_i, \text{IN1}_i, \text{IN2}_i \)). It should be noted that, \( \text{PIN} = \text{PIN0} \lor \text{PIN1}, \text{IN} = \text{IN0} \lor \text{IN1} \lor \text{IN2} \), and finally, \( \text{SURR} = \text{PIN} \lor \text{IN} \).

![Diagram](image)

**Figure 6-2** Distinction between visual and topological surroundedness. Although \( A \) is not topologically surrounded by \( B \), it is visually surrounded by \( B \) because there is no opening in \( B \) that is visible from \( A \), and we have \( \text{IN2}(A, B) \). Compare with the last configuration in Figure 6-1 (labelled IN2), where \( A \) is both visually and topologically surrounded by \( B \).

<table>
<thead>
<tr>
<th>name</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>divides0</td>
<td>( \exists \theta, F_d^{AB}(\theta) = 0 )</td>
</tr>
<tr>
<td>divides1</td>
<td>( \exists \theta, (F_d^{AB}(\theta) \neq 0 \land F_d^{AB}(\theta) \neq</td>
</tr>
<tr>
<td>divides2</td>
<td>( \exists \theta, F_d^{AB}(\theta) =</td>
</tr>
</tbody>
</table>

Table 6-1 Basic \( \Phi \) relations to model surroundedness
Table 6-2 A model of surroundedness. It is defined in terms of PIN0, PIN1, IN0, IN1 and IN2, which constitute with OUT0 and OUT1 a set of seven JEPD $\Phi$ relations.

<table>
<thead>
<tr>
<th>name</th>
<th>OUT0</th>
<th>OUT1</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUT</td>
<td>$\neg$disjointInteriors &amp; $\neg$divides1 &amp; $\neg$divides2</td>
<td></td>
</tr>
<tr>
<td>PIN</td>
<td>PIN0</td>
<td>PIN1</td>
</tr>
<tr>
<td>PIN</td>
<td>disjointInteriors &amp; divides0 &amp; divides1 &amp; $\neg$divides2</td>
<td></td>
</tr>
<tr>
<td>PIN</td>
<td>divides0 &amp; divides1 &amp; $\neg$divides2</td>
<td></td>
</tr>
<tr>
<td>IN</td>
<td>IN0</td>
<td>IN1</td>
</tr>
<tr>
<td>IN</td>
<td>disjointInteriors &amp; divides0 &amp; divides1 &amp; divides2</td>
<td></td>
</tr>
<tr>
<td>IN</td>
<td>divides0 &amp; divides1 &amp; divides2</td>
<td></td>
</tr>
<tr>
<td>IN</td>
<td>divides1 &amp; divides2</td>
<td></td>
</tr>
</tbody>
</table>

6.1.2 Quantitative models

The following sections give the quantitative models. Section 6.1.2.1 describes the model based on convex hull $\Phi$ relations. Sections 6.1.2.2 and 6.1.2.3 present the models derived from the area functions of the $\Phi$-descriptor.

6.1.2.1 First model

Consider the fuzzification of the JEPD relations OUT0, OUT1, PIN0, PIN1, IN0, IN1 and IN2 (Section 6.1.1). Each of these relations is defined as a conjunction of basic $\Phi$ relations (possibly negated) in which each basic relation appears at most once. Moreover, the conjunctions reflect the JEPD property. The general expectations listed in Section 4.2.1 can, therefore, be met by fuzzifying the basic $\Phi$ relations and replacing the negation $\neg$ with the function $x \mapsto 1 - x$ and the conjunction $\wedge$ with the function $(x, y) \mapsto xy$. See Section 4.2.2 for explanation. To fuzzify the other relations, OUT, PIN, IN and Surr, the disjunction $\vee$ can be replaced with the function $(x, y) \mapsto x + y$. For example, let $S\overline{U}R\overline{R}$, $I\overline{N}$, $P\overline{I}N$, etc. represent the fuzzy counter-parts of Surr, IN, PIN, etc., then:

$$S\overline{U}R\overline{R}(A, B) = P\overline{I}N(A, B) + I\overline{N}(A, B),$$
\[ \tilde{I}(A, B) = I_0(A, B) + I_1(A, B) + I_2(A, B), \]

\[ I_2(A, B) = \text{disjointInteiors}(A, B) \times \left( 1 - \text{divides}_1(A, B) \right) \times \text{divides}_2(A, B) \]

It should be noted that \( \tilde{O}(0), \tilde{O}(1), \tilde{P}(0), \tilde{P}(1), I_0, I_1 \) and \( I_2 \) are JEPD, the relations \( \tilde{O}(0), \tilde{P}(0) \) and \( I_0 \) are JEPD, and the relations \( \tilde{O}(0) \) and \( S \) are JEPD as well.

Thus the problem of finding fuzzy models of \( S \), \( I(A, B) \), etc. then comes down to finding a way to fuzzify the basic \( \Phi \) relations \( \text{DISJOINTINTERIORS}, \text{DIVIDES}_0, \text{DIVIDES}_1, \) and \( \text{DIVIDES}_2 \). A method to fuzzify these relations is given in Figure 6-3. The fuzzification functions use the quantities \(|A|, |B|\) and \(|A \cap B|\), which are either elements of the \( \Phi \)-descriptor or are measures that can be derived from it. Threshold other than those shown in the figure could be chosen for the triangular functions, or trapezoidal functions instead of triangular ones, but the basic principle remains the same. The fuzzification of \( \text{divides}_0 \), \( \text{divides}_1 \) and \( \text{divides}_2 \) is based on the following observations: if there is a direction \( \theta \) such that \( F_{AB}(\theta) \) is close to 0 then \( \min F_{AB}(\theta) \) is close to 0 (and vice versa); if there is a direction \( \theta \) such that \( F_{AB}(\theta) \) is far from 0 and far from \(|A|\) then \( F_{AB}(\theta) \) and \( |A| - F_{AB}(\theta) \) are far from 0, i.e., \( \min \left( F_{AB}(\theta), |A| - F_{AB}(\theta) \right) \) is far from 0, i.e., \( \max \min(F_{AB}(\theta), |A| - F_{AB}(\theta)) \) is large (and the largest possible value is \(|A|/2\)); if there is a direction \( \theta \) such that \( F_{AB}(\theta) \) is close to \(|A|\) then \( \max \theta F_{AB}(\theta) \) is close to \(|A|\).

\[ \text{disjointInteiors}(A, B) = \max \left( 0, 1 - 2 \times \frac{|A \cap B|}{\min(|A|, |B|)} \right) \]  
\[ \text{divides}_0(A, B) = \max \left( 0, 1 - 2 \times \frac{\min F_{AB}(\theta)}{|A|} \right) \]  
\[ \text{divides}_1(A, B) = \frac{2}{|A|} \max \min \left( F_{AB}(\theta), |A| - F_{AB}(\theta) \right) \]  
\[ \text{divides}_2(A, B) = \max \left( 0, 2 \times \frac{\max F_{AB}(\theta)}{|A|} - 1 \right) \]
Fuzzifying the relations PIN0, PIN1, IN0, IN1 and IN2 as in Section 6.1.2.1 amounts to considering that there is no clear boundary between the different ways an object may be surrounded by another. However, the proposition “A is surrounded by B” may give rise to the question, “to what extent?” instead of “in which way?” To answer this question, there is a need for more satisfactory fuzzy models of surroundedness. A second model of surroundedness, the \( SURROUND_0 \), is defined by:

\[
SURROUND_0(A, B) = \frac{1}{2\pi \max_{\theta} F_d^{AB}(\theta)} \int_0^{2\pi} F_d^{AB}(\theta) d\alpha 
\]

(6-5)

where \( \theta_\alpha = [\cos \alpha \sin \alpha]^T \). In this model, the surroundedness degree is derived from the average value of \( F_d^{AB}(\theta) \) for all directions. The presence of \( \max_{\theta} F_d^{AB}(\theta) \)—which is the maximum possible value for \( F_d^{AB}(\theta) \)—ensures that the result belongs to the interval \([0,1]\). There is an obvious similarity between the expression in equation (6.5) and the degree of visual surroundedness introduced in [79]. The truth value in Figure 6-4(a) should seem reasonable to the reader. On the other hand, the configurations in Figure 6-4(a) and
Figure 6-4(b) give exactly the same value, which may negate reader’s perception; the configuration in Figure 6-4(c) gives 0, which may also go against the reader’s perception. Unfortunately, it is not possible to design a satisfactory model, SURROUNDED, such that \( \text{SURROUNDED}(A, B) \neq 0 \) when \((A, B)\) is the pair in Figure 6-4(c). This is a limitation of all \(F\)-histogram based relative position descriptors and is a consequence of the way disconnected objects are handled by these descriptors. Consider Figure 6-4(c) (d); in both cases, \(|A| = |B|\). Consider any area or length function value at direction \([\cos \alpha \sin \alpha]^t\), where \(\alpha\) is \(-\pi/2\) (resp. \(\pi/6, 5\pi/6\)); this value is exactly the same in Figure 6-4(c) as in Figure 6-4(d). Consider any area or length function value at direction \([\cos \alpha \sin \alpha]^t\), where \(\alpha\) belongs to the interval \([-\pi/2 + \pi/12, \pi/6 - \pi/12], [\pi/6 + \pi/12, 5\pi/6 + \pi/12], \) or \([5\pi/6 + \pi/12, 3\pi/2 - \pi/12]\); this value is 0 in Figure 6-4(c) as in Figure 6-4(d). In the end, it is hard for the \(\Phi\)-descriptor to distinguish between the two configurations and determine which object surrounds which.

Figure 6-4 Surroundedness and disconnected objects. \(\text{SURROUNDED}_0\) gives the same truth value in (a) and (b). It is hard for the \(\Phi\)-descriptor to distinguish between (c) and (d). It should be noted that the object in light gray is the object A (argument object) and that in dark gray is the object B (reference object).
6.1.2.3 Third model

The model SURROUNDED_0 presented in Section 6.1.2.2 may be overly pessimistic in its assessment of surroundedness. It gives reasonable truth values in many cases, but it may also give lower than expected truth values of surroundedness, especially when the surrounding object is disconnected. This section introduces another model, SURROUNDED_1, of surroundedness, which exhibits the opposite behavior, i.e., it may be overly optimistic. Thus for any objects A and B, the condition, $0 \leq \text{SURROUNDED}_0(A,B) \leq \text{SURROUNDED}_1 \leq 1$ may hold. This gives rise to the possibility of considering SURROUNDED_0 as a necessity measure and SURROUNDED_1 as a possibility measure [62].

The model SURROUNDED_1 is defined as follows:

$$\text{SURROUNDED}_1(A,B) = \frac{1}{2\pi|F|} \int_0^{2\pi} \frac{H^{AB}_{tf}(\theta)}{\max(H^{AB}_{tf}(\theta),H^{AB}_{tf}(-\theta))} F^{AB}_{\text{cla}}(\theta) d\alpha$$  \hspace{1cm} (6-6)

where

$$H^{AB}_{tf}(\theta) = H^c_t(\theta) + H^f_t(\theta) + H^o_A(\theta)$$

$$F^{AB}_{\text{cla}}(\theta) = F^c_{\text{AB}}(\theta) + F^l_{\text{AB}}(\theta) + F^a_{\text{AB}}(\theta)$$

$$F^{AB}_{\text{cla}}(-\theta) = F^c_{\text{cla}}(\theta) + F^a_{\text{cla}}(-\theta)$$

and $0/0$ is 0. The quantity $F^{AB}_{\text{cla}}(\theta)$ is the area of a subregion of the region of interaction in direction $\theta$: the subregion covered by A, and A only. The expression in equation (6.6) is basically a weighted average of all the $F^{AB}_{\text{cla}}(\theta)$ areas. $|F|$ is the maximum possible value for $F^{AB}_{\text{cla}}(\theta)$, and its presence ensures the result is restricted to the interval [0,1].

The weight has the following explanation. The region of interaction in direction, say, $\theta_0$ (left-to-right), or $\theta_\pi = -\theta_0$, is considered to be as in Figure 6-5. The area $F^{AB}_{\text{cla}}(\theta_0)$ is that of the light grey region (not dark or medium grey), and is equal to $F^{AB}_{\text{cla}}(\theta_\pi)$. The whole light grey region is within the height $H^{AB}_{tf}(\theta)$ when $\theta = \theta_0$, but only about
The five sets of images used for experiments are shown in Figure 6-6 through Figure 6-10. It should be noted that the dark grey object represents object B whereas the light-grey object, the object A. In the figures, \textit{SURR}2 denotes \textit{SURR}2 or \textit{SURR}2\_i depending on whether A or B is the argument. Likewise, \textit{SURR}5 denotes \textit{SURR}5 or \textit{SURR}5\_i. Moreover, \textit{SURR}2\_i = \textit{IN}i + \textit{PIN}i (resp. \textit{SURR}2 = \textit{IN} + \textit{PIN}), \textit{SURR}5\_i = \textit{IN}0i + \textit{IN}1i + \textit{IN}2i + \textit{PIN}0i + \textit{PIN}1i (resp. \textit{SURR}5 = \textit{IN}0 + \textit{IN}1 + \textit{IN}2 + \textit{PIN}0 + \textit{PIN}1). \textit{SURR}0 denotes \textit{SURROUND}0 (second model) and \textit{SURR}1 denotes \textit{SURROUND}1 (third model). When A is the argument object, the notation \textit{IN}, \textit{PIN}, etc. is used. When B is the argument object, \textit{IN}i, \textit{PIN}i, etc. is used. As apparent from the configurations in the figures, the results meet the expectations. For example, in Figure 6-6(a), all the surroundedness types have been reported to be 1.0, which matches the definition of visual surroundedness. For Figure 6-6(b), a higher truth value is produced for \textit{IN}1i and a lower truth value for \textit{IN}0i because the opening in the object B is not wider enough for object A to pass through. This changes in Figure 6-6(c)(d) because now the opening is wide enough for the object A to go through and consequently a higher truth value is reported for \textit{IN}1i. In Figure 6-6(a)-(h), the truth value for \textit{SURR}1 and \textit{SURR}2 decreases gradually as the condition for visual surroundedness continuously weakens with \textit{SURR}1 reaching the zero truth degree earlier because it is based on the pessimistic assessment of surroundedness. In Figure 6-7 are the configurations, in which the two objects intersect. The truth-degrees are influenced...
accordingly. For example, in 6.7(a), all surround types are reported to have truth degrees 1.0 but the degrees quickly decrease as the objects intersect. Figures 6.8 and 6.10 include experiments on objects, which are in PIN/IN - PINi/INi relationship. Figure 6.9 includes miscellaneous configurations. As can be seen, results correspond to the definition of surroundedness.

### 6.3 Summary

Surroundedness is an important category of spatial relationships and has been understood in terms of visual and topological surroundedness in literature. Like other spatial relationships, it plays an important role in image analysis and description. In this chapter, it was demonstrated that models of surroundedness as understood visually can be derived from the $\Phi$-descriptor in a variety of ways. Moreover, both qualitative and quantitative definitions of surroundedness can be obtained. The work in this chapter suggests that many more efficient models of surroundedness may likely be derived from the $\Phi$-descriptor.
Figure 6.6 Experimental results for Surroundedness: I

Figure 6.7 Experimental results for surroundedness: II
Figure 6-8 Experimental results for surroundedness: III
Figure 6-9 Experimental results for surroundedness: IV

Figure 6-10 Experimental results for surroundedness: V
7 Computation of the $\Phi$-descriptor

The $\Phi$ –descriptor has a low computation time for both raster and vector objects (Section 3.3.1). For raster objects, the computation time depends on the number of pixels and number of reference directions to produce the descriptor. For vector objects, the computation time is determined by the number of vertices of polygons into which the objects are partitioned. In this chapter, an algorithm for computing the $\Phi$ –descriptor for raster objects is presented. Section 7.1 describes the algorithm, Section 7.2 presents the results, and Section 7.4 gives the summary. Section 7.3 includes a short discussion of algorithm for computing the $\Phi$-descriptor for vector data (which is the future work).

7.1 An algorithm for computing the $\Phi$-descriptor for raster objects

The $\Phi$-descriptor is computed in $O(nN)$ time, where $N$ is the number of pixels in the image and $n$ is the number of directions $\theta$ considered. The higher the number of direction $n$ used to compute the descriptor, the richer is the collected histogram data, but the longer is the processing time. Practically, there generally does not arise a need to consider more than a few hundred directions when computing $F$-histograms, and $n$ can thus be chosen between 4 and 360 [90].

All the area and length histograms are handled simultaneously. For every direction $\theta$, the image is partitioned into parallel raster lines. The pixels in a line are examined one by one and all the $F^A_B(\theta)$ and $F^A_B(-\theta)$ values are updated on the fly; basically, calculating the descriptor comes down to counting the number of pixels between every two consecutive entry or exit points (i.e., boundary pixels, see Section 3.2).

Actually, only a limited number of raster lines and pixels are usually considered. As shown in Figure 7-1, a projection parallel to $\theta$ transforms the MBR (Minimum Bounding Rectangle) of A and the MBR of B into two aligned segments, which are either vertical (if
Figure 7-1 Relevant raster lines and pixels. Here, the dark grey (resp. medium grey, light grey) rectangle is the MBR of $B$ (resp. $A$, $A \cup B$). There are 3 relevant raster lines in
direction $\theta$, and each is composed of 12 relevant pixels (labelled 1, 2, and 3, resp.). These
36 pixels are the only ones to be considered when computing the $F_{i}^{AB}(\theta)$ and $F_{t}^{AB}(\theta)$
values.

$|\cos \theta| \geq |\sin \theta|$ or horizontal (if $|\cos \theta| < |\sin \theta|$). The intersection of these segments
determines the raster lines in direction $\theta$ that cross both MBRs. These relevant lines may
cross both objects and must therefore be considered, while the other lines can be ignored.
Likewise, the relevant pixels in each raster line are those contained in the MBR of $A \cup B$, and the other pixels of the line can be ignored.

The algorithm below computes the 10 basic area histograms as well as the corresponding
length histograms. For the sake of clarity, reverse directions are not processed
simultaneously, and it is assumed that the first and last rows and columns of the image
do not intersect any object.

7.2 Experiments

7.2.1 Model descriptors for comparison

Four descriptors are considered in this section: a 10-tuple $\Phi$-descriptor, a 28-tuple $\Phi$-
descriptor, the force histogram, and the skeleton F-histogram. The 10-tuple $\Phi$-descriptor
is the basic descriptor defined as in Section 3.2.5. The 28-tuple Φdescriptor is the tuple whose elements are the 10 basic elements as well as the 18 additional elements (2 areas, 3 area histograms and 13 length histograms) listed in the introductory paragraph of

**ALGORITHM 1: 20 tuple Φ-decriptor**

1. **for** each reference direction θ

2.    **for** each histogram index i (other than e, d, w)

3.       set the area histogram value \( F_i^{AB}(\theta) \) to 0

4.       set the counter value \( C_i \) to 0

5.    **for** each raster line \( L \) in direction θ

6.       set point1Category to “nil”

7.       set state1 to “not in I and not in J”

8.       set length to 0

9.    **for** each pixel of \( L \) (from first to last)

10.      set state2 to the pixel’s state

11.      increase length by 1

12.      **if** state2 is not equal to state1

13.         **if** point1Category is not “nil”

14.            **set** i to the appropriate histogram index (depending on point1Category and point2Category)

15.            **increase** \( F_i^{AB}(\theta) \) by length, length/2 or 0

16.            **increase** \( C_i \) by 1, 1/2 or 0

17.      **set** point1Category to point2Category

18.      **set** state1 to state2

19.      **set** length to 0

20.    **for** each histogram index i (other than e, d, w)

21.       **set** the length histogram value \( F_i^{AB}(\theta) \) to \( F_i^{AB}(\theta)/ \left( C_i \max\{\cos(\theta), \sin(\theta)\} \right) \)

**Line 7:** The symbol \( I \) denotes the set of pixels that belong to both \( A \) and \( L \), while \( J \) denotes the set of pixels that belong to both \( B \) and \( L \). — **Line 13:** For example, if state1 is “not in I and not in J” and state2 is “in I but not in J” then point2Category is set to “\( I \)-entry not in J”; if state1 is “in I but not in J” and state2 is “in I and in J” then point2Category is set to “\( J \)-entry in I”.* — **Line 15-17:** For example, if point1Category is “\( I \)-exit J-exit” and point2Category is “\( J \)-entry not in I”, then \( F_\theta^{AB}(\theta) \) is increased by length and \( C_\theta \) is increased by 1; if point1Category is “\( J \)-entry not in I” and point2Category is “\( J \)-exit not in I”, and if \( L \) crosses both objects, then \( F_\theta^{AB}(\theta) \) is increased by length/2 and \( C_\theta \) is increased by 1/2; if point1Category is “\( I \)-exit J-exit”, then \( F_\theta^{AB}(\theta) \) is increased by 0 and \( C_\theta \) is increased by 0 (i.e., nothing is done).
Section 3.2.6.2. The force histogram is an F-histogram; it is probably the relative position descriptor backed up with the most theoretical and applied results [99]. The skeleton F-histogram is a void descriptor (see Algorithm 2); it leads to the shortest processing times that can be expected from a descriptor computed according to the general principle of F-histogram computation. The Φ-descriptor and skeleton F-histogram are computed in \(O(nN)\) time. As for the force histogram, two algorithms are available: one runs in \(O(nN\sqrt{N})\) time and the other in \(O(N \log N)\) time [91].

**Algorithm 2: Skeleton F-histogram**

1. for each reference direction \(\theta\)
2. 
   for each raster line \(L\) in direction \(\theta\)
3.   
   for each pixel of \(L\) (from first to last)
4.     set state to the pixel’s state

### 7.2.2 The test objects

Each object is represented by a square 8-bit grayscale image. The pixels that belong to the object are those with gray level 255. The number of pixels in an image goes from 100,000 to 1,000,000 with an increment of 100,000. Consider a pair of objects represented by two images of the same size and such that all the raster lines and pixels are relevant: the shortest possible processing times (best-case scenario) are obtained when the gray level of each pixel in each image is set to 255 (square objects); the longest possible processing times (worst-case scenario) are obtained when the gray level of each pixel in each image is chosen at random from the set \{0,255\} (randomly-built objects).

### 7.2.3 Results

Figure 7-2 and Figure 7-3 compare the processing times for the different descriptors. The unit of image size \((N)\) is 100,000 pixels and the unit of processing time \((t)\) is 1 second. In all cases, \(n = 100\) reference directions were considered: the directions \(2\pi k/100\) with \(k\) an integer between 0 and 99. The processing times obtained with the linearithmic time algorithm for force histogram computation are basically independent of \(n\); all the other processing times are exactly proportional to \(n\). The algorithms were implemented in C
and run on a MacBook Pro equipped with 2.8 GHz Intel Core i7, 8 Go 1067 MHz DDR3, OS X 10.8.5 and GCC 4.2.1.

Here are a few points worth noting: the processing times for the 28-tuple $\Phi$-descriptor are only about 75% longer than those for the skeleton $F$-histogram and only about 25% longer than those for the 10-tuple $\Phi$-descriptor (Figure 7-2, graph on the left); in the best-case scenario, the processing times for the 10-tuple $\Phi$-descriptor are practically the same as those for the skeleton $F$-histogram (Figure 7-2, graph on the right); the processing times for the force histogram in the best-case scenario are practically the same as those for the 28-tuple $\Phi$-descriptor in the worst-case scenario (Figure 7-3); in the best-case scenario, the processing times indicated in Figure 7-3 for the force histogram were obtained with the $O(N \log N)$ algorithm (the other algorithm gives higher processing times, e.g., 15 seconds for a 1000×1000 image); in the worst-case scenario, they were obtained with the $O(N \log N)$ algorithm (the other algorithm gives much higher processing times, e.g., 270 seconds for a 1000×1000 image).

In the end, computation of the $\Phi$-descriptor is extremely fast for a descriptor that is based on the concept of the $F$-histogram and that carries so much spatial relationship information. In particular, it is much faster than that of the force histogram.
7.3 A note on computing the $\Phi$-descriptor for vector data

An algorithm for computing the $\Phi$-descriptor for raster data was proposed above. An algorithm for calculating the descriptor for vector data will be investigated in future. From a theoretical point of view, however, algorithm for computing the $\Phi$-descriptor for vector data is possible through the method of polygon partitioning. Polygon is a geometric primitive in vector data format (besides point and line) and is used in GIS (Geographic Information System) to represent 2D geospatial entities with areas and discrete boundaries such as, cities, lakes, and forests [125] and in the field of interactive computer.
graphics for designing simple, hardware embedded rendering algorithms [126]. They are also the principle tool in the method of polygonal approximation for capturing the outline of a shape for use in shape analysis and recognition [127, 128].

The principle of polygon partitioning is illustrated in figures 7.5 and 7.5. Directed lines going through the objects vertices partition the objects into trapezoidal pieces. The trapezoids are cut into further pieces by lines passing through the intersection points of the polygons. The sub-pieces are sorted along the considered direction and associated with specific sub-regions in the region of interaction of A and B, e.g., into sub-regions occupied by only A or only B or both A and B, etc. The elements of the \( \Phi \)-descriptor (i.e., \( F^A_{aB} \), \( F^A_{cB} \), \( F^A_{fB} \), etc.) are then calculated from the areas of the sub-pieces of trapezoids. Thus in this method, computation time of the descriptor will depend on the number of polygon intersections. In the worst-case scenario when the number of intersections of two polygons would be \( pq \) (\( p \) and \( q \) number of vertices of the polygons) [129], the computation time can be \( \mathcal{O}(K \eta^3 \log \eta) \) where \( \eta \) denotes the total number of vertices and \( K \), the number of directions. In the best case scenario (when the objects are convex), two polygons will intersect at most in \( \min(2p, 2q) \) points [130] and the time will be \( \mathcal{O}(K \eta^2) \).

Figure 7-4  Partitioning of the objects. (a) Each object is divided into trapezoidal pieces. (b) The trapezoids \( p_1p_2p_3p_4 \) and \( q_1q_2q_3q_4 \) are cut along the dotted lines into three pieces each. (c) \( p_1p_2p_3p_4 \) is broken into \( p_1q_2q_3p_4 \) and \( q_2p_2p_3q_3 \), and \( q_1q_2q_3q_4 \) is broken into \( q_1p_1q_4p_4 \) and \( p_1q_2q_3p_4 \).
Figure 7-5 Tuples of trapezoids sorted in direction $\theta$. The tuple shown here is $(t_1, t_2, t_3, t_4, t_5)$, where $t_1$ is a piece of $A - B$, $t_2$ is a piece of $A \cap B$, $t_3$ and $t_5$ are pieces of $B - A$ and $t_4$ is a piece of $A \cup B$. The area of $t_1$ contributes to $F_{c}^{AB}(\theta)$, the area of $t_2$ to $F_{c}^{AB}(\theta)$, the area of $t_3$ to $F_{u}^{AB}(\theta)$, the area of $t_4$ to $F_{v}^{AB}(\theta)$ and $F_{v}^{AB}(-\theta)$, etc.

7.4 Summary

An algorithm for the $\Phi$-descriptor was developed and tested for the 10-tuple and 28-tuple variants of the descriptor using grayscale images of realistic sizes. The processing times were compared with those for the skeleton F-histogram and force histogram. It was established on the basis of the above experiment that the $\Phi$-descriptor is extremely fast and outperforms force histogram in terms of running time.
Chapter VIII

8 Summary and future work

A new relative position descriptor, i.e., the $\Phi$-descriptor, was introduced in this work based on the original idea of associating spatial relationship categories with pairs of aligned boundary points of objects [131]. The idea of aligned points was combined with the modeling principles from $F$-histogram, Radon transform, and Allen interval relations to derive an efficient model of relative position. The approach has many advantages. The incorporation of features from Radon transform, $F$ – histogram, and Allen relations with a well-developed theory gives the new descriptor a potentially strong theoretical basis. The use of a parallel beam of directed lines as in the $F$ – histogram and Radon transform to reduce the handling of $nD$ objects to the handling of more primitive and simpler 1D entities gives the new descriptor better computational efficiency and the ability to handle complex objects (e.g., objects with holes or objects with multiple disconnected components). For example, the quantitative evaluation of relative position is reduced to the simpler tasks of measuring areas and volumes of directed regions. A consequence of this feature may be that certain problems such as the behavior of the descriptor to geometric transformations can be easily modeled. Last, the use of Allen interval relation algebra allows the descriptor to capture a richer variety of spatial relationship information.

The new descriptor has other nice properties. This includes the ability to handle both raster and vector objects, retrieval of semantic inverse, and an easy normalization procedure. A detailed description of the properties of the $\Phi$-descriptor was given in Chapter III. A feature of the descriptor is the capability to represent a wide range of spatial relationships. As shown in Chapters IV-VI, topological, directional, and surround relationships can be interpreted from the descriptor. The extraction is straight forward and the same relationship type can be derived from the descriptor in many ways. Furthermore, both quantitative and qualitative models of spatial relationships can be derived from the descriptor. In Chapter IV, qualitative models of the RCC8 and RCC23 topological
relationships based on the $\Phi$-descriptor were derived. The models are built on first-order logic applied to the basic $\Phi$-relations after some fine-tuning. The derivation is thus straightforward and the models take the form of simple logical expressions. The quantitative models of the RCCn (RCC8 and RCC23) relations involves the fuzzification of the basic $\Phi$-relations. In addition to the RCCn relations, qualitative models of DE-4IM relations were also derived. In Chapter V, models of directional relationships obtained from the $\Phi$-descriptor were proposed. It was established that new models of directional relationships can be derived from the $\Phi$-descriptor as well as existing models can be emulated. Surroundedness is an important spatial relationship. In Chapter VI, some methods for deriving qualitative and quantitative models of surroundedness were presented. It was shown that qualitative definition of visual surroundedness based on RCC23 convex hull relations can be obtained from the $\Phi$-descriptor. Last, an algorithm for computing the $\Phi$-descriptor was presented in Chapter VII. The computation time for the 10-tuple and 28-tuple $\Phi$-descriptor was compared with that of the force histogram for realistic sized images and was found to be much lower.

Finally, Chapter II gave a comprehensive review of the literature on relative position in which a first classification of relative position descriptors was proposed. The descriptors were classified as generic or specific, point-based or pixel, segment, or core-based, and region or boundary based. A list of properties for comparing relative position descriptors was also suggested. The work in this segment may contribute to making the study and understanding of relative position descriptors more systematic.

**8.1 Conclusions**

The idea of aligned boundary points of objects combined with modeling principles from Radon transform, F-histogram, and Allen relations gave rise to an efficient relative position descriptor. The descriptor defines a region of interaction comprising directionally aligned regions of the argument and reference objects and divides the region into sub-region based on the position of the 1D segments of the objects relative to each other.
With the new approach, it is possible to retrieve a variety of spatial relationships from the descriptor. Moreover, the descriptor is fast, can handle raster as well as vector objects, and has known behavior to affine geometric transformations.

### 8.2 Future work

The work on the Φ-descriptor can be expanded into many directions. For example, modeling of Allen relations, extraction of more complex spatial relationships such as “among” and “between”, handling of 3D raster and vector objects, and extension to fuzzy object are some areas for future research on the descriptor. Other potential directions for future research on the Φ-descriptor are the recovery of the objects associated with a descriptor (from the descriptor) and investigation of the shape description capabilities of the descriptor. Information preserving shape descriptors [132, 133] have been proposed in the area of shape representation [134]. The Φ-descriptor encapsulates information about the shape geometry of objects and re-construction of the objects from the descriptor may be a potential future work on the descriptor. Investigation of the shape description capability may be another research topic. In particular, it would be interesting to explore whether shape parameters can be obtained from the Φ-descriptor. Shape parameters (e.g., rectangularity, convexity, etc.) are simple geometric features of shape and can be used to eliminate false hits in content-based image retrieval [CBIR] from large databases or can be coupled with other shape descriptors in shape matching and retrieval tasks [29]. In future, work on the Φ-descriptor will be expanded in the following directions.

**Generation of linguistic description of relative position from the Φ-descriptor.** Linguistic description of relative position refers to the communication of quantitative or qualitative relative position information through linguistic expression. Linguistic description of relative position has application in many areas, e.g., navigation tasks in mobile robots, target recognition and acquisition, digital medical image diagnostics, landmark navigation, image database search and retrieval, and numerous GIS applications [106]. Although, not widely studied in the case of other descriptors, extraction of linguistic description of
relative position from force histogram has attracted a good amount of attention in literature. For example, the use of linguistic description of relative position based on force histogram has been studied in mobile robot navigation [95, 135, 136] and digital image analysis [66, 137]. A strength of Φ-descriptor is that it encapsulates a wider range of spatial relationships. Consequently, high quality linguistic description of a variety of spatial relationships like topological, directional, and surround relationships can be obtained from the Φ-descriptor (See chapters IV-VI) to be used in tasks such as mobile robot navigation, scene analysis, etc. A focus of the future work will be the extraction of linguistic description of relative position from the Φ-descriptor.

Handling of vector objects: An algorithm to implement the Φ-descriptor for raster objects was presented in Chapter VII. A target of the future work is the extension of the Φ-descriptor to vector objects. Vector data has extensive applications in a number of areas, e.g., in the domain of GIS (Geographic Information Systems), vector data can represent 2D geospatial entities with areas and discrete boundaries such as cities, lakes, or forests. Vector format is based on points, lines, and polygons, and can be modeled by the concept of 1D entities used in the Φ-descriptor. An approach in a vector-based algorithm for the Φ-descriptor would be to partition the objects into polygons and calculate the descriptor in terms of the areas of the polygons (See Chapter VII). Methods to perform this efficiently have been proposed in the area of computational geometry [129].

Extraction of affine transformations: The Φ-descriptor has a known behaviour to affine transformations, i.e., when objects undergo affine transformation, the associated descriptor is scaled by the absolute value of the determinant of the linear part of the transformation (see Section 3.3.4.1). So, given an affine transformation applied to an object-pair and the Φ-descriptor associated with the object pair, the resulting Φ-descriptor can be determined in a straightforward manner (called the direct problem in this work). An interesting problem is the inverse problem, which deals with the question that, given the descriptors before and after the affine transformation, is it possible to determine the affine transformation from that information? Inferring affine parameters from images is an
important problem in image processing and is studied in areas like image registration [138] and motion estimation. For this purpose, techniques such as search optimization and parameter estimation methods, etc. are used. Extraction of affine transformations from relative position descriptors have been considered in a few studies only [99]. A target of the future work will be Inferring affine parameters from the \( \Phi \)-descriptor.

8.3 Summary

A summary of the work in this thesis was presented in this chapter and a number of areas for future work on the \( \Phi \)-descriptor were identified. Finally, linguistic description of relative position, extension to vector data, and retrieval of affine transformations from the descriptor were identified as the intended areas of future research.
9 References


[54] P. Matsakis and M. Naeem, "Fuzzy models of topological relationships based on the phi-descriptor," in *IEEE World Congress on Computational Intelligence (Accepted)*, Vancouver, Canada, 2016.


(IEEE World Congress on Computational Intelligence). IEEE International Conference on, Hong Kong, 2008.


147


Appendix A  Modeling of the distance relationships from $\Phi$-descriptor

Two approaches are used in existing methods for the fuzzy modeling of distance spatial relationship between objects. The first group, i.e. the *relative distance approach*, computes the distance spatial relationships (“far”, “near”) from the geometric properties of the objects and/or the scene into which the objects are embedded. In this approach, no prior knowledge of the mapping between the distance and the “far” and “near” degrees is required. The truth-degrees are thus computed from information intrinsic to the scene or objects. In the second approach, i.e. the *absolute distance approach*, the modeling of distance relationships is based on a prior knowledge of the limits of metric distance when distance map to maximum or minimum “far” and “near” degrees. The methods thus rely on extrinsic information. Additionally, some methods consider intersecting objects; others non-intersecting. In this work, the relative distance approach is considered. However, the absolute distance approach will also be explored cursorily. Furthermore, only two types of distance relationships are considered; “far” and “near”. The quantitative models proposed for these relationships is based on the following considerations.

A.1  Considerations

The following properties are sought in the “far” and “near” relationships. Both “far” and “near” are symmetric, i.e., if “A if far from B” then “B is far from A” or when “A is near to B” then “B is near to A”. Extension to both infinite and finite spaces, i.e., if $D \in [L, U]$ where $1 \leq L < U$ and $U < \infty$, $\lim_{D \to U} far(A, B) \equiv 1$; if $D \in [1, \infty)$, $\lim_{D \to \infty} far(A, B) \equiv 1$. Invariance to scale change, i.e., scale change doesn’t affect the truth-degree of the relationship.

A.2  Modeling the far relationship

Let A and B be two crisp objects with A as the argument object and B the referent object. Then the “far” relationship between A and B is expressed as the proposition “A is far from B”. The solver system will associate a value in [0,1] with the proposition that will represent its truth degree. The truth degree is determined from the size of the region between
objects A and B. Let $D$ be the average area of this region. Thus, to model “far”, a function $f(x): \mathbb{R}^+ \rightarrow [0,1]$ that is constant over the interval $x \in [0,L)$ and then monotonically increasing over the interval $x \in [L,\infty)$ is required. Many known functions have such a growth behavior and can be used to model the above behaviour. Some choices may be exponential, logistic, and arctangent functions. However, in the present work, a simple function defined as follows will be used. Let $A - \text{far} - B$ be the fuzzy “far” spatial relation. Let $\mu$ be the function that gives to $D$ a membership degree in $A - \text{far} - B$. Then $\mu$ is defined as,

$$
\mu_{A - \text{far} - B}(D) = \max \left[ 0, \frac{D - bL}{|A| + D + |B|} \right] \quad (A.1)
$$

where $|A|$ and $|B|$ denote areas of the objects A and B and $b$ and $L$ are two constants whose values on the context in which the above definition is used. In the present case, $b$ is set as $b = 1$ and $L = |A| + |B|$. When A and B are disconnected, $|A| + D + |B|$ can be substituted with $F_{wAB}$. For $D$, three definitions, $D_{\text{avg}} = \frac{\sum_0^{2\pi} F_{t}^{AB}(-\theta) \times F_{|ar|}^{AB}(\theta)}{\sum_0^{2\pi} F_{|ar|}^{AB}(\theta)}$, $D_{\text{min}} = \max(F_{t}^{AB}(\theta)_{\theta \in [0,2\pi]})$, and $D_{\text{min}} = \min(F_{t}^{AB}(\theta)_{\theta \in [0,2\pi]})$ are used.

For the modeling of the “far” based on absolute distance, $L$ can be set to some fixed value (e.g., $L = k$) defined $a - \text{priori}$. It can be some limit of distance between A and B when the perception of “A is far from B” starts crystallizing. Alternatively, the following simple definition can be used for $\mu_{A - \text{far} - B}(D)$. Let $l$ and $u$ be some limits of $D$ such that $\lim_{D \rightarrow l} \mu_{A - \text{far} - B}(D) \approx 0$ and $\lim_{D \rightarrow u} \mu_{A - \text{far} - B}(D) \approx 1$, then,

$$
\mu_{A - \text{far} - B}(D) = \begin{cases} 
0, & \text{if } D < l \\
1, & \text{if } D \geq u \\
\frac{D}{u - l}, & \text{otherwise}
\end{cases} \quad (A.2)
$$
where $D$ is defined in the same way as in (A.1). For the computation of $\mu_{A-far-B}(D)$ based on absolute distance $L$ is set to $k \in \mathbb{R}$ (such that $k > 0$) where $k$ is a pre-defined value. Finally, the truth-degree of “A is far from B” is defined by,

$$far(A, B) = \mu_{A-far-B}(D) \times \omega$$  \hspace{1cm} (A.3)

Where $\omega = 1 - (A \cap B)/\min(|A|, |B|)$ represents the influence of the intersection of the objects on the truth degree of the proposition “A is far from B”.

**A.3 Modeling the “near” relationship**

Let $near(A, B)$ represent the truth degree of the proposition “A is near to B”. Then $near(A, B)$ is defined as,

$$near(A, B) = \mu_{A-near-B}(D) \times \omega$$ \hspace{1cm} (A.4)

Where $D$ and $w$ are calculated in the same way as in (A.3). The function $\mu_{A-near-B}(D)$ is given by,

$$\mu_{A-near-B}(D) = \begin{cases} 0, & \text{if } D < h \\ \frac{L + h}{|A| + D + |B|}, & \text{Otherwise} \end{cases}$$ \hspace{1cm} (A.5)

Where $L$ and $h$ are two constants whose values depend on the context in which the above definition is used. In the experiments that follow, $L$ is set to $|A| + |B|$ and $h$ to a real value.

**A.4 Experiments**

Results for “far” and “near” are given in Figure A.1 and Figure A.2. Figure A.1 includes configurations that represent a transition from “near” to “far” relationships between A (argument object) and B (reference object). Figure A.2 includes two sets of images. In the first set are the configurations that show the influence of the distance between A and the massive parts of B on “far” and “near”; The second set includes images
in which A and B intersect. Truth-degrees are given with each configuration. Meanings of terms, Near0 and Far0, are given in Table A.1.

Table A.1 Meaning of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Assessment type</th>
<th>Calculated from**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near1, Far1</td>
<td>Average assessment of “near” and “far”</td>
<td>$D_{avg}$</td>
</tr>
<tr>
<td>Near0, Far0</td>
<td>Optimistic assessment of near, pessimistic of “far”</td>
<td>$D_{min}$</td>
</tr>
<tr>
<td>Near2, Far2</td>
<td>Pessimistic assessment of near, optimistic of “far”</td>
<td>$D_{max}$</td>
</tr>
</tbody>
</table>

**See Section A.2 for description of $D_{avg}$, $D_{min}$, and $D_{max}$

In Figure A.1, as object A moves away from object B, the average truth-degree Near0 of “A is near B” (i.e., $\mu_{A-\text{near-B}}(D_{avg})$) decreases whereas that of “A is far from B” (i.e., $\mu_{A-\text{near-B}}(D_{avg})$) increases. The optimistic and pessimistic values (Near1, Far2 and Near2, Far1) show the same behavior, e.g. Near0 is the highest for Figure A.1(a) and lowest for Figure A.1(h) and vice versa for Far0. The results match the expectation.

In A.2(a), although A is very close to B but as it is away from the massive part of B, the truth-degree of “near” is calculated to be lower whereas that of “far” relatively higher. In (b) and (c) A moves farther from B and the truth degree of “far” further increases (A is far from both the massive and lighter parts of B). In (d), the truth degree of near is higher compared to that (a) because A is closer to the massive part now. In (e) – (h), the truth-degree of “near” decreases because A and B intersect. The relationship is more topological than distance.
Figure A-1 Near and Far configurations. Configurations representing NEAR to Far transition. In (a)-(h), object A moves away from object B object until it is far from B. It should be noted that Near$^0$ denotes optimistic assessment of “near”, Far$^0$ pessimistic assessment of “far”, Near$^2$ pessimistic assessment of “near”, and Far$^2$ denotes optimistic assessment of “far”. Near$^1$ and Far$^1$ represent average assessment of “near” and “far” respectively.
Figure B-1. Quantitative RCC8 relations. Object A is in light-gray, object B in white. The background is black.
Figure B-2. Quantitative RCC8 relations. A is in light-gray, B in white. The background is black.
Figure B.3. Quantitative RCC8 relations. Object A is in light-gray and B in white, background is white.

B.2 Directional relations
Figure B-4 Directional relations. Comparison of models M, K', F0', and F2' (Chapter V).
Object A is in light-gray, object B in white.