Spatial relations and spatial reasoning for the interpretation of Earth observation images using a structural model.

Maria Carolina Vanegas Orozco

Thèse

présentée pour obtenir le grade de docteur
de Télécom ParisTech

Spécialité : Signal et Images

Maria Carolina Vanegas Orozco

Relations spatiales et raisonnement spatial pour l'interprétation des images d'observation de la Terre utilisant un modèle structurel

Spatial relations and spatial reasoning for the interpretation of Earth observation images using a structural model

Soutenue le 13 janvier 2011 devant le jury composé de

Monique Thonnat	Président
Anca Ralescu	Rapporteurs
Florence Le Ber	
Céline Hudelot	Examinateurs
Henri Maître	
Isabelle Bloch	Directeurs de Thèse
Jordi Inglada	

Abstract

High resolution remote sensing images allow discriminating between different objects that compose a scene. However, due to the large quantities of information it is difficult to differentiate the meaningful characteristics or regions necessaries for the description of the scene. Thus, the interpretation of these image requires the introduction of new tools which allows us to distinguish the objects of interest from the rest of the image.

First we study the spatial relations which can be useful for the interpretation of satellite images. We focused on the following spatial relations: surround, relations between linear objects and regions, alignment and parallelism. For each of these relations we developed formal models within the fuzzy set framework, which take into account the semantics, the perception and the context of use of these relations. These relations were evaluated on real objects, obtaining satisfaction degrees which fit well with the intuition, even in the case of complex objects.

Then we propose an application of spatial relations for higher level tasks. We introduce an interpretation system which is capable of finding the instantiations of an structural model in an image. The interpretation problem is described as a flexible constraint satisfaction problem. We proposed adapted propagation algorithms for flexible constraint satisfaction problems in order to cope with complex relations and to take into account the difficulties of properly detecting the objects in the image. We tested our algorithm in scenes containing harbors and airports and the results show the interest of incorporating this methodology in a more complete image interpretation system.

Résumé

L'amélioration de la résolution des images satellites optiques permet de distinguer les différents objets qui composent une scène. Néanmoins il reste difficile d'extraire les caractéristiques ou les régions qui sont pertinentes pour la description d'une scène. L'interprétation de ce type de données requiert donc l'introduction d'outils qui permettent de discriminer les objets d'intérêt du reste de l'image. Dans cette thèse nous proposons des outils de raisonnement spatial qui aident à l'interprétation des images satellites.

D'abord nous nous intéressons aux relations spatiales qui peuvent être utiles pour l'interprétation des images satellites. Nous nous concentrons sur les relations spatiales suivantes : entourer, alignement, parallélisme et des relations entre lignes et régions. Pour chacune de ces relations nous introduisons des modèles formels, qui considèrent la sémantique des relations et le leur contexte d'utilisation.

Ensuite nous proposons une utilisation des modèles de relations spatiales pour des tâches de haut niveau : nous introduisons un système d'interprétation qui est capable de trouver les instanciations d'un modèle structurel dans une image. Le problème d'interprétation d'une image est formulé comme un problème de satisfaction de contraintes floues. Nous proposons des algorithmes de propagation adaptés aux relations complexes telles que l'alignement, et qui prennent en compte les difficultés de détection des objets dans les images. Ce système a été testé sur des scènes contenant des ports et des aéroports et les résultats montrent l'intérêt d'incorporer cette méthodologie dans un système d'interprétation d'image plus complet.

Acknowledgements

I am grateful to my advisors Isabelle Bloch and Jordi Inglada for their warm supervision, their support, their advises and their availability.I really enjoyed working with you, and I feel very fortunate to have had this opportunity.

Special thanks to Anca Ralescu and Florence Le Ber for having accepted to review this work. I am also thankful to the other committee members: Monique Thonnat (president of the jury), Celine Hudelot and Henri Maître, for their comments and evaluation of my work.

I acknowledge CNES and the former Center of Competence (CNES/DLR/Télécom ParisTech) for financing this thesis. I especially want to thank Alain Giros and Mihai Datcu.

I will like to thank whole the TSI department for providing a pleasant working environment. Specially Marine for her good humor and her interest on my work, Soufiane for his advices and Patricia for her warmness. Thanks to those who have shared part of these three years: Camille, Jeremie, Olivier, David, Darwin, Emi, Nicolas, Juan Pablo, Pierre, Ana and Kevin. I am thankful to my office mates Julien and Geoffroy for always listening to me, helping me with my French grammar, explaining me the French culture and for their advices.

I wish to thank my friends Carlos, Daniel, Patricia, Olivia, Rodrigo and Nancy for all the good times, which obliged me to make a pause.

Thanks to my family, specially to my mother, who supported my decision of studying abroad.

I am very grateful to Felipe for being by my side all these three years, for his support, patience and encouragements.

Contents

Résumé des travaux **13**

Introduction **27**

I Spatial representations and reasoning **31**

1 Spatial relations for satellite image interpretation: state of the art **33**
1.1 Cognitive aspects of spatial representations . . . 34
1.2 Classification of spatial relations . . . 36
1.3 The set of relations . . . 37
1.3.1 Topological relations . . . 38
1.3.2 Metric relations . . . 44
1.3.3 Structural relations . . . 54
1.4 Conclusion . . . 54

2 New spatial relations **55**
2.1 Surround . . . 55
2.1.1 Related work . . . 56
2.1.2 Definition of surroundedness as a fuzzy landscape . . . 58
2.1.3 Properties . . . 63
2.1.4 Computational complexity . . . 63
2.1.5 Illustrative example . . . 63
2.1.6 Discussion . . . 66
2.2 Alignment . . . 66
2.2.1 Related work . . . 67
2.2.2 Considerations for modeling alignment . . . 68
2.2.3 Definition and identification of aligned groups of objects . . . 71
2.2.4 Stability with respect to segmentation errors . . . 81
2.2.5 Complexity analysis . . . 82
2.2.6 Discussion . . . 83
2.3 Parallelism . . . 84
2.3.1 Considerations . . . 84
2.3.2 Related work . . . 86
2.3.3 Parallelism with (fuzzy) linear objects . . . 86
2.3.4 Properties . . . 88
2.3.5 Parallelism with a *globally* aligned group of objects . . . 89
2.3.6 Discussion . . . 90

2.4 Line-region relations 90
2.4.1 Set of line-region relations 91
2.4.2 Go Across 94
2.4.3 “Enter” 103
2.4.4 Illustrative example 105
2.4.5 Comments about “along” and “go around” 107
2.4.6 Discussion 108
2.5 Conclusion 109

3 Spatial reasoning **111**
3.1 Qualitative spatial reasoning (QSR) 112
3.2 Reasoning with fuzzy spatial relations based on fusion operators 113
3.3 Illustrative examples: Spatial reasoning for object recognition 115
3.4 Conclusion 117

II Satellite image interpretation **121**

4 Knowledge based image interpretation: an overview **123**
4.1 Introduction 123
4.2 Knowledge based systems (KBS) 126
4.2.1 Knowledge representation schemes 127
4.2.2 Sources and types of information imperfections 130
4.2.3 Reasoning under imperfections 132
4.3 Discussion 136
4.4 Conclusion 137

5 A new mapping approach for satellite image interpretation using a structural model **139**
5.1 Structural model representation 140
5.1.1 Choice of the representation framework: conceptual graphs vs. description logics 140
5.1.2 Nested conceptual graphs 141
5.1.3 Vocabulary 145
5.2 Mapping a conceptual graph to an image: general principle 146
5.3 Simple case 147
5.3.1 Constraint Satisfaction Problems (CSP) 150
5.3.2 Illustration 152
5.3.3 Discussion 155
5.4 Including fuzzy spatial relations 156
5.4.1 Fuzzy constraint satisfaction problems (FCSP) 156
5.4.2 Dealing with alignments 163
5.4.3 Dealing with groups (not necessarily aligned) 168
5.4.4 Proposed algorithm 170
5.4.5 Complexity and improvements 178
5.4.6 Finding a solution 180
5.4.7 Discussion 182
5.5 Interpreting an unlabeled image 185
5.5.1 Image partition 185

5.5.2 Region labeling 185
5.5.3 Spatial reasoning over the tree of regions 187
5.5.4 Finding a solution 187
5.5.5 Results 189
5.6 Towards the introduction of priorities and uncertainties 200
5.7 Conclusion 201

6 Conclusion **203**
6.1 Main contributions 203
6.1.1 Novel definitions of spatial relations 203
6.1.2 Spatial reasoning and image interpretation guided by a model 204
6.2 Perspectives 205
6.2.1 Short-term perspectives 205
6.2.2 Long-term perspectives 206

Notations **208**

A List of publications **211**

Bibliography **213**

Résumé des travaux

Introduction

L'amélioration de la résolution des images satellites optiques, telles que celles obtenues par le capteur Quickbird et bientôt par le capteur Pléiades, permet d'identifier des objets qui font partie d'un objet complexe. Les objets complexes sont caractérisés par un ensemble d'objets qui satisfont des relations spatiales entre eux et qui forment une structure, par exemple les aéroports, les ports et les gares. Bien que l'image porte beaucoup d'information qui permet la reconnaissance d'objets complexes, il reste difficile d'extraire les caractéristiques ou les régions qui sont pertinentes pour la description d'une scène. L'interprétation de ce type de données requiert donc l'introduction d'outils qui permettent de discriminer les objets d'intérêt du reste de l'image.

Dans cette thèse nous proposons des outils de raisonnement spatial qui aident à l'interprétation des images satellites. Dans une première partie nous nous intéressons aux relations spatiales qui se trouvent dans les images satellites. Nous faisons une étude des relations spatiales qui sont utilisées en traitement d'images et dans d'autres domaines liés à l'imagerie satellitaire, tels que la cartographie. Nous avons déterminé un ensemble de relations, avec les modèles formels correspondants, qui sont utiles pour l'interprétation des images. Cet ensemble comporte des relations qui existent déjà dans la littérature, ainsi que des relations et modèles originaux que nous introduisons, qui sont bien adaptés aux images satellites.

Dans une deuxième partie nous proposons une utilisation des modèles de relations spatiales introduits pour des tâches de haut niveau. Cette application consiste en la proposition d'un système d'interprétation qui est capable de trouver les instanciations d'un modèle structurel dans une image, en lui donnant une quantité de connaissance minimale sur l'extraction des objets du modèle. Ce système a été testé sur des images réelles.

Chapitre 2 : Les relations spatiales dans des images de télédétection

Pour déterminer les relations spatiales entre des objets dans des images satellites, nous considérons les relations spatiales qui ont été proposées dans les domaines de l'intelligence artificielle, l'interprétation des images, les systèmes d'information géographique (GIS en anglais) et la généralisation cartographique. Ces domaines nous intéressent car ils contribuent à produire des modèles pour des relations que l'on peut observer dans des images qui prennent en compte l'intuition. Les domaines des systèmes d'information géographique et de la généralisation cartographique s'intéressent à l'utilisation des relations spatiales pour décrire l'agencement des objets géographiques. Cependant il y a des relations pour lesquelles il n'existe pas de modèles ou les modèles proposés ne s'adaptent pas à la nature des images.

Les relations spatiales entre des objets dans des images satellites peuvent être réparties en trois familles :

Relations topologiques : dans cette famille la relation d'adjacence est très importante en analyse d'images. La figure 1.6 illustre les relations de cette famille.

Relations métriques : cette famille est composée des relations de distance telle que « proche », « loin », et des relations de direction comme « au nord ». La figure 1.15 illustre l'ensemble des relations qui appartiennent à cette famille.

Relations de groupement : ce type de relations est utilisé dans la généralisation cartographique [Steiniger and Weibel, 2007]. Cette famille fait référence aux relations qui forment des motifs entre plusieurs objets, par exemple les alignements d'objets, les objets qui sont agencés en forme d'étoile comme les rues dans un carrefour.

Différents auteurs [Rosenfeld and Klette, 1985, Bloch, 2005, Aksoy et al., 2003, Miyajima and Ralescu, 1994a, Takemura et al., 2005] se sont intéressés à la modélisation des relations topologiques et des relations métriques pour l'interprétation et l'analyse des images. Lorsque l'on veut représenter une relation spatiale entre des objets de l'image il faut prendre en compte les imprécisions dues aux traitements de bas niveau et au processus d'acquisition, ainsi que les imprécisions liées à la sémantique de la relation. Par exemple la relation « proche » est par nature imprécise. La logique floue est bien adaptée pour la modélisation des relations spatiales car elle permet de prendre en compte ces imprécisions. Dans le cadre des relations spatiales floues, deux questions peuvent se poser par rapport à une relation [Bloch, 2005] :

(i) quel est le degré de satisfaction de la relation entre deux objets ?

(ii) étant donné un objet de référence, quelle est la région de l'espace (paysage flou) où la relation est satisfaite (et à quel degré) ?

Selon la question à laquelle on veut répondre, la relation est modélisée d'une façon différente. L'avantage de modéliser une relation qui répond à la deuxième question est que, à partir d'un objet de référence A, on peut construire une région d'intérêt pour chercher des objets qui sont en relation avec A. Un autre avantage est que pour répondre à la première question, il suffit de trouver le paysage flou une seule fois et après on peut évaluer le degré de satisfaction avec plusieurs objets. Nous proposons des relations qui répondent à la première question et d'autres à la seconde. Dans la suite nous allons présenter les principales caractéristiques des modèles des relations que nous avons construits.

Chapitre 3 : Nouvelles relations spatiales

« Entourer »

La relation « entourer » est fréquemment trouvée dans des images satellites, par exemple entre des structures dans la mer et la mer. Pour définir cette relation, nous allons dire qu'un objet A est « entouré » par un objet B si l'objet A est observé par l'objet B dans presque toutes les directions. [Rosenfeld and Klette, 1985] et [Miyajima and Ralescu, 1994b] ont proposé des approches pour modéliser cette relation qui sont en accord avec la relation que nous voulions définir. Ces deux approches proposent un modèle qui détermine le degré de satisfaction de la relation A est « entouré » par B, en utilisant des mesures de couverture angulaire des points de A par les points de B. Par rapport aux approches de [Rosenfeld and Klette, 1985] et [Miyajima and Ralescu, 1994b], nous proposons une définition de la relation comme une région floue de l'espace où la relation est satisfaite pour un objet de référence donné. Les avantages de cette approche ont été soulignés dans la section précédente. De plus, notre définition prend en compte la morphologie de l'objet de référence ainsi que la notion de distance par rapport à l'objet cible.

Lorsqu'un objet B « entoure » un objet A, il y a une portion de la frontière de B qui contourne A. [Mathet, 2000] suggère que la relation « un chemin C contourne un objet A » est satisfaite si et seulement si C n'intersecte pas A et l'enveloppe convexe de C intersecte A. Cette définition a été conçue en considérant les aspects linguistiques de la relation. Donc en prenant en compte la définition de la relation « contourner » on peut dire que la relation « entouré » a un sens uniquement dans le cas où l'objet de référence a des concavités. Nous utilisons donc les points de la frontière de B qui se trouvent dans une concavité pour déterminer si A est vu par B dans presque toutes les directions. Nous nous appuyons sur le travail de [Rosenfeld and Klette, 1985, Miyajima and Ralescu, 1994b] pour définir la notion de visibilité dans « presque toutes » les directions. Nous définissons la couverture angulaire d'un point p par les points qui se trouvent dans les concavités de ∂B, la frontière de B, comme :

$$\theta_{coverage\_CH}(B, \mu_n)(p) = \int_0^{2\pi} \tilde{r}_\theta(B)(p) d\theta \tag{1}$$

où $\tilde{r}_\theta(B) : \mathfrak{I} \longrightarrow [0, 2\pi]$ est une fonction qui détermine s'il existe un point dans une concavité de B qui voit p dans la direction θ, c'est-à-dire :

$$\tilde{r}_\theta(B)(p) = \begin{cases} 1 & \text{si } \exists b \in \partial B \setminus \partial CH(B) \text{ tel que } \angle(\vec{pb}, \vec{u}_x) = \theta \text{ et } [p, b[\cap B = \emptyset, \\ 0 & \text{sinon.} \end{cases} \tag{2}$$

où $\angle(\vec{pb}, \vec{u}_x)$ représente l'angle entre le vecteur $\vec{pb}$, joignant p et b et le vecteur unitaire dans la direction de l'axe x.

Maintenant que nous avons défini la couverture angulaire, la région de l'espace qui est « entourée » par un objet B est définie par l'ensemble flou $\mu_{surround}(B) : \mathfrak{I} \longrightarrow [0, 1]$:

$$\mu_{surround}(B)(p) = f(\theta_{coverage\_CH}(B)(p)) \tag{3}$$

où $f : [0, 2\pi] \longrightarrow [0, 1]$, $f(0) = 0$ et $f(2\pi) = 1$. L'objectif de la fonction f est de définir la fonction d'appartenance de la variable linguistique « presque tous » de la couverture angulaire. La figure 2.5 montre la région floue où la relation « entourer » est satisfaite en utilisant comme objets de référence ceux de la figure 2.3. Les points qui se trouvent à l'intérieur des concavités de l'objet de référence ont un fort degré d'appartenance à $\mu_{surround}(B)$.

La définition de l'équation 3 est bien adaptée si l'on considère que les objets visés par la relation ont une taille comparable à celle des concavités de l'objet de référence. Néanmoins quand le point q dans les concavités de ∂B qui voit l'objet cible A dans une direction θ est très « loin » de A, tout se passe comme si q ne le voyait pas. La notion de « loin » est étroitement liée aux caractéristiques intrinsèques de A, comme la taille de A [Hernandez et al., 1995]. Maintenant, nous considérons la fonction d'appartenance de « entourer » par B en prenant en compte la distance, pour cela nous avons besoin uniquement d'adapter l'équation 2 selon :

$$\tilde{r}_\theta(p, B, \mu_n) = \begin{cases} \mu_n(d_E(p,q)) & \text{si } \exists q \in \partial B \setminus \partial CH(B) \text{ tel que } \angle(\vec{pq}, \vec{u}_x) = \theta \text{ et } [p,q[\cap B = \emptyset, \\ 0 & \text{sinon.} \end{cases} \tag{4}$$

où $\mu_n(d_E(p,q))$ est une fonction de $\mathbb{R}^+$ dans $[0,1]$ qui représente la notion de « près » selon l'objet cible.

L'extension de ces définitions au cas flou est faite de manière directe en utilisant la méthode de traduction formelle. La définition de « entourer » de l'équation 3 est un cas particulier de la définition qui prend en compte la distance quand μ_n est une fonction constante égale à 1. La définition de « entourer » est invariante par rapport aux transformations géométriques (pour les homothéties, cela suppose que μ_n soit invariant). La définition de « entourer » est croissante par rapport à μ_n.

Les deux définitions (avec et sans prise en compte de la distance) ont été testées dans des images réelles pour les objets de la figure 2.7 en utilisant les différentes fonctions de distance μ_n de la figure 2.8. Les résultats se trouvent dans la figure 2.9. Ils montrent que quand on considère la fonction de distance on peut obtenir des régions plus précises. Les régions obtenues sont en accord avec l'intuition.

« Alignement »

La relation d'alignement est souvent observée dans des images satellites pour des objets construits par l'homme. Par exemple dans les aires de stationnement telles que stationnements de voitures, aéroports, ports ou stationnements de camions, les véhicules forment des ensembles alignés. Ainsi l'identification des ensembles de véhicules alignés peut aider à l'identification de ces objets. De plus trouver les ensembles d'objets alignés peut aider à la réduction de l'information qui se trouve dans les images. Par exemple, pour une cartographie il est nécessaire d'identifier les ensembles de bâtiments alignés pour la généralisation des cartes [Steiniger and Weibel, 2007]. Donc les alignements donnent de l'information structurelle qui peut être utilisée pour l'identification des zones urbaines, rurales ou résidentielles.

En vision l'alignement entre points ou segments a été largement étudié [Lowe, 1987, Christophe and Ruas, 2002, Ortner et al., 2007, Likforman-Sulem and Faure, 1994, Desolneux et al., 2008, Ralescu and Shanahan, 1999] car il correspond à des caractéristiques saillantes pour l'interprétation des images conformément à la théorie de la Gestalt. Dans [Desolneux et al., 2003, Christophe and Ruas, 2002] l'extension aux objets a été faite en regardant si leurs centres de masse sont alignés. Cette méthode est appropriée pour déterminer les groupes d'objets qui sont alignés par leur centre de masse, mais il y a d'autres types d'alignement, comme l'alignement par rapport aux extrémités, qui sont moins étudiés. Par rapport aux travaux existants nous proposons une approche qui utilise des mesures de positions relatives où l'objet n'est pas réduit à un seul point, comme son centre de masse, mais tout l'objet est pris en considération. Dans l'exemple de la figure 2.10 nous montrons un exemple d'objets réels qui sont alignés et pourtant leurs centres de masse ne le sont pas. L'algorithme que nous proposons permet d'identifier des ensembles d'objets qui sont alignés même si les objets ont des tailles différentes ou des formes complexes, ce que ne permet pas un algorithme qui s'appuie sur l'alignement des centres de masse.

Nous disons qu'un ensemble d'objets est aligné s'il a au moins trois éléments et s'il existe un angle θ pour lequel chaque membre du groupe est capable de voir les autres membres du groupe dans une direction θ ou $\theta + \pi$. Pour mesurer la direction entre objets nous définissons l'histogramme d'orientations qui est fondé sur l'histogramme d'angles de [Miyajima and Ralescu, 1994a]. Étant donné deux objets A et B dans $\mathfrak{I}$, l'histogramme d'orientation est :

$$O(A,B)(\theta) = \frac{|\mathbf{1}_{\{(p,q)\in A\times B | mod(\angle(\vec{pq},\vec{u}_x),\pi)=\theta\}}|}{\max_{\phi\in[0,\pi)} |\mathbf{1}_{\{(p,q)\in A\times B | mod(\angle(\vec{pq},\vec{u}_x),\pi)=\phi\}}|}, \tag{5}$$

où $\mathbf{1}_X$ est la fonction indicatrice de l'ensemble X, $\angle(\vec{pq},\vec{u}_x)$ est l'angle entre le vecteur joignant p et q et le vecteur $\vec{u}_x$ qui représente l'orientation de l'axe x. L'histogramme d'orientations est un ensemble flou de $[0,\pi[$ qui représente l'orientation entre deux objets. Il correspond à un histogramme d'angles [Miyajima and Ralescu, 1994a] où les angles sont calculés modulo π, il préserve les mêmes propriétés que l'histogramme d'angles, et de plus il est symétrique.

Nous définissons la similarité entre deux histogrammes d'orientations $O(A,B)$ et $O(C,D)$ comme le maximum de l'intersection de $D_{\mu_0}(O(A,B))$ et $D_{\mu_0}(O(C,D))$, où $D_{\mu_0}(O(X,Y))$ est une dilatation morphologique floue de $O(X,Y)$ par un élément structurant flou μ_0 [Bloch and Maître, 1995]. Dans notre cas μ_0 représente l'imprécision attachée à la comparaison d'angles qui sont presque égaux, et cette fonction peut être modélisée par une fonction en trapèze.

Pour déterminer les ensembles d'objets qui sont alignés nous déterminons d'abord les ensembles d'objets qui sont alignés localement, et ces ensembles vont être les candidats pour les ensembles d'objets alignés, que nous appelons alignés globalement. Étant donné un ensemble d'objets $\mathcal{A} = \{a_1, \ldots a_n\}$, nous construisons un graphe de voisinage $G_N = \{V, E\}$, où les nœuds représentent les objets de l'ensemble, et il y a une arête entre deux nœuds si et seulement si les objets correspondants sont voisins selon une relation de voisinage (Voronoï, distance tronquée). Chaque arête est attribuée avec l'histogramme d'orientations entre les deux objets qui sont représentés par les nœuds de ses extrémités. A partir du graphe de voisinage nous construisons son graphe dual. Le graphe dual est noté $\tilde{G}_N = \{\tilde{V}, \tilde{E}\}$. Chaque nœud $\tilde{V}_i \in \tilde{V}$ représente une arête du graphe G_N. Il existe une arête entre deux nœuds $\tilde{V}_i$ et $\tilde{V}_j$ de $\tilde{G}_N$ s'il y a un nœud commun de G_N entre les arêtes correspondantes de G_N. Chaque arête $\tilde{e}_{ij}$ est étiquetée avec la valeur de similarité entre les histogrammes d'orientation de $\tilde{V}_i$ and $\tilde{V}_j$ (équation 5). La figure 2.20 montre un exemple de graphe de voisinage et son graphe dual. Dans le graphe dual il est possible de comparer directement les orientations entre paires d'objets avec un objet en commun : les arêtes du graphe dual qui ont une valeur haute (respectivement faible) représentent des triplets d'objets d'orientations similaires (respectivement différentes).

Nous disons alors qu'un ensemble est localement aligné avec un degré α, si pour le sous-ensemble $\tilde{A} \subseteq \tilde{V}$ correspondant à $\tilde{G}_N$ toutes les arêtes qui relient ses nœuds ont une valeur supérieure à α.

Un groupe localement aligné $\mathcal{A} = \{A_0, \ldots, A_n\}$ est un candidat pour être globalement aligné. Pour déterminer s'il est globalement aligné nous considérons le groupe comme un tout. Nous définissons le degré d'alignement global par la valeur de similarité entre les histogrammes d'orientation $O(A_i, \mathcal{A} \setminus A_i)$, pour $i = 0, \ldots, n$, où $O(A_i, \mathcal{A} \setminus A_i)$ est l'histogramme d'orientations entre A_i et le reste du groupe.

Cette méthode nous permet d'obtenir les ensembles d'objets alignés dans une image satellite d'une façon originale. Les figures 3.3 et 3.5 montrent des applications de cette relation pour la classification des morphologies urbaines et pour la désambiguïsation de l'information dans la détection de routes et de bâtiments.

« Parallélisme »

Comme l'alignement, la relation de parallélisme est importante lorsque nous considérons des structures qui ont été faites par l'homme. Dans le cadre de la vision, le parallélisme entre segments et lignes a été largement étudié [Lowe, 1987, Kang and Walker, 1992, Rouco et al., 2007, Ralescu and Shanahan, 1999, Toh, 1992]. Néanmoins le parallélisme entre un groupe d'objets alignés et un objet linéaire ou un autre groupe d'objets alignés a été moins étudié. De plus, les relations d'alignement et de parallélisme se trouvent fréquemment ensemble. Nous proposons une définition du parallélisme qui est bien adaptée pour évaluer le parallélisme quand l'objet de référence ou l'objet cible de la relation est un groupe d'objets alignés ou un segment flou. Notre définition repose aussi sur la notion de visibilité (floue) et pas uniquement sur l'orientation comme dans la plupart des travaux de la littérature.

Dans le contexte de cette thèse nous considérons que la relation de parallélisme n'est pas symétrique, car les longueurs des objets dans l'image sont finies et peuvent être différentes. Lorsque l'on veut évaluer la relation de parallélisme entre des objets de différentes longueurs, comme A et B de la figure 2.24(a), la propriété de symétrie est discutable. La relation « B est parallèle à A » peut être considérée comme vraie : pour tous les points de la frontière de B qui sont en face de A, il est possible de voir (dans la direction normale à l'axe principal de A) un point de A, et les orientations de A et B sont similaires. En revanche la relation « A est parallèle à B » est ambiguë par rapport à la position dans A : depuis le point d il est possible de voir un point de B dans la direction normale à l'axe principal de B, alors que depuis le point c nous ne verrons aucun point de B. Cet exemple permet d'illustrer aussi que la relation de « parallélisme » doit être évaluée comme une relation floue plutôt que par une réponse oui/non. Dans ce cas, la relation « B est parallèle à A » aura une plus forte valeur de satisfaction que « A est parallèle à B ». Avec des argumentes similaires nous pouvons montrer aussi que la relation n'est pas transitive.

Nous traitons d'abord le cas de parallélisme entre objets linéaires flous. Soient A et B deux objets linéaires flous avec des orientations θ_A et θ_B, respectivement. Soit $\vec{u}_{\theta_A+\frac{\pi}{2}}$ le vecteur normal à l'axe principal de A. Donc, le degré de satisfaction « A est parallèle à B » dépend des conditions suivantes :

(i) Il y a une grande proportion de la frontière de A qui voit B dans la direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$.

(ii) L'orientation de A et l'orientation de la frontière de B qui est en face de A et qui est vue par A dans la direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$ sont similaires.

Ces deux conditions sont reliées à la notion de visibilité. Le sous-ensemble de X qui est visible par Y dans la direction θ est défini comme un ensemble flou avec la fonction d'appartenance suivante :

$$\mu_{Xvis(Y,\theta)}(x) = \mu_X(x) \wedge D_{\nu_\theta}(\mu_Y)(x). \tag{6}$$

où $D_{\nu_\theta}(\mu_Y)(x)$ est la dilation directionnelle morphologique [Bloch, 1999] et $\wedge$ est une t-norme. Cette notion est illustrée dans la figure 2.26. En utilisant la définition de visibilité et en traduisant les conditions de parallélisme, nous définissons le degré de satisfaction de la relation « A est parallèle à B » comme :

$$\mu_{parallel}(A,B) = \frac{V_n(\mu_{Avis(B,\theta_A-\frac{\pi}{2})})}{V_n(\mu_A)} \bigwedge \nu_0(\theta_{\delta Bvis(A,\theta_A+\frac{\pi}{2})} - \theta_A), \tag{7}$$

où la première partie utilise l'hyper-volume flou V_n [Bloch, 2005] pour mesurer la proportion de A qui voit B. Dans la deuxième partie nous évaluons si l'orientation de A et celle de la frontière de B vue par A dans la direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$ sont similaires, pour cela nous utilisons une fonction ν_0 qui évalue si les deux angles sont égaux.

Considérons que l'objet de référence est un groupe d'objets alignés $\mathcal{S} = \{A_0, \ldots, A_N\}$ et que l'objet cible est un objet linéaire B. Pour évaluer les conditions de visibilité nous utilisons $\mathcal{S} \cup \beta_S$, où β_S est la région « entre » deux éléments consécutifs du groupe. En prenant en compte β_S pour l'évaluation de la visibilité nous considérons le groupe comme un objet linéaire qui peut être facilement remplacé par A dans l'équation 7 :

$$\mu_{parallel}(\mathcal{S}, B) = \frac{V_n(\vee_i \mu_{A_i vis(B, \theta_s - \frac{\pi}{2})})}{V_n(\vee \mu_{A_i})} \bigwedge \nu_0(\theta_{\delta B vis(\beta_s \cup \mathcal{S}, \theta_s + \frac{\pi}{2})} - \theta_S), \quad (8)$$

où θ_S correspond à l'angle d'alignement du groupe par rapport à l'axe x. De façon similaire nous étendons la définition du parallélisme pour traiter les cas où l'objet cible est un groupe aligné ou ceux où les deux objets impliqués dans la relation sont des groupes alignés. Les figures 3.3 et 3.5 montrent des exemples d'évaluation de la relation entre deux groupes alignés et entre un groupe et une ligne.

Relations entre lignes et régions

Les images satellites comportent une grande quantité d'objets linéaires (réseaux routiers et fluviaux), et nous nous sommes donc intéressés aux relations spatiales entre une région et un objet linéaire. Nous les appelons « relations entre lignes et régions ». Ces relations ont été étudiées dans la communauté des SIG [Roussopoulos et al., 1988, Egenhofer and Herring, 1990, Mark and Egenhofer, 1994a, Shariff et al., 1998, Kurata and Egenhofer, 2007, Schwering, 2007] et dans le domaine de la cognition [Landau and Jackendoff, 1993, Talmy, 1983, Herskovits, 1997]. Dans le domaine des SIG, plusieurs modèles binaires ont été proposés pour ce type de relations. Ces modèles sont très limités, car il est par exemple difficile de dire si un chemin traverse ou pas une région quand la région a une forme complexe. Par rapport aux travaux précédents nous nous sommes attaché à modéliser les relations qui prennent en compte l'imprécision par rapport à la relation elle-même et certains des modèles que nous proposons sont valables pour des régions floues.

Nous considérons quatre types de relations : « traverser », « rentrer », « longer » et « éviter ». Trois de ces relations ont été identifiées comme les plus significatives des relations entre lignes et régions [Landau and Jackendoff, 1993, Schwering, 2007]. [Takemura et al., 2005] propose un modèle pour la relation « longer » qui prend en compte les imprécisions et qui est bien adapté à nos besoins. La relation « éviter » peut être vue comme un cas particulier de la relation « entourer » (présentée précédemment), pour un objet de référence linéaire. Ainsi nous nous concentrons uniquement sur la modélisation des relations « traverser » et « rentrer ».

D'abord nous étudions la relation « traverser ». Pour mieux comprendre sa sémantique, nous avons effectué une étude sur 32 personnes. Dans cette étude nous avons montré les 8 situations de la figure 2.31 aux sujets, qui devaient indiquer si, selon eux, le chemin traverse ou pas la région. Nous avons laissé un espace pour faire des commentaires. A partir des résultats nous observons que la relation a trois types de sens :

(i) la ligne rentre et après sort de la région ;

(ii) les points par lesquels la ligne rentre et sort de la région sont localisés sur des bords opposés de la région ;

(iii) la ligne rentre et après sort de la région, et elle s'enfonce vers le centre de la région.

Ces trois types de sens montrent que la relation « traverser » ne peut pas être définie comme une relation binaire, car elle est très ambiguë. Le premier sens est très permissif, il prend en compte uniquement les aspects topologiques de la relation. Néanmoins dans

[Talmy, 1983] et [Mark and Egenhofer, 1994a] la similitude entre les relations « passe par » (« go through » en anglais) et « traverser » (« go across » en anglais) est discutée. Elles sont même parfois utilisées comme des synonymes. Cependant, la relation « passe par » est fondée sur des contraintes topologiques : « entrer et sortir d'un milieu » [Talmy, 1983], tandis que la relation « traverser » prend en compte des aspects géométriques. C'est pourquoi nous avons choisi d'appeler le premier sens « passe par », et les deux autres « traverser (i) » et « traverser (ii) », respectivement.

Pour la définition de « passe par » nous nous appuyons sur la définition proposée par [Mark and Egenhofer, 1994a]. Un objet linéaire L « passe par » une région R si ses extrémités L_a et L_b n'intersectent pas l'intérieur de la région R°, et L intersecte R° :

$$\mu_{go\_through}(L, R) = t\left(\mu_{int}(L, R^\circ), \mu_{\neg int}(L_a \cup L_b, R^\circ)\right). \tag{9}$$

où μ_{int} et $\mu_{\neg int}$ sont des degrés d'intersection et non-intersection [Bloch, 2005].

Pour la définition « traverser (i) » nous avons étudié la signification de « aller d'un côté à l'autre » d'un point de vue cognitif. Nous nous appuyons sur les travaux de [Herskovits, 1997, Landau and Jackendoff, 1993] pour définir quand deux points p et q sur la frontière de R sont sur des côtés différents de R. Si p et q ne sont pas dans une concavité de R, alors nous allons dire qu'ils sont sur des côtés opposés de R si leurs vecteurs tangents sont déphasés d'environ 180°. Supposons que p se trouve dans une concavité, dans ce cas nous ne regardons pas son vecteur tangent mais le vecteur tangent du segment qui ferme la concavité dans laquelle se trouve p, dans l'enveloppe convexe (voir la figure 2.33). Finalement, le degré de satisfaction de « L traverse R » en utilisant la notion d'aller d'un côté à l'autre est :

$$\mu_{go\_across1}(L, R) = t\left(\mu_{go\_through}(L, R), \mu^R_{opposite\_sides}(p_1, p_2)\right) \tag{10}$$

où p_1 et p_2 sont les points de L qui intersectent la frontière de R et $\mu^R_{opposite\_sides}(p_1, p_2)$ est une fonction floue qui mesure le degré avec lequel p_1 et p_2 sont sur des côtés opposés, en s'appuyant sur l'argumentation exposée auparavant.

Dans la définition de « traverser (ii) » nous nous intéressons à la notion de « s'enfoncer ». Pour cela nous proposons une définition fondée sur la morphologie mathématique. Le degré avec lequel une « ligne s'enfonce dans une région R » est mesuré en regardant à quel point la ligne passe près d'un point d'érosion ultime de la région. Finalement la définition de « traverser (ii) » est donnée par la conjonction du degré avec lequel la ligne « passe par » la région et du degré avec lequel elle s'enfonce dans la région.

En utilisant les outils définis pour la création des modèles de « traverser (ii) » et « passe par » nous proposons deux définitions pour la relation « rentre ». Une première définition que nous appelons « rentrer » exprime uniquement les contraintes topologiques : il y a au moins une extrémité de la ligne qui n'intersecte pas la région, et la ligne intersecte l'intérieur de la région. Une deuxième définition que nous appelons « pénétrer » tient compte des contraintes géométriques : elle repose sur la conjonction du degré avec lequel la ligne « rentre » dans la région et le degré avec lequel la ligne s'enfonce dans la région.

Les modèles de relations entre lignes et régions sont illustrés dans la figure 2.38 et le tableau 2.5. Ces résultats montrent que les modèles représentent bien les définitions que nous proposons. Nous avons comparé les résultats de l'enquête avec les modèles, montrant ainsi que les modèles que nous avons proposés sont conformes à la perception des relations.

Les relations spatiales qui sont fréquentes dans les images satellites vont être utiles pour définir un vocabulaire des relations spatiales pour décrire le contenu des images. La plupart des modèles que nous avons choisis ou proposés reposent sur des éléments du langage naturel. Donc ils peuvent être utilisés par des personnes non-expertes en interprétation des images.

Chapitre 4 : Raisonnement spatial

Nous pouvons combiner les relations spatiales en utilisant des opérateurs de fusion floue. Cela nous permet de faire des requêtes comme « où se trouve la région qui est proche et au nord de A ? ». Néanmoins les requêtes que nous pouvons faire de cette façon restent très simples, donc il est nécessaire d'utiliser des systèmes à base de connaissances pour permettre des requêtes plus complexes. Dans la deuxième partie de la thèse nous proposons un système d'interprétation qui a l'architecture d'un système à base de connaissances.

Chapitres 5 et 6 : Interprétation des images satellites

L'interprétation des images satellites est une tâche difficile même lorsqu'elle est effectuée par une personne. Il existe plusieurs facteurs qui peuvent donner lieu à des interprétations différentes :

- le niveau conceptuel de connaissance de la personne influence le niveau de détail ou la différenciation de différents objets dans l'image ;
- l'information contextuelle, c'est-à-dire l'information qui n'est pas observable dans l'image. Cette information permet de mieux comprendre les situations observées et peut ajouter des éléments dans la description ;
- le vocabulaire utilisé détermine les mots de la description ;
- l'objectif de la description oblige à la focalisation sur certaines caractéristiques ou certains endroits de l'image.

Lorsque nous voulons interpréter une image d'une façon automatique il est donc nécessaire de donner toutes ces connaissances pour pouvoir obtenir l'interprétation désirée. Dans cette thèse nous utilisons la connaissance sur l'agencement spatial d'un objet complexe pour l'identifier dans une image satellite. L'architecture que nous utilisons est un système à base de connaissances. Un système à base de connaissances (KBS) a trois composantes [Le Ber et al., 2006] :

La base de connaissances : contient tout les connaissances sur le domaine. Dans notre cas, cette base contient trois types des connaissances :

Connaissances de traitement d'images : elles sont utilisées pour extraire et décrire les primitives de bas niveau, par exemple les régions ;

Connaissance du domaine : c'est la connaissance par rapport à la sémantique du domaine ;

Connaissance pour lier les primitives de l'image aux concepts du domaine : c'est la connaissance permettant d'établir la correspondance entre les primitives de bas niveau de l'image et les concepts du domaine.

La base d'observations : elle contient les données qui caractérisent le problème que nous voulons résoudre. Dans notre cas cette base contient l'image.

Un moteur d'inférence : il est en charge de traiter l'information de la base d'observations en utilisant la base de connaissances pour résoudre le problème.

L'utilisation d'un KBS implique la simplification de l'information, ce qui donne lieu à des imperfections de l'information. Dans le cas particulier où nous représentons l'information spatiale pour l'interprétation d'images, il y a plusieurs types d'imperfections qu'il faut considérer :

Imprécision des relations spatiales : comme nous l'avons vu précédemment il existe plusieurs relations spatiales qui sont par nature imprécises, par exemple la relation « entourer ».

Imprécision des objets dans l'image : le processus de la création de l'image et les traitements utilisés pour extraire des objets dans l'image conduisent à des frontières d'objets mal définies.

Incertitude par rapport à l'étiquetage de régions : l'étiquetage des régions après un processus de segmentation peut être entaché d'incertitude.

Incertitude par rapport au modèle : nous ne savons pas si tous les objets et relations qui apparaissent dans le modèle sont présents dans l'image.

Méconnaissance de nombre d'instanciations : dans le cas particulier des images satellites, le nombre de fois où un modèle ou les objets du modèle apparaissent dans l'image est inconnu.

Dans un KBS le moteur d'inférence doit être adapté pour prendre en compte ces imperfections. Dans le domaine de l'interprétation des images utilisant un modèle structurel, les travaux prennent en compte, en général, un ou deux types de ces imperfections. [Deruyver and Hodé, 1997, Perchant, 2000, Saathoff and Staab, 2008] proposent des méthodes pour trouver dans l'image des structures décrites par un modèle. Pour cela ils partent d'une (sur)segmentation de l'image et ils étiquettent les régions en s'appuyant sur la connaissance spatiale. Ces méthodes considèrent l'incertitude par rapport à l'étiquetage. Récemment le travail de [Deruyver and Hodé, 1997] a été élargi pour prendre en compte l'incertitude par rapport au modèle [Deruyver et al., 2009]. Dans les travaux de [Bloch et al., 2003, Colliot et al., 2006, Nempont, 2009] dans le domaine médical, la segmentation et l'interprétation sont faites de façon simultanée en prenant en compte les imprécisions liées aux relations spatiales et à l'extraction d'objets dans l'image. Cette approche évite de se poser la question sur l'incertitude par rapport à l'étiquetage de régions. Néanmoins il est compliqué de la mettre en œuvre pour l'interprétation des images satellites, à cause de l'hypothèse de monde ouvert dans ces images.

Nous proposons un système d'interprétation fondé sur l'étiquetage de régions issues d'une segmentation. Ce système prend en compte l'imprécision liée aux relations spatiales, l'incertitude par rapport à l'étiquetage de régions et l'ignorance du nombre d'instanciations, et nous donnons quelques idées afin d'adapter le système pour prendre en compte l'incertitude par rapport au modèle. Avant d'entrer dans les détails du système, nous allons présenter le modèle utilisé pour décrire l'agencement spatial.

Représentation du modèle

Le modèle de la scène qui nous intéresse dans l'image doit permettre la représentation de relations d'arité supérieure à deux aussi bien que des groupes d'objets alignés. Pour la représentation des groupes d'objets les considérations suivantes doivent être prises en compte :

- le nombre d'objets dans le groupe est en général inconnu ;

- le groupe peut satisfaire des relations avec d'autres objets (par exemple un groupe d'arbres alignés qui est parallèle à une route) ;
- il peut y avoir des relations spatiales entre chaque membre du groupe et d'autres objets (par exemple un groupe de maisons où chaque maison a une relation spatiale avec son ombre) ;
- il peut y avoir des relations spatiales entre les membres du groupe. Par exemple, les objets qui sont consécutifs dans un groupe doivent être « proches ».

Par suite le formalisme de représentation utilisé pour représenter le modèle doit permettre la représentation de toutes ces situations. Parmi les divers formalismes de représentation des connaissances (graphe d'adjacence attribué, réseaux sémantiques, frames, ...), les logiques de description et les graphes conceptuels peuvent être généralisés pour gérer ces contraintes.

Nous avons choisi les graphes conceptuels pour leur simplicité et leur représentation graphique. Plus particulièrement, nous utilisons des graphes conceptuels emboîtés pour représenter les structures qui contiennent un groupe. Un graphe conceptuel est un graphe bipartite où un ensemble de nœuds, appelés nœuds concepts, représente des entités et l'autre ensemble de nœuds, appelés nœuds relations, représente les relations entre les nœuds concepts. Dans notre cas les nœuds concepts sont les objets du modèle et les nœuds relations correspondent aux relations spatiales. Les concepts et les relations qui sont utilisées dans le graphe viennent d'une hiérarchie de concepts et de relations, respectivement. Dans un graphe conceptuel emboîté un nœud concept peut contenir un graphe conceptuel à l'intérieur, qui sert à représenter des groupes d'objets (pas forcement alignés). Ces nœuds sont appelés nœuds emboîtés. Un exemple de graphe conceptuel est illustré dans la figure 5.2. Comme nous ne connaissons pas a priori le nombre d'éléments d'un groupe d'objets alignés, nous posons par convention qu'un groupe d'objets alignés est représenté par trois éléments et pour chaque nœud emboîté nous assignons une propriété qui indique si le groupe est aligné ou pas. Pour représenter des relations entre les nœuds à l'intérieur d'un nœud emboîté et des nœuds à l'extérieur nous utilisons des liens de coréférence.

Système d'interprétation des images satellites

Le système que nous proposons est présenté dans la figure 5.19. Les entrées du système sont :

- l'image initiale I que nous voulons interpréter ;
- le graphe conceptuel G qui contient le modèle que nous cherchons dans l'image ;
- la hiérarchie de concepts H_C sur laquelle est construit le graphe conceptuel. Pour chaque concept de la hiérarchie il y a un champ qui qualifie la taille de l'objet en utilisant des valeurs linguistiques telles que « petite », « grande », etc. ;
- les fonctions floues qui définissent les variables linguistiques de « taille ». Ces fonctions dépendent de la sémantique du domaine.

La procédure d'interprétation est composée des étapes suivantes :

(1a) Une étape de segmentation où nous faisons une segmentation multi-échelles en utilisant un algorithme de « mean-shift » hiérarchique. Cette étape nous permet d'extraire des régions de différentes tailles et homogénéités.

(1b) Nous extrayons de l'image les classes des objets dans H_C pour lesquels nous connaissons un algorithme d'extraction. Nous ne connaissons pas forcément toutes les méthodes d'extraction pour toutes les classes de la hiérarchie. Par exemple, supposons que les classes « végétation » et « jardin » se trouvent dans H_C. Nous savons comment extraire la végétation en utilisant des indices de radiométrie, mais ne nous ne pouvons pas extraire la classe « jardin » car nous ne connaissons pas une méthode pour l'extraire directement.

(2) Pour chaque nœud concept n_c de G nous construisons une fonction d'appartenance μ_c sur le domaine des régions obtenues à l'étape **(1a)**. La fonction μ_c indique quel est le degré avec lequel une région appartient à la classe de l'objet de n_c. Pour chaque région r, la valeur de $\mu_c(r)$ est égale à la conjonction du degré de satisfaction de la fonction de taille correspondant à la taille de l'objet représenté par n_c et les degrés de satisfaction des classes de H_c qui contiennent le concept de n_c. Par exemple si G contient un objet de la classe « jardin », nous utilisons la connaissance que « jardin » appartient à la classe « végétation » dans H_C et que les « jardins » ont une taille « petite ».

(3) Pour chaque nœud concept n_c de G nous réduisons les domaines en éliminant les régions qui ne satisfont pas les relations spatiales qu'elles devraient satisfaire selon G.

(4) Nous trouvons les régions qui satisfont le mieux le modèle G.

Les étapes **(1a)** et **(1b)** peuvent être réalisées en parallèle, puis les autres étapes sont faites de manière séquentielle. Dans la suite nous allons expliquer plus en détail les étapes **(3)** et **(4)**.

Réduction des domaines

En nous inspirant du travail fait dans le domaine des graphes conceptuels pour trouver les homomorphismes entre graphes [Chein and Mugnier, 2008], nous formulons le problème de trouver des régions dans l'image qui correspondent aux objets décrits par un graphe conceptuel comme un problème de satisfaction de contraintes floues (FCSP) [Dubois et al., 1996]. Tout d'abord nous allons traiter uniquement le cas où le graphe conceptuel n'a pas de nœuds concepts emboîtés, puis nous traiterons ce cas. Pour cela nous donnons la définition d'un FCSP $\mathcal{P} = \{\mathcal{X}, \mathcal{D}, \mathcal{C}\}$ où [Dubois et al., 1996] :

- $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ est un ensemble de n variables ;
- $\mathcal{D} = \{D_1, D_2, \ldots, D_n\}$ est un ensemble de n domaines. Chaque domaine D_i est associé à une variable x_i et représente l'ensemble de valeurs que x_i peut prendre ;
- $\mathcal{C} = \{C_1, \ldots, C_t\}$ est un ensemble de t contraintes floues. Chaque contrainte floue C_k est définie par une paire $\langle R_k, S_k \rangle$ où $S_k \subset \mathcal{X}$ est l'ensemble de variables impliquées dans C_k, et R_k est une relation floue sur le produit cartésien des domaines $D_{k_1} \times \ldots \times D_{k_n}$ entre les variables de S_k. R_k est défini par sa fonction d'appartenance $\mu_{R_k} : D_{k_1} \times \ldots \times D_{k_n} \to [0, 1]$.

Le degré de « consistance » d'un FCSP $\mathcal{P}$ est donné par :

$$Cons(v_1, \ldots, v_n) = \min_{\tilde{C}_k \in \mathcal{C}} \mu_{R_k}((v_1, \ldots, v_n) \downarrow_{S_k}) \tag{11}$$

où $(v_1, \ldots, v_n) \downarrow_{S_k}$ représente la projection de $(v_1, \ldots, v_n)$ sur l'ensemble des variables S_k. Un des problèmes associés à un FCSP $\mathcal{P}$ est de trouver les sous-ensembles des ensembles de $\mathcal{D}$ qui maximisent le degré de consistance (équation 11).

Dans notre cas $\mathcal{X}$ représente les nœuds concepts du graphe conceptuel, chaque $D_i \in \mathcal{D}$ représente l'ensemble de régions obtenues après la segmentation pour lesquelles la valeur d'appartenance à la classe représentée par x_i est supérieure à 0, et $\mathcal{C}$ représente les nœuds relations. Donc, si $\{v_1, \ldots, v_n\}$ sont des régions de l'image, la valeur $Cons(v_1, \ldots, v_n)$ détermine le degré avec lequel les régions $\{v_1, \ldots, v_n\}$ satisfont le modèle G. Malheureusement le problème de trouver les sous-ensembles de D qui ont un degré de consistance supérieur à zéro est un problème NP complet. Cependant, il existe des algorithmes qui permettent de réduire les domaines de $\mathcal{D}$ en appliquant des critères de consistance locale. Une fois que les ensembles de $\mathcal{D}$ sont réduits, il est possible de chercher les solutions de $\mathcal{P}$. Donc, nous utilisons la consistance locale.

Il existe plusieurs critères de consistance locale, et nous avons décidé d'utiliser le critère de consistance d'arc car il permet de bien réduire les domaines et il est facile à calculer. La définition de consistance d'arc existait uniquement pour des contraintes binaires floues, alors nous l'avons étendue pour des contraintes floues de n'importe quelle arité. Nous disons que $\mathcal{P}$ est arc-consistant si pour chaque contrainte C_k, chaque $x_i \in S_k$ et chaque $v \in D_i$ satisfont :

$$\mu_{x_i}(v) \leq \sup_{A=(a_{k_1},\dots,a_{k_n}):A\downarrow_i=v} \min[\mu_{R_k}(a_{k_1},\dots,a_{k_n}), \min_{j=k_1,\dots,k_n} \mu_{x_j}(a_{k_j})]. \tag{12}$$

où $\mu_{x_i}(u)$ indique le degré avec lequel $u \in D_i$ est approprié pour représenter x_i. Nous disons qu'une contrainte est arc-consistante si ses variables satisfont l'équation 12 pour chaque valeur dans son domaine. Pour trouver la fermeture arc-consistante de $\mathcal{P}$ l'algorithme FAC-3 est couramment utilisé, néanmoins cet algorithme traite uniquement des contraintes binaires. Donc nous l'étendons pour traiter le cas des contraintes avec n'importe quelle arité. Essentiellement, l'algorithme modifie de façon itérative les fonctions μ_{x_i} et réduit les domaines D_i jusqu'à ce que chaque contrainte C_k soit arc-consistante.

Si le graphe conceptuel a des nœuds concepts emboîtés il faut modifier le formalisme des FCSP pour prendre en compte les groupes d'objets. Nous proposons de voir les nœuds concepts emboîtés de manière duale, car ils peuvent être vus comme des objets ou comme des relations. Quand ils sont considérés comme des objets ils peuvent avoir des relations spatiales avec d'autres objets, et quand ils sont vus comme des relations nous évaluons les propriétés spatiales que doit satisfaire le groupe, par exemple l'alignement. Donc nous proposons de représenter des groupes dans le formalisme des FCSP à la fois comme une variable et comme une contrainte. Un groupe a deux fonctions d'appartenance sur l'ensemble des groupes de régions : une fonction qui indique le degré avec lequel un groupe représente bien les propriétés internes du groupe, et l'autre fonction qui indique le degré des propriétés internes et externes. Les propriétés internes traitent les degrés avec lesquelles les propriétés spatiales du groupe sont satisfaites (cette fonction d'appartenance est utilisée quand le groupe est vu comme une contrainte). L'autre fonction d'appartenance indique à quel degré un groupe représente bien le groupe vu comme un objet dans le modèle, et la valeur de cette fonction prend en compte le degré de satisfaction des relations spatiales avec d'autres objets, ainsi que les propriétés internes. Nous proposons un algorithme pour trouver la consistance d'arc d'un problème représentant un graphe emboîté. Cet algorithme considère la relation entre les deux fonctions d'appartenance qui caractérisent les nœuds concepts emboîtés.

Trouver une solution

Une fois que nous avons utilisé la consistance d'arc pour réduire les domaines, nous utilisons un algorithme de « back-tracking » pour trouver les solutions. Nous utilisons l'ordre leximin pour ordonner les résultats en partant du résultat qui satisfait le mieux le graphe. Cet ordre a été utilisé par [Fargier, 2006, Möller and Näth, 2008, Saathoff and Staab, 2008] et il maximise le nombre des contraintes qui sont satisfaites.

Conclusions

Dans la première partie de cette thèse nous avons introduit de nouvelles relations spatiales pour l'interprétation des images. Ces relations ont été modélisées dans le cadre de la logique floue, afin de bien représenter leur sémantique.

Dans une deuxième partie nous avons proposé un système d'interprétation qui s'appuie principalement sur la connaissance spatiale. Les principales contributions dans cette partie sont :

- Nous avons adapté le modèle de graphe conceptuel pour permettre la représentation de groupes d'objets alignés.
- Dans le cadre des FCSP nous avons deux contributions :
  - l'extension de la définition de consistance d'arc pour des contraintes avec une arité supérieure à 2,
  - la proposition d'un algorithme pour déterminer la fermeture d'arc pour les problèmes qui représentent un graphe conceptuel avec des nœuds concepts emboîtés.

Ce travail ouvre plusieurs perspectives. Dans le court terme les perspectives sont :

- Voir comment il est possible d'étendre les définitions des relations entre lignes et régions pour pouvoir les appliquer aux cas où nous appliquons les relations sur un groupe aligné au lieu d'une ligne.
- Considérer l'incertitude par rapport au modèle dans le système d'interprétation.
- Optimiser l'algorithme de la fermeture d'arc pour les problèmes représentant un graphe conceptuel avec des nœuds concepts emboîtés.
- Utiliser les connaissance des autres ontologies (par exemple celles du projet Differential and Formal Ontology Editor [1]) dans l'étape **(2)** du système d'interprétation.

Dans le long terme les perspectives sont :

- L'intégration du système d'interprétation à une plateforme de requêtes qui utilise une boucle de pertinence avec un utilisateur. L'idée est de créer une plateforme où l'utilisateur donne un graphe qui décrit la scène qu'il veut trouver dans l'image. A partir du résultat l'utilisateur pourra ajouter de nouvelles relations spatiales ou de nouveaux objets pour mieux cibler sa requête.
- Étudier différentes façons de créer automatiquement des modèles de graphes conceptuels à partir d'une base d'images étiquetées. Pour cela il faut étudier quelles sont les relations spatiales qui sont les plus pertinentes pour une description d'une classe d'objets complexes.

[1] http://dafoe4app.fr/

Introduction

The apparition of high resolution remote sensing (HRRS) images opens new challenges to the satellite image interpretation domain. HRRS images have more details than low resolution images, making it difficult to distinguish useful details from trivial ones. Therefore low level features such as texture or edges are not sufficient to identify a complex scene. Complex scenes are characterized by a group of objects having a spatial structure that is ruled by its function. Some examples of complex scenes are airports, harbors, train stations, malls and toll gates.

Considering only individual objects is not sufficient to determine the semantic of a complex scene. To describe, interpret and recognize a complex scene the spatial arrangement of these objects has to be considered, and thus it is necessary to make use of spatial structural representations. In the area of human cognition, a spatial representation is composed of two parts [Landau and Jackendoff, 1993]: *what* and *where*. The *what* deals with the geometric properties of the objects, such as shape, color, texture. The *where* deals with the position and spatial distribution of objects, represented as spatial relations. This representation is used to understand how humans encode a spatial description of an environment. When adapting this model for image description, imprecision and uncertainty should be considered, and more precisely, the imprecision of images due to acquisition processes and processing steps for object extraction, as well as the imprecision associated to the semantics of the spatial relations (such as "near", "east of") [Bloch, 2005], and the uncertainty of giving the correct label to a region.

In this thesis we address the problem of applying spatial reasoning for the interpretation of Earth observation images. To accomplish this objective we focus on the following questions:

1. What are the relations we can find in Earth observation images?

2. How can we represent the spatial knowledge in the image?

3. How can we represent the spatial structure of a complex scene?

4. How can we reason with the spatial knowledge to validate it ?

The first question deals with the domain of representation. To answer this question we determine the spatial relations that take place among objects in satellite images and which can be useful for their interpretation. The second question addresses the problem of properly representing the spatial relations in the image, and to capture their semantics. The third question searches for an appropriate structure to represent the spatial knowledge of the scene, and the last question deals with the use of spatial reasoning techniques over the image to eliminate inconsistencies in its interpretation. We concentrate on applying spatial reasoning on the problem of object recognition using a model. The model describes the spatial arrangement of the objects of the scene that we want to find in the image. To represent this knowledge we propose to use a conceptual graph which allows to represent relations for groups of objects like alignment which cannot be simply represented in a graph. We formulate the problem of finding the instantiations of a model in an image as a graph homomorphism, which is further modeled as a fuzzy Constraint Satisfaction Network. We extend the algorithms for determining the arc-consistency closure of a Constraint Satisfaction Network to cope with complex relations and take into account the difficulties of properly detecting the objects on the image. Finally, we test the proposed algorithm in scenes containing harbors and airports, obtaining satisfying results.

This thesis was performed at the TSI department at Telecom ParisTech in collaboration with the French Spatial Agency (CNES). This work was done within the Center of Competence on Information Extraction and Image Understanding for Earth Observation(CoC) involving CNES, DLR and Telecom ParisTech. It also contributed to the ANR project DAFOE.

Document structure

This document is divided into two parts: in the first one we concentrate on the definition of spatial relations, and in the second part we apply these relations for the interpretation of satellite images using a model.

Chapter 1 studies in detail the spatial relations which are found in Earth observation images. For each relation we present the definitions which have been proposed in the literature, in the case where they exist, focusing primarily on the definitions which have been defined within the fuzzy set framework and allows capturing the imprecision in the semantics of the relations. At the end of the chapter we present a hierarchy of spatial relations which summarizes the spatial relations, which are then used for the description of Earth observation images in this thesis.

Some of the relations presented in Chapter 1 are not well adapted for Earth observation satellite images, or have not been defined. Thus, in Chapter 2 we concentrate on these relations. These relations include the relations “surround”, “alignment”, “parallelism” (involving groups of aligned objects) and relations between linear objects and regions. For all these relations we study in detail their semantics, and propose definitions which are in accordance with the intuition.

To conclude this part, in Chapter 3, we present some of the reasoning methods which can be used to reason with the spatial relations presented in the preceding chapters, in order to combine, detect inconsistencies or acquire new spatial knowledge. We finish this chapter by presenting an example, where we combine several spatial relations by using simple rules. Nevertheless, when we want to find a complex scene containing several spatial relations among its parts, it is necessary to structure that knowledge, which is the aim of the second part of the thesis.

In Chapter 4 we review knowledge based systems, which allow to structure the knowledge needed for the task of image interpretation. We review the problems which have to be considered for automating the image interpretation task, in particular the image interpretation of Earth observation images. We detail some of the works which addressed the problem of image interpretation by considering the spatial relations between the objects, and moreover which consider the possible sources of information imperfections which can be present in this type of task.

Finally, in Chapter 5 we develop a formal framework for performing satellite image interpretation using a model which describes the spatial structure of the scene. The model extends nested conceptual graphs to allow us to represent all the relations defined in Chapter 2. To solve the mapping problem we propose new reasoning tools based on fuzzy constraint satisfaction problem approaches. Some examples illustrate the interest of the proposed approach for complex interpretation in high resolution satellite images.

Description of the image base

The algorithms proposed in this thesis were tested on images of an image base of *QuickBird* images created for the ORFEO project. The images base is constituted of fused multispectral very high resolution images of 61cm, obtained from the fusion of panchromaric images of 0.61m and multispectral (red, blue, green and near infrared bands) of 2.44m resolution. The image base contains images from Arles, Boumerdès, Boigneville, Marseille, Cannes, Cevennes, Durance, Genève, Saint Gilles, Massif de Maures, Peyrat le Château, Salon de Provence, Villamblain and Yard.

Part I

Spatial representations and reasoning

Chapter 1

Spatial relations for satellite image interpretation: state of the art

Semantical scene understanding involves the assessment of the spatial arrangement of objects. Using spatial relations does not only help us to discriminate the objects in the scene [Bloch, 2005], it also allows us to distinguish between different interpretations of two scenes with similar objects having different spatial arrangements [Aksoy et al., 2003]. Spatial relations are used in several applications in various domains: in medical images to recognize different brain structures [Bloch et al., 2003, Colliot et al., 2006], in image interpretation to provide linguistic scene descriptions [Keller and Wang, 2000], in remote sensing images mining of different types of objects or landcover [Aksoy et al., 2003, Guo et al., 2009], in Geographical Information Systems (GIS) applications to monitor land use [Mota et al., 2009] and cover changes [Le Ber and Napoli, 2002], and in robotics [Kuipers, 1978]. Their importance in describing the spatial organization of objects has been highlighted in several works. When developing a theory of spatial relations it is necessary to determine the minimal set of spatial relations needed to describe the spatial organization of objects [Abler, 1987]. In [Freeman, 1975] a set of fundamental relations for image interpretation is proposed: "left of", "right of", "beside", "above", "below", "behind", "in front of", "near", "far", "touching", "between", "inside" and "outside". Although this set of relations is appropriate for describing the content of an image, when considering spatial relations that take place among objects extracted from a satellite image, we should have in mind that we are observing geographical objects from above, and relations such as "below" should be redefined within this context or should not even be considered. The definition of an adequate set of spatial relations is important for defining a vocabulary to be used for the description of images' contents. By defining a set of relations, we mean establishing the relations that take place among objects in satellite images and also developing models which are adapted to the images. The evaluation of spatial relations that take place among objects in an image should consider the intrinsic imprecision of images due to acquisition processes and processing steps for object extraction, as well as the imprecision associated to the semantics of the spatial relations (such as "near") [Bloch, 2005]. The fuzzy set framework is appropriate for modeling spatial relations since it is able to capture these imprecisions.

Spatial relations should be modeled in such a way that they are able to answer one of the following questions [Bloch, 2005]:

(i) which is the region of space where the relation is satisfied?

(ii) to which degree is the relation between two objects satisfied?

According to the question that one wants to answer, the relation is modeled in a different manner. In section 1.3, we show how this issue has been addressed in literature. Besides the mathematical aspects, the cognitive, psychological and linguistic considerations must also be examined when developing models for spatial relations [Egenhofer and Franzosa, 1991].

In summary, to describe the spatial organization of objects it is necessary to specify the set of spatial relations that take place among objects in satellite images. In addition, models of spatial relations should consider the inherent imprecision of them and of images. Therefore, the primary goal of this chapter is to introduce the models of the spatial relations that we consider useful for satellite image interpretation. We focus mainly on the spatial relations that are usually used in image interpretation, and in geography, particularly in the fields of cartographic map generalization and of GIS, since satellite images contain a huge amount of geographical information. We would like to present the spatial relations in an structured manner as a hierarchy that can be further used for reasoning. First, in Section 1.1, we give an overview of some of the cognitive aspects of spatial relations. Then in Section 1.2, we discuss the different ways in which spatial relations have been classified in the literature, and finally we present some relations in Section 1.3.

1.1 Cognitive aspects of spatial representations

One of the branches of linguistics has focused on the link between language and spatial representations. In this section, we draw on the works of [Landau and Jackendoff, 1993, Talmy, 1983, Mathet, 2000]. This section is devoted to present some of the considerations that should be raised when dealing with the spatial representation in a general sense. We briefly examine these considerations and explain their link with computer vision. According to [Landau and Jackendoff, 1993] a spatial representation is composed of two parts:

what: deals with the geometric properties of objects,

where: deals with the spatial relations among objects.

In the following we will explain the characteristics of these parts.

The *what*

In language, objects are described using nouns. Among the objects properties, shape is one of the most characteristic ways to describe an object. This is translated in the computer vision domain by a large number of works dealing with finding an appropriate way to describe shape. Shape is usually described by a decomposition of an object into simpler parts. For instance Marr's representation of objects is based on a decomposition into several cones [Marr, 1982], or a boundary can be described by means of Freeman's chain code [Freeman, 1961, 1974]. The *what* has been an extensively studied problem in the computer vision community, and for which cognitive aspects have been considered.

The *where*

On the other hand, the *where* is a less mature problem which has not been extensively studied in computer vision. The *where* deals with objects' position. This position is defined using three elements [Talmy, 1983]:

1. the object to be located, that we refer to as target objects,
2. the reference object, and
3. the spatial relation.

Usually, spatial relations are encoded using spatial prepositions. Table 1.1 [1] shows some of the common prepositions used in English. It is surprising to realize that the number of prepositions in English is almost negligible with respect to the number of nouns. This would imply that we need much less words to describe the *where* than the *what*. [Landau and Jackendoff, 1993] suggest that one of the possibilities for this difference is that most of the spatial relations do not encode shape properties of the target nor the reference objects within the relation. This can explain why there are certain relations which are by nature ambiguous, for instance "near" or "between", and the evaluation of which differs according to the shape or the size of the reference and/or target objects [Mathet, 2000]. The spatial relations which are restricted to target and/or reference objects of certain shape are for example the relations "along" or "across" which require that the reference objects are elongated [Landau and Jackendoff, 1993, Talmy, 1983]. However, for the case of the relations "across", there are several possibilities of evaluation according to the shape of the target object [Landau and Jackendoff, 1993].

about	**behind**	during	**off**	till
above	**below**	except	**on**	**to**
across	**beneath**	for	**onto**	**toward**
after	**beside**	**from**	**out**	**under**
against	**between**	**in**	**outside**	**underneath**
along	**beyond**	**inside**	**over**	until
among	but	**into**	past	**up**
around	**by**	like	since	upon
at	despite	**near**	**through**	with
before	**down**	of	throughout	**within**

Table 1.1: Common English prepositions, spatial prepositions are in bold letters.

One might think that the use of target and reference objects is symmetrical in a binary relation. Nevertheless if we exchange the roles of reference and target objects in a phrase involving a symmetrical relation, then the sentence produces a strange phrase, for example [Talmy, 1983]:

(a) The bike is near the house.

(b) The house is near the bike.

The second sentence sounds odd, since the reference object is an object which is moving, while the house has a permanent location. [Talmy, 1983] argues that reference objects should be larger, easy to characterize and easy to perceive. This coincides with the approach in computer vision suggested by [Fouquier, 2010] for guiding the interpretation of an image, where the saliency as defined by [Itti et al., 2002] is used as a criterion for choosing the reference objects for the evaluation of the spatial relations.

[1] Taken from `http://www.englishlanguageguide.com/english/grammar/preposition.asp`

However, we should have in mind that the relations presented in this section are sufficient for expressing position in every day language, but when considering spatial relations that take place among objects extracted from a satellite image other relations should be considered. Nevertheless, the remarks about the separation of the *what* and *where*, and their characteristics, are still applicable.

1.2 Classification of spatial relations

A variety of spatial relations can be grouped into two families [Kuipers, 1978, Kuipers and Levitt, 1988]:

- topological relations which are invariant under topological transformations of the reference objects [Egenhofer and Herring, 1990, Randell et al., 1992], and
- metric relations in terms of distance and direction [Bloch and Ralescu, 2003, Krishnapuram et al., 1993, Bloch, 1999].

These categories have been used in the ontology of spatial relations proposed by [Hudelot et al., 2008] for image interpretation, and in [Kuipers, 1978] for navigation. The ontology of spatial relations for image interpretation of [Hudelot et al., 2008] gives a generic framework for spatial relations that appear in images. The models proposed in this ontology are adapted to represent structural knowledge in medical images. They are presented as fuzzy representations which allow considering the imprecision of the objects in the images as well as of the semantics of the relations. However, when dealing particularly with satellite images, these relations are not sufficient for describing the spatial distribution of objects, since there are relations which are very common among geographical objects which were not considered in the ontology, for instance alignment. Therefore, it is necessary to consider other relations which are dedicated only to geographical objects, for instance in the areas of map generalization or GIS.

In [Steiniger and Weibel, 2007] a classification of relations that appear among objects in a map is proposed. The objective of this enumeration is to determine the relations used in the process of map generalization. This work does not present any model of the proposed relations, nevertheless it includes an interesting catalog of relations. The relations used for map generalization can be seen as relations which try to extract the meaningful information of the spatial structure of the objects, since when passing from one cartographic scale to a smaller one, the map should remain equally informative. Among the proposed relations the authors dedicate a category to structural relations. Structural relations describe patterns that are perceived in maps, such as star-like or grid patterns (see Fig. 1.1). They can be considered as an extension of simple directional relations of two objects to more complex spatial relations. They are appropriate for describing man-made structures which appear in satellite images and should be considered for their interpretation.

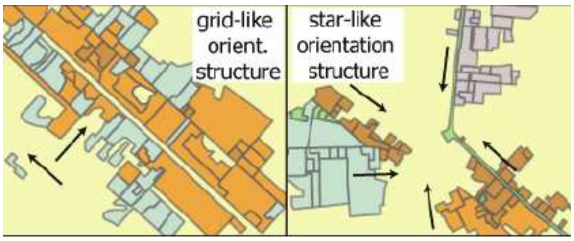


Figure 1.1: Example of structural relations, image taken from [Steiniger and Weibel, 2007].

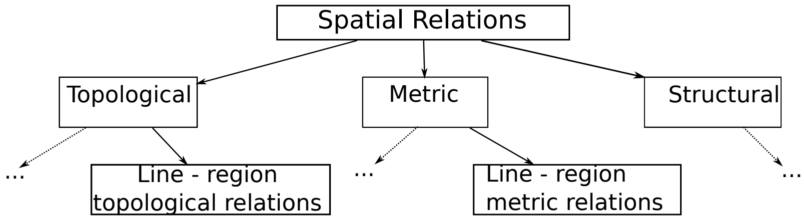


Figure 1.2: Main categories of spatial relations.

In the area of GIS [Liu et al., 2008b, Mark and Egenhofer, 1995] and in the area of linguistics [Mathet, 2000], relations have been discriminated according to the shape of their reference and target objects. As highlighted by [Landau and Jackendoff, 1993, Talmy, 1983] there are relations for which it only makes sense to evaluate them when the reference object is linear, for example "go across", "go into", "parallel to". In satellite images there is a strong presence of linear structures such as rivers or transport networks which makes spatial binary relations between a linear object and a region very frequent. Therefore, a classification of spatial relations for satellite image interpretation should also consider this category of relations.

A number of classifications have been proposed in the areas of artificial intelligence, GIS and map generalization. Each of these classifications has proved to be adapted to its targeted applications. Spatial relations which have been proposed for image interpretation consider the imprecision inherent to images and propose models for the relations. They have been applied in the area of medical images obtaining satisfying results for the segmentation and interpretation of brain images. However, the classification does not consider the structural patterns that appear in satellite images, nor the relations that only take place among objects of a particular shape. The spatial relations proposed for map generalization are meant to be identified by humans, and therefore they are not rigorously defined [Steiniger and Weibel, 2007]. In GIS we observe the introduction of relations depending on the shape of the objects involved in the relation. These types of relations do not appear in the previous classifications, since they are very specific to geographical objects. Hence, it is necessary to adapt and fuse the previously proposed classifications to derive a new classification which is adapted to satellite image interpretation. In the remainder of this chapter we present such a set of relations and their formal models when they exist.

1.3 The set of relations

We propose the classification shown in Fig. 1.2 which integrates the classification used in image interpretation by [Hudelot et al., 2008] and the structural relations of [Steiniger and Weibel, 2007]. In each of the categories we add the relations considering the kinds of objects (line or region). We dedicate a subsection to each of these categories where we present the relations and the models that have been proposed in the literature, and make a selection of them according to our needs. In the following, $\mathfrak{I}$ represents the image space and A, B denote two objects defined by two regions in $\mathfrak{I}$.

1.3.1 Topological relations

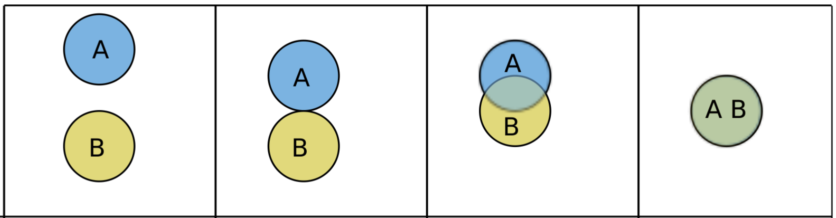
1.3.1.1 Region connection calculus RCC

Topological relations deal with the connection between two objects. Randell et al. [1992] defined the region connection calculus (RCC). The RCC8 framework is composed of 8 jointly exhaustive and pairwise disjoint (JEPD) base topological relations between two spatial regions A and B. These relations are based on the binary primitive $\mathbf{C}(A, B)$, which means A *is connected to* B. In the RCC8 context, $\mathbf{C}(A, B)$ is interpreted as being true when the closure of A and B share a point, where A and B are viewed as sets of points. The only requirement for the relation $\mathbf{C}$ is that it is reflexive and symmetric. Using $\mathbf{C}(A, B)$ a large number of relations can be defined [Cohn et al., 1997]. Table 1.2 contains some of these relations. The set of eight relations $\{\mathbf{DC}, \mathbf{EC}, \mathbf{PO}, \mathbf{EQ}, \mathbf{TPP}, \mathbf{NTPP}, \mathbf{TPPi}, \mathbf{NTPPi}\}$ constitutes the set of the RCC8 relations (see Fig. 1.3). They are invariant with respect to geometric transformations. It is possible to add more expressiveness to the RCC relations by introducing additional primitives. In [Cohn et al., 1995] 23 relations are defined by adding the convex hull as another primitive. This extension allows distinguishing different types of "inside" a region. In the context of [Cohn et al., 1995], a region is said to be inside another one when it is connected to its convex hull, but the regions do not overlap. This type of "inside" should not be confused with the "inside" of the *9-intersection model* of [Egenhofer and Herring, 1990] that is presented in Section 1.3.1.2.

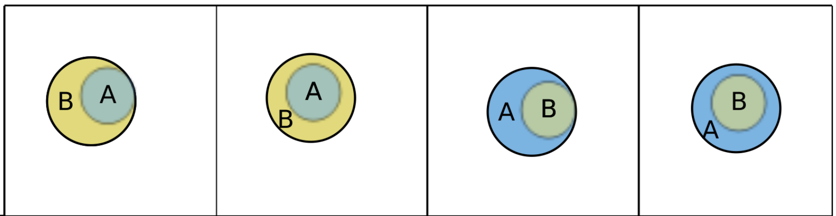
Relation	Interpretation	Definition of $\mathbf{R}(A, B)$
$\mathbf{DC}(A, B)$	A is disconnected from B	$\neg\mathbf{C}(A, B)$
$\mathbf{P}(A, B)$	A is a part of B	$\forall D[\mathbf{C}(D, A) \rightarrow \mathbf{C}(D, B)]$
$\mathbf{PP}(A, B)$	A is a proper part of B	$\mathbf{P}(A, B) \wedge \neg\mathbf{P}(B, A)$
$\mathbf{EQ}(A, B)$	A is identical with B	$\mathbf{P}(A, B) \wedge \mathbf{P}(B, A)$
$\mathbf{O}(A, B)$	A overlaps B	$\exists D[\mathbf{P}(D, A) \wedge \mathbf{P}(D, B)]$
$\mathbf{DR}(A, B)$	A is discrete from B	$\neg\mathbf{O}(A, B)$
$\mathbf{PO}(A, B)$	A partially overlaps B	$\mathbf{O}(A, B) \wedge \neg\mathbf{P}(A, B) \wedge \neg\mathbf{P}(B, A)$
$\mathbf{EC}(A, B)$	A is externally connected to B	$\mathbf{C}(A, B) \wedge \neg\mathbf{O}(A, B)$
$\mathbf{TPP}(A, B)$	A is a tangential proper part of B	$\mathbf{PP}(A, B) \wedge \exists D[\mathbf{EC}(D, A) \wedge \mathbf{EC}(D, B)]$
$\mathbf{NTPP}(A, B)$	A is a non-tangential proper part of B	$\mathbf{PP}(A, B) \wedge \neg\exists D[\mathbf{EC}(D, A) \wedge \mathbf{EC}(D, B)]$
$\mathbf{TPPi}(A, B)$	B is a tangential proper part of A	$\mathbf{PP}(B, A) \wedge \exists D[\mathbf{EC}(D, B) \wedge \mathbf{EC}(D, A)]$
$\mathbf{NTPPi}(A, B)$	B is a non-tangential proper part of A	$\mathbf{PP}(B, A) \wedge \neg\exists D[\mathbf{EC}(D, B) \wedge \mathbf{EC}(D, A)]$

Table 1.2: Some of the relations defined by $\mathbf{C}(A, B)$, taken from [Cohn et al., 1997].

The RCC relations are defined as crisp relations between crisp regions. However, in the context of geographical regions, as in satellite images, it is possible to have a region with imprecise boundaries, for example a city. Moreover, there are situations where two relations, for instance **EC** and **DC**, are difficult, or even undesirable to differentiate. This happens when the two regions are very close together, even though the two relations are mutually exclusive. Thus, in such cases, it is desirable to have a gradual transition from **EC** to **DC**. Several works have been dedicated to the extension of these relations to fuzzy relations. [Clementini and Felice, 2001, Cohn and Gotts, 1996] represent an imprecise region A by two sets of crisp regions; $\overline{A}$ consisting of the points which definitely belong to the region and $\underline{A}$ for the other points of the region. Then the topological relations between two objects A and B are obtained by observing the RCC8 relations for crisp objects between the pair of regions $\overline{A}$ and $\overline{B}$, $\overline{A}$ and $\underline{B}$, $\underline{A}$ and $\overline{B}$, and $\underline{A}$ and $\underline{B}$. Giving a total of 44 crisp RCC relations, the resulting relations do not have a semantic meaning and



RCC8	**DC**(A,B)	**EC**(A,B)	**PO**(A,B)	**EQ**(A,B)
9 Intersection	Disjoint $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Meet $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Overlap $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Equal $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$



RCC8	**TPP**(A,B)	**NTPP**(A,B)	**TPPi**(A,B)	**NTPPi**(A,B)
9 Intersection	Covered by $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	Inside $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	Covers $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$	Contains $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Figure 1.3: Illustration of the eight JEPD topological relations between two objects proposed by Randell et al. [1992] and Egenhofer and Herring [1990]. In the first row, a diagram of the relation is displayed (each image represents a different granularity of connection between two regions), with its corresponding RCC8 relation in the second row and the *9 intersection model* of Egenhofer and Herring [1990] relation in the third row.

only deal with the imprecision of the objects, but relations are still crisp with an abrupt passage between them. Other approaches search to extend the relations as fuzzy ones. [Schockaert et al., 2008] extended the relations by defining a fuzzy connection relation $\mu_C(A, B)$ between two regions A and B (which can be fuzzy or not). Such a connection relation should satisfy the requirement of being reflexive and symmetric. The obtained relations have the same properties as the RCC8 relations, which can be further used for spatial reasoning (see Chapter 3).

1.3.1.2 Interior, boundary and complement

Alternatively, in the area of GIS, Egenhofer and Herring [1990] also developed topological relations by observing the intersection between the interior, the complement and the boundary of the two objects. Each relation is uniquely identified through an intersection matrix. This matrix is called the *9-intersection* matrix and is defined as:

$$R(A, B) = \begin{bmatrix} A^\circ \cap B^\circ & A^\circ \cap \partial B & A^\circ \cap B^c \\ \partial A \cap B^\circ & \partial A \cap \partial B & \partial A \cap B^c \\ A^c \cap B^\circ & A^c \cap \partial B & A^c \cap B^c \end{bmatrix} \tag{1.1}$$

where A°, ∂A and A^c represent the interior, the boundary and the complement of the region A, respectively. In the following, we refer to these relations as the *9-intersection model* relations. Figure 1.3 shows the corresponding matrices for each relation for simple regions with no holes. Notice that these relations have a semantical meaning, which makes them suitable for a description. In total there are 512 matrices of 3×3, with elements belonging to $\{0, 1\}$. However, there are only 8 possible combinations of the $R(A, B)$ matrices which can be realized between simple regions with no holes in 2D. These matrices correspond to the same relations as the ones defined by the RCC8 model. By redefining the concept of region in this model it is possible to define other sets of JEPD relations. For example there are 33 relations between two lines, 2 between two points, 3 between a point and a line, and 31 between a line and a 2D region [Egenhofer and Herring, 1990]. Unlike the original *9-intersection model* relations (the ones shown in Fig. 1.3), the set of relations obtained between two lines or a line and a 2D region do not have a semantical meaning; therefore they cannot be used for image interpretation. There was an effort in [Mark and Egenhofer, 1994b] to obtain semantic relations between lines and 2D regions by adding metrical measures to the topological relations. The relations between lines and regions will be further discussed in Chapter 2.

In [Hudelot et al., 2008] the following fuzzy set theoretical relations have also been proposed as topological relations for image processing and interpretation: "degree of intersection", "degree of non-intersection" and "degree of inclusion". The "degree of intersection" corresponds to the topological "connection", the "degree of non-intersection" to "outside", and the "degree of inclusion" to "inclusion". For a complete review on these relations, refer to [Bloch, 2005]. Aside from observing if two fuzzy objects intersect, defining a degree of intersection is of great importance for image interpretation, since it can be used to define other relations. It can be often applied as a fusion operator, used as a measure of conflict or information, which are important components for spatial reasoning [Bloch, 2005]. The fuzzy intersection is directly extended from the crisp case. The degree of intersection between two fuzzy sets A and B with the membership functions μ_A and μ_B, respectively, is for instance given as:

$$\mu_{int}(A, B) = \sup_{x \in \mathfrak{I}} t\left[\mu_A(x), \mu_B(x)\right], \tag{1.2}$$

where t is a t-norm, and $\mathfrak{I}$ represents the image domain. This degree of intersection corresponds to the maximum of the intersection between two fuzzy sets, which is represented as a t-norm [Dubois and Prade, 1980]. The degree of intersection defined in Equation 1.2 does not account for different overlapping situations. To take into account the notion of spatial overlapping the following degree is proposed:

$$\mu_{int}(A,B) = \frac{V_n\left[t\left(\mu_A(x), \mu_B(x)\right)\right]}{\min\{V_n(\mu_A), V_n(\mu_B)\}}, \tag{1.3}$$

where $V_n(\mu_A)$ is the fuzzy hyper-volume of μ_A defined as: $V_n(\mu_A) = \sum_{x\in\mathfrak{I}} \mu_A(x)$.

Using the degree of intersection, the degree of non-intersection can be obtained as the fuzzy complementation of the degree of intersection:

$$\mu_{\neg int}(A,B) = c\left[\mu_{int}(A,B)\right], \tag{1.4}$$

where c is a fuzzy complementation, for example $c(x) = 1 - x$.

The degree of inclusion $\mu_{inclusion}(A,B)$ can be defined as:

$$\mu_{inclusion}(A,B) = \mu_{\neg int}(\mu_A, c(\mu_B)), \tag{1.5}$$

which expresses a degree of non intersection between A and the fuzzy complement of B. This forbids the intersection between A and the exclusion of B, as in the crisp case. In [Bloch, 2006] the author uses mathematical morphology to define the boundary and the interior of an object (fuzzy or not). Moreover the author shows how it is possible to extend the *9-intersection model* relations by performing a conjunction of the constraints of intersection and non-intersections of the terms in the *9-intersection model*. In this approach an extension of the relation set is performed which can be used in several applications, and can be used to construct new spatial relations. One very useful and interesting application is their use as comparison measures of a fuzzy set μ_B with a fuzzy set μ_R by using fuzzy pattern matching [Dubois et al., 1988]. The comparison is given by two values: a possibility degree Π, and a necessity degree N:

$$\Pi(\mu_B, \mu_R) = \mu_{int}(\mu_B, \mu_R) \tag{1.6}$$

$$N(\mu_B, \mu_R) = \mu_{inclusion}(\mu_B, \mu_R) \tag{1.7}$$

These measures have been used in the evaluation of directional relations in [Bloch, 1999, Bloch and Ralescu, 2003] which will be presented in Section 1.3.2.1.

1.3.1.3 Adjacency

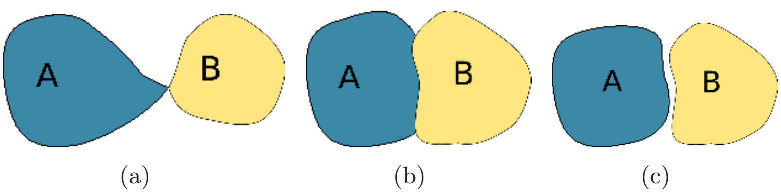


Figure 1.4: Illustrations where the relations of adjacency should be considered as a matter of degree.

In computer vision the adjacency relation is of great importance. This relation corresponds to the topological relation "meet". Two regions A and B are adjacent in an image if they do not intersect and if there exists at least a pixel of A which is a neighbor of a pixel of B. Thus, for the adjacency relation to be satisfied between two sets A and B, it is only necessary that there is one point in A which is a neighbor of a point in B. However, sometimes it is more appropriate to express the adjacency relation as a matter of degree. For example, one may want to give a higher degree to the situation shown in Fig. 1.4(b) than to the one in Fig. 1.4(a), since the common boundary is larger in the first case, and the relation does not depend on only one point. Moreover, in the situation shown in Fig. 1.4(c) one would like to say that the regions are almost adjacent, even if they do not touch. This remark led to a fuzzy definition of adjacency by [Rosenfeld and Klette, 1985] which depends on how nearly the borders of the two objects touch and for how long they do so. To do this [Rosenfeld and Klette, 1985] first introduced the notion of admissible segments. A line segment $]c, d[$ is admissible with respect to the crisp objects A and B if:

$$c \in \partial A \text{ and } d \in \partial B \text{ and }]c, d[\subseteq (A \cup B)^c, \tag{1.8}$$

where ∂A corresponds to the boundary of A. Let l_c be the shortest length of the admissible segment having $c \in \partial A$ as endpoint (if such a segment does not exist then $l_c = \infty$), and let $ADMI_{A,B}$ denote the set of admissible segments from A to B, then a degree of adjacency between A and B can be defined as:

$$\mu_{adj}(A, B) = \begin{cases} \frac{1}{|\partial A|} \sum_{d \in \partial A} \frac{1}{l_d + 1} & \text{if } ADMI_{A,B} \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{1.9}$$

This definition was extended to fuzzy regions in [Rosenfeld and Klette, 1985]. Unfortunately this definition is not symmetrical, and furthermore it is equal to zero only when $A \subseteq B$, which is not always desired. Another possibility is to express the degree of adjacency considering the intersection between boundaries. In [Bloch et al., 1997] the adjacency condition is expressed by means of morphological dilation:

$$A \cap B = \emptyset \text{ and } D_{V_c}(A) \cap B \neq \emptyset \text{ and } D_{V_c}(B) \cap A \neq \emptyset, \tag{1.10}$$

where $D_{V_c}(A)$ denotes the morphological dilation of A by the structuring element V_c. The first condition implies that both regions do not intersect. The neighborhood constraint of the adjacency definition is represented by the last two conditions, and the structuring element is chosen so that it represents the connectivity between the pixels. The last condition ensures the symmetry of the relation for the cases when the center of the structuring element V_c does not belong to V_c. This definition is directly extended to a fuzzy definition [Bloch et al., 1997], where the degree of adjacency is given as a conjunction of the constraints of Definition (1.10):

$$\mu_{adj}(A, B) = t\left[\mu_{\neg int}(\mu_A, \mu_B), \mu_{int}(D_{V_c}(\mu_A), \mu_B), \mu_{int}(D_{V_c}(\mu_B), \mu_A)\right]. \tag{1.11}$$

When Equations 1.2 or 1.3 are used as degrees of intersection, the obtained relation is symmetric, consistent with the crisp definition, decreasing with respect to the distance between both sets, and invariant to geometrical transformations. One advantage of this definition is its flexibility to allow different representations by using different definitions of μ_{int} or different structuring elements. For instance, by using a fuzzy structuring element one can represent the spatial imprecision and have a degree greater than zero when the objects are very close to each other [Bloch, 2005]. By using the hyper-volume of the intersection as μ_{int} (Equation 1.3) overlap between the intersecting sets is considered.

1.3.1.4 Topological surround

The topological relations of the *9-intersection model* are two point-based relations. This means that they can only be computed between regions that can intersect. They can be used in multi-scale analysis in image processing, for example [Inglada and Michel, 2009], or when objects are extracted in an independent manner. However, when dealing with regions that cannot intersect, for example when objects are extracted from a partition of an image, the only two *9-intersection* model relations that can take place are "disjoint" and "meet". [Aksoy et al., 2003] developed spatial relations based on the intersection of the boundary of the objects to quantify the different possibilities of "meet". These relations are referred to as "perimeter class" relations and correspond to the following set $\{\mathrm{BORDERING}, \mathrm{INVADED\_BY}, \mathrm{SURROUNDED\_BY}\}$ (see Fig. 1.5). These relations are characterized by fuzzy membership functions over the domain of r_{ij}, where r_{ij} corresponds to the ratio of the common perimeter of both regions over the perimeter of the reference region. [Liu et al., 2008a] compared the "perimeter class" and the *9-intersection model* relations. The only relations for which there is a one to one correspondence are the relations "inside" and "contains" of the *9-intersection model* which correspond to "surrounded by" and "surrounds" of "perimeter class", respectively.

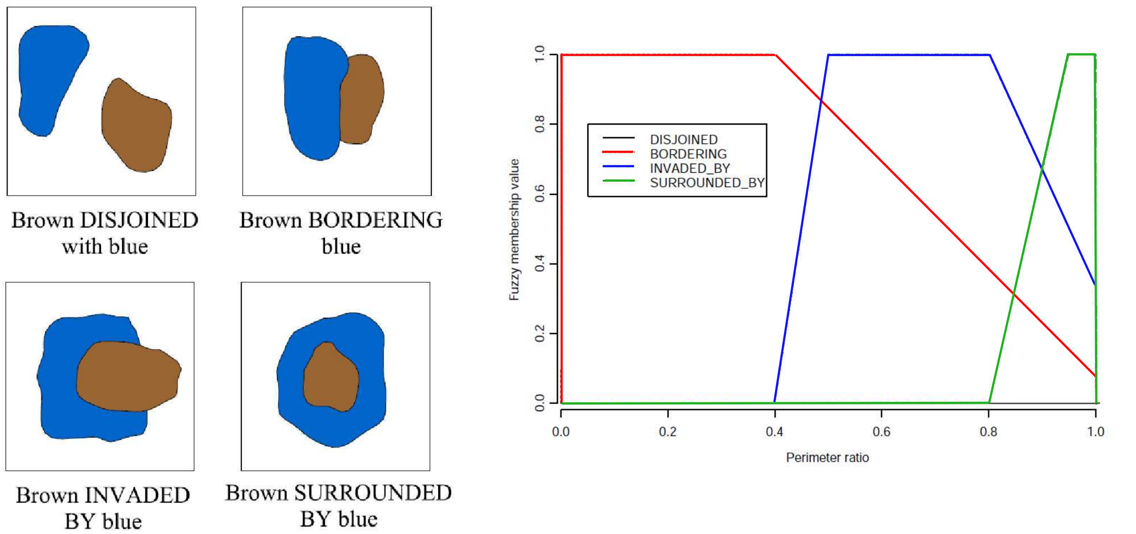


Figure 1.5: Perimeter relations taken from [Aksoy et al., 2003].

From the degree of adjacency of Equation 1.11, it is possible to think of an extension of "surrounds" and "surrounded by" relations for fuzzy objects. The ratio r_{ij} of the "perimeter class" relations can be redefined as:

$$r_{ij} = \frac{V_n[t(D_{V_c}(\mu_A), \mu_B)]}{V_n[t(D_{V_c}(\mu_A), c(\mu_A))]}, \tag{1.12}$$

where the numerator represents the volume of the intersection between the dilation of μ_A and μ_B, which can be considered as their common boundary. The denominator represents the volume of the boundary of μ_A. The "surrounded by" relation is defined through a membership function:

$$\begin{aligned} \mu_{SURROUNDS} : [0,1] &\rightarrow [0,1] \\ r_{ij} &\longmapsto f(r_{ij}) \end{aligned}$$

where f is an increasing function of $[0,1]$ into $[0,1]$, such that $f(1) = 1$ and $f(0) = 0$, for instance a trapezoid function. We refer to this type of surroundness as "topological surround" that should not be confused with the "topological surroundedness" of [Rosenfeld and Klette, 1985]. Among the relations presented by [Aksoy et al., 2003], we decided to only extend the "surrounded by" relation as a fuzzy relation, since it is the only one for which there is a one to one correspondence with the *9-intersection model* relations.

1.3.1.5 Discussion

We have presented the topological relations from the perspective of computer vision and from the one of GIS. Figure 1.6 presents a hierarchy of topological relations which are suitable for satellite image interpretation. This hierarchy is based on the hierarchy proposed in [Hudelot et al., 2008], which includes the RCC8 topological relations as well as the relations used for image processing. In addition we incorporate the refinement of the adjacency relation, the "surrounded by" relation proposed by [Aksoy et al., 2003], and add a class of line region topological relations that will be studied in Chapter 2.

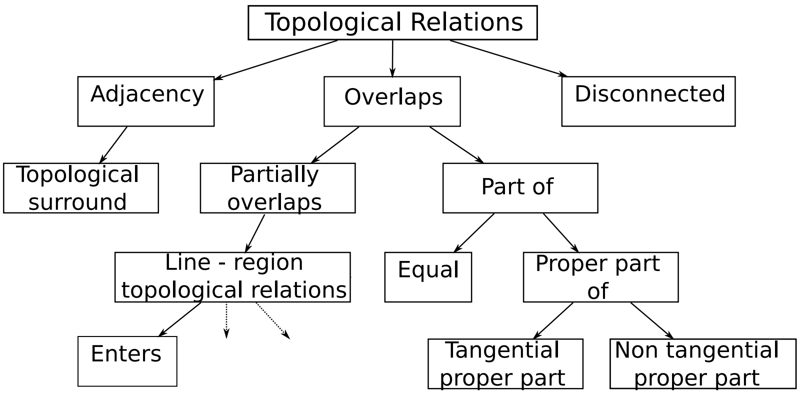


Figure 1.6: Topological relations.

In the following chapters we will use the models used in [Hudelot et al., 2008] for topological relations of intersection, interior and complement, because their extension to fuzzy objects is natural, and in the case of crisp objects the result is the same as for the RCC8 relations.

1.3.2 Metric relations

Metric relations are divided into two classes: distance, and directional relations. Unlike topological relations which express the position of an object with respect to another one, the metrical relations express the position of one object with respect to another one within a reference frame. For expressing a metric relation it is thus necessary to specify the reference frame. Three types of reference frames are distinguished [Retz-Schmidt, 1988, Hernandez et al., 1995]:

intrinsic reference frame: the orientation or distance is given by some internal property of the reference object, like its topology, size or shape,

extrinsic reference frame: the orientation or distance is determined by an external factor like motion, arrangement of objects,

deictic reference frame: the orientation or distance is given by an external point of view.

In the following we see that according to the reference frame different spatial relations can be specified.

1.3.2.1 Directional relations

The directional relations express the orientation of one object with respect to another one within a reference frame. When it is possible to differentiate the three axes of the reference object which define its orientation, one can use an intrinsic reference frame and distinguish directional relations such as: "to the left of", "to the right of", "behind of', "in front of", "above", and "below". However, in the case of satellite images, due to the point of view from which objects are observed it is often difficult to identify the intrinsic reference frame of them. These relations are therefore not very frequently used. In the context of geographic space it is common to use the extrinsic reference frame of cardinal directions: "North of", "East of" "South of", and "West of". Another possibility is to express direction by using the reference frame of the observer of the image, a deictic reference frame, and have directional relations as "to the left of", "to the right of", "behind of', "in front of', "above", and "below". As highlighted by [Retz-Schmidt, 1988] the same vocabulary of relations is used when using a deictic or an intrinsic reference frame, therefore it is necessary to clearly state the reference frame to avoid confusions.

Despite of the reference frame, directional relations are by nature imprecise [Miyajima and Ralescu, 1994a, Bloch and Ralescu, 2003]. In the following we define and give examples of directional spatial relations using the reference frame of the observer. Nevertheless these relations can be defined in any of the three frames just by specifying how the orientation is measured.

The three situations illustrated in Fig. 1.7 satisfy the relation "B is to the right of A". If we consider the two objects in Fig. 1.7(a) it is easy to identify that "B is to the right of A" since all the points of B are to the right of the points of A. However, in the situation presented in Fig. 1.7(b) the object B is still "to the right of A", but it can also be considered to be to some extent "above"A. Finally, for the third case, B is strongly "to the right of" A and "above" A. This example shows that even when the objects are crisp, the evaluation of the directional relations is ambiguous. It illustrates the interest of evaluating the directional relations by a degree of satisfaction (through a fuzzy definition) rather than a crisp evaluation. In the following we only review the fuzzy definitions of the directional relations.

Angle histograms

Angle histograms were introduced in [Miyajima and Ralescu, 1994a]. They can be interpreted as a function that captures the directional position between two objects. The angle histogram from A to B is obtained by computing, for each pair of points $p_a \in A$ and

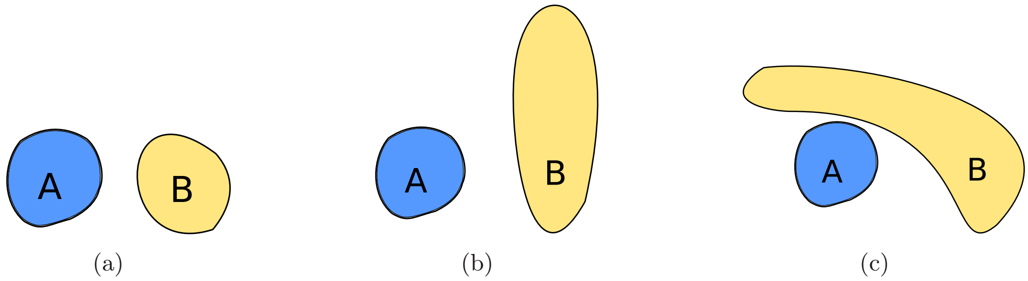


Figure 1.7: Ambiguity of the relation "B is to the right of A" (adapted from [Miyajima and Ralescu, 1994a, Bloch and Ralescu, 2003]).

$p_b \in B$, the angle between the segment joining them and the horizontal axis. Angles are organized in a histogram normalized by the largest frequency:

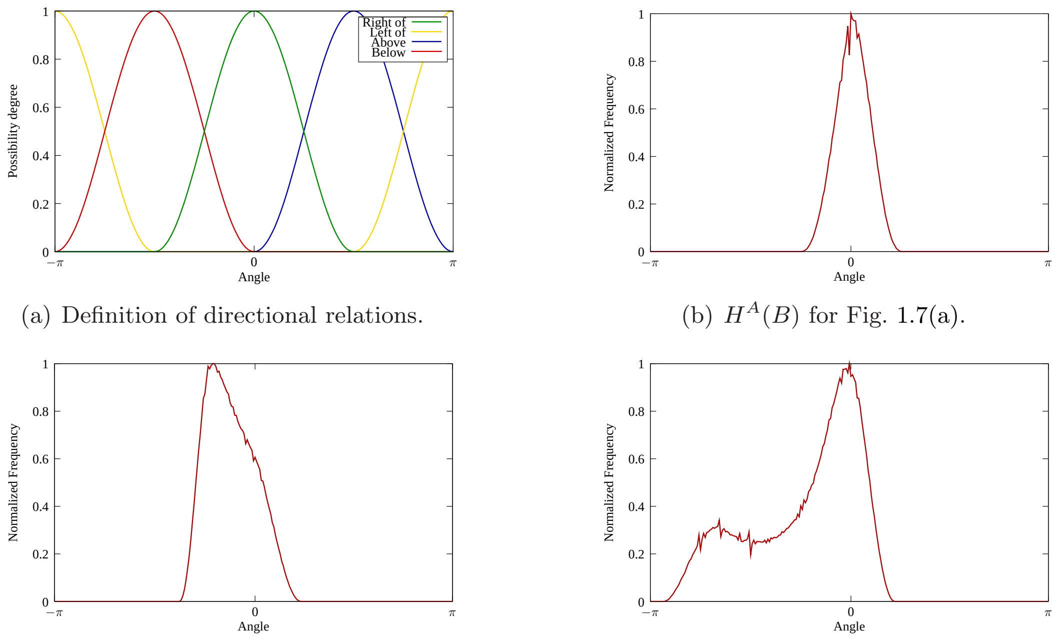
$$H^A(B)(\theta) = \frac{\sum_{p_a \in A, p_b \in B | \angle(\vec{p_a p_b}, \vec{u}_x) = \theta} 1}{\max_{\phi \in [0,2\pi)} \sum_{p_a \in a, p_b \in b | \angle(\vec{p_a p_b}, \vec{u}_x) = \phi} 1}. \tag{1.13}$$

where $\vec{u}_x$ is an unitary vector in the direction of the x-axis, $\vec{p_a p_b}$ is the vector between p_a and p_b, and $\angle(\vec{p_a p_b}), \vec{u}_x)$ denotes the angle between them. To determine if an object A is in a given direction with respect to an object B (for example "to the left of"), we can compute the angle histogram $H^A(B)$ and compare it with a template for the relation "to the left of". This can be done by using the compatibility between the computed histogram and the template [Miyajima and Ralescu, 1994a], or by using a fuzzy pattern matching approach [Bloch, 2005] and give the evaluation as a necessity(N)/ possibility (Π) pair (see Equations 1.6 and 1.7). Figure 1.8(a) shows an example of membership functions to define the directional relations. Figures 1.8(b), 1.8(c), and 1.8(d) show the angle histograms for the objects in Fig. 1.7. The angle histograms can be viewed as possibility functions which express the relative position between objects. For Fig. 1.7(a) the angle histogram has a well defined maximum with a strong main direction in accordance with the definition of "to the right of". The angle histogram for Fig. 1.7(b) still fits with the definition "to the right of", however it shows a tendency towards the function "above". The angle histogram for Fig. 1.7(c) shows the strong presence of both directions.

Angle histograms are also formalized for fuzzy objects. Given two fuzzy objects A and B defined through their membership functions $\mu_A : \mathfrak{I} \to [0,1]$ and $\mu_B : \mathfrak{I} \to [0,1]$, respectively, it is possible to define the angle histogram between A and B by considering the membership of each point to the fuzzy set:

$$H^A(B)(\theta) = \frac{\sum_{p_a, p_b \in \mathfrak{I} | \angle(\vec{p_a p_b}, \vec{u}_x) = \theta} t\,[\mu_A(p_a), \mu_B(p_b)]}{\max_{\phi \in [0,2\pi)} \sum_{p_a, p_b \in \mathfrak{I} | \angle(\vec{p_a p_b}, \vec{u}_x) = \phi} t\,[\mu_A(p_a), \mu_B(p_b)]}. \tag{1.14}$$

When μ_A and μ_B are crisp, Equations 1.14 and 1.13 are equivalent. Angle histograms are invariant to simultaneous translation, scaling and rotation of both objects. They are not symmetrical, but they satisfy: $H^A(B)(\theta) = H^B(A)(\theta + \pi)$. In addition, they have proved to be an adequate way for evaluating the directional spatial relation between two objects [Miyajima and Ralescu, 1994a] since they take into account the shape of the regions. However, angle histograms do not consider the distance between the points during the computation. [Matsakis and L.Wendling, 1999] developed the notion of histogram of forces which takes into account the distance information.



(a) Definition of directional relations.

(b) $H^A(B)$ for Fig. 1.7(a).

(c) $H^A(B)$ for Fig. 1.7(b).

(d) $H^A(B)$ for Fig. 1.7(c).

Figure 1.8: Angle histograms for objects in Fig. 1.7.

Angle histograms allow us to answer the question of to which degree two objects A and B satisfy one of the directional relations. They can also be used to characterize the relative direction between two objects, and to define more complex spatial directions as "between" or "alignment". In the next section we present an approach that allows us to define the region of space where a directional relation is satisfied.

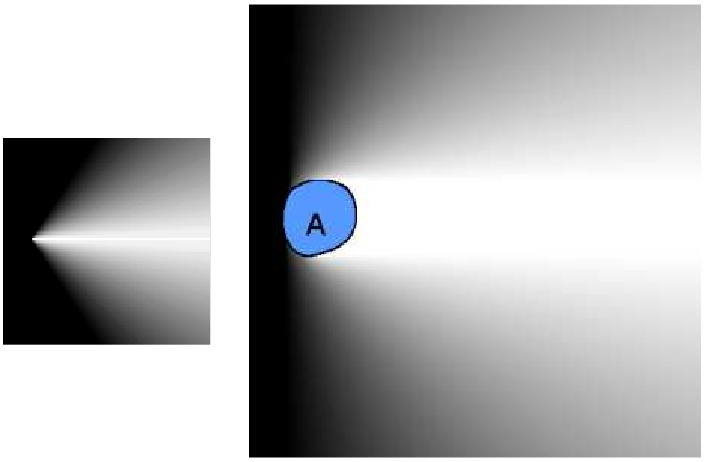
Fuzzy directional dilations

Fuzzy directional dilations were proposed by [Bloch, 1999] to construct a fuzzy landscape. A fuzzy landscape is defined as a fuzzy set in $\mathfrak{I}$ which represents for each pixel the degree of satisfaction of the directional relation with respect to a reference object A. Let $\vec{u}_\theta$ denote the normal unitary vector representing the direction of interest. The fuzzy landscape representing the objects is computed by performing a fuzzy morphological dilation of the object A by a fuzzy structuring element ν_θ chosen so that its has a high membership values in the direction $\vec{u}_\theta$. Its value at a point $x = (r, \alpha)$ (in polar coordinates) is a decreasing function of $|\theta - \alpha|$ modulo 2π, for example:

$$\forall p \in \mathfrak{I},\, \nu_\theta(p) = \max\left[0, 1 - \frac{2}{\pi}\arccos\frac{\vec{op}\cdot\vec{u}_\theta}{||\vec{op}||}\right], \tag{1.15}$$

where $\vec{op}$ is the vector between the center of the structuring element o and the point p. Figure 1.9 shows the structuring element defining the relation "to the right of" and the corresponding fuzzy landscape when the object A of Fig. 1.7 is used as reference object. From Fig. 1.9(b) it is possible to see how the obtained region which represents "to the right of" A is in accordance with the intuition.

The evaluation of to which degree an object B satisfies a directional spatial relation is done by comparing how well B matches the region having a high membership value in the



(a) Structuring element. (b) Fuzzy landscape.

Figure 1.9: Structuring element and fuzzy landscape representing the direction “to the right of” object A of Fig. 1.7.

fuzzy landscape. Several measures can be used for this comparison (see [Bouchon-Meunier et al., 1996] for a complete review on comparison measures for fuzzy sets). [Bloch, 1999] proposes to use a necessity(N)/ possibility (Π) pair (see Equations 1.6 and 1.7) or a mean measure given by:

$$M(B,R) = \frac{\sum_{p\in\mathfrak{I}} t\left[\mu_B(p), \mu_R(p)\right]}{\sum_{p\in\mathfrak{I}} \mu_B(p)} \tag{1.16}$$

where μ_R represents the fuzzy landscape and μ_B represents the membership function of B. In the case where B is a crisp object, then μ_B is an indicator function.

Representing a relation by a fuzzy set in the image has several advantages. For instance, if we need to evaluate a spatial relation between a reference object A and several objects $\{B_1, \ldots, B_n\}$ it is necessary to compute the fuzzy landscape one time, and evaluate the compatibility of the fuzzy landscape with every object. Another advantage is that it is possible to combine various relations which can be represented in this manner by fusioning their fuzzy landscapes (see Chapter 3).

The above presented methods allow us to represent the directional position of an object with respect to another one. However, if we are searching for a structural pattern in a satellite image and an extrinsic reference frame is used, the information of “object A is to the North of object B” does not give a lot information. Since most of the times spatial structures are orientation independent. For instance, in an airport, the oil storage zone can be to the North, East, South or West of the runaway. If we use a deictic reference frame, the directional relations “object A is to the right object B”, has the same inconvenient as the cardinal relations, since it does not express a particular arrangement. However, we can use the directional relations by not using the common directions, but as “object A is to a degree α of object B, with respect to the x-axis”. These kinds of relations are useful for identifying shadows of objects, since we know a priori the angle of the direction in which the shadow is projected (see Fig. 1.10). Similarly, as mentioned at the beginning of this section, in the case of satellite images, it is very difficult to distinguish the intrinsic reference frame of the reference object. However, in the case of elongated objects, we can

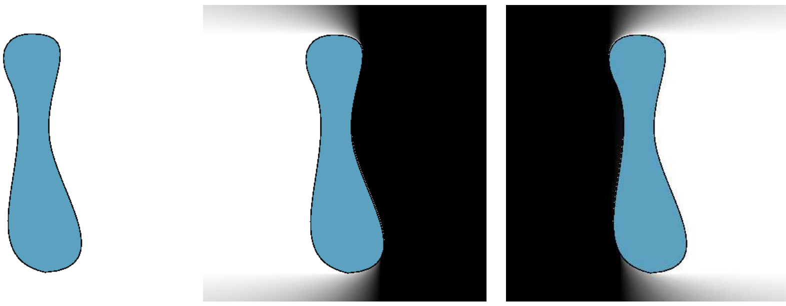
(a) Original image. (b) Building with fuzzy landscape. (c) Building and shadow (in blue) with fuzzy landscape.

Figure 1.10: Illustration of the relation "object A is in a direction α of object B, with respect to the x-axis" where α is equal to the angle of direction to which the shadow is projected. The shadow in (c) intersects the regions of the fuzzy landscape which have a high membership degree.

distinguish their sides from their ends, and then it is possible to express relations such as "on one side of". This relation is obtained by computing the orientation θ of the principal axis of the reference object with respect to the horizontal and performing two directional dilations, in the directions $\theta + \frac{\pi}{2}$ and $\theta - \frac{\pi}{2}$ (see Fig. 1.11). Notice that this relation produces two fuzzy landscapes R_1 and R_2. To evaluate if a target object B is "on one side of" an elongated reference object A, we evaluate to the degree to which B matches one of the two fuzzy landscapes. For instance if we use a mean measure (Equation 1.16), then the degree of satisfaction is given by:

$$M(B, \{R_1, R_2\}) = T\left[M(B, R_1), M(B, R_2)\right]. \tag{1.17}$$

This equation represents a disjunction (performed with a t-conorm T) of the degrees to which the object B intersects the fuzzy landscape R_1 or the fuzzy landscape R_2.



(a) (b) (c)

Figure 1.11: (a) Reference object. (b) and (c) correspond to the two fuzzy landscapes representing the relation "on one side of".

Between

The spatial relation "between" can have several meanings and can vary depending on the shape of the object. In [Bloch et al., 2006] several definitions for the relation "between" are proposed, given as "fuzzy landscapes". In this section we only discuss some of these definitions.

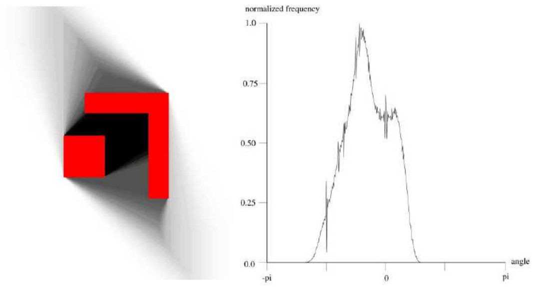


Figure 1.12: Illustration of the "between" relation based on dilation by a structuring element derived from the angle histogram (Equation 1.18). Objects A_1 and A_2 are displayed in red, and the membership values to $\beta(A_1, A_2)$ vary from (white) 0 to (black) 1. The angle histogram is shown on the right. Figure taken from [Bloch et al., 2006].

Let A_1 and A_2 be the reference objects. When A_1 and A_2 do not have concavities, or concavities which are not "facing each other", a definition based on directional dilation can be used. Let $\nu_1(\theta)$ be the angle histogram $H^{A_1}(A_2)(\theta)$ computed using Eq. 1.14 and let $\nu_2(\theta) = H^{A_2}(A_1)(\theta)$ (remember that according to the properties of angle histograms $H^{A_1}(A_2)(\theta) = H^{A_2}(A_1)(\theta + \pi)$). Then $\nu_1(\theta)$ and $\nu_2(\theta)$ represent the main directions between A_1 and A_2 and vice-versa. The region between A_1 and A_2 is given by:

$$\beta_{dil1}(A_1, A_2) = D_{\nu_2}(A_1) \cap D_{\nu_1}(A_1) \cap A_1^c \cap A_2^c, \tag{1.18}$$

where $D_{\nu_2}(A_1)$ represents the directional dilation of A_1 by the structuring element ν_2, and A_1^c represents the complement of A_1, and $\cap$ is an intersection which can be modeled by a t-norm. Equation 1.18 represents the intersection of the region which is at the same direction of A_1 as is A_2, the region which is in the same side of A_2 as is A_1 and the complement of $A_1 \cup A_2$. Figure 1.12 shows an illustration of this definition. When objects have concavities that are not "facing each other" , these concavities can be removed by using the following definition:

$$\begin{aligned}\beta_{dil2}(A_1, A_2) =& D_{\nu_2}(A_1) \cap D_{\nu_1}(A_1) \cap A_1^c \cap A_2^c \\ & \cap [D_{\nu_1}(A_1) \cap D_{\nu_1}(A_2)]^c \\ & \cap [D_{\nu_2}(A_1) \cap D_{\nu_2}(A_2)]^c \,,\end{aligned} \tag{1.19}$$

where the first line represent the region of (1.18) which is intersected with the regions corresponding to the complements of the concavities, represented in the second and third lines. Definition (1.19) is able to deal with simple convex objects and at the same time deal with complex objects having several concavities. The computational complexity is of the order of $O(NN_\nu + N_1N_2)$ where N denotes the image space, N_1 and N_2 denote the cardinality of A_1 and A_2 respectively, and N_ν denotes the cardinality of the support of the structuring elements ν_1 and ν_2.

Another definition proposed in [Bloch et al., 2006] is based on the concept of admissibility segments (see Section 1.3.1.3). This definition is appropriate to deal with objects having concavities. The region between A_1 and A_2 is defined as the union of the admissible segments. However the obtained region is very restricted, therefore the notion of admissible segments is extended to fuzzy semi-admissible segments to obtain a more flexible definition. This definition also deals with convex and complex objects, however the computational time is $O(NN_1N_2)$. Therefore, in our work we use Equation 1.19 to define the region between A_1 and A_2 when they are convex or complex objects. Furthermore, from Equation 1.18 one can think of the relation "on the same side", which can be seen as a sort of inverse of the relation "between". The fuzzy landscape $\gamma(A_1, A_2)$ corresponding to the relation "on the same side of A_1 as A_2" can be constructed by performing a morphological dilation of A_1 by $\nu_2(\theta)$, and intersecting it with the complement of A_1:

$$\gamma_{dil1}(A_1, A_2) = D_{\nu_2}(A_1) \cap A_1^c. \tag{1.20}$$

To extend this definition to the case where A_1 has concavities, we can simply eliminate the concavities of A_1 from the region γ as in Equation 1.19:

$$\begin{aligned}\gamma_{dil2}(A_1, A_2) = D_{\nu_2}(A_1) &\cap A_1^c \\ &\cap [D_{\nu_1}(A_1) \cap D_{\nu_1}(A_2)]^c \,.\end{aligned} \tag{1.21}$$

The definition of the relation "between" presented in Equation 1.19 is inappropriate when one of the reference elements has an infinite size with respect to the other one, such as the region that is found between a house and a road. Suppose A_2 has an infinite size with respect to A_1, then the region "between" A_1 and A_2 should be formed by the region between A_2 and the part of A_1 which is the closest to A_2. This region is obtained by approximating the part of A_2 which is closest to A_1 by a segment denoted by $\vec{u}$, and performing a fuzzy directional dilation of A_1 in the direction orthogonal to $\vec{u}$ (see Figure 1.13).

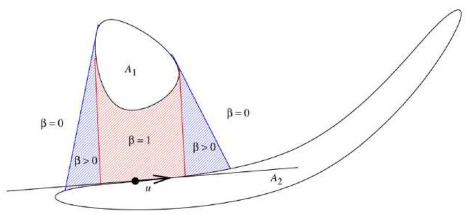


Figure 1.13: Illustration of the region between A_1 and A_2, when A_2 has very large size with respect to A_1. Figure taken from [Bloch et al., 2006].

One should notice that the relation "between" is independent of the reference frame.

Along

In [Takemura et al., 2005] a definition for the along relation is proposed, based on the relation "between". The "along" relation is evaluated by measuring the degree of elongation of the region between the target and the reference objects. To measure the degree of elongation several definitions are proposed according to the distance between the objects or their parts and the shape of the region between the objects. This approach gives satisfying results for objects with simple shapes.

1.3.2.2 Distance relations

Several works have addressed distances between fuzzy sets. The definitions can be classified into two classes [Bloch, 1999]:

- distances that compare membership functions, and
- distances that combine the membership functions and spatial distances.

The second type of distances is more appropriate to describe the spatial arrangement between objects in an image. In this section we only focus on the second type of distances. For a complete review refer to [Bloch, 1999]. [Bloch, 1999] proposes to use fuzzy mathematical morphology to construct the fuzzy landscape representing the satisfaction of the relations: “at a distance less than d”, “at a distance greater than d”, and “at a distance between d_1 and d_2”. The fuzzy landscape corresponding to the relation “at a distance between d_1 and d_2” is constructed by defining a membership function μ_n over $\mathbb{R}^+$ with core equal to $[d_1, d_2]$, for example a trapezoid function. From this membership function two fuzzy structuring elements are defined:

$$\nu_1(x) = \begin{cases} 1 - \mu_n(d_E(x,0)) & \text{if } d_E(x,0) \leq d_1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.22}$$

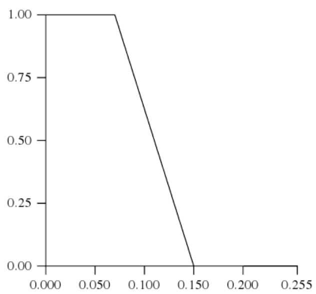
$$\nu_2(x) = \begin{cases} 1 & \text{if } d_E(x,0) \leq d_2, \\ \mu_n(d_E(x,0)) & \text{otherwise.} \end{cases} \tag{1.23}$$

where d_E is the Euclidean distance in $\mathfrak{I}$ and $d_E(x,0)$ is the Euclidean distance between x and the origin of the structuring element. Finally, the fuzzy landscape representing the relation “at a distance between d_1 and d_2 from A” is defined as:

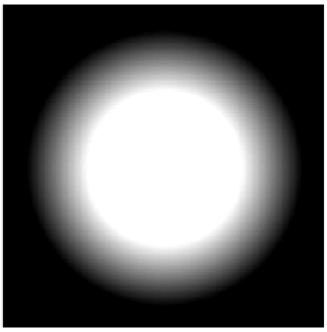
$$\mu_{distance}(x) = t\left[D_{\nu_2}(\mu_A), 1 - D_{\nu_1}(\mu_A)\right], \tag{1.24}$$

which is the conjunction of a distance inferior to d_2 represented by the first term and a distance greater than d_1 represented by the second term. The distance greater than d_1 is expressed as the complement of a distance inferior to d_1. Figure 1.14 shows an example of the relation “at a distance less than d” by replacing d_1 by 0, and d_2 by d in Equations 1.22 and 1.23.

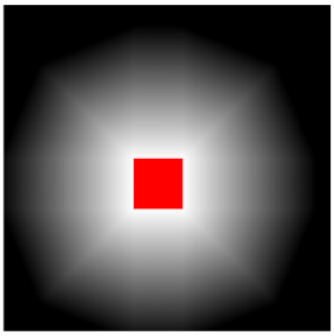
Distance relations can also be named with respect to a granularity of the space. For example, we can define values of d_1 and d_2 in Equations 1.22 and 1.23 to define distances such as “near” or “far” as in [Hudelot et al., 2008]. This type of relations are referred to as *qualitative* distance relations [Hernandez et al., 1995]. For these distances [Hernandez et al., 1995] makes the distinction of the different ways of how “near” and “far” can be defined according to the reference frame which is used. In the context of distances, in an intrinsic reference frame, the distance is determined by the intrinsic characteristics of the reference objects. For example, distance can be defined as a function of the size of the reference object. In an extrinsic reference frame, the distance is determined by some external factor. For instance, in a subway system, we may say that two metro stops are close if the time from going to one station to the other one is less than 2 minutes. In a deictic reference frame a distance "near" can make reference to a mental map of the observer. In the context of satellite image interpretation, it is not always possible to have information to construct an extrinsic or deictic reference frame to define the relation “near”. Therefore we will use only the intrinsic reference frame, where the parameters d_1 and d_2 used to construct the membership function μ_n are proportional to the size of the reference object.



(a) Function μ_n over $\mathbb{R}^+$.



(b) Structuring element $\nu_1(x)$ over $\mathfrak{I}$.



(c) Fuzzy landscape.

Figure 1.14: Representation of the fuzzy relation "at a distance less than d" with $d = 0.1$. Image taken from [Hudelot et al., 2008].

1.3.2.3 Discussion

Figure 1.15 shows the hierarchy of metric relations. This hierarchy has the standard classification distance vs. directional relations. For the distance relations we add the two distance relations initially presented in [Freeman, 1975] as well as the relation "at a distance between d_1 and d_2 from A", which allow us to cover different spatial configurations. The proposed directional relations do not include the cardinal relations, which are frequently used in GIS, because they cannot be used to discriminate a specific spatial arrangement. The directional relations which use an intrinsic or deictic reference frame such as "to the left of" are not considered because they depend on the objects' reference system which is usually unknown. Nevertheless, we include relations such as "at an angle of" or "on one side" which can be seen as different interpretations of classical directional relations and are better adapted to the satellite image interpretation context. The spatial relations "parallelism", "across" and "surround" are introduced in Chapter 2.

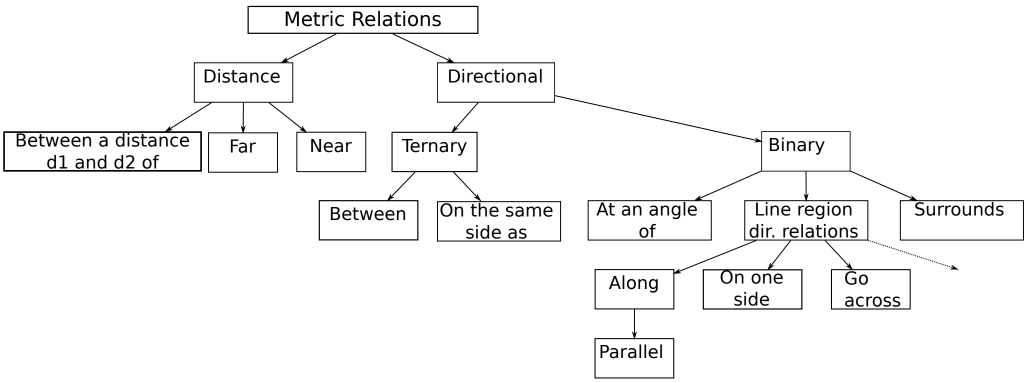


Figure 1.15: Metric relations.

1.3.3 Structural relations

Structural relations have not been used so far in image interpretation, although they can play an important role in image description. The structural relations which are proposed in [Steiniger and Weibel, 2007] are grid-like, star-like and alignment. In the following chapter we will concentrate on the definition of the alignment relation within the context of image interpretation, the other two relations will not be treated in this work. Although they are important relations, they are less frequent than the alignment relation.

One important characteristic of the structural relations is that they describe a group of objects satisfying a spatial property, which in turn can satisfy a spatial relation as a group.

1.4 Conclusion

In this chapter we presented a hierarchy of spatial relations which can be used for the interpretation of satellite images. A schema of the hierarchy is illustrated in Fig. 1.16. This selection of relations considered the relations necessary for the image interpretation taking into account the geographical nature of the objects which are observed in this type of images. Although there has been a large amount of work devoted to the formal definitions of spatial relations, in the crisp case and for fuzzy objects, some complex relations, that are useful for satellite image interpretation have not been addressed and require specific developments. This is a contribution of this Thesis, detailed in Chapter 2.

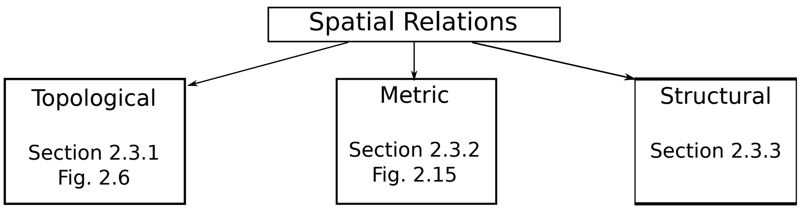


Figure 1.16: Hierarchy of spatial relations

In chapter 3 we present some mechanisms used to perform spatial reasoning with the presented relations. In part II we present an application of how these relations can be used for image interpretation.

Chapter 2

New spatial relations

In the previous chapter we presented a set of spatial relations that are of interest in satellite images. Some of the presented relations have not been addressed in the literature, or the models which have been proposed are not well adapted to the context of satellite image interpretation. Hence, in this chapter we propose to model some of these relations. In Section 2.1, we study the directional relation "surround". Then in Section 2.2 we study the structural relation of "alignment" and we study its link with parallelism, which is further developed in Section 2.3. Finally in Section 2.4 we study the line-region relations.

2.1 Surround

In satellite images it is frequent to observe that one object is surrounded by another one. For example, the structures on the sea are surrounded by the sea (see Fig. 2.1). We can say that "A is surrounded by B" if A is able to see B in almost all the directions.

Figure 2.1: Example of surround.

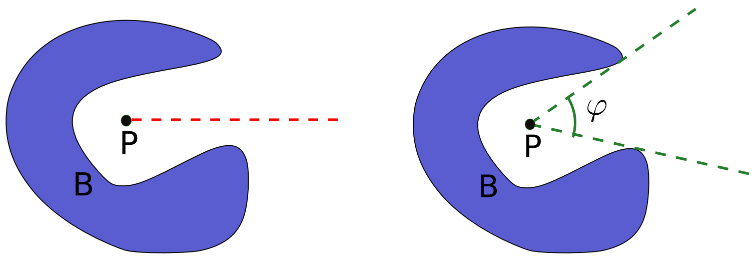
The notion of B being in almost all directions with respect to A is by definition imprecise. It is imprecise in the sense that "almost all" is not a well defined quantity. Additionally, as explained in Section 1.3.2.1, the notion of being "in a direction α" from another object is also imprecise. Therefore, the fuzzy set theory is appropriate for modeling this relation.

In Section 2.1.1 we review some of the definitions proposed in the literature for the surround relation. Then, inspired by some of these works, we propose a novel definition of fuzzy surround in Section 2.1.2. The contributions are twofold: first we propose a definition as a spatial fuzzy set (a fuzzy landscape), which defines the degree of satisfaction of the relation to a reference object for each point of the space; secondly, we propose to include several important pieces of information in the definition, such as the potential imprecision of the objects (i.e. dealing with fuzzy sets), the distance to the target object, and shape information. In Section 2.1.5 an example with a complex shaped reference object illustrates that the proposed definition provides adequate results, in accordance with intuition.

2.1.1 Related work

[Rosenfeld and Klette, 1985] distinguish between two types of fuzzy surround: visual and topological surround. These definitions are extensions of the crisp definition of surroundedness. In the crisp case, a region A is surrounded by a region B if for every point $p \in A$ all the linear paths from p to the background (the background is considered as the outside of the image) intersect B. It is equivalent to say that B separates A from the background.

One extension of surroundedness proposed by [Rosenfeld and Klette, 1985] is the topological surroundedness. This definition is first proposed for a point and then extended to the case of objects. The degree to which a point p is topologically surrounded by an object B depends on the angular change followed by a path from p to the background. [Rosenfeld and Klette, 1985] give as an example a point which is in the center of a spiral, so the point is "topologically surrounded by" the spiral. This notion is extended to the case of objects, by considering the "topological surroundedness" of the points on the boundary of the target object. The degree to which "A is topologically surrounded by B" is equal to the infimum of the degree of "p is topologically surrounded by B" for all the $p \in \partial A$, where ∂A is the boundary of A. Nevertheless, this definition of fuzzy surroundedness does not express the relation we want to define mainly because it is a crisp notion which is not flexible enough. Fig. 2.2(a) shows a situation for which a point is not topologically surrounded by an object, although we would like to have a high degree of "surround".



(a) Topological surroundedness is not satisfied.

(b) Visual surroundedness has a non-zero degree of satisfaction.

Figure 2.2: (a) In red a direct path from the point to the background is shown, demonstrating that the topological surroundedness is not satisfied. (b) In green the angular interval φ for which $r_\theta(p, B) = 0$.

The other extension of surround proposed by [Rosenfeld and Klette, 1985] is better adapted to our needs. The visual surroundedness deals with the angular coverage of p by a set B. [Rosenfeld and Klette, 1985] first introduce the function r_θ:

$$r_\theta(p, B) = \begin{cases} 1 & \text{if } \exists b \in B \text{ such that } \angle(\vec{pb}, \vec{u}_x) = \theta, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

where $\vec{u}_x$ is the unitary vector in the direction of the x-axis, and $\angle(\vec{pb}, \vec{u}_x)$ denotes the angle between the segment $[p, b]$ and x-axis. The function $r_\theta(p, B)$ is equal to one, if there exists a ray from p in the direction θ which intersects B. The degree to which a point p is visually surrounded by an object B is then given by:

$$\mu_{visual\_surround}(p, B) = \frac{1}{2\pi} \int_0^{2\pi} r_\theta(p, B) d\theta \tag{2.2}$$

This degree measures the portion of angular coverage of p by B. Fig. 2.2(b) shows the same situation which did not satisfy the topological surroundedness, and we see that for the visual surroundedness the degree of satisfaction is $\mu_{visual\_surround}(p, B) = 1 - \frac{\varphi}{2\pi}$, where φ is equalt to the angular interval for which $r_\theta(p, B)$ is equal to zero. The extension to see whether an object A is surrounded by an object B is straightforward:

$$\mu_{visual\_surround}(A, B) = \min_{p \in \partial A} \mu_{visual\_surrounds}(p, B) \tag{2.3}$$

The definition of visual surroundedness is in agreement with the relation we want to define. If an object A sees B in almost all directions, then there is a large angular coverage of A by B and therefore $\mu_{visual\_surround}(A, B)$ is high. However, the extension to non punctual objects is not a trivial task.

[Miyajima and Ralescu, 1994b] propose a definition of surround also based on the angular coverage. To evaluate the degree to which an object A is surrounded by B the maximum angle φ between two tangent lines of B passing through a point of A is computed (this angle is the same as the one illustrated in green in Fig. 2.2(b)). The degree of satisfaction of the relation decreases as φ gets larger. So the membership function defining the surround relation is expressed by:

$$\mu_{surrounds}(\varphi) = \begin{cases} cos^2\left(\frac{\varphi}{2}\right) & \text{if } 0 \leq \varphi \leq \pi, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Then the degree to which a region B surrounds a region A is obtained by the center of gravity of the compatibility distribution of the angle histogram $H^B(A)$ (Equation 1.13) and $\mu_{surrounds}(\varphi)$. This definition is also in agreement with the definition of surroundedness. A high satisfaction value is obtained for the situation shown in Fig. 2.2(b).

[Matsakis and Andréfouet, 2002] also present a definition of surround based on angle calculation. The normalized angle or force histogram $H^A(B)$ is computed. Then for every α-cut $H^A(B)_\alpha$ of $H^A(B)$ the angle z_1 is defined as the largest angle interval which is not included in $H^A(B)_\alpha$. The degree of surround for an α-cut is given by a decreasing function of z_1, and in [Matsakis and Andréfouet, 2002] the function $\max(0, 1 - \frac{z_1}{\pi})$ is used. Finally, the degree of surround is given by an integration over all the α-cuts. This approach does not measure the angle coverage, but it only observes the largest angular interval which is not covered. For instance, if B is a disconnected object composed of three points which are arranged to form an equilateral triangle, and A is just a point located in the center of the triangle, then for every α-cut the angle z_1 is equal to $\frac{\pi}{3}$ and the degree of surround is 0.67. This degree is not in accordance with the definition of surroundedness we want to define, since B only sees A in very few directions.

Other works define surroundedness as a conjunction of the degrees of relative position in all the directions [Bloch, 1999, Miyajima and Ralescu, 1994b]. These approaches assume that an object is in all directions with respect to the other object. The relation which we want to define here does not satisfy this condition, since we are interested in "almost" all directions.

The surround relations studied in [Miyajima and Ralescu, 1994b, Rosenfeld and Klette, 1985] are in agreement with the relation which we want to define. However, these definitions have focused on assessing the degree to which the relation is satisfied between two objects. Here we concentrate on defining a region of space, as a fuzzy landscape, where the relation is satisfied. The advantages of using this approach were highlighted in Section 1.3.2.1 when directional relations were introduced. Moreover, in order to deal with complex shapes, we propose as an additional contribution to include information on the morphology of the objects as well as their distance to the target objects, which were not directly taken into account in existing definitions.

2.1.2 Definition of surroundedness as a fuzzy landscape

We first define the relation for a crisp reference object, and then we extend it to the case of fuzzy objects.

2.1.2.1 Crisp reference object

Using the same idea as the visual surroundedness of [Rosenfeld and Klette, 1985], we define the angular coverage of a point by a region B. For every point $p \notin B$ the angular coverage is equal to the total angular length of the angular intervals for which p is able to see B, that is:

$$\theta_{coverage}(B)(p) = \int_0^{2\pi} r_\theta(p, B)d\theta \tag{2.5}$$

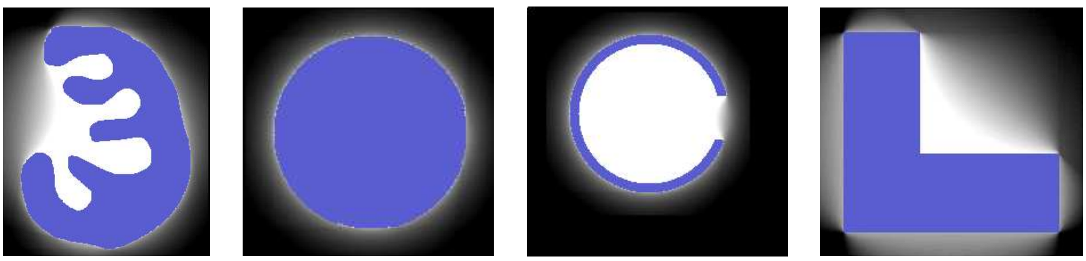
The region representing "surrounded by B" can then be defined as a fuzzy region of space with membership function equal to:

$$\mu_{surround}(B)(p) = f(\theta_{coverage}(B)(p)) \tag{2.6}$$

where $f : [0, 2\pi] \to [0, 1]$, $f(0) = 0$ and $f(2\pi) = 1$. The purpose of f is to define the membership function defining the semantics of "almost all" for the angular coverage.

Figure 2.3 shows some examples of this definition of surroundedness applied to synthetic objects. In these examples f was defined as:

$$f(x) = \begin{cases} 1 & \text{if } x \geq \frac{5\pi}{4} \\ \frac{x-\frac{\pi}{2}}{\frac{3\pi}{4}} & \text{if } \frac{\pi}{2} \leq x < \frac{5\pi}{4} \\ 0 & \text{if } x < \frac{\pi}{2} \end{cases} \tag{2.7}$$



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

Figure 2.3: Fuzzy landscapes obtained with the definition of Equation 2.6 when the blue object is used as reference object. The brightest grey level corresponds to the higher membership values. We note grey levels "around" the objects which do not correspond to the wanted property "to be surrounded by" the object.

The results shown in Figure 2.3 are not in complete agreement with the notion of angular coverage. Some non zero membership areas appear around convex parts of objects, which is counter-intuitive (see e.g. the convex object in Figure 2.3(b)), since a convex object cannot surround another object.

On the other hand, one can interpret the surround relation by considering that A is "surrounded by" B, if there is a portion of the boundary of B that goes around A. In [Mathet, 2000] the linguistic aspects of the relation "go around" are analyzed, which leads to consider that a path C "goes around" an object A if A and C do not intersect, and if A intersects the convex hull of C. Consequently, the "surround" relation should only take place when the reference object has concavities. Moreover, the portion of boundary belonging to the concavities is the one which "goes around" the other object. As a result, the angular coverage should be computed only using the points on the boundary of B which lie in a concavity. A point lies in a concavity if it is in $\partial B \setminus \partial CH(B)$, where $CH(B)$ and ∂B are the convex hull and boundary of B, respectively. A point b that lies in a concavity sees p if the segment $[p, b[$ does not intersect B. Thus, to take into account the concavities as well as the notion of visibility, we redefine the function $r_\theta(p, B)$ of [Rosenfeld and Klette, 1985] (Equation 2.1). We only consider the segments which have an endpoint in $\partial B \setminus \partial CH(B)$ and do not intersect B, using the same idea as for the admissible segments of [Rosenfeld and Klette, 1985]:

$$\tilde{r}_\theta(p, B) = \begin{cases} 1 & \text{if } \exists b \in \partial B \setminus \partial CH(B) \text{ such that} \\ & \angle(\vec{pb}, \vec{u}_x) = \theta \text{ and } [p, b[\cap B = \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

Figure 2.4 shows an example of the angular interval where $r_\theta(p, B) = 0$ and $\tilde{r}_\theta(p, B) = 0$. We can see that $\varphi < \tilde{\varphi}$, and therefore it is necessary to add the notion of visibility to penalize the points which are not really seen by a concavity.

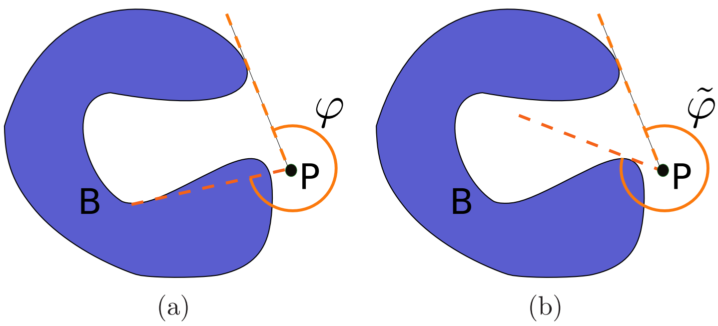


Figure 2.4: In orange the angular interval φ and $\tilde{\varphi}$ for which $r_\theta(p, B) = 0$ in (a) and $\tilde{r}_\theta(p, B) = 0$ in (b).

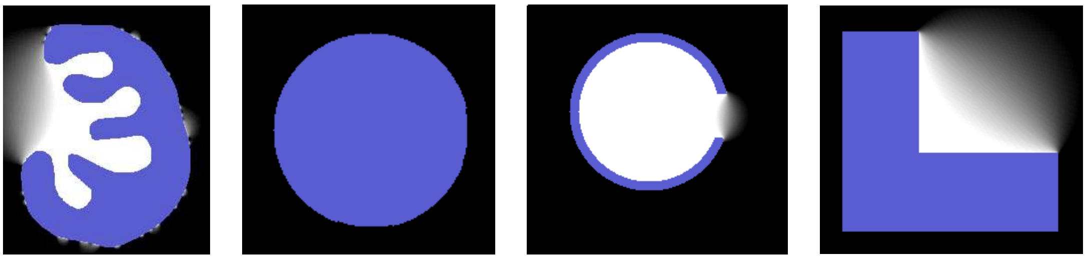
As a result, the angular coverage should be computed using $\tilde{r}_\theta$:

$$\theta_{coverage\_CH}(B)(p) = \int_0^{2\pi} \tilde{r}_\theta(p, B) d\theta \tag{2.9}$$

Again the fuzzy landscape representing the region surrounded by B is obtained by applying $f(\theta_{coverage\_CH})$ to each point in $\mathfrak{I}$, where f has the same properties as the function presented in Equation 2.6:

$$\mu_{surround}(B)(p) = f(\theta_{coverage\_CH}(B)(p)) \tag{2.10}$$

Figure 2.5 shows the fuzzy landscapes for the reference objects in Figure 2.3. These fuzzy landscapes have high membership values for the points which are in the interior of the concavities of the reference object and lower ones for the points which are away from the concavities. Clearly, there is no region of space which satisfies the surrounded relation when the reference object is convex (Figure 2.5(b)).



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

Figure 2.5: Fuzzy landscapes obtained with Equation 2.10. The defect presented in Figure 2.3 no longer exists.

Similarly to the metric relations defined through fuzzy landscapes in Section 1.3.2, we can evaluate the degree of satisfaction to which an object is surrounded by B by using a necessity(N)/ possibility (Π) pair (see Equations 1.6 and 1.7) or a mean measure M (see Equation 1.16).

2.1.2.2 Considering the distance to target object

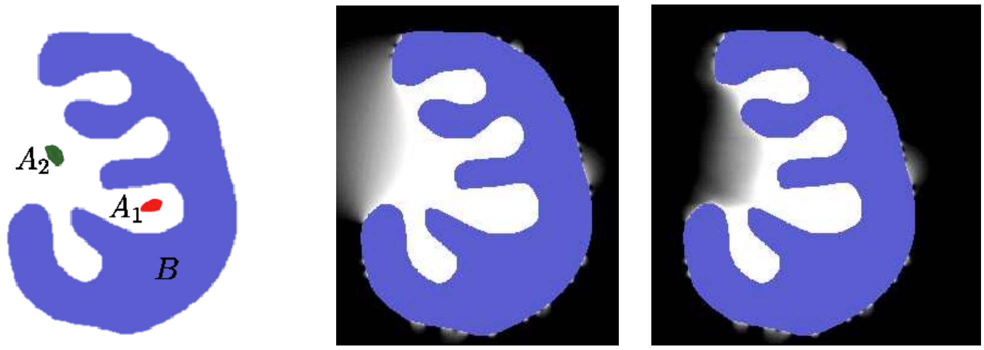
The fuzzy landscape defined by Equation 2.10 is adequate when the size of the target object is comparable to the size of the concavity. However, when this is not the case, the obtained fuzzy landscape is not appropriate. For instance, if B is the reference object in Figure 2.6(a), A_1 should have a high satisfaction degree for the relation surrounded, since there is a portion of the boundary of $\partial B \setminus \partial CH(B)$ which goes around the object and which is close to it. On the contrary, intuitively, the object A_2 should not have a high satisfaction degree of the relation surround. Even if there is a portion of the boundary of $\partial B \setminus \partial CH(B)$ which goes around the object, some of the points belonging to this portion of boundary are so "far" from A_2, that they are not considered visible from A_2. The interpretation of "far" and "near" according to an object can depend on its intrinsic characteristic, for instance on its size [Hernandez et al., 1995].

The second and third columns of Table 2.1 show the results of evaluating the relation "surround by" using Equation 2.10. Both the $[N, \Pi]$ and the M measures give similar results for A_2 and A_1, although, intuitively, A_1 should better satisfy the relation since it is more inside the concavity of B (hence more surrounded).

Target object	$\mu_{surround}(B)$ Equation 2.10		$\mu_{surround}(B, \mu_n)$ Equation 2.13	
	$[N, \Pi]$	M	$[N, \Pi]$	M
A1	[1.0,1.0]	1.0	[1.0,1.0]	1.0
A2	[0.73,1.0]	0.85	[0.21,0.64]	0.48

Table 2.1: Evaluation of the surround relation using Equations 2.10 and 2.13, for objects in Figure 2.6(a).

In order to include the "near" relation of the target object with respect to the reference



(a) (b) (c)

Figure 2.6: Illustration of the fuzzy landscapes when not considering the distance of the target object and when considering it. (a) The reference object is denoted by B, and two target objects by A_1 and A_2. (b) $\mu_{surround}(B)$ using Equation 2.10. (c) $\mu_{surround}(B, \mu_n)$ using Equation 2.13.

object, the measure of angular coverage should be modified. Let $p \in A$ and α an angle in $[0, 2\pi[$. Suppose that there exists $q \in \partial B \setminus \partial CH(B)$ which intersects the ray emanating from p in the direction α. However, if q is not "near" p according to A, then it should be considered as if there was no point in $\partial B \setminus \partial CH(B)$ in the direction α.

Let μ_n be a function over $\mathbb{R}^+$ which represents the "near" distance according to the target object. It can defined as a trapezoid function, where the parameters are adjusted according to the size of the target object. If we want to evaluate the relation for several target objects of comparable size, then we can define μ_n according to this size. As a result, we define the function $\tilde{r}_\theta(p, B, \mu_n)$ as:

$$\tilde{r}_\theta(p, B, \mu_n) = \begin{cases} \mu_n(d_E(p,q)) & \text{if } \exists q \in \partial B \setminus \partial CH(B) \text{ such that } \angle(\vec{pq}, \vec{u}_x) = \theta \\ & \text{ and } [p, q[\cap B = \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{2.11}$$

The function $\tilde{r}_\theta(p, B, \mu_n)$ is equal to the membership value of the "near" relation of the point in $\partial B \setminus \partial CH(B)$ which lies on the ray emanating from p in the direction θ. The point in $\partial B \setminus \partial CH(B)$, which satisfies this, is uniquely defined. Using $\tilde{r}_\theta(p, B, \mu_n)$ we can define the angular coverage which takes into account the size of the desired target object or objects, as well as considering the points of the convex hull of the reference object B:

$$\theta_{coverage\_CH}(B, \mu_n)(p) = \int_0^{2\pi} \tilde{r}_\theta(p, B, \mu_n) d\theta. \tag{2.12}$$

Finally, the degree of surround is given by:

$$\mu_{surround}(B, \mu_n)(p) = f(\theta_{coverage\_CH}(B, \mu_n)(p)). \tag{2.13}$$

Figure 2.6(c) shows the fuzzy landscape defined by Equation 2.13. For this experiment we modeled μ_n as a trapezoid function:

$$\mu_n(x) = \begin{cases} 1 & \text{if } x \leq d_1 \\ \frac{d_2 - x}{d_2 - d_1} & \text{if } d_1 < x \leq d_2 \\ 0 & \text{otherwise} \end{cases} \tag{2.14}$$

The parameters d_1 and d_2 are associated with the imprecision of defining the relation "near". For instance we can use a very restrictive function were:

$$d_1 = 4l_{average} \tag{2.15}$$
$$d_2 = 5l_{average} \tag{2.16}$$

where $l_{average}$ is equal to the average of the lengths of the maximum diameter. In this particular case, $l_{average} = 25.3$, and so $d_1 = 101.2$ and $d_2 = 126.5$. We have chosen these parameters just to illustrate the influence of considering the distance when evaluating the "surround" relation. However, for a real application these parameters can be learned. The fuzzy landscape of Figure 2.6(c) has high membership values in the regions where the concavities are small, and can go around an object of the same size as A_1 or A_2. The last two columns of Table 2.1 show the results of evaluating the satisfaction of the relation "surrounded by" defined by Equation 2.13. Both the mean and the necessity/possibility measures give results which fit with the intuition, indicating that A_1 completely satisfies the relation, while A_2 has a much lower satisfaction, as expected.

Attaching the fuzzy landscape of surround to the characteristics of the target object, such as size, might seem very restrictive, since one of the advantages of computing a fuzzy landscape is that one has to compute the landscape only once, and then it can be used to evaluate the relation for several objects. However, other advantages of the fuzzy landscape, such as determining the region of space where it is possible to find a particular target object that satisfies the relation, remain to be valid. Other applications, such as evaluating the relation for several objects of similar size can also be envisaged.

2.1.2.3 Extension to fuzzy objects

When B is a fuzzy object with membership μ_B, the function $\tilde{r}_\theta(p, B, \mu_n)$ and the notion of angular coverage should be changed. When B is a fuzzy set, their corresponding convex hulls are also a fuzzy set. The α-cuts of a fuzzy set μ are nested, as well as its convex hulls. So, the convex hull of a fuzzy set can be defined as the convex hull of its α-cuts [Bloch et al., 2006]:

$$(CH(\mu_B))_\alpha = CH((\mu_B)_\alpha) \tag{2.17}$$

Then the notion of angular coverage $\theta_{coverage\_CH}(B, \mu_n)(p)$ is directly extended by replacing $\partial B \setminus \partial CH(B)$ by the fuzzy set $\mu_{\partial B\setminus\partial CH(B)}$ defined as:

$$\mu_{\partial B\setminus\partial CH(B)}(p) = t\left[\mu_{\partial B}(p), c(\mu_{\partial CH(B)}(p))\right] \tag{2.18}$$

Where c denotes a fuzzy complementation and $\mu_{\partial B}$ represents the fuzzy boundary of μ_B, which can be computed as in [Bloch, 2006] by using mathematical morphology. For instance, the membership function of the internal boundary of a fuzzy set μ_X is defined as:

$$\mu_{\partial X}(p) = t[D_{V_c}(c(\mu_X))(p), \mu_X(x)] \tag{2.19}$$

where V_c is a fuzzy structuring element representing the notion of neighborhood, as the one used in Equation 1.11.

Now the function $\tilde{r}_\theta(p, B, \mu_n)$ is adapted for the case where B is fuzzy:

$$\tilde{r}_\theta(p, B, \mu_n) = \max_{\{q\in\Im|\angle(\vec{pq},\vec{u}_x)=\theta\}} t\left[\mu_n(d_E(p,q)), \mu_{\partial B\setminus\partial CH(B)}(q), \mu_{inclusion}([p,q[, B^c)\right] \tag{2.20}$$

where $\mu_{inclusion}$ corresponds to the degree of inclusion of Equation 1.5. Let $\Im$ represent the image domain, then for every $q \in \Im$, for which the segment $[p, q]$ makes an angle θ with the horizontal axis, we take the value which better satisfies conjunctively the three conditions used for the definition of this function in the crisp case. The first condition refers to the fact that p and q should be "near" according to the function μ_n. The second one establishes that q should belong to B and finally the third one states that the segment $[p, q[$ should be included in the fuzzy complement of B. When B is crisp, Equations 2.20 and 2.11 give the same result.

To extend the definition in Equation 2.10, where the distance to the target object is not taken into account, the function $\tilde{r}_\theta(p, B)$ is changed to:

$$\tilde{r}_\theta(p, B) = \max_{\{q \in \Im | \angle(\vec{pq}, \vec{u}_x) = \theta\}} t\left[\mu_{\partial B \setminus \partial CH(B)}(q), \mu_{inc}([p, q[, B^c)\right] \tag{2.21}$$

As for the aforementioned case, this function represents the conjunction of the two conditions which define $\tilde{r}_\theta(p, B)$ in the crisp case.

2.1.3 Properties

The fuzzy landscape obtained by the definition for surround that does not take into account the distance to the object (Equation 2.10) is particular case of the definition that takes into account the distance to the object (Equation 2.13). Indeed, for every $x \in \Im$ and μ_n, $\mu_{surround}(B)(x) \geq \mu_{surround}(B, \mu_n)(x)$ with equality when μ_n is the constant function equal to 1. It follows that all the properties that are satisfied by $\mu_{surround}(B, \mu_n)$, are also satisfied by $\mu_{surround}(B)$.

The proposed surround relation is invariant with respect to the geometrical transformations of translation and rotation. It is invariant with respect to scaling if and only if μ_n is invariant with respect to scaling.

The membership function of surround is increasing with respect to μ_n. For every μ_{n1}, μ_{n2} such that $\mu_{n1}(x) \leq \mu_{n2}(x)$ for every $x \in \Im$, then $\mu_{surround}(B, \mu_{n1})(x) \leq \mu_{surround}(B, \mu_{n2})(x)$.

2.1.4 Computational complexity

Assume that the image is a square with sides $\sqrt{N}$ with N the number of points. Let N_B be the number of points in the reference object and N_d the length of the core of μ_n. The complexity of computing the convex hull is $O(N_B \log N_B)$[1]. The complexity of computing the function $\tilde{r}_\theta(p, B)$ of Equation 2.9 is of the order of N^2. Therefore the complexity of computing $\mu_{surround}(B)$ defined by Equation 2.10 is $O(N_B \log N_B + N^2)$.

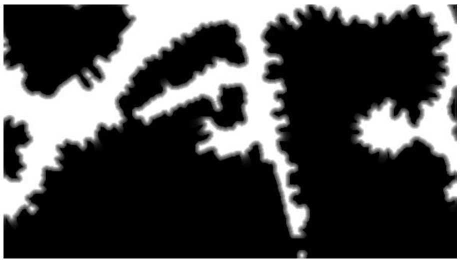
The complexity of computing $\mu_{surround}(B, \mu_n)$ defined by Equation 2.13 is $O(N_B \log N_B + N_d N)$, since the complexity of computing the function $\tilde{r}_\theta(p, B, \mu_n)$ is of the order of $N_d N$.

2.1.5 Illustrative example

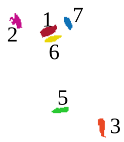
In this section we present an example to illustrate the defined relation. We compute the fuzzy landscapes for the relation surround using the sea, as reference object, shown in Fig. 2.7(b). This reference object is a fuzzy region, where the white region represents the points of space with a high membership value, while the black region represents the points with

[1] We assume that the q-hull algorithm is used for computing the convex hull [Barber et al., 1996].

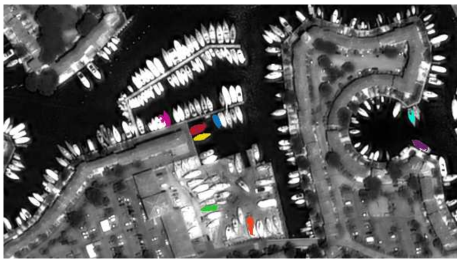
(a) Original image.



(b) Reference object: the sea white region.



(c) Target objects.



(d) Target objects superimposed on the original image.

Figure 2.7: Reference and target objects used to evaluate the relation surround.

low membership value. The target objects are the boats shown in Fig. 2.7(c). The sizes of the target objects are comparable, therefore the relation can be evaluated by using the same fuzzy landscape for all objects.

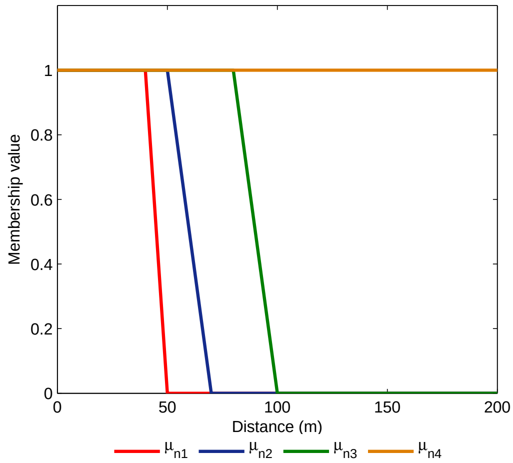
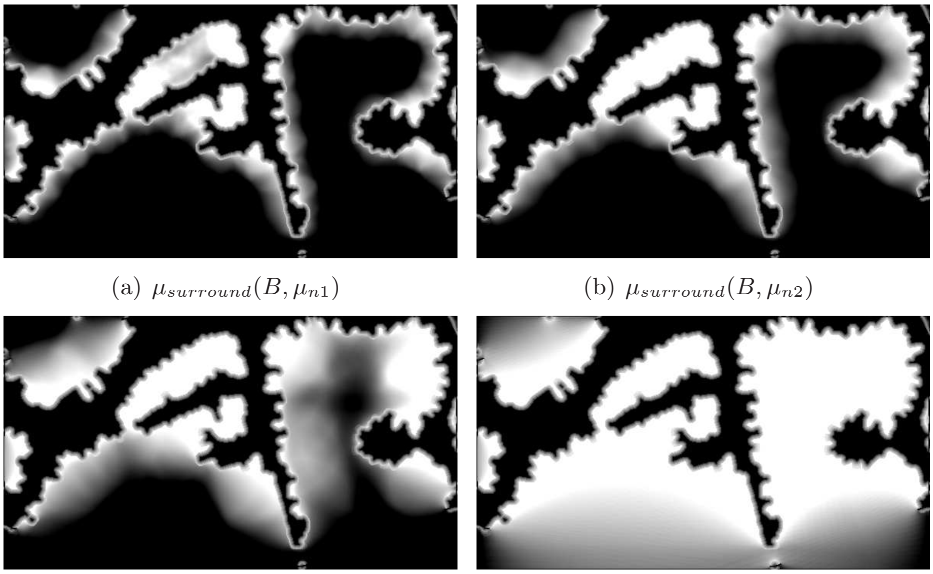


Figure 2.8: Membership functions for μ_n.

To observe the influence of the distance relation involved in the computation of $\mu_{surround}$, we used different membership functions to represent the "near" notion. The membership functions for the "near" functions are illustrated in Fig. 2.8. There are four membership functions representing different degrees of permissiveness of "near". The function μ_{n1} is very strict. The functions μ_{n2} and μ_{n3} are more in accordance with the size of the target objects, since the average lengths of the object's main directions is $l_{average} = 10.3$. Finally, μ_{n4} is a constant membership function equal to one, which gives a fuzzy landscape equivalent to the one in Equation 2.10.



(a) $\mu_{surround}(B, \mu_{n1})$ (b) $\mu_{surround}(B, \mu_{n2})$

(c) $\mu_{surround}(B, \mu_{n3})$ (d) $\mu_{surround}(B, \mu_{n4})$

Figure 2.9: Fuzzy landscapes representing the relation surround for different distance membership functions shown in Fig. 2.8. (a) Using μ_{n1}. (b) Using μ_{n2}. (c) Using μ_{n3}. (d) Using μ_{n4}.

The corresponding fuzzy landscapes are shown in Fig. 2.9 and the evaluation for the target objects is shown in Table 2.2. The fuzzy landscape of Fig. 2.9(a) is very restrictive, and it is not well adapted to evaluate the relation for the selected target objects. It represents a situation where the target objects should be very small, and the reference object must be almost touching them. This fuzzy landscape does not allow us to cope with the imprecision linked to the segmentation of the reference object. The restrictions imposed by the function "near" are not appropriate for the situation. For instance, in the evaluation of the yellow boat (6), the large interval $[N, \Pi]$ exhibits an ignorance with respect to the satisfaction of the relation. Even though there is no ambiguity that this boat is surrounded by the sea.

The second and third landscapes (Figs. 2.9(b) and 2.9(c)) are more suitable for evaluating the relation. We obtain high satisfaction degrees for all the boats, except the green (5) and the orange (3) ones, as expected. The satisfaction values obtained for this relation are in better accordance with the intuition. The intervals $[N, \Pi]$ are shorter, showing less ambiguity in the satisfaction of the relation for these target objects. Moreover, the fuzzy landscape of Fig. 2.9(b) corresponds to the μ_n obtained using the parameters of Equations 2.15 and 2.16. This suggests that these parameters are an appropriate choice.

The last fuzzy landscape (Fig.2.9(c)) is very permissive. According to the first property listed in Section 2.1.3, this fuzzy landscape corresponds to the case where the size of the targets is not taken into account. The results obtained for this situation are not really significant, since a high satisfaction degree M is obtained for all the target objects, while, intuitively, the green (5) and orange (3) boats should have a low satisfaction value.

In this example we showed the influence of the notion of "near" introduced in the

Target Objects Fig. 2.7(c)	$\mu_{surround}(B, \mu_{n1})$ Fig. 2.9(a)		$\mu_{surround}(B, \mu_{n2})$ Fig. 2.9(b)		$\mu_{surround}(B, \mu_{n3})$ Fig. 2.9(c)		$\mu_{surround}(B, \mu_{n4})$ Fig. 2.9(d)	
	$[N, \Pi]$	M	$[N, \Pi]$	M	$[N, \Pi]$	M	$[N, \Pi]$	M
1 (red)	$[0.4, 1]$	0.75	$[0.4, 1]$	0.91	$[0.4, 1]$	0.91	$[0.4, 1]$	0.91
2 (pink)	$[0.29, 0.83]$	0.54	$[0.5, 0.90]$	0.72	$[0.5, 1]$	0.94	$[0.5, 1]$	0.95
3 (orange)	$[0, 0]$	0.00	$[0, 0]$	0.00	$[0, 0.28]$	0.09	$[0.68, 0.90]$	0.81
4 (purple)	$[0.6, 1]$	0.98	$[0.6, 1]$	0.98	$[0.6, 1]$	0.98	$[0.6, 1]$	0.98
5 (green)	$[0, 0]$	0.00	$[0, 0]$	0.00	$[0, 0]$	0.00	$[0.82, 0.89]$	0.86
6 (yellow)	$[0.37, 1]$	0.78	$[0.7, 1]$	0.94	$[0.7, 1]$	0.96	$[0.7, 1]$	0.96
7 (blue)	$[0.6, 1]$	0.94	$[0.6, 1]$	0.94	$[0.6, 1]$	0.94	$[0.6, 1]$	0.94
8 (cyan)	$[0.7, 1]$	0.96	$[0.7, 1]$	0.96	$[0.7, 1]$	0.96	$[0.7, 1]$	0.96

Table 2.2: Results obtained for the target objects shown in Fig. 2.7(c) with respect to the reference objects of Fig. 2.7(b), using the membership function distance of Fig. 2.8.

definition of Equation 2.13. We demonstrated the limitations of using very strict or loose functions. The examples in Figs. 2.9(b) and 2.9(c) show fuzzy landscapes which are in accordance with the situation. These two situations exhibit similar results, showing that there continues to be a flexibility in the decision of the function used to represent the notion of "near". Moreover, this function could be learned as in [Colliot et al., 2006]. This would allow an automated use of this approach.

2.1.6 Discussion

In this section we have presented a fuzzy definition of the relation "surround". We first presented a definition that only considered the morphology of the reference objects. However, we saw that this definition is not adapted to evaluate the relation when the target objects are much smaller than the concavities and far of them. Therefore, we adapted this definition to take into account the shape of the reference object, as well as the size of the target objects and their distance to the concavities of the reference object. The proposed definition has nice geometric properties, it is in accordance with the intuition and is modeled as a fuzzy landscape which has shown interesting properties for image interpretation.

2.2 Alignment

Determining the groups of aligned objects is crucial for image interpretation. According to the Gestalt theory, the human perceptual vision system groups objects together using certain rules. Among these rules there is one called *continuity of direction* which groups together objects in the same direction, and one particular case is the constancy of direction that refers to alignments [Desolneux et al., 2008]. An aligned group of objects has the characteristic that it should be seen as a whole, since, if its elements are observed in an independent manner, then the alignment property is lost. However, having to look it as a whole makes alignment detection a difficult task.

Identifying the aligned groups of objects in satellite images is important for several applications. Satellite images provide a huge amount of geographical information, and aligned groups of objects can be seen as a way to reduce this information in a meaningful way. For example in cartography, it is necessary to find groups of aligned buildings for map generalization [Steiniger and Weibel, 2007]. Observing if a group of buildings is aligned can give information about the structure of their arrangement, and whether they belong to a urban, rural or residential area [Dogrusoz and Aksoy, 2007]. In object detection, complex semantic classes such as parking areas (car parking lots, ports, truck parkings lots or airports) comprise aligned groups of transport vehicles. Therefore, the identification of aligned groups of transport vehicles can be useful for detecting instantiations of these complex classes, and is meaningful for the description of this kind of scenes.

This section is organized as follows. Section 2.2.1 reviews some of the models of alignment. In Section 2.2.2, we discuss some properties our model should satisfy. Section 2.2.3 introduces the definitions of *local* alignment and *global* alignment. A method for extracting the *locally* aligned groups is then proposed. In this method we construct a neighborhood graph of the objects of the image, and its dual graph where we incorporate information about the relative direction of the objects, evaluated using fuzzy measures of relative position. The groups of objects satisfying the fuzzy criterion of being *locally* aligned are extracted from the dual graph. These groups are the candidates for being *globally* aligned. In Section 2.2.4 we discuss the behavior of the method with respect to segmentation errors. In Section 2.2.5 we discuss its computational complexity. Part of the work presented in this section was presented in [Vanegas et al., 2010a,b].

2.2.1 Related work

Alignment between low level features has been widely studied in computer vision. For instance, alignment between groups of points [Desolneux et al., 2008, Ortner et al., 2007] and alignments between linear segments [Lowe, 1987, Desolneux et al., 2008] as collinearity on digital images. Most of these works are inscribed within the framework of perceptual organization. Their first objective is to find how to organize low level features, such as edge segments, into groups, such as aligned segments. The groups are evaluated according to their perceptual significance based on the grouping laws of the Gestalt theory [Wertheimer, 1938, Koffka, 1935]. Their second objective is to differentiate the groupings that arise from the structure of a scene from those that arise due to accidents of view point or position [Lowe, 1987]. The objective of perceptual organization in computer vision and our objective are different. Perceptual organization deals with how these relations take place among low level features and their meaningfulness according to the image's structure, while we are interested in defining these relations for objects in the image taking into account their semantics. Nevertheless, it is important to review the work that has been done on studying these relations as grouping laws between low level features in digital images, in order to understand how these relations have been modeled.

Several methods to determine the alignment between points methods relying on the Hough transform [Desolneux et al., 2008] or the Radon transform [Likforman-Sulem and Faure, 1994] have been proposed. Other examples are the identification of aligned segments which have the same orientations as the alignment [Desolneux et al., 2008, Ortner et al., 2007, Ralescu and Shanahan, 1999, Lowe, 1987, Kang and Walker, 1994]. However, alignment extraction as a high level feature has been less studied. One example is the work of [Christophe and Ruas, 2002], where an algorithm to detect aligned groups of buildings in maps is presented. In this algorithm buildings with aligned barycenters are extracted, and the quality of the alignments is evaluated based on the criteria of proximity and similarity laws of Gestalt theory.

The above mentioned methods for determining alignment focused on extracting groups of objects with aligned barycenters. When these methods are applied to a group of heterogeneous objects (according to shape and size) we obtain an alignment of barycenters which does not correspond to the perceived alignment of the group, since neither the size nor the shape of the objects are considered. An example of this is given in Section 2.2.2.1. Also, the previously proposed alignment relations have not been extended to deal with fuzzy objects.

Alignment in computer vision has been studied to organize low level features but it has not been combined with other spatial relations. However, when dealing with objects and observing these relations as spatial relations, it is interesting to combine alignment with other relations, for instance to determine when two groups of aligned objects are parallel, or when a group of aligned objects is parallel to another object. These kinds of combinations give us more information about the scene and can be meaningful for the description. This point will be further developed in Section 2.3.

2.2.2 Considerations for modeling alignment

Alignment can be defined as the spatial property possessed by a group of objects arranged in a straight line[2]. As it was highlighted by the Gestalt theory, alignment should be seen as a whole [Desolneux et al., 2008]: if we observe each element of the group individually, then the alignment property is lost. Having to look at it as a whole makes alignment detection a difficult task, since in order to detect an aligned group of objects we have to identify its members, but to know if an object belongs to an aligned group the alignment has to be already identified. In the following we study the different approaches to model alignment, and introduce the notions of *local* and *global* alignment.

2.2.2.1 Different approaches to model alignment

For the case of a group of points in $\mathbb{R}^2$ there are two equivalent strategies for verifying whether they are aligned. Let $\mathcal{A} = \{a_1, \ldots, a_n\}$ be a group of points in $\mathbb{R}^2$. $\mathcal{A}$ is aligned if and only if:

(i) there exists a line L that intersects all the points, or

(ii) there exists an angle $\theta \in [0, \pi)$ such that for every pair of points a_i and a_j $(i \neq j)$, a_j is located in direction θ or $\theta + \pi$ from a_i with respect to the horizontal axis.

The first strategy is a linear regression problem and it has been used to identify if a group of points is aligned in an image, by considering that points should fall into a strip [Desolneux et al., 2003], the thinner the strip the better the alignment. Extending the first definition to identify a group of aligned objects can be done by using objects' barycenters. Unfortunately this will only be appropriate for objects of similar sizes (see Figure 2.10) and approximately convex. Another possibility is to search a thin strip where all the objects fall into, but the width of the strip will depend on the objects' sizes. Thus the notion of falling into a thin strip is not appropriate for objects with different sizes.

The difficulty of extending the second strategy comes from determining the angle between two objects. One alternative, which is the one proposed in this work, is to use measures of relative position used in spatial reasoning. Before entering into the details of the method, we introduce the notion of *orientation histogram*, inspired by the angle histogram (see Section 1.3.2.1). This notion is a fundamental concept for our method. In the following sections we present two definitions of alignment and an algorithm to extract

[2]Definition taken from ThinkMap Visual Thesaurus http://www.visualthesaurus.com/

(a) Original image. (b) Segmented objects. (c) Barycenters of objects in (b).

Figure 2.10: Example of an aligned group of objects of different sizes and with non-aligned barycenters. However, this objects are aligned according to their lower boundaries

these alignments which is flexible enough to extract groups of aligned object which are not only aligned by their barycenters.

2.2.2.2 Orientation histograms

To measure the orientation between two objects, we define the orientation histogram, which is simply an angle histogram where the angles are computed modulo π and its support has a length equal to π. The orientation histogram between two fuzzy objects A and B, with membership function μ_A and μ_B is given by:

$$O(A,B)(\theta) = \frac{\sum_{p,q\in\Im | mod(\angle(\vec{pq},\vec{u}_x),\pi)=\theta} \mu_A(p) \wedge \mu_B(q)}{\max_{\phi\in[0,\pi)} \sum_{p,q\in\Im | mod(\angle(\vec{pq},\vec{u}_x),\pi)=\phi} \mu_A(p) \wedge \mu_B(q)}, \tag{2.22}$$

where $\angle(\vec{pq},\vec{u}_x)$ is the angle between the vector joining p and q, and the vector $\vec{u}_x$, which represents the orientation of the x-axis. The orientation histogram is a fuzzy subset of $[0,\pi[$ that represents the orientation between two objects. It preserves the same properties as the angle histogram and, in addition, is symmetrical.

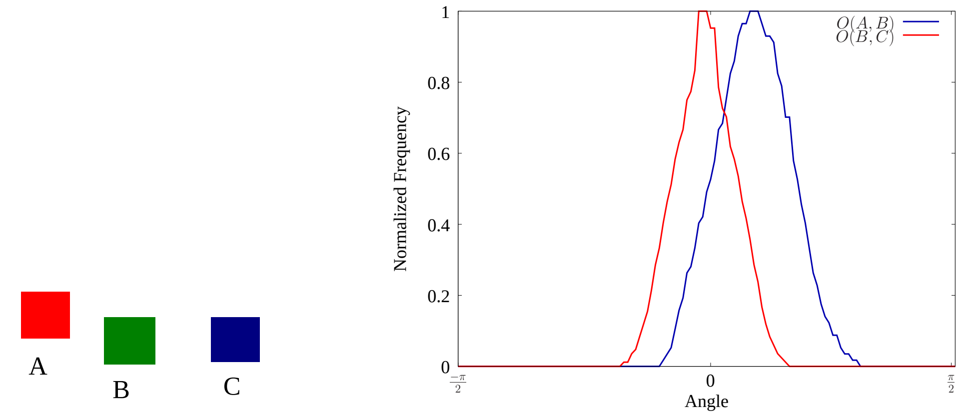
Similarity degree for orientation histograms There are several ways to define a degree of similarity between orientation histograms. One possibility is to interpret the orientation histograms as fuzzy sets and use similarity measures based on aggregation [Zwick et al., 1987]. Another possibility is to interpret the orientation histogram as a function and use a similarity degree based on distance between functions. We use here a simple similarity degree based on intersection, but other similarity measures are possible. More information on distances between histograms and similarity measures between fuzzy sets can be found in [Zwick et al., 1987, Bloch, 1999, Rabin et al., 2008].

A degree of similarity between two orientation histograms can be defined as the maximum height of the fuzzy intersection of the two orientation histograms [Zwick et al., 1987]. Let $O(A,B)$ and $O(C,D)$ be two orientation histograms. The degree of similarity between them is given by:

$$sim(O(A,B),O(C,D)) = \max_{\theta\in[0,\pi)} \left[O(A,B)(\theta) \wedge O(C,D)(\theta)\right]. \tag{2.23}$$

In this measure, the possibility for each angle is aggregated in a conjunctive way, and we take its maximum value. In the case were the orientation histograms do not intersect then the similarity value is 0.

Figure 2.11(b) shows the two orientation histograms computed for the objects of Figure 2.11(a). Both histograms have a well defined maximum, showing a strong main orientation,



(a) (b)

Figure 2.11: (a) Objects. (b) Orientation histograms of the objects in (a).

and these maximum values are close to each other. Therefore a high similarity value is expected. However, if the *min* operator is used as a t-norm in Equation 2.23, the similarity value between the two histograms is 0.67. This non-high value is due to the comparison of the value for each angle separately, since the aggregation is done for every angle. However, in this context referring to an orientation equal to an angle θ actually represents the quantity "approximately θ". Therefore, in order to compare if two orientation histograms are similar, it is important to consider the imprecision that is linked to the comparison of two angles that are approximately the same. When a fuzzy morphological dilation is performed on an orientation histogram using a structuring element ν_0, then the high values of the histogram will be propagated to the similar angle values according to ν_0. The structuring element ν_0 is designed such that $\nu_0(\theta - \tilde{\theta})$ represents the degree to which $\tilde{\theta}$ and θ are "approximately" equal, and in our experiments we modeled ν_0 as a trapezoid function:

$$\nu_0(\theta) = \begin{cases} 1 & \text{if } |\theta| \leq t_1 \\ \frac{t_2 - |\theta|}{t_2 - t_1} & \text{if } t_1 < |\theta| \leq t_2 \\ 0 & \text{if } |\theta| > t_2, \end{cases} \tag{2.24}$$

where t_1 and t_2 represent the parameters of a trapezoid function. The parameters t_1 and t_2 are related to the imprecision linked to computing angles in a discrete grid, as is the case for images. This imprecision is a function of the distance between the points for which the angle is computed: the greater the distance between the points, the more precise is the calculation. However, in our experiments, since most of the time we are dealing with objects which have a similar distance between them, we have chosen to use:

$$t_1 = \arcsin\left(\frac{1}{0.5 d_{average}}\right) \tag{2.25}$$

$$t_2 = \arcsin\left(\frac{1}{0.25 d_{average}}\right) \tag{2.26}$$

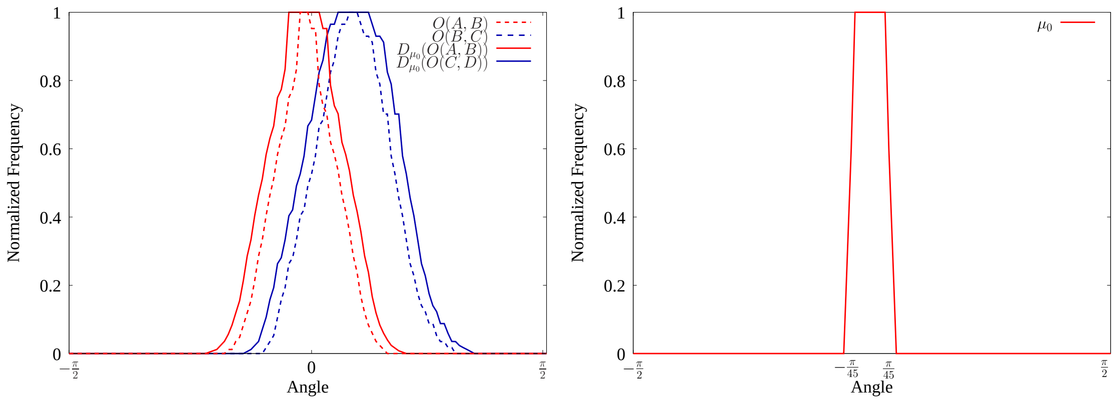
where $d_{average}$ is the average distance in pixels between the barycenters of "neighboring" objects. This values have been used in all experiments, and did not require any fine tuning of the parameters.

By performing a morphological dilation of $O(A,B)$ with the structuring element ν_0 we introduce the imprecision to all the angles in $O(A,B)$ (as in [Bloch et al., 1996]). For this reason, the similarity should be computed using the dilated orientation histograms, and Equation 2.23 becomes:

$$sim(O(A,B),O(C,D)) = \max_{\theta\in[0,\pi)} \left[D_{\nu_0}(O(A,B))(\theta) \wedge D_{\nu_0}(O(C,D))(\theta)\right], \tag{2.27}$$

where $D_{\nu_0}(O(X,Y))$ is the dilation of $O(X,Y)$ by a structuring element ν_0.

Figure 2.12(a) shows the result of performing a morphological dilation of the orientation histograms of Figure 2.11(b). The parameters of Equations 2.25 and 2.26 are $t_1 \approx 0.07$ and $t_2 \approx 0.12$ for the structuring element. The similarity value between the histograms obtained with Equation 2.27 is 0.93 which is more consistent with the perceived orientation of the objects than the value obtained by applying Equation 2.23.



(a) (b)

Figure 2.12: (a) Orientation histograms of the objects in Figure 2.11(a) and dilated orientation histograms. (b) Structuring element used for the dilation.

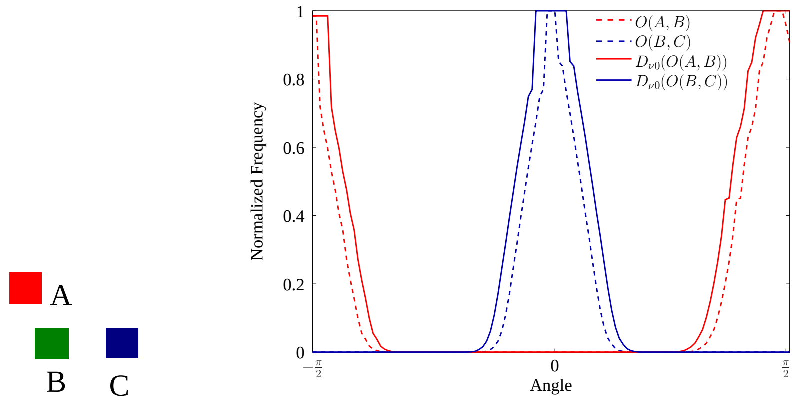
When orientation histograms are not similar (see Figure 2.13) we obtain a zero similarity value, as desired.

This measure of similarity can be extended to compare several orientation histograms. Let $\{O(A_i,A_j)\}_{i=0,j\neq i}^{N}$ be a set of orientation histograms, the similarity degree between them is equal to:

$$sim(O(A_0,A_1),\ldots,O(A_i,A_j),\ldots O(A_N,A_{N-1})) = \max_{\phi\in[0,\pi[} \bigwedge_{i=0,j\neq i}^{N} D_{\nu_0}(O(A_i,A_j))(\phi) \tag{2.28}$$

2.2.3 Definition and identification of aligned groups of objects

We consider a set of objects $\mathcal{A} = \{A_0,\ldots,A_n\}$ and want to determine the subsets of aligned groups of $\mathcal{A}$. For the sake of clarity we will call them *globally* aligned groups. First we identify the *locally* aligned groups (defined in Section 2.2.3.2), and the set of *locally* aligned groups will be denoted by $\mathfrak{L}$. For each *locally* aligned group $\mathcal{S}_i \in \mathfrak{L}$ we measure the



(a) (b)

Figure 2.13: (a) Objects. (b) Orientation histograms of the objects (a) and dilated orientation histograms.

degree of *global* alignment (defined in Section 2.2.3.1). If this degree is lower than a user acceptance value α, then elements of the group will be deleted until the degree is equal or greater than α or until the group has less than 3 elements. In the case where the degree is greater than α, the group $\mathcal{S}_i$ will be added to the set of aligned groups $\mathfrak{G}$. A last step of addition and fusion of the groups is done to obtain the largest possible *globally* aligned groups. This method is illustrated in Figure 2.14. In the following sections we detail each step of the method.

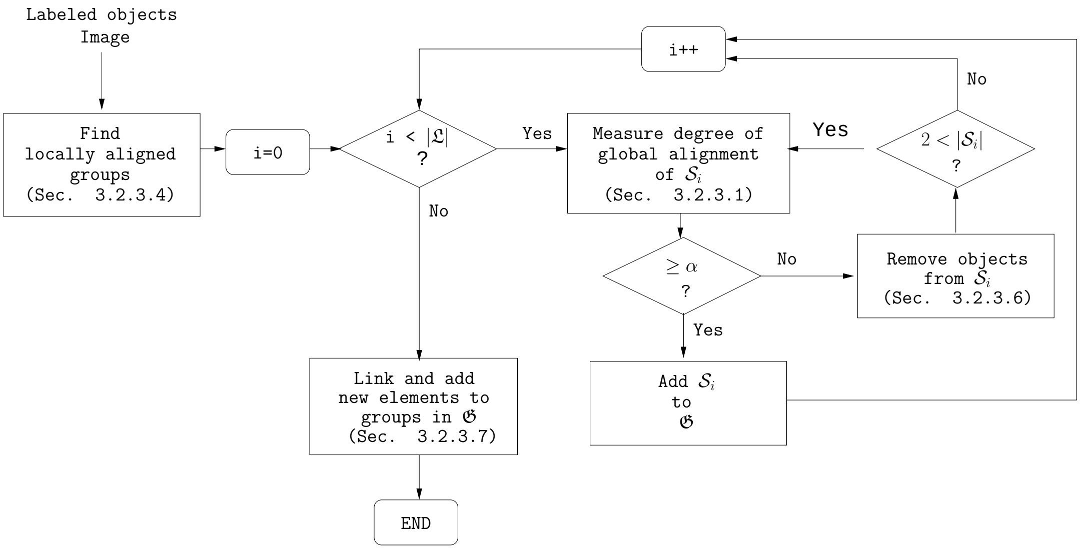


Figure 2.14: Proposed method for determining aligned groups of objects.

2.2.3.1 Globally aligned groups

Before defining a *globally* aligned group of objects, we define two preliminary concepts. Let $\mathcal{S}$ be a group of objects, and $A, B \in \mathcal{S}$, we define the $Neigh(A, B)$ relation as being satisfied if and only if $B \cap A = \emptyset$ and $B \cap N(A) \neq \emptyset$ where $N(A)$ is defined as a neighborhood of A. One possible choice for the neighborhood of an object is $N(A) = N_d(A)$ where $N_d(A)$ is the Voronoi neighborhood constrained by a distance d. Other possible choices are discussed in Section 2.2.3.3.

Definition 2.1. *A group $\mathcal{S}$ is called connected by the $Neigh$ relation if for every $A, B \in \mathcal{S}$, there exist $C_0, \ldots C_M$ objects in $\mathcal{S}$, such that $C_0 = A$, $C_M = B$ and for every $m = 0, \ldots, M-1$, the relation $Neigh(C_m, C_{m+1})$ is satisfied.*

Returning to the discussion of Section 2.2.2.1, the group $\mathcal{S}$ is *globally* aligned if the following conditions are satisfied:

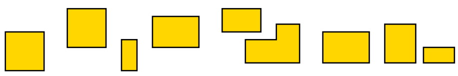
(i) $\mathcal{S}$ is connected by the $Neigh$ relation,

(ii) $|\mathcal{S}| \geq 3$, and

(iii) there exists $\theta \in [0, \pi[$ such that for every $A, B \in \mathcal{S}$, A is able to "see" B in direction θ or $\theta + \pi$ with respect to the horizontal axis.

The first condition ensures that the group is not "divided", for instance the red objects in Figure 2.15(a) do not satisfy this condition and can be considered as two groups. The second condition states that an aligned group should have at least 3 elements. To verify the third condition it would be necessary to compare all the orientation histograms between any two objects of $\mathcal{S}$. Unfortunately, this measure is very restrictive, and a more flexible measure is to consider that all the orientation histograms of $O(A_i, \mathcal{S} \setminus \{A_i\})$ are similar for all $A_i \in \mathcal{S}$. Figure 2.16 shows the dilated orientation histograms $D_{\nu_0}(O(A_i, \mathcal{S} \setminus \{A_i\}))$ and the dilated orientation histograms between every pair of objects of the group in Figure 2.15(b). For each orientation histogram we used the same structuring element ν_0. We can notice that for the dilated histograms $D_{\nu_0}(O(A_i, \mathcal{S} \setminus \{A_i\}))$ it is possible to see a tendency towards a similar angle, while for the dilated orientation histograms $D_{\nu_0}(O(A_i, A_j))$ this is not the case. This is reflected when histograms are aggregated using the Lukasiewicz t-norm in Figure 2.16(c) where the aggregation of $D_{\nu_0}(O(A_i, \mathcal{S} \setminus \{A_i\}))$ results in a function with a maximum of 0.81, while the aggregation of $D_{\nu_0}(O(A_i, A_j))$ produces a constant function equal to zero, which is not meaningful here. One should notice that when using the conjunction of the dilated orientation histograms $D_{\nu_0}(O(A_i, A_j))$, if two pairs of objects do not have a similar orientation then the whole conjunction is equal to zero. However, when using the dilated histograms $D_{\nu_0}(O(A_i, \mathcal{S} \setminus \{A_i\}))$ the dissimilarity between the orientations of two pairs of objects will not affect the whole conjunction, since it is a comparison between the orientations of the whole group with respect to its members.

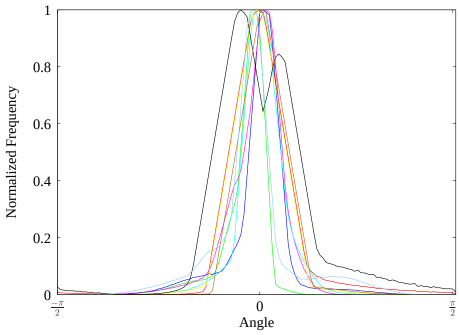
(a) Groups should be connected, the red group should be considered as two groups.



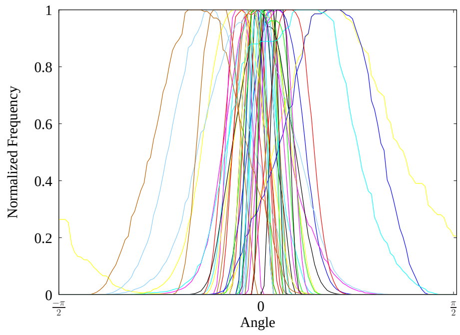
(b) An object of the group should see the other members in the alignment orientation.

Figure 2.15: Considerations for an aligned group.

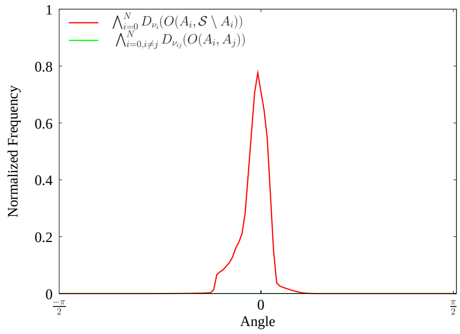
Thus, it is possible to define the degree of *global* alignment as follows:



(a) $D_{\nu_0}(O(A_i, \mathcal{S} \setminus \{A_i\}))$



(b) $D_{\nu_0}(O(A_i, A_j))$



(c) Aggregated histograms

Figure 2.16: Dilated orientation histograms for objects of Figure 2.15(b) and their aggregation using Lukasiewicz t-norm.

Definition 2.2. *Let $\mathcal{S} = \{A_0, \dots, A_N\}$, with $N \geq 3$, be a group of objects in $\mathfrak{I}$, connected by the Neigh relation. Then, the degree of global alignment of $\mathcal{S}$ is given by:*

$$\mu_{ALIG}(\mathcal{S}) = sim\left(O(A_0, \mathcal{S} \setminus \{A_0\}), \dots, O(A_N, \mathcal{S} \setminus \{A_N\})\right). \tag{2.29}$$

Notice that the measure of *global* alignment presented above is independent of the order since the function used to measure the similarity between several orientation histograms (Equation 2.28) is symmetric.

2.2.3.2 Locally aligned groups

We can say that a group $\mathcal{S}$ is *locally* aligned if it satisfies:

(i) for every $A \in \mathcal{S}$ the elements in the neighborhood $N(A)$ in $\mathcal{S}$ are aligned, and

(ii) it is connected by the $Neigh$ relation.

For the first condition, we will only verify for simplicity that for every pair of elements $B, C \in \mathcal{S}$ belonging to $N(A)$, the orientations $O(A, B)$ and $O(A, C)$ are similar. Thus, we can define the degree of *locally* alignment as follows:

Definition 2.3. *Let $\mathcal{S} = \{A_0, \dots, A_N\}$, with $N \geq 3$, be a group of objects in $\mathfrak{I}$, connected by the Neigh relation. The degree of locally alignment of $\mathcal{S}$ is given by:*

$$\mu_{LA}(\mathcal{S}) = \min_{X,Y,Z:Neigh(X,Y) \wedge Neigh(Y,Z)} sim(O(X,Y), O(Y,Z)). \tag{2.30}$$

We will say that a group of objects $\mathcal{S}$ is *locally* aligned to a degree β if $\mu_{LA}(\mathcal{S}) \geq \beta$. The preceding definition can be summarized by saying that a group $\mathcal{S}$ with $|\mathcal{S}| \geq 3$ is *locally* aligned to a degree β if it satisfies the following relations:

$$R1 : \forall X, Y, Z \ \ (Neigh(X, Y) \wedge Neigh(Y, Z)) \Rightarrow (sim(O(X, Y), O(Y, Z)) \geq \beta) \tag{2.31}$$

$$R2 : \forall A, B, \ \exists X_0, \ldots, X_m \text{ such that } X_0 = A, X_m = B \text{ and } \left(\bigwedge_{i=0}^{m-1} Neigh(X_i, X_{i+1}) \right) \tag{2.32}$$

2.2.3.3 Other neighborhood choices

The choice $N_d(A)$ in Section 2.2.3.1 as the Voronoi neighborhood of A constrained by a distance d was motivated by the fact that in a group, any subgroup of aligned objects should be "consecutive" (see Figure 2.17), and that successive objects should be close to each other. Nevertheless, there are other possibilities for choosing the neighborhood $N(A)$. For instance, we can consider the neighborhood of objects that are "near" A using the notion of a distance "less than d" (Equation 1.24). For a fixed d the neighborhood $N(A)$ is also

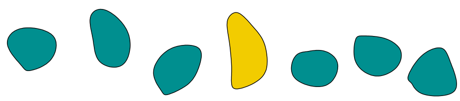


Figure 2.17: Illustration of a non-consecutive group. In a consecutive group their should be a succession of elements, which is not the case for the blue group, since the yellow object breaks the succession.

univocally defined. This choice gives a more restrictive condition of *locally* alignment since the neighborhood contains more objects that should verify the alignment conditions, and it can also be used in the case where A and B are fuzzy objects. In the following of this subsection we assume that we are dealing with fuzzy objects A and B defined through their membership functions μ_A and μ_B, since all equations that will be presented are valid for crisp and fuzzy objects. When considering a fuzzy neighborhood, the $Neigh$ relation becomes also fuzzy, and we denote by $\mu_{Neigh}(A, B)$ its degree of satisfaction. Similarly to the degree of adjacency (Equation 1.11), the degree of μ_{Neigh} is defined as a conjunction of the degree of intersection between μ_A and $\mu_{N(B)}$, and μ_B and $\mu_{N(A)}$. Therefore, the degree of neighborhood is defined by:

$$\mu_{Neigh}(A, B) = \mu_{\neg int}(\mu_A, \mu_B) \wedge \mu_{int}(\mu_{N(A)}, \mu_B) \wedge \mu_{int}(\mu_{N(B)}, \mu_A) \tag{2.33}$$

where μ_{int} and $\mu_{\neg int}$ are a degree of intersection and non-intersection (Equations 1.2 and 1.4. To have a symmetrical relation we consider the intersection between μ_A and $\mu_{N(B)}$, and between μ_B and $\mu_{N(A)}$.
Consequently, the relation "connected by $Neigh$" also becomes a fuzzy relation denoted by μ_{conn} and the degree of connectedness in a group $\mathcal{S}$ between two objects A, B in $\mathcal{S}$ is defined as in [Rosenfeld, 1979] by:

$$\mu_{conn}(A, B) = \max_{\mathcal{P} \in \mathfrak{P}_{AB}} \left[\min_{1 \leq i \leq l_p} \mu_{Neigh}(C_{i-1}^{(p)}, C_i^{(p)}) \right] \tag{2.34}$$

where $\mathcal{P}$ denotes a list of objects $\langle C_0^{(p)} = A, C_1^{(p)}, \ldots, C_{l_p}^{(p)} = B \rangle$ in $\mathcal{S}$, called path, and $\mathfrak{P}_{AB}$ is the set of all the paths from A to B in $\mathcal{S}$. The degree of connectedness of a group can be defined as the minimum degree of connectedness between its elements:

$$\mu_{conn}(\mathcal{S}) = \min_{A,B \in \mathcal{S}} \mu_{conn}(A, B) \tag{2.35}$$

When using a fuzzy neighborhood, the definitions of *globally* aligned and *locally* aligned have to be revised. The degree of *global* alignment of a group of objects $\mathcal{S} = \{A_0, \dots, A_N\}$ and with $N \geq 3$ becomes:

$$\mu_{ALIG}(\mathcal{S}) = \mu_{conn}(\mathcal{S}) \wedge sim\left(O(A_0, \mathcal{S} \setminus \{A_0\}), \dots, O(A_N, \mathcal{S} \setminus \{A_N\})\right) \tag{2.36}$$

This definition represents a conjunctive combination of the condition of being connected by the neighborhood relation and the similarity among the orientation histograms. When using a crisp neighborhood this definition is equivalent to the one given in Definition 2.1, where the condition of satisfying the relation of connection by the neighborhood is implicit in the definition.
In a similar way, it is possible to extend the definition of *local* alignment to:

$$\mu_{LA}(\mathcal{S}) = \mu_{conn}(\mathcal{S}) \wedge \left[\min_{X,Y,Z} \left(sim(O(X,Y), O(Y,Z)) \wedge \mu_{conn}(\{X, Y, Z\})\right)\right]. \tag{2.37}$$

Again, the equation represents the conjunction of combining two conditions, the first one is the condition of being connected and the second one represents that objects X and Z belong to the neighborhood of Y, and that the orientation histograms $O(X, Y)$ and $O(Y, Z)$ should be similar. In the case where a crisp neighborhood is used, we obtain the same degree as in Definition 2.3.

2.2.3.4 Identification of *locally* aligned groups

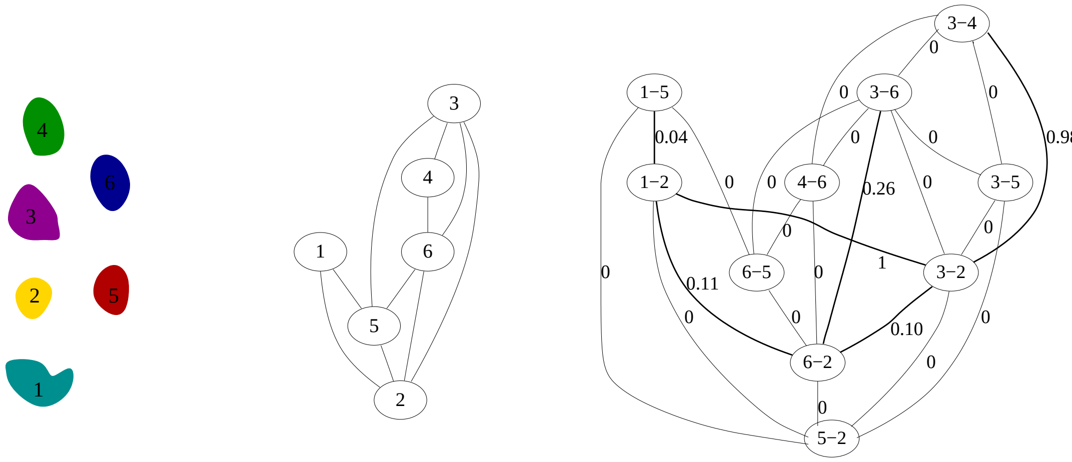
In this section we explain how it is possible to extract the *locally* aligned subgroups from a group of objects. For clarity purposes we first use a crisp neighborhood, and in the next section we explain how the algorithm is extended to cope with fuzzy neighborhoods.
As discussed in Section 2.2.3.2, the notion of *local* alignment strongly depends on the notion of neighborhood, since an aligned group should be connected by the $Neigh$ relation. Therefore, we propose to construct a neighborhood graph G_N to obtain the information of which objects are connected via the $Neigh$ relation. In a neighborhood graph $G_N = (\mathcal{V}, E)$ the vertices represent the objects of the group, and there is an edge between two vertices if and only if the corresponding objects are neighbors. Notice that only the connected subsets of three vertices X, Y and Z in G_N which share a common vertex, for example Y, satisfy:

$$Neigh(X, Y) \wedge Neigh(Y, Z) \tag{2.38}$$

These connected subsets are called *triplets*. According to $R1$, only the *triplets* $\{X, Y, Z\}$ that satisfy (2.27):

$$sim(O(X, Y), O(Y, Z)) \geq \beta \tag{2.39}$$

are aligned and can belong to a group which is *locally* aligned to a degree β. *Triplets* can be easily identified as the edges of the dual graph, when the dual graph is constructed in the following manner. The dual graph is denoted by $\tilde{G}_N = \{\tilde{\mathcal{V}}, \tilde{E}\}$ where each vertex $\tilde{V}_i$ represents an edge in the graph G_N. An edge exists between two vertices $\tilde{V}_i$ and $\tilde{V}_j$ of $\tilde{G}_N$ if the two corresponding edges of the graph G_N have a common vertex. If, additionally, we attribute to each edge (i, j) the similarity degree between the orientation histograms of $\tilde{V}_i$ and $\tilde{V}_j$ that we denote by $\tilde{s}_{ij}$, then it is possible to verify if relation $R1$ holds for its corresponding *triplet*. Figure 2.18 shows an example of a neighborhood graph and its dual graph. Notice that the edges of $\tilde{G}_N$ with a high value represent the *triplets* of objects with a similar orientation histograms. For instance, in the dual graph the edge between the nodes (1 - 2) and (2 - 3) has a similarity value of 1, this edge corresponds to the objects labeled 1, 2 and 3 of Figure 2.18(a). In a similar way, edges with a low value represent objects which are not aligned, for example in the dual graph the edge between the nodes (1 - 2) and (6 - 2) has a similarity value of 0.11 and corresponds to the objects labeled 1, 2 and 6, which do not form a *globally* aligned *triplet*.



(a) Labeled objects. (b) Neighborhood graph of (a). (c) Dual graph of (b).

Figure 2.18: Neighborhood graph and dual graph of a group of objects.

Returning to the conditions of *local* alignment $R1$ (2.31) and $R2$ (2.32), the first one states that *triplets* should be *globally* aligned, and the second one that the group should be formed by connected objects according to the $Neigh$ relation. Then a group $\mathcal{S}$ satisfies these relations if and only if the subset $\tilde{\mathcal{S}} \subseteq \tilde{\mathcal{V}}$ which represents the dual of $\mathcal{S}$ satisfies the following relations:

$$\tilde{R}1 : \forall \tilde{V}_i, \tilde{V}_j \; Connected(\tilde{V}_i, \tilde{V}_j) \Rightarrow (\tilde{s}_{ij} \geq \beta) \tag{2.40}$$

$$\tilde{R}2 : \forall \tilde{V}_i, \tilde{V}_j \; \exists \tilde{U}_0, \ldots \tilde{U}_K \text{ for } K \geq 1 \text{ such that } \tilde{U}_0 = \tilde{V}_i, \tilde{U}_N = \tilde{V}_j \text{ and } \bigwedge_{k=0}^{K-1} Connected(\tilde{U}_0, \tilde{U}_k), \tag{2.41}$$

where $Connected(\tilde{U}, \tilde{V})$ is true if there exists an edge between $\tilde{U}$ and $\tilde{V}$. Condition $\tilde{R}2$ expresses that $\tilde{\mathcal{S}}$ should be connected, since if $\tilde{\mathcal{S}}$ is not connected then $\mathcal{S}$ is not connected. Therefore, a *locally* aligned group is a subset $\mathcal{S} \subseteq \mathcal{V}$ for which its dual set $\tilde{\mathcal{S}} \subseteq \tilde{\mathcal{V}}$ is connected in $\tilde{G}$ and the value of all the edges joining the vertices within $\tilde{\mathcal{S}}$ is greater than or equal to β.

Algorithm 1 can be used to extract the $\tilde{\mathcal{S}}_i \subset \tilde{\mathcal{V}}$ corresponding to the dual sets of the *locally* aligned sets $\mathcal{S}_i \subset \mathcal{V}$. First the connected components of a graph $\tilde{G}_{TH}$ are computed and stored in $\mathfrak{C}$. $\tilde{G}_{TH}$ is a non-attributed graph containing the same vertices as $\tilde{G}$ and there is an edge between two vertices if the edge in $\tilde{G}$ has a degree greater than or equal to β. Then, for each component $\mathcal{C}_k$ we obtain the minimum value of its edges in $\tilde{G}$ that we call consistency degree of C_k and is denoted by $cons(\mathcal{C}_k)$:

$$cons(\mathcal{C}_k) = \min\{\tilde{s}_{ij} | \tilde{V}_i, \tilde{V}_j \in \mathcal{C}_k\}$$

If $cons(\mathcal{C}_k) < \beta$ then $\mathcal{C}_k$ does not satisfy $\tilde{R}1$, thus vertices are removed until $cons(\mathcal{C}_k) \geq \beta$. If in the process of vertex removal $\mathcal{C}_k$ becomes disconnected, then each of the connected components of $\mathcal{C}_k$ is treated separately. The vertices which are removed are the ones having more conflict with their neighbors in $\mathcal{C}_k$. We say that two connected vertices $\tilde{V}_i$ and $\tilde{V}_j$ are in conflict if $\tilde{s}_{ij}$ is close to zero, that is if the corresponding orientation histograms of both vertices are not similar. The conflict of a vertex $\tilde{V}_t$ with its neighbors in $\mathcal{C}_k$ is measured by using what we call the degree of the vertex in $\mathcal{C}_k$:

$$deg(\tilde{V}_t) = \frac{\sum_{\tilde{V}_j \in \mathcal{C}_k} \tilde{s}_{tj}}{|\{(i,j) | \tilde{V}_j \in \mathcal{C}_k\}|}. \tag{2.42}$$

This degree represents the average edge value over all the edges connected to $\tilde{V}_t$. It is clear that if $\tilde{V}_t$ is in conflict with several of its connected vertices in $\mathcal{C}_k$ then $deg(\tilde{V}_t)$ will be close to 0, and it will be close to 1 if there is no conflict. Then the conflict of a vertex will be given by $1 - deg(\tilde{V}_t)$.

Figure 2.19 shows an example where there is a conflict between the vertices of a connected component for $\beta = 0.8$. Figure 2.19(b) shows the dual graph of the objects, and Figure 2.19(c) shows the thresholded graph, where there is a connected component with three vertices that we denote by C_0. The consistency degree of C_0 is $cons(C_0) = 0.05$, which is inferior to β. The conflict of the nodes (1-2), (2-3) and (2-4) are 0.64, 0.49 and 0.08, respectively. Therefore, the nodes (1-2) and (2-3) have a conflict with their neighbors in C_0. To reduce the conflict we remove the node (1-2) since it is the one having the higher conflict. By removing this edge the conflict is solved and the consistency degree of C_0 becomes $cons(C_0) = 0.89$, which is higher than β.

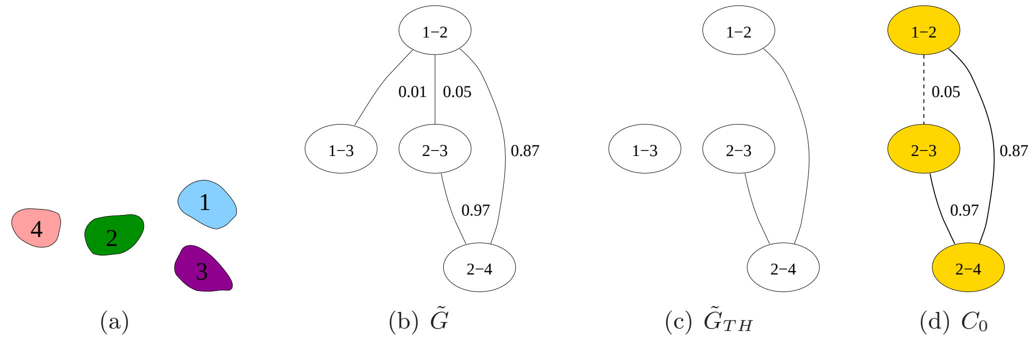


Figure 2.19: (b) Dual graph of objects in (a). (c) Thresholded dual graph for $\beta = 0.8$. (d) Vertices of the connected component of $\tilde{G}_{TH}$ seen by $\tilde{G}$.

2.2.3.5 Extension for fuzzy neighborhoods

If instead of having a crisp $Neigh$ relation, we have a fuzzy relation μ_{Neigh}, the procedure should be adapted for the extraction of the *locally* aligned groups.
When constructing the neighborhood graph G_N each edge should be attributed the degree of satisfaction of μ_{Neigh}. The notion of *triplets* also becomes fuzzy, and the degree to

Input: Dual graph $\tilde{G}$, β
Output: $\mathfrak{L}$

Create $\tilde{G}_{TH} = (\tilde{\mathcal{V}}, \tilde{E}_{TH})$ where $E_{TH} = \{(i,j) \in E | \tilde{e}_{ij} \geq \beta\}$;
Find $\mathfrak{C}$ the set of connected components of $\tilde{G}_{TH}$;
foreach $\mathcal{C}_k \in \mathfrak{C}$ **do**
  Let $cons = \min\{\tilde{s}_{ij} | \tilde{V}_i, \tilde{V}_j \in \mathcal{C}_k\}$;
  while $cons < \beta$ *and* $|\mathcal{C}_k| \geq 2$ **do**
    foreach $\tilde{V}_t \in \mathcal{C}_k$ **do**
      $deg_t = \frac{\sum_{\tilde{V}_j \in \mathcal{C}_k} \tilde{s}_{ij}}{|\{(i,j) | \tilde{V}_j \in \mathcal{C}_k\}|}$;
    end
    Delete from $\mathcal{C}_k$ the $\tilde{V}_j$ for which $\tilde{V}_j = \min_{\tilde{V}_i \in \mathcal{C}_k} deg_i$;
    if C_k *is disconnected in* $\tilde{G}_{TH}$ **then**
      Let $\mathfrak{D} = \{\mathcal{D}_0, \ldots, \mathcal{D}_L\}$ the connected components of $\mathcal{C}_k$;
      $\mathcal{C}_k = \mathcal{D}_0$;
      for $l = 1$ **to** L **do**
        Add $\mathcal{D}_l$ to $\mathfrak{C}$;
      end
      Update $cons = \min\{\tilde{s}_{ij} | \tilde{V}_i, \tilde{V}_j \in \mathcal{C}_k\}$;
    end
  end
  if $cons \geq \beta$ **then**
    Add $\mathcal{C}_k$ to $\mathfrak{L}$;
  end
end

Algorithm 1: Algorithm for finding locally aligned groups $\mathfrak{L}$ from $\tilde{G}$.

which three vertices X, Y and Z form a *triplet* is given by the degree of connectedness of $\{X, Y, Z\}$ using Equation 2.35.
The degree of connectedness is taken into account in the construction of the dual graph $\tilde{G}_N$, and only the *triplets* with a connectedness value greater than the user defined acceptance value β will be considered. Each edge (i, j) between the vertices $\tilde{V}_i$ and $\tilde{V}_j$ of the dual graph will be attributed with the degree:

$$\tilde{s}_{ij} = sim(O(X,Y), O(Y,Z)) \wedge \mu_{conn}(\{X,Y,Z\}), \tag{2.43}$$

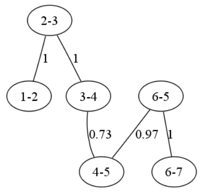
where X, Y and Z build the vertices *triplet* represented by the vertices $\tilde{V}_i$ and $\tilde{V}_j$.

The extraction of *locally* aligned groups is performed by applying Algorithm 1 on the dual graph. Let $\mathcal{S}$ be a resulting group. Due to the choice of construction of the dual graph, we can guarantee that, for every pair of elements $A, B \in \mathcal{S}$ the degree $\mu_{Neigh}(A, B)$ is greater than β, since only the edges satisfying this condition were used for the construction of the dual graph. Therefore, the degree of connectedness of the group is $\mu_{conn}(\mathcal{S}) \geq \beta$. It is straightforward to see that $\mathcal{S}$ satisfies the second condition of Equation 2.37 since this condition is imposed by the choice of $\tilde{s}_{ij}$. Hence, the resulting groups are *locally* aligned to a degree greater than β according to Equation 2.37.

2.2.3.6 Candidates for *globally* aligned groups

The *locally* aligned groups $\mathfrak{L}$ to a degree β are the possible candidates for being *globally* aligned groups to a degree α, for $\alpha \leq \beta$. The evaluation is performed by measuring the degree of *global* alignment using Equation 2.30. Usually the *locally* aligned groups are *globally* aligned. However there are cases as the one shown in Figure 2.20 where a *locally* aligned group is not *globally* aligned.
To increase the degree of *global* alignment of a group $\mathcal{S}$ we divide the group by eliminating the vertices in $\tilde{\mathcal{S}}$ with the minimum vertex degree (Equation 2.42) in $\tilde{\mathcal{S}}$, we repeat this step until $\mu_{ALIG}(\mathcal{S}) \geq \alpha$. If the degree of all vertices in $\tilde{\mathcal{S}}$ is equal to one, and $\mu_{ALIG}(\mathcal{S}) < \alpha$ it means that a lot of imprecision was introduced for the similarity computation, and the measurement of similarity is very permissive, thus the whole process should be repeated using a ν_0 with a tighter support in Equation 2.28.

(a) *Locally* aligned group, that is not *globally* aligned.



(b) Corresponding dual graph.

(c) Resulting groups after group division.

Figure 2.20: (a)*Locally* aligned group which is not *globally* aligned. (b) Its dual graph, where the objects are labeled from 1 to 7 (left to right). (c) Resulting groups obtained after solving the conflict.

2.2.3.7 Adding more elements to the group

Once the *globally* aligned groups of objects are identified it is possible to add new objects to the group or fuse two *globally* aligned groups to obtain a larger *globally* aligned group. For each group $\mathcal{S}_i$ we perform two morphological directional dilations of the group in the directions θ and $\theta + \pi$, where θ is the orientation of the alignment (the angle which maximizes the conjunction of the orientation histograms $O(A_i, \mathcal{S} \setminus \{A_i\})$). These dilations will be denoted by $D_{\nu_\theta}(\mathcal{S}_i)$ and $D_{\nu_{\theta+\pi}}(\mathcal{S}_i)$. The directional dilation of a fuzzy set μ in a direction $\vec{u}_\theta$ is defined in Section 1.3.2.1.
The fuzzy sets $D_{\nu_\theta}(\mathcal{S}_i)$ and $D_{\nu_{\theta+\pi}}(\mathcal{S}_i)$ represent the regions of space that are in direction θ and $\theta+\pi$ of $\mathcal{S}_i$. An object A which satisfies the $Neigh$ relation with one of the members of $\mathcal{S}_i$ and which is included in $D_{\nu_\theta}(\mathcal{S}_i)$ or $D_{\nu_{\theta+\pi}}(\mathcal{S}_i)$ with a degree greater than or equal to β (that is $\mu_{include}(A, D_{\nu_\theta}(\mathcal{S}_i) \cup D_{\nu_{\theta+\pi}}(\mathcal{S}_i)) \geq \beta$, where $\mu_{include}$ denotes a degree of inclusion of Equation 1.5) is added to $\mathcal{S}_i$, since is in the same direction as the orientation alignment

and it is connected to the group. If a whole group $\mathcal{S}_j$ is included in $D_{\nu_\theta}(\mathcal{S}_i)$ or $D_{\nu_{\theta+\pi}}(\mathcal{S}_i)$ with a degree greater than β and one of the elements of $\mathcal{S}_i$ is connected to one of the members S_j and both groups have a similar orientation, then both groups are fused into one.

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

Figure 2.21: (a) Labeled image. (b) Locally aligned group. (c) The region seen by the group of (b) in the direction of the alignment (white = high value of visibility). (d) Group obtained after adding new elements.

Figure 2.21 shows an example of extracting the *locally* aligned groups of Figure 2.21(a). The resulting group is *locally* aligned to a degree 0.9, and it is *globally* aligned to a degree 0.85. This group is extended to add more elements to the group resulting in a larger group with a degree of *global* alignment of 0.8.

2.2.4 Stability with respect to segmentation errors

One interesting feature of our approach is that it is robust to the quality of the segmentation of the objects. This property is particularly important in real applications where it is difficult to guarantee that all objects have been segmented and that the segmentation is accurate. Figure 2.22 shows two examples of the stability of the algorithm with respect to segmentation errors such as the absence of an object, or the merging of an undesired region to one of the objects. Figure 2.22(b) shows a segmentation of the houses of Figure 2.22(a), which is almost correct, except for some false detections and a missing house. Two of the aligned groups of objects extracted from this segmentation are shown in Figure2.22(c). Actually more groups are extracted by the algorithm but for the sake of clarity only a few groups are shown. Figure 2.22(d) shows another segmentation which was manually modified to introduce errors: two missing houses, and one of the houses is merged with other regions. Figure 2.22(e) shows two of the retrieved aligned groups, which correspond to the groups found in Figure2.22(c). We can see that the blue group in Figure 2.22(e) is retrieved even with the absence of two objects, and that the orange group is retrieved although the center of mass of one of its members has been displaced.

For this example we used a Voronoi neighborhood constrained by a distance d, where d was larger than the separation between the objects. However, if we had used a smaller distance, then the algorithm would not have retrieved the blue group, since there would have been a disconnection between the upper and the lower part of the blue group of Figure 2.22(e), and therefore only the lower part of the group would have been retrieved. This is an expected behavior.

In both experiments we used the same parameters for the extraction of the *locally* aligned groups, which are $\beta = 0.8$ and $d = 50$ pixels. The degrees of *global* alignment in both cases are very similar, for the blue groups the degree was $\mu_{ALIG} = 1.0$ in the two cases, and for the orange groups it was $\mu_{ALIG} = 1.0$ in Figure 2.22(c), and $\mu_{ALIG} = 0.94$

in Figure 2.22(e). The degrees of *global* alignment remained almost the same, since the missing objects and the modification of the object are local changes which do not affect the *global* orientation of the groups. Nevertheless if one of the objects is severely modified as in Figure 2.22(f) then it is not possible to retrieve the same aligned groups as before, which is again an expected behavior.

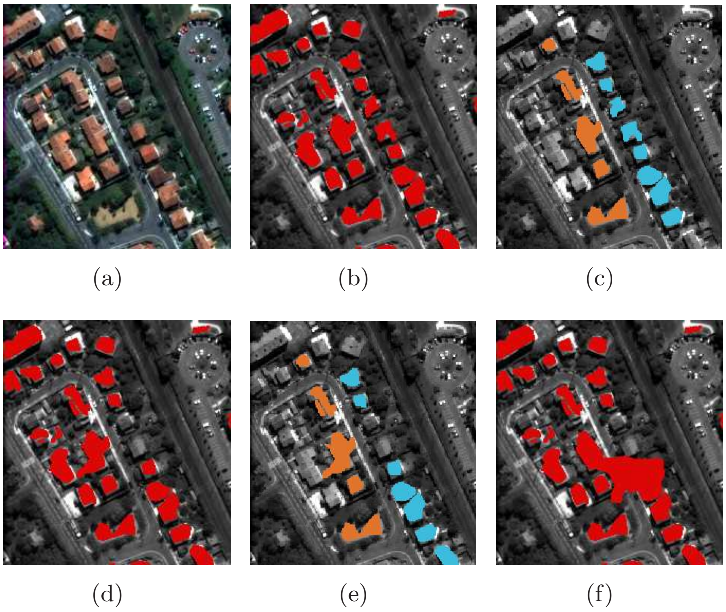


Figure 2.22: (a) Original image. (b) Segmented houses (red) using algorithm from [Poulain et al., 2009]. (c) Some of the extracted *globally* aligned groups from objects of (b). (d) Segmented houses with errors of missing objects and some merged regions. (e) Some of the extracted *globally* aligned groups from objects of (d). (f) Segmented houses with errors that do not allow the recovery of the *globally* aligned groups.

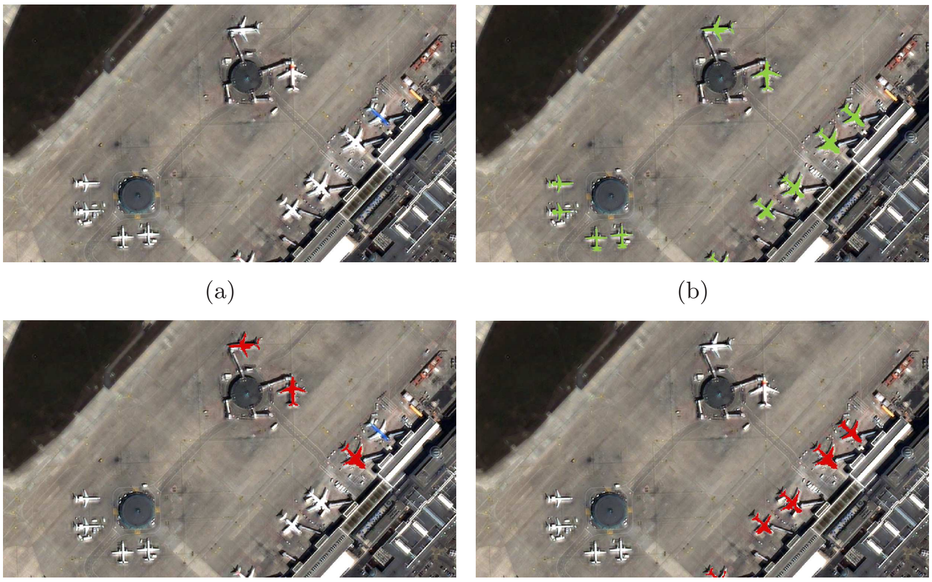
2.2.5 Complexity analysis

In this section we deal with the cost of the basic operations of the algorithm for extracting *locally* aligned groups and *globally* aligned groups.
First, we consider the complexity of extracting *locally* aligned groups. Consider we have N objects each with at most n_o points. The complexity of the algorithm is $O(N^2)$ since most of steps of the algorithm deal with operations over the graph or its dual. It should be noticed that the step which corresponds to the construction of the orientation histograms has a complexity of $O(N^2 n_o^2)$, since there are at most $N(N-1)$ edges in the graph and for each edge an orientation histogram is constructed with a complexity of $O(n_o^2)$.
The complexity of finding a *globally* aligned group from a *locally* aligned group with N_A elements each having at most n_o points lies in the following steps. The first step consists in evaluating the degree of *global* alignment and division of the group in the case where it is not aligned, and this step has a complexity of $O(N_A^2 n_o^2)$. The second step consists in performing the morphological directional dilations of the group in the directions of alignment θ and $\theta + \pi$, and has a complexity of $O(N_I)$ [Bloch, 1999], where N_I is the number of points in the image (see [Bloch, 1999] for the implementation of the directional

morphological dilation using a propagation method). Finally, the complexity of the step of evaluating the degree of inclusion of each object not belonging to the group into the directional dilations of the group is $O((N-N_A)n_o^2)$, where N is the total number of objects. Hence, summing the three steps we obtain that the total complexity is $O(N_A^2 n_o^2 + N_I)$.

2.2.6 Discussion

In this section we have introduced the definitions of *globally* aligned groups and *locally* aligned groups of objects, and gave a method to extract alignments from an image of labeled objects. Both definitions are appropriate to determine alignments of objects of different sizes. Therefore, the extraction method can be used to find alignments of objects of the same type or class, for instance buildings in a urban scene, since all the objects do not necessarily have the same size. However, not all the obtained groups are meaningful for the description of the scene, since the subsets which are found only satisfy the conditions of alignment. For example Figure 2.23 shows two *globally* aligned subsets of airplanes extracted using the proposed algorithm, the group of Figure 2.23(c) is *globally* aligned but does not give any information about the arrangement of the airplanes, while the group of Figure 2.23(d) gives us more information about the arrangement of the airplanes. Hence it is necessary to use additional information to put the aligned groups into context, for example find whether the alignments are parallel between them or parallel to a linear structure. In the case of Figure 2.23 it would be interesting to find the alignment parallel to the buildings. This point will be further discussed in the next section.



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

Figure 2.23: (a) Airport image. (b) Manually segmented airplanes (green). (c) Extracted *globally* aligned group in red with degree 0.97. (d) Extracted *globally* aligned group in red with degree 0.99.

The proposed method for alignment extraction is very flexible. One should notice that it is possible to incorporate more information according to the type of alignment. For instance, if we are searching for an alignment where the objects of the alignment have the same orientation as the alignment, this could be incorporated in the weight $\tilde{s}_{ij}$ attributed to the edges of the dual graph. The new weight $\tilde{s}_{ij}$ of the edge between the vertices $\tilde{V}_i$ and $\tilde{V}_j$, which represents a *triplet* $\{X, Y, Z\}$, is given by a conjunction combining the condition that $O(X, Y)$ and $O(Y, Z)$ should be similar, and the condition that every member of the *triplet* should have a similar orientation to the orientation histograms between itself and the other members of the triplet which are connected to it. For instance, the condition for Y is expressed as:

$$sim(O(X, Y), \delta_{\theta_Y}) \wedge sim(O(Z, Y), \delta_{\theta_Y}),$$

where δ_{θ_Y} is the Dirac function at the angle θ_Y which represents the orientation of Y. For X and Z the condition is verified by observing the degree of $sim(O(X, Y), \delta_{\theta_X})$ and $sim(O(Z, Y), \delta_{\theta_Z})$, respectively. Since X and Z are connected to Y in the *triplet*, and the only histograms that involve them are $O(X, Y)$ and $O(Z, Y)$, respectively. Finally, when combining all the conditions the weight $\tilde{s}_{ij}$ is given by:

$$\begin{aligned}\tilde{s}_{ij} =& sim(O(X, Y), O(Y, Z)) \wedge \\ & [\min [sim(O(X, Y), \delta_{\theta_Y}) \wedge sim(O(Z, Y), \delta_{\theta_Y}), sim(O(X, Y), \delta_{\theta_X}), sim(O(Z, Y), \delta_{\theta_Z})]]\end{aligned} \tag{2.44}$$

In this equation we use the *min* to ensure that the second condition is satisfied by all the members of the *triplet*.

In Chapter 3 we will present a reasoning example using the alignment relation, that can be used to better illustrate this relation.

2.3 Parallelism

Assessing the parallelism between objects is an important issue when considering man-made objects such as buildings, roads, etc. Parallelism takes place among linear objects. Furthermore it can be evaluated between a linear object and a group of aligned objects, or even between groups of aligned objects.

This section is structured as follows. In Section 2.3.1 we discussed some considerations that should be taken into account when modeling the parallel spatial relation. Section 2.3.2 reviews some of the previous works on parallelism. A model for fuzzy segments is proposed in Section 2.3.3 this model considers the imprecision of the objects as well as the imprecision of the relation. The relation is modeled based in two condition the "visibility" of the target object to the reference object and a similar orientation of both objects. Considering the "visibility" yields a non-symmetric relation which is appropriate for defining situations such as "the house is parallel to the road" which is not equivalent to the situation described by "the road is parallel to the house". This non-symmetric relation is well adapted to the semantics of the parallel relation. The properties of this non-symmetric relation are discussed in Section 2.3.4, as expected this properties differ from the properties of the parallelism defined in geometry. A model for *globally* aligned groups is presented in Section 2.3.5. Part of this work was presented in [Vanegas et al., 2009b].

2.3.1 Considerations

For linear objects to be parallel, we expect a constant distance between them, or that they have the same normal vectors and the same orientation. Although classical parallelism in

Euclidean geometry is a symmetric and transitive relation, these properties are subject to discussion when dealing with linear objects of finite length. When objects have different extensions as in Figure 2.24(a), where B can be a house, and A a road, the symmetry becomes questionable. The statement "B is parallel to A" can be considered as true, since from every point on the boundary of B that faces A it is possible to see (in the normal direction to A's principal axis) a point of A, and the orientations of A and B are similar. On the other hand, the way we perceive "B parallel to A" will change depending on our position: from point d it is possible to see a point of B in the normal direction of B, while this is not possible from point c. In both cases (symmetrical and non symmetrical ones) the transitivity is lost. For example, in Figure 2.24(b) and 2.24(c) the statements "A is parallel to B" and "B is parallel to C" hold, but "A is not parallel to C" since it is not possible to see C from A in the normal direction to C. This example also illustrates the interest of considering the degree of satisfaction of the relation instead of a crisp answer (yes/no). Then the relation "B parallel to A" will have a higher degree than "A parallel to B" in Figure 2.24(a).

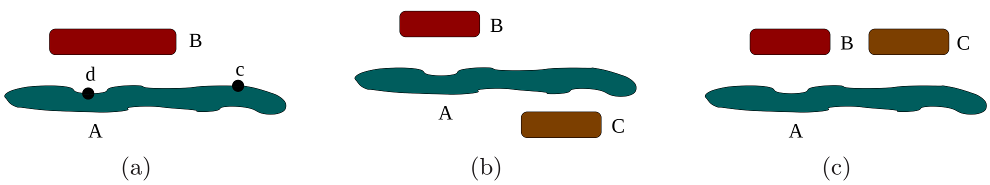


Figure 2.24: Examples where parallelism should preferably be considered as a matter of degree, and should not be necessarily symmetrical and transitive.

The parallel relation can also be considered between a group of objects $\mathcal{A} = \{A_i\}$ and an object B, typically when the objects in the group are *globally* aligned and B is elongated. For example a group of boats and a deck in a port. When evaluating the relation "$\mathcal{A}$ is parallel to B", we are actually evaluating whether the whole set $\mathcal{A}$ and the boundary of B that faces $\mathcal{A}$ have a similar orientation, and whether there is a large proportion of $\cup_i A_i$ that sees B in the normal direction to the group. Similar considerations can be derived when considering the relation "B is parallel to $\mathcal{A}$" or the relation between two groups of objects. All these considerations form the basis for the formal models provided in this section.

For "A to be parallel to B" it is only necessary that A is a linear object, while B can be a non linear object, and in this case A would be parallel to the boundary of B which is facing A. The same idea is also applicable for the parallelism between a *globally* aligned group of objects parallel to an object (see Figure 2.25).

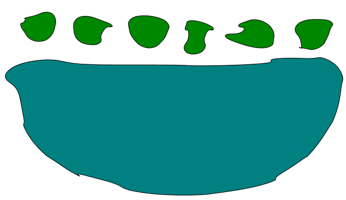


Figure 2.25: A group of *globally* aligned objects parallel to a non-linear object.

2.3.2 Related work

In the same case as for the alignment relation, the parallel relation also been widely studied in computer vision, as a low level feature operator. Parallelism between two linear segments is usually modeled as a relation that should satisfy three constraints. The first one is that both segments should have a small angular difference. The second one is that the perpendicular distance between the two segments should be small and the last one is that there should be a high overlap between them [Mohan and Nevatia, 1989]. There exist different approaches to integrate these three constraints into a model and measure the meaningfulness of the relation. Lowe [Lowe, 1987] was one of the first to model parallelism to perform perceptual grouping: for every couple of linear segments having a perpendicular separation d and an angular difference θ, he assigns a significant value to establish that it has not been originated by an accident of viewpoint. The significant value is used to determine the expected number of lines for a given perpendicular separation and an angular difference. In fact, this value is proportional to the prior probability of appearance and it is inversely proportional to the angular difference and the perpendicular separation. In [Mohan and Nevatia, 1989] the constraint about overlapping is introduced and it is determined by the orthogonal projection of one segment over the other and vice-versa. The meaningful segments are then obtained by applying a threshold on the measured values of the constraints. Fuzzy approaches have been proposed in [Kang and Walker, 1992, 1994, Rouco et al., 2007, Ralescu and Shanahan, 1999, Toh, 1992], leading to a measure of the degree of parallelism between two linear segments, in which trapezoid type functions are used to evaluate to which degree the three constraints are satisfied. The three degrees are combined in a conjunctive way.

Parallelism between curves has been modeled as a type of symmetry [Kang and Walker, 1992, Mohan and Nevatia, 1992]. In [Mohan and Nevatia, 1992] the authors consider that two curves are parallel if their respective orthogonal distances from the symmetry axis are almost equal. In [Ip and Wong, 1997] and [Shen et al., 1999] parallelism is modeled as a matching problem where two curves are parallel if there is a point-wise correspondence. In both works the parallelism is treated as a shape matching problem.

Previous works focused on parallelism between crisp linear segments. However, when dealing with objects extracted from images, it is important to consider parallelism between fuzzy linear objects. Indeed, the object extraction processes can introduce imprecision, and therefore we are not always dealing with crisp linear objects or segments. In Section 2.3.3 we will give some definitions about parallelism between fuzzy linear objects. The definitions of parallelism presented in this work differ from those of the previous works in the sense that we are defining the relation considering its semantics, and we are not worried about whether or not the detected parallelism is an accident of the view point or position. We are interested in the conditions that have to be satisfied to decide whether two objects are parallel or not.

2.3.3 Parallelism with (fuzzy) linear objects

In this section we propose a definition of parallelism between a fuzzy linear object[3] and a fuzzy object, including the particular case of crisp linear segments, and taking into account

[3] We consider that an object is linear if the ratio of its principal axis given by $\frac{c_{yy}+c_{xx}-\sqrt{(c_{xx}+c_{yy})^2-4(c_{xx}c_{yy}-c_{xy}^2)}}{c_{yy}+c_{xx}+\sqrt{(c_{xx}+c_{yy})^2-4(c_{xx}c_{yy}-c_{xy}^2)}}$ where $C = \begin{pmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{pmatrix}$ corresponds to the second inertial moments matrix, is high [Peura and Iivarinen, 1997].

the above mentioned considerations. Suppose A is a linear object and B an object that is not necessarily linear, let θ_A be the orientation of A and $\vec{u}_{\theta_A+\frac{\pi}{2}}$ be the normal unit vector to the principal axis of A. Then, according to the considerations of the previous section, the degree of satisfaction of the relation "A is parallel to B" depends on two conditions:

(i) There should be a large proportion of A that sees B in the direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$.

(ii) The orientation of A and the orientation of the boundary of B that is facing A and that is seen by A in the direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$ should be similar.

Both conditions deal with the notion of visibility. Let p be a point, X be a fuzzy object with membership function μ_X and $\vec{u}_\theta$ a vector with angle θ with respect to the x-axis. Then the subset of X that is seen by p in the direction $\vec{u}_\theta$, that we denote by $X_{vis(p,\theta)}$, is equal to the intersection of X with the visual field of p, when p observes in the direction $\vec{u}_\theta$. The visual field is represented as a morphological directional dilation in the direction θ of p. The set $X_{vis(p,\theta)}$ is a fuzzy set with membership function $\mu_{Xvis(p,\theta)}$, where $\mu_{Xvis(p,\theta)}(x)$ represents the degree to which x is in X and is seen by p in the direction $\vec{u}_\theta$:

$$\mu_{Xvis(p,\theta)}(x) = \mu_X(x) \wedge D_{\nu_\theta}(p)(x), \tag{2.45}$$

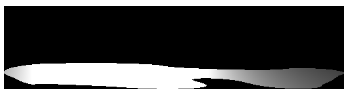
where $D_{\nu_\theta}(p)(x)$ is the morphological directional dilation defined in Section 1.3.2.1.
Let Y be a fuzzy object with membership function μ_Y not intersecting X. We denote by $X_{vis(Y,\theta)}$ the subset of X that is seen by the points on the boundary of Y, i.e the subset of X that is seen by Y, and it is defined by:

$$\mu_{Xvis(Y,\theta)}(x) = \mu_X(x) \wedge D_{\nu_\theta}(\mu_Y)(x). \tag{2.46}$$

Figure 2.26 shows two objects A and B, and $B_{vis(A,\theta_A+\frac{\pi}{2})}$, where θ_A is the orientation of A. When A and B are linear segments, $A_{vis(B,\theta_A+\frac{\pi}{2})}$ can be interpreted as the projection of B onto A.

(a) Objects

(b) $D_{\nu\theta_A+\frac{\pi}{2}}(A)$



(c) $\mu_{Bvis(A,\theta_A+\frac{\pi}{2})}$

Figure 2.26: Illustration of the notion of visibility for the objects A and B of (a). (b) Visual field of A in the direction of $\theta_A + \frac{\pi}{2}$ (the white pixels have a high membership value of being observed by A in the direction θ_A). (c) Membership function of $B_{vis(A,\theta_A+\frac{\pi}{2})}$. The white pixels are the points which have a high membership value.

For the first condition of parallelism, we are interested in the proportion of A that sees B in the direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$. This subset is equivalent to $A_{vis(B,\theta_A-\frac{\pi}{2})}$, since the degree to which a point $x \in A$ sees B in the direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$ is equivalent to the degree to which the point is seen by B in the direction $\vec{u}_{\theta_A-\frac{\pi}{2}}$. Therefore, the proportion of A that sees B in the normal direction $\vec{u}_{\theta_A+\frac{\pi}{2}}$ is equal to the fuzzy hypervolume of $A_{vis(B,\theta_A-\frac{\pi}{2})}$ over the fuzzy hypervolume of A, where the fuzzy hypervolume V_n of a fuzzy set μ is given by: $V_n(\mu) = \sum_{x \in \mathfrak{I}} \mu(x)$ [Bloch, 2005]. Hence, the proportion is equal to:

$$\frac{V_n(\mu_{Avis(B,\theta_A - \frac{\pi}{2})})}{V_n(\mu_A)}$$

For the second condition we are interested only in the subset of the boundary of B that faces A and that is seen by the boundary of A in the direction θ_A. The boundary of B which faces an object A corresponds to the points on the boundary of B that delimit the region between A and B, and are defined as the extremities of the admissible segments [Bloch et al., 2006]. These are the points $b \in B$ for which there exists a point $a \in A$ such that the segment $]a, b[$ is included in $A^C \cap B^C$. In the case were A and B are fuzzy objects, we will be interested in the points which are the extremities of a segment with a high degree of admissibility [Bloch et al., 2006]. Therefore, the subset of the boundary of B that faces A and that is seen by the boundary of A in the direction $\theta_A + \frac{\pi}{2}$ is a fuzzy subset were the membership of a point $x \in \mathcal{I}$ is equal to the conjunction between its membership to B, the degree of being the extremity of an admissible segment and the degree of being seen by A. We denote this subset by $\delta B_{vis(A,\theta_A + \frac{\pi}{2})}$ and its membership function by $\mu_{\delta Bvis(A,\theta_A + \frac{\pi}{2})}$:

$$\mu_{\delta Bvis(A,\theta_A + \frac{\pi}{2})}(x) = \mu_{adm}(x) \wedge \mu_{Bvis(A,\theta_A + \frac{\pi}{2})}(x) \tag{2.47}$$

where μ_{adm} represents the degree of being the extremity of an admissible segment.

The relation "A is parallel to B" is given by the following measure:

$$\mu_{parallel}(A, B) = \frac{V_n(\mu_{Avis(B,\theta_A - \frac{\pi}{2})})}{V_n(\mu_A)} \bigwedge \nu_0(\theta_{\delta Bvis(A,\theta_A + \frac{\pi}{2})} - \theta_A), \tag{2.48}$$

where $\nu_0(\theta)$ is the same as in Equation 2.24 and it evaluates the degree to which $\theta_{\delta Bvis(A,\theta_A + \frac{\pi}{2})}$, the normal angle to $\delta B_{vis(A,\theta_A + \frac{\pi}{2})}$, and θ_A are "approximately" equal.

In some contexts a symmetrical relation is needed (for example in perceptual organization), and is then expressed as "A and B are parallel". In such cases, we verify that at least one of the sets is visible from the other in the normal direction and that the orientations of both sets are similar. Then, the degree of satisfaction of the symmetrical relation, "A and B are parallel" is expressed by:

$$\begin{aligned} \mu_{parallel_S}(A, B) = & \left[\frac{V_n(\mu_{Avis(B,\theta_A - \frac{\pi}{2})})}{V_n(\mu_A)} \bigvee \frac{V_n(\mu_{Bvis(A,\theta_B - \frac{\pi}{2})})}{V_n(\mu_B)} \right] \bigwedge \nu_0(\theta_{\delta Bvis(A,\theta_A + \frac{\pi}{2})} - \theta_A) \\ & \bigwedge \nu_0(\theta_{\delta Avis(B,\theta_B + \frac{\pi}{2})} - \theta_B) \end{aligned} \tag{2.49}$$

2.3.4 Properties

Both relations (Equations 2.48 and 2.49) are invariant with respect to geometric transformations (translation, rotation, scaling). None of the relations is transitive, as discussed previously. However, if A, B, C are linear crisp segments, and if $\mu_{parallel}(A, B) = 1$, $\mu_{parallel}(B, C) = 1$ and $\theta_A = \theta_B = \theta_C$, then $\mu_{parallel}(A, C) = 1$. This result shows that in the crisp case we have transitivity. To have the transitivity property, it is necessary that $\theta_A = \theta_B = \theta_C$, since $\nu_0(\theta_A - \theta_B) = 1$ and $\nu_0(\theta_B - \theta_C) = 1$ do not imply $\nu_0(\theta_A - \theta_C) = 1$ due to the tolerance value t_1 of the function ν_0 (see Equation 2.24). To have the transitivity without imposing the condition $\theta_A = \theta_B = \theta_C$, it is necessary that ν_0 is a linear function (i.e. $t_1 = 0$). But this is restrictive.

It is clear that both relations are reflexive. However, depending on the context we may not want to consider intersecting objects as parallel. In this case, it is necessary to combine in a conjunctive way the previous degree (Equations 2.48 or 2.49) with a degree of non-intersection between the two sets.

2.3.5 Parallelism with a *globally* aligned group of objects

When considering parallelism with a *globally* aligned group of objects, the group has a similar role as the linear object in the definitions introduced in the previous section. When defining the relation with a group there is a modification, with respect to the case of a linear object, in the way the visibility constraint is computed, this modification will be discussed in the following.

2.3.5.1 A group of *globally* aligned objects parallel to an object

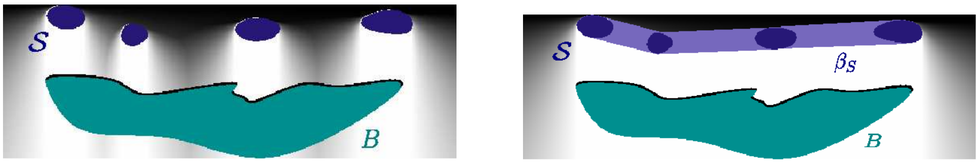
Let $\mathcal{S} = \{A_0, \ldots, A_N\}$ be a group of globally aligned objects, as defined in Section 2.2.3.1, and let B be another object. For $\mathcal{S}$ to be parallel to B it is necessary that there is a large portion of $\mathcal{S}$ that sees B, and this is computed in the same way as for the case of parallelism between a linear object and an object. For the second condition we need to create the fuzzy set β_S which is composed of the union of the regions between two consecutive elements of $\mathcal{S}$. β_S can be constructed using the definition that involves the convex hull presented in [Bloch et al., 2006]. In Figure 2.27(b) an example of the region β_S of a group S is shown in light purple. From Figure 2.27(a) we can see that the boundary of B that faces $\mathcal{S}$ and that is visible by $\mathcal{S}$ depends on the separation between the members of the group. However, it is desirable that the degree of parallelism of a group to an object is independent of the separation of its members, since if we add more elements to a group without changing its orientation or its extension the degree of alignment to the object should remain the same. Therefore, in order to have a degree of parallelism independent of the separation of the members of the group we should use β_S in the second condition of parallelism. Then, the second condition becomes that the boundary of B that faces $\mathcal{S}$ and that is visible by β_S or by $\mathcal{S}$ should have the same orientation as the orientation of the alignment of $\mathcal{S}$. Finally, the degree of satisfaction of the relation "$\mathcal{S}$ is parallel to B" is given by:

$$\mu_{parallel}(\mathcal{S}, B) = \frac{V_n(\vee_i \mu_{A_i vis(B,\theta_s - \frac{\pi}{2})})}{V_n(\vee \mu_{A_i})} \bigwedge \nu_0(\theta_{\delta B vis(\beta_s \cup \mathcal{S}, \theta_s + \frac{\pi}{2})} - \theta_S), \tag{2.50}$$

where θ_s is the orientation of the alignment of the group $\mathcal{S}$. In this definition, the first part of the equation represents that there should be a large portion of the union of all the $A_i \in \mathcal{S}$ that see B, and the second part evaluates the degree of similarity between the orientation of the group and the orientation of the boundary of the object seen by the group in the direction $\theta_s + \frac{\pi}{2}$.

2.3.5.2 A linear object parallel to a *globally* aligned group of objects

Using the same notations as above, suppose B is a fuzzy linear object. Then for "B is parallel to $\mathcal{S}$" to be true, it is necessary that B has a similar orientation to the orientation of the alignment of $\mathcal{S}$, and that there is a large proportion of B that sees the group of objects or β_S. As in Equation 2.50 it is necessary to use $\mathcal{S} \cup \beta_S$ in order to assure that



(a) (b)

Figure 2.27: (a) Visibility field of the group of objects $\mathcal{S}$, and the boundary of the object facing $\mathcal{S}$ (in black). (b) Visibility filed of the region between two consecutive members of the group β_S (light purple) and the group $\mathcal{S}$, and the boundary of B facing $\mathcal{S} \cup \beta_S$ (in black).

the parallel relation is independent of the separation of the element of $\mathcal{S}$. The degree of satisfaction of the relation "B is parallel to $\mathcal{S}$" is given by:

$$\mu_{parallel}(B, \mathcal{S}) = \frac{V_n(\vee_i \mu_{Bvis(\mathcal{S} \cup \beta_S, \theta_B - \frac{\pi}{2})})}{V_n(\vee \mu_B)} \bigwedge \nu_0(\theta_S - \theta_B), \tag{2.51}$$

where θ_B is the orientation of B.

2.3.5.3 Parallelism between two *globally* aligned groups of objects

Using the same notation as in Equation 2.51, we can define the parallelism between two *globally* aligned of fuzzy sets $\mathcal{S} = \{A_0, \dots, A_N\}$ and $\mathcal{T} = \{B_0, \dots, B_M\}$. The degree of satisfaction of the relation "$\mathcal{S}$ is parallel to $\mathcal{T}$" is given by:

$$\mu_{parallel}(\mathcal{S}, \mathcal{T}) = \frac{V_n(\vee_i \mu_{A_i vis(\mathcal{T} \cup \beta_T, \theta_s - \frac{\pi}{2})})}{V_n(\vee \mu_{A_i})} \bigwedge \nu_0(\theta_T - \theta_S), \tag{2.52}$$

where β_T is the region formed between two consecutive elements of $\mathcal{T}$ and θ_T is the orientation of alignment of the group $\mathcal{T}$.

2.3.6 Discussion

In this section we discussed the considerations that should be taken into account when modeling the parallel relation. We highlighted that the parallel relation should be modeled as a fuzzy relation represented as the conjunction of two conditions, one dealing with visibility and the other with similarity of orientation. Using the directional morphological dilation to model the visibility condition allows us to identify the region on the image where it is possible to find objects to which the object of interest or groups of objects are parallel to. We presented a definition of parallelism between two fuzzy linear objects, and between *globally* aligned groups.

In Chapter 3 we illustrate an example of combining parallelism and alignment for spatial reasoning.

2.4 Line-region relations

This section is motivated by the importance of the relations between linear objects and regions when creating natural language descriptions of spatial scenes of satellite images.

These types of relations are frequently observed between linear structures such as rivers or roads and spatial regions such as cities, agricultural fields, water surfaces, etc. Some examples of these relations are “to go across”, “to go through”, “to bypass”, “to intersect”, “to go along”, “to enter”, “to go into”, etc. These relations depend on the shape of the region and their definition can be sometimes vague or imprecise [Herskovits, 1997]. For instance the relation “to go across” can have the following meaning: a line goes across a region if it goes from one side of the region to the opposite one. However, when the region has a complex shape or imprecise boundaries, the notion of opposite side becomes vague. Figure 2.28 exemplifies a case where the house settlement has a complex shape, and therefore it is difficult to say which are the routes that “go across” it. Thus, the fuzzy sets framework is appropriate for modeling these kinds of relations since it captures the imprecision inherent to the spatial information and to the semantics of the relations.

(a)

(b)

Figure 2.28: Illustration of when the notion of across is not well defined by a crisp relation. If we want to determine if the roads in (a) go across the house settlement shown in (b), then it would be difficult to give an answer due to the complex shape and the fuzzy boundaries of the region shown in (b).

In this section we concentrate on the binary spatial relations between a region and a linear object. For simplicity we refer to these relations as “line-region” relations. By linear object we mean an elongated structure. To determine whether an object is elongated we can measure its elongation[4].

In Section 2.4.1 we briefly present the “line-region” relations which have been studied in the literature. From these relations we have chosen 4 relations which are studied in the following sections (“go across” and “go through” in Section 2.4.2, “enter” and “go into” in Section 2.4.3) and are included in the classification of Figure 1.16.

2.4.1 Set of line-region relations

Relations between linear objects and regions have been studied in the framework of spatial relations for GIS [Roussopoulos et al., 1988, Egenhofer and Herring, 1990, Mark and Egenhofer, 1994a, Shariff et al., 1998, Kurata and Egenhofer, 2007] and in the spatial cognition community [Landau and Jackendoff, 1993, Talmy, 1983].

[4]The elongation can be measured by using $E = \frac{P^2}{A}$, where A is the area and P is the geodesic diameter

[Roussopoulos et al., 1988] proposed six relations that can take place among a region and a line: “intersect”/ “not-intersect”, “within”/ “not within” and “cross”/ “not cross”. However, the first two pairs of relations can take place among any two types of entities, they are not exclusive to the line-region configuration. Moreover, there is no definition given in [Roussopoulos et al., 1988].

[Egenhofer and Herring, 1990] focused on the topological properties of the relations. They proposed 19 spatial relations between a region and a line using the *9-intersection model* (Section 1.3.1.2). However, these relations are not directly linked to natural-language expressions, and therefore they cannot be easily used for image interpretation. In [Mark and Egenhofer, 1994a] the link between topological constraints and natural language expressions is studied. The authors conclude that topology alone is not sufficient to define the relations between a linear object and a region. [Shariff et al., 1998] incorporated metrical properties to the topological models and suggested 59 crisp natural language spatial relations. The model of each relation is obtained by first extracting the topological relation between the line and the region using the *9-intersection model*. Afterwards it is refined by adding measurements of metric properties. The metric properties which are measured are:

Splitting: determines how a region's interior, boundary and exterior are divided by a line's interior and boundary, and vice-versa.

Closeness: determines how far the region's boundary is apart from the line.

Approximate alongness: assesses the length of the section where the line's interior runs parallel to the region's boundary. It is a combination of the closeness and splitting measures.

Figure 2.29 shows a schema of the measures used. Basically, the measures refer to ratios of distances, length or size differences of the region and the line intersections. For each model of each relation, the values of these measures were calibrated from the results obtained from a human experiment. Despite of using human studies to define the models, the proposed relations do not consider the shape of the region.

Among the 59 relations proposed by [Shariff et al., 1998], there are relations which can be applied to any two kinds of objects, such as “in”. Additionally, several relations share the same model, for instance “go across”, “cut across”, “split” and “divide”. This indicates that within the 59 relations there are terms that correspond to synonyms. In [Schwering, 2007] a subset of these relations is automatically grouped into three clusters using a similarity measures among the models. To evaluate the validity of the clusters a human experiment was conducted. The following list shows some of the members of each cluster:

1. “split”, “traverse”, “cross”, “intersect”, “bisect”,
2. “in”, “inside”, “within”, “enclosed by”
3. “avoids”, “bypass”, “along edge”, “near”,

Apart from these clusters, there is the “enter” relation, which did not fit into any of the clusters, since there was a disagreement between the results obtained from the automatic clustering and the human experiment. From the work of [Schwering, 2007] we can conclude that it is possible to simplify the set of 59 relations to only 4 relations (the three clusters and the “enter” relation).

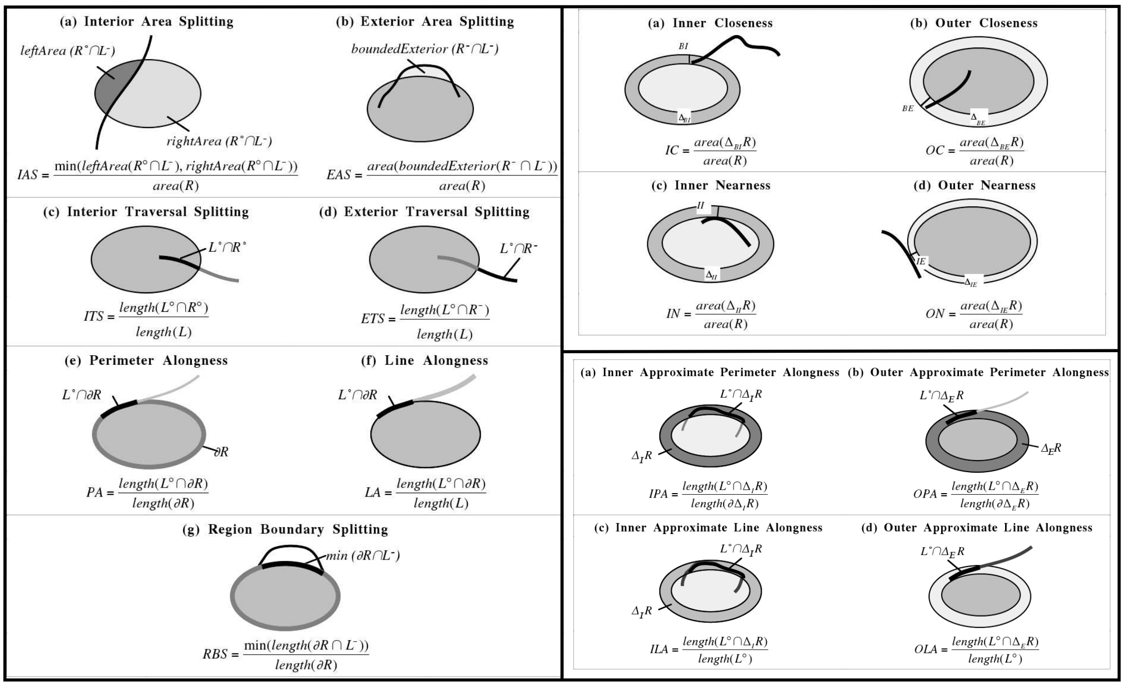


Figure 2.29: Measures of the metric properties used in [Shariff et al., 1998]. The left box represents the measures used for the splitting property. The upper right box contains the measures for the closeness property, and the lower right box the measures for the approximate alogness property. Images taken from [Shariff et al., 1998].

In the cognition area, [Landau and Jackendoff, 1993] identified four English prepositions which express the relations between a reference region and a linear structure. These prepositions are: “along”, “in line with”, “across” and “around”. Notice that the three clusters identified by [Schwering, 2007] coincide with three of these relations. The relation “in line with” requires that the linear object is a straight line and that the region is also linear, therefore it does not fit exactly into the category of line - region relations.

Taking into account the relations listed by [Shariff et al., 1998] and [Landau and Jackendoff, 1993] we decided to concentrate on the following relations: “go across”/ “go through”, “enter”/ “go into”, “along” and “around”. The relation “go across”, “along” and “around” where identified by [Landau and Jackendoff, 1993] and are present in two of the clusters proposed by [Schwering, 2007]. The relation “enter” was the relation which did not fit into any cluster in [Schwering, 2007]. The other cluster of relations proposed by [Schwering, 2007], which includes the relation “inside” is not studied since it is not exclusive to the case of “line-regions”, moreover it can computed by using a degree of inclusion (see Equation 1.5).

This group of relations includes the significant relations found in the cognitive area and in GIS, as well as the relation “enter”. Nevertheless, this set of relations is not exhaustive.

First, we study the “go across” relation (Section 2.4.2) because it is one of the most complex relations in our list. To model this relation we require to introduce notions that are useful for the definition of the models of the relations “enter” and “go into”. These two relations are studied in Section 2.4.3. Finally, we discuss the models of the relations “along” that are found in the literature, and make the connection with “around” and the relation "surround" studied in Section 2.1. This work was presented in [Vanegas et al., 2009a].

In the following L is a linear object represented by its skeleton. The extremities of L are noted L_a and L_b.

2.4.2 Go Across

The relation "go across" has been studied in different contexts, producing different meanings. In GIS, the relation proposed by [Shariff et al., 1998] satisfied two topological conditions; the line should intersect the interior of the region, and its extremities should intersect the exterior of the region. The metrical properties which are measured are the splitting and the closeness. The splitting is measured through the interior area splitting (see Figure 2.29). The interior area splitting is based on the proportion between the area obtained by the division of the region and the original region. It is calculated as the ratio of the smallest area in which the line splits R, and the total area of R. In [Shariff et al., 1998], the relation is considered as satisfied if this value is inside the interval $[0.18, 0.5]$. Nevertheless, this value is not informative, since this measure has a range of [0,0.5]. The outer closeness measure is based in the distance between the endpoints of the line and the boundary of the region. However, this measure is not critical, since the importance of the relation depends on the way the line "moves" inside the region as we will see in the following paragraphs. Therefore, none of the measures are significant. Thus, from this definition we can only retain the topological constraints.

[Tellex and Roy, 2006] use the relation "go across" to guide the movement of a robot in a room. The relation is satisfied when the trajectory of the robot goes from a point on the boundary of the region to another point on the boundary following a straight path which intersects the region's center of mass. This interpretation of the relation deals with the center of mass. It is also more flexible in the topological constraints, since the extremities of L are allowed to be on the boundary of the region.

In the spatial cognition community this relation has been studied by [Landau and Jackendoff, 1993, Talmy, 1983, Herskovits, 1997]. [Landau and Jackendoff, 1993, Talmy, 1983, Herskovits, 1997] highlighted that the preposition is linked with the notion of going from "side to side" of the reference object. For [Talmy, 1983] the reference object must be ribbonal in order to have sides. For [Landau and Jackendoff, 1993] the reference object must be a surface with sides, so that we can go across "from one side to the other". The definition proposed by [Herskovits, 1997] is more flexible and allows reference objects to have any shape. The notion of "opposite sides" is further discussed in Section 2.4.2. Additionally, [Herskovits, 1997] proposed 4 other configurations for the preposition "across". Some of the 5 configurations have a three-way distinction between motion, target object's disposition and vantage point, giving a total of 10 definitions. Figure 2.30 shows the senses of "across" according to [Herskovits, 1997]. From the ten definitions that are proposed, we are only interested in the senses that involve target's disposition, since it is the only one that takes place in the context of satellite image interpretation. However, in [Herskovits, 1997] only the definitions dealing with the motion sense are discussed. The first configuration is the one concerning the notion of going from "side to side". Additionally the line's path must be close to a straight line. The second configuration takes place when the reference object is an unbounded region and the linear object has salient segments of a straight line. For instance, "they walked across the sand for hours". The third configuration refers to a reference object with an intrinsic directionality, and the target object "goes across" if it

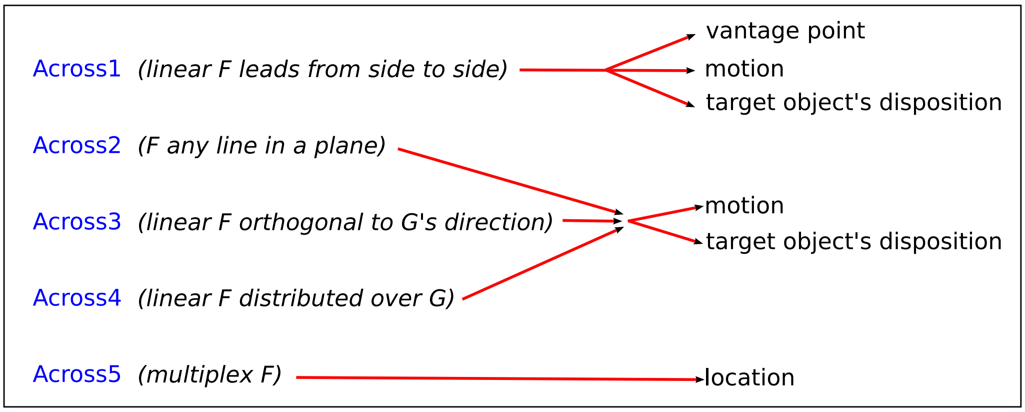


Figure 2.30: The senses of across according to [Herskovits, 1997]. Figure adapted from [Herskovits, 1997].

moves in an orthogonal direction to it. For example, "swimming across the flow". The fourth and fifth configurations refer to the relation "all over", in "across" the target object is a linear object distributed all over the reference object. For example, "For a whole year, I traveled across Colombia". In the fifth configuration the target object is a set of "points". Not all the configurations can be used in the sense of target's position in image interpretation. Additionally, we are interested in studying the relation "go across", which refers to the French verb "traverser". Only the first configuration is applicable to the our context. The third one involves the motion of the reference object, which is not possible in the case of satellite images. The fourth and fifth configurations refer to the relation "all over" which is out of the scope of the relation that we want to study, and differs from the French meaning. The second definition proposed by [Herskovits, 1997] means "à travers" in French, which is also different from the relation we want to study.

The relation "across" has several definitions, and concerns notions which are imprecisely defined. Therefore, to better understand the relation, we conducted an experiment with several volunteers.

The experiment

To understand the usual perception of the considered relation, 8 line-region configurations were proposed to 32 French-speaking persons. The configurations are shown in Figure 2.31. Each configuration had a different shape and the line had a different trajectory. The persons were asked if they agree or disagree with the statement "the line goes across the region"[5]. Some space was left for comments.

Results

The obtained results were very different across the subjects, showing that it is difficult to achieve a unique consensus for this relation. This confirms the idea that it is more appropriate to evaluate it as a matter of degree rather than a crisp relation.

There were 14 persons, who agreed with the statement in all the eight cases. These persons made comments like "for me, to go across is to enter and then to exit", "In all the 8 cases the road enters into the region and then exits".

[5]equivalent to "La ligne traverse la région"

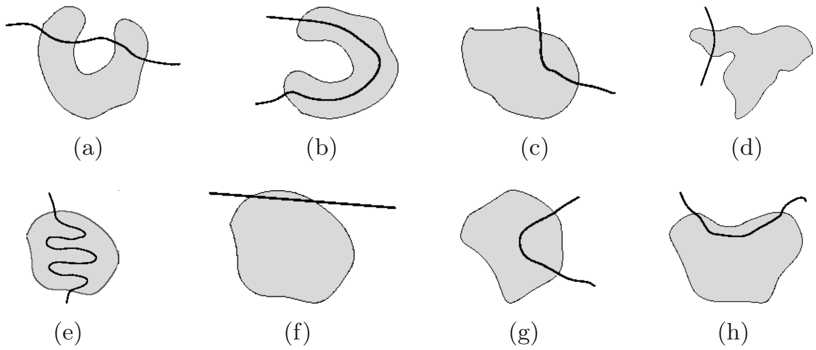


Figure 2.31: Configurations used in the experiment.

To analyze the answers we observed the answer for each question separately, taking into account the comments made by the subjects. Table 2.3 shows the results for each configuration with their respective comments.

From the results we observe a three typed meaning for the relation "go across".

(i) The line enters and then leaves the region.

(ii) The points on which the line enters and leaves region are located on "opposite" sides.

(iii) The line enters and leaves the region and its trajectory goes deeply into the region.

The first meaning comes from the fact that 43.8% of the persons answered this way. The second meaning was obtained from the acceptance of the configurations of Figures 2.31(a), 2.31(d) and 2.31(e), and the disagreement of 2.31(b) and 2.31(g). The third meaning is obtained because of the agreement with Figure 2.31(g) and the disagreement with Figures 2.31(c), 2.31(f) and 2.31(h). Figure 2.31(g) has a high acceptance of going across, even if the trajectory of the path is not straight, contrary to what was suggested by [Herskovits, 1997].

The first meaning is very permissive. It only takes into account the topological aspects of the relation. However, taking into account the remarks of [Talmy, 1983] and [Mark and Egenhofer, 1994a], the relation "to go through" and "to go across" are very similar. In some dictionaries they are even used as synonymous. However, "go through" relies on topological aspects, it refers to entering and leaving a medium [Talmy, 1983], while "go across" or "cross" have a more geometrical aspect: go from one side to the other. Therefore we are going to call the first model "to go through". The other two meanings are called "to go across (i)" and "to go across (ii)". It is not surprising to specify the relation using different definitions, since as we saw in the work of [Herskovits, 1997], this relation can have more than one sense.

"Go through" model

This model is based on topological constraints. It verifies that there is an intersection between the interior of the region and the linear object, and that the linear object does not start nor ends inside the region. This definition is similar to the one proposed by [Mark and Egenhofer, 1994a], the only difference is that we admit that the extremities of the linear object are on the edge of the region. The degree to which L "goes through" R is

Figure	Agree %	Agree Comments	Disagree Comments
2.31(a)	87	• The line goes across two times.	
2.31(b)	90	• It is acceptable, due to the shape of the region. • The areas on the right and on the left of the line are almost the same.	• The line starts and ends on the same side. • The entering and exiting vectors are not collinear.
2.31(c)	74	• I agree even if the region is not split into two regions of comparative size. • Although the part that is crossed by the line is small, the line touches two segments that are almost opposite.	• The line does not meet the center of the region. • It superficially goes across. • The line passes but does not go across.
2.31(d)	58	• The line touches two segments which are almost opposite, even if the part that is crossed by the line is small. • The line goes across an arm of the region.	• It superficially goes across. • The line does not meet the center. • The line goes across a part of the region, but does not go across the whole region. • The line passes but it does not go across.
2.31(e)	94	• It goes across even if it is not straight. • The line goes from one end to the other despite the sinusoidal movement.	• The line explores the region.
2.31(f)	58		• It superficially goes across. • The line goes along or tangent to the region. • The line passes by the region. • The part that is crossed by the line is not significant enough.
2.31(g)	65	• It is similar to case 2.31(c) but the line enters more deeply into the region. • It is sufficient at sight, even if it does not split the region into comparable parts.	• It does not go across since the line does not arrive at the other side of the region. • The entering and exiting vectors are not collinear.
2.31(h)	52		• It superficially goes across. • The line passes, or goes along the region. • The line does not introduce enough into the region. • The part that is crossed by the line is not significant enough.

Table 2.3: Table showing the answers and the comments of the questionnaire.

given by:

$$\mu_{go\_through}(L, R) = t\left(\mu_{int}(L, R^\circ), \mu_{\neg int}(L_a \cup L_b, R^\circ)\right). \tag{2.53}$$

where R° is the interior of R.

This definition is composed of a conjunction, expressed as a t-norm, of a degree of intersection μ_{int} of L and R° and a degree of non intersection $\mu_{\neg int}$ between the endpoints of L and R°. For the degree of intersection and non intersection we used the definitions introduced in Section 1.3.1.2. R° can be expressed as a fuzzy set, in order to take into account the imprecision on defining the boundaries of the region. This allows a more flexible definition. Using fuzzy mathematical morphological operators, the membership function of R° denoted as μ_{R° is given by:

$$\mu_{R^\circ}(x) = E_\nu(\mu_R)(x) = \inf_{x \in \mathfrak{I}} T\left[c(\nu(y-x)), \mu_R(y)\right] \tag{2.54}$$

where T is a t-conorm, c is a fuzzy complementation function, ν is a fuzzy structuring element and μ_R represents the membership function of R (when R is crisp μ_R is the indicator function of R). The structuring element ν is chosen in a way that it represents the imprecision that we want to model. For example, it can be chosen to represent a fuzzy neighborhood, as in the definition of fuzzy adjacency (Equation 1.11).

"Go across" from one side to the other

For the "go across (i)" relation the "go through" relation should be satisfied and in addition the linear object should go from one side to the other one. There are three approaches for defining what is meant by opposite sides using the studies developed in the area of spatial cognition. The first approach is the one from [Talmy, 1983], where the opposite sides of a region are defined only when the region is ribbonal. In that case the opposite sides are the two parallel sides that define the region. Another approach is based on the studies of Williams cited in [Landau and Jackendoff, 1993]. An adult considers that a linear object "goes from one side to the other", if it goes through two different segments of the boundary region. These two segments can be the consecutive sides of a rectangle or two segments of the boundary of a circle. This definition is more permissive than the one suggested by Talmy, and agrees with the acceptance of the relation of Figure 2.31(c) in the experiment. Finally, [Herskovits, 1997] proposes a definition to determine whether two points of the boundary of a regular region, without concavities, are on "opposite sides". The two points are on "opposite sides" of the region if the angle between the tangents is 180 degrees (when all the tangents are oriented counterclockwise). As the angle decreases, the relation is no longer satisfied.

Using the same idea as in [Herskovits, 1997, Landau and Jackendoff, 1993], we define the degree to which two points of the boundary of a region are on opposite sides. Let $p_1, p_2 \in \partial R$, and $\theta_1, \theta_2 \in [0, 2\pi]$ be the orientation of the vector defining the tangent of the boundary at p_1 and p_2, respectively. Then the degree to which p_1 and p_2 belong to opposite sides is given by:

$$\mu^R_{opposite\_sides}(p_1, p_2) = f(|\theta_1 - \theta_2|), \tag{2.55}$$

where f is a continuous function such that $f : [0, 2\pi] \to [0, 1]$, $f(0) = f(2\pi) = 0$ and $f(\pi) = 1$. The value $f(|\theta_1 - \theta_2|)$ represents the tolerance of considering θ_1 and θ_2 as "approximately" opposite. This definition is appropriate when the region is convex. However, when the region is not convex then we can accept cases as the one shown in Figure 2.32. In this example, the line does not go across since intuitively it does not go from one side to the opposite one, but according to Equation 2.55 the intersection points of the line with the boundary of the region are on opposite sides. When objects have concavities we use

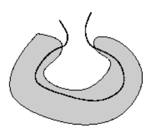


Figure 2.32: Inconvenient of Equation 2.55 for objects with concavities.

the convex hull of the object to determine if the line goes from one side to the other one. The convex hull of a region is the intersection of the minimum set of half planes that include the object. In the 2D space, if p_1 or p_2 are in a concavity, then there is a linear segment which closes that concavity. We use the orientation of that segment, instead of the orientation of the tangent at the point to determine if the points are on opposite sides. We refer to this segment as the closest segment of $CH(R)$ to the points. Let $p \in \partial R$, to find the closest segments of $CH(R)$ to p we represent $CH(R)$ as a polygon. If p coincides with one of the vertices of the polygon, then there are two segments. These segments correspond to the two segments adjacent to the vertex. Else if p is not a vertex of the polygon, then there is only one segment close to it. This segment corresponds to the segment which bounds the portion of ∂B that contains p. In the following we formalize the concept of closest segment in the case of a 2D crisp region.

Let R be a 2D region, and $CH(R)$ be its convex hull. Let V and E be the set of vertices and edges corresponding to the polygon which represents $CH(R)$. $V = \{C_1, \dots, C_n\}$ and $E = \{(C_i, C_{i+1}) | i = 1, \dots, n-1\} \cup (C_n, C_1)$, where $C_i \in \partial R$ for $i = 1, \dots, n$. For $p \in \partial R$ the closest segments of $CH(R)$ to p are defined as:

(i) If $p \in V$:

- If $\exists i \in]1, n[$ such that $p = C_i$. Then the closest segments are: (C_{i-1}, C_i) and (C_i, C_{i+1}).
- If $p = C_1$, then the closest segments are: (C_n, C_1) and (C_1, C_2).
- If $p = C_n$, then the closest segments are: (C_{n-1}, C_n) and (C_n, C_1).

(ii) If $p \notin V$.
Let $a, b \in \mathbb{R}$ and $\phi(t) : [a, b] \to \mathbb{R}^2$ be the continuous curve that describes ∂R, such that $\phi(a) = C_1$ and $\phi(b) = C_1$. Let $t_1 \in]a, b[$ such that $\phi(t_1) = p$. Define $i = \min\{j \in [1, n] : t < t_1 \text{ and } \phi(t) = C_j\}$:

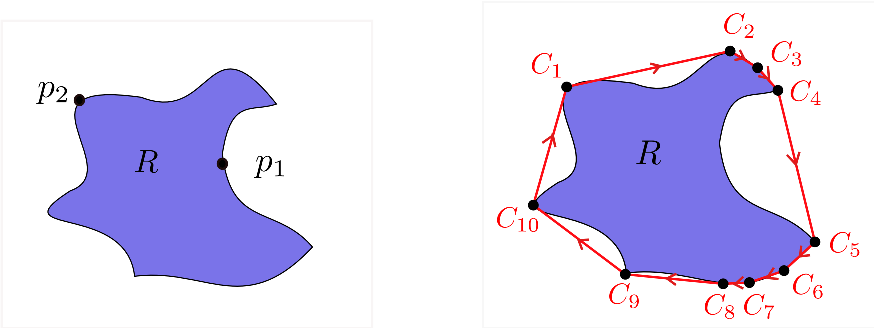
- If $i \in [1, n[$, then the closest segment is: (C_i, C_{i+1})
- If $i = n$, then the closest segment is: (C_n, C_1)

Figure 2.33 shows an illustration of the definition of closest segments. Figure 2.33(a) shows a region R and two points p_1 and p_2 on its boundary, for which we want to find the closest segments. Figure 2.33(b) shows the convex hull of R represented as a polygon, with vertices $V = \{C_1, \dots, C_{10}\}$. For the point p_2 the closest segments are (C_{10}, C_1) and (C_1, C_2) because $p_2 = C_1$. The point p_1 only has one closest segment, corresponding to segment (C_4, C_5).

To determine whether two points on the boundary of a region are on opposite sides we observe the difference of orientation of their closest segments of $CH(R)$. When one of the points has more than one closest segment, then the degree of opposite sides is calculated as the disjunction of the different orientations of all its segments with the segments of the other point. Let $p_1, p_2 \in \partial R$, and $\{\theta_{1_i}\}_{i=1}^{N_1}, \{\theta_{2_j}\}_{j=1}^{N_2}$ be the orientations of the closest segments of $CH(R)$ to p_1 and p_2 respectively, where N_1 and N_2 are the numbers of closest segments to p_1 and p_2. Then the degree to which p_1 and p_2 belong to opposite sides of R is given by:

$$\mu^R_{opposite\_side}(p_1, p_2) = T_{i \le N_1, j \le N_2} f(|\theta_{1_i} - \theta_{2_j}|). \tag{2.56}$$

where f is the same as in Equation 2.55 and T is a t-conorm.



(a) Region and points

(b) In red polygon representing $CH(R)$

Figure 2.33: Illustration of definition of closest segments of $CH(R)$ to p_1 (segments (C_1, C_2) and (C_{10}, C_1))and p_2 (segment (C_4, C_5)).

The degree of satisfaction of the relation "L goes across R" using the meaning of going from one side to the opposite one is defined as:

$$\mu_{go\_across1}(L, R) = t\left(\mu_{go\_through}(L, R), \mu^{R}_{opposite\_sides}(p_1, p_2)\right) \tag{2.57}$$

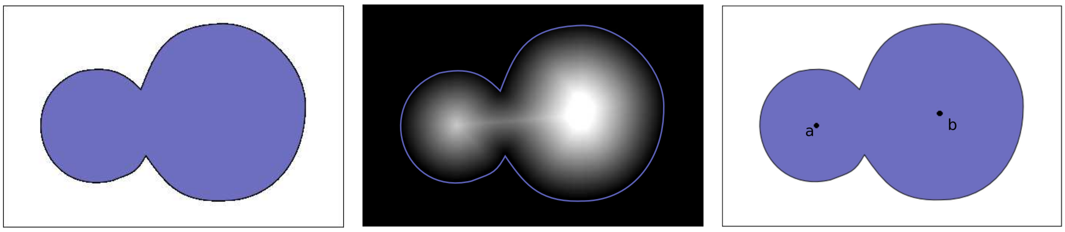
where p_1 and p_2 are the entering and exiting points.

This definition is adequate to represent the membership function of going across in the sense of going from one side to the other. This definition takes into account the shape of R. However, in the case where the region has a significant linear elongation, the sides are distinguished from the ends [Landau and Jackendoff, 1993]. Moreover, the relation only involves the sides, and if the linear object goes from one end to the other end of the region, then it is more appropriate to describe the relation as going along.

"Go across" model (ii)

A first approach to the idea of going deeply into a region is to think of passing near the center of mass of the region, like in [Tellex and Roy, 2006]. However, the center of mass can be outside the region, as in Figure 2.31(b). Therefore, it is more convenient to perceive how deep does the line "go into" a region by observing how far it is from the boundary. Hence, if we construct the distance map of the complement of the region, then the point which has the maximum value in the distance map is the deepest point of the region. Therefore, it is better to consider the ultimate erosion points (UEP) rather than the center of mass. The UEP are defined as the local maxima of the distance map in the interior of the region. To determine how deeply a line goes into a region we will measure how close it passes to one UEP. A region can have several UEPs. For instance Figure 2.34 shows a region and its UEPs. The UEP labeled with a is closest to the boundary of R than the one labelled c. Thus it is necessary to also consider how deep is the UEP according to the maximum possible depth.

First, we define a measure to determine whether L goes deep into R. Let $S = \{p_1, \ldots, p_n\}$ be the set of UEPs of the region R. Let $Vor(S)$ be the Voronoi partition associated to S. Let J be the set of indices of the Voronoi polygons that are intersected by the line. Let DM be the distance map defined in the interior of the region, where for each $p \in R$, $DM(p) = dist(p, R^c)$. Then the degree to which the line goes deeply into the region is given by:



(a) Original object (b) Internal distance map (c) Ultimate erosion points

Figure 2.34: Ultimate erosion points. For visualization purposes we dilated the ultimate erosion points.

$$\mu_{deep}(L,R) = \max_{j\in J}\left\{ g\left(\frac{DM(p_j) - dist(L,p_j)}{DM(p_j)}\right) g\left(\frac{DM(p_j)}{M_{max}}\right)\right\} \tag{2.58}$$

where M_{max} is the maximum value of the distance map for all the UEPs, that is $M_{max} = \max_{p_i\in S} DM(p)$. g is an increasing function from $[0,1]$ to $[0,1]$. The role of function $g(\frac{x}{y})$ is to evaluate to which degree the quantity x is almost equal to y, by observing its ratio. Equation 2.58 is composed of two terms. The first one indicates how close does the line pass to a UEP. The second one measures if the chosen UEP is close to the deepest point of the region. We decided to use two terms since we want to evaluate the two conditions. They are not grouped into one term because the function g might not be necessarily linear.

The degree of satisfaction of the relation "L goes across R" using the meaning of going deeply into the region is then defined as the conjunction of the relations "going through" the region and of the line "going deep" into the region:

$$\mu_{go\_across2}(L,R) = t\left(\mu_{go\_through}(L,R), \mu_{deep}(L,R)\right) \tag{2.59}$$

To avoid situations where the line goes deeply into the region and returns basically using the same path we can combine in a conjunctive way Eq. 2.59 with Eq. 2.56, or measure that the angle made between the entering point, the closest UE and the leaving point is not acute.

When R is a fuzzy set, it can be envisaged to build the UEPs as fuzzy sets by using a morphological approach. Further we can consider a fuzzy Voronoi partition as the one proposed in [Jooyandeh et al., 2009] which is adapted to fuzzy objects. And then think of a fuzzy distance measure that can help us to quantify the degree to which the line passes close to a UEP and how this UEP is close to the deepest part of the region. A review on fuzzy distance measures is given in [Bloch, 1999]. This is just a possibility for extending the relation, however it has to be further investigated.

2.4.2.1 Computational complexity

Let N denote the number of points in the image, N_R the number of points in the region, $N_{\partial R}$ the number of points in the boundary of R, and N_L the number of points in the line. The complexity of computing $\mu_{go\_through}(L,R)$ is $O(N_L + N_k N_R)$, where N_k is the size of the structuring element used to compute the interior of R. The complexity of computing $\mu_{go\_across1}(L,R)$ is $O(N_L + N_k N_R + N_{\partial R} \log N_{\partial R})$. The most consuming computation step in $\mu_{go\_across1}$ is the computation of the convex hull which is of the order of $O(N_{\partial R} \log N_{\partial R})$[6].

The computation of $\mu_{go\_across2}(L, R)$ is obtained through several steps: first the computation of $\mu_{go\_through}(L, R)$ which is $O(N_L + N_k N_R)$; then the complexity of computing the distance map is of $O(N)$[7]. The complexity of computing the UEPs using the distance map is $O(N N_V)$, where N_V is the size of the neighborhood that is used to determine the regional maxima of the distance map, where $N_V = 4$ or 8. Computing the Voronoi partition from the UEPs is of the order of $N_u \log N_u$[8], where N_u is the number of UEPs. Moreover, the maximum number of UEPs in an object with N_R points, is less or equal than N_R. Therefore the Voronoi computation is $O(N_R \log N_R)$, and finally finding the polygons of the Voronoi partition which intersect L is $O(N_L)$. Hence, the computational complexity of $\mu_{go\_across2}(L, R)$ is $O(2N_L + N_R(N_k + N_V + \log N_R) + N_{\partial R})$.

2.4.2.2 Discussion

We proposed three fuzzy models to represent the natural language relations of "to go across" and "to go through". These models are based on the results of a human subject test. They consider the perception of the relation. The three models are invariant with respect to geometrical transformations.

Depending on the form of the region or on the context, it can be more appropriate to use one of the models rather than another one. Subjects use different criteria to evaluate the degree depending on the region. Therefore, the question is to know which model to use in which situation. Unfortunately, this question remains open, and it should be necessary to perform more experiments to help answering this question, but this is out of the scope of this thesis.

We briefly discussed some extensions of these relations to the case where R is a fuzzy set, but there is still a lot of work to do on this subject.

We performed an evaluation of the three models for the synthetic figures used in the questionnaire and compared them with the human subject tests. The results are shown in Table 2.4. For these experiments we used the following function:

$$f(x) = \begin{cases} 0 & \text{if } x < \frac{\pi}{4} \\ \frac{4x-\pi}{2\pi} & \text{if } \frac{\pi}{4} \leq x < \frac{3\pi}{4} \\ 1 & \text{if } x \geq \frac{3\pi}{4} \end{cases} \tag{2.60}$$

in Equation 2.56 to quantify when two angles are considered in opposite directions. For the function g in Equation 2.58 we used:

$$g(x) = \begin{cases} 1 & \text{if } x \geq 0.8 \\ \frac{x-0.5}{0.3} & \text{if } 0.5 \leq x < 0.8 \\ 0 & \text{if } x < 0.5 \end{cases} \tag{2.61}$$

We can see how the three models coincide with a high membership in the configurations which had a high acceptance for most observers (Figures 2.31(b) and 2.31(e)). The "go through" model has a satisfaction degree of 1, since in all the cases the line enters and the

[6] Using qhull algorithm [Barber et al., 1996]
[7] Using Danielsson's algorithm [Danielsson, 1980]
[8] Using Fortune's algorithm [Fortune, 1987]

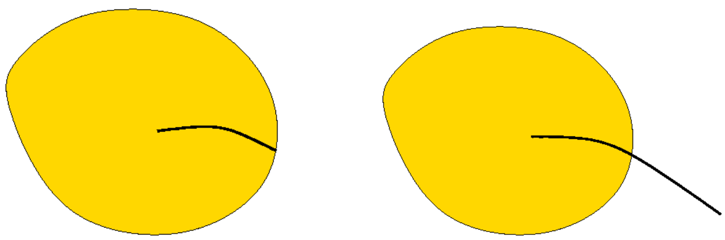
exits. The "go across 1" model has high membership values for the cases 2.31(a), 2.31(d) and 2.31(e), this coincides with the experiment that we performed. In a similar way, the model "go across 2" has a high membership value for the cases 2.31(e) and 2.31(g), as expected. More examples of these relations are given in Section 2.4.4.

Model	Figure 2.31(a)	Figure 2.31(b)	Figure 2.31(c)	Figure 2.31(d)	Figure 2.31(e)	Figure 2.31(f)	Figure 2.31(g)	Figure 2.31(h)
"go through"	1	1	1	1	1	1	1	1
"go across (i)"	1	0.2	0.33	1	1	0.33	0.2	0.53
"go across (ii)"	0.7	1	0.68	0.32	1	0	1	0
"human-subject test"	.87	0.9	0.74	0.58	0.94	0.58	0.65	0.52

Table 2.4: Results of the three models evaluated on the 8 configurations of Figure 2.31. Where "human-subject test" refers to the percentage of subjects which answered "yes".

2.4.3 "Enter"

In [Shariff et al., 1998] the relation "enters" is also considered, producing two different models. The topological constraints satisfied by the first model are that the interior of the line should not intersect the exterior of the region, and that the boundary of the linear object should intersect the boundary of the region. In the second model, the linear object should intersect the boundary of the object and the interior of the linear object should intersect the interior of the object. Additionally, there should be a large distance between the extremity of the line that is inside the region and the boundary of the region. This is measured using the inner nearness measure (see Figure 2.29), which is the ratio between the area of the eroded region until it reaches the end point of the line and the area or the region. Also, the line cannot be completely inside nor outside the region. This suggests that the line should go deeply into the region. Figure 2.35 shows two examples of the definitions proposed in [Shariff et al., 1998].



(a) Model 1 (b) Model 2

Figure 2.35: Example of models of the relation "enters" proposed by Shariff et al. [1998].

As for the relation "go across" we can think of two meanings for the relation "enters". The first one is based on topological constraints and the second one is more restrictive than the definition proposed by [Shariff et al., 1998].
A first meaning is based on the idea that, to enter a region, a line should go from the exterior or the boundary of the region to its interior. A second meaning is that the line should also satisfy the toplogical constraints but additionally it should penetrate the region. This second meaning is useful since it helps us to describe, for instance, the relation between artificial man made structures constructed in sea (See Figure 2.36(a)) or

the relation between the boarding bridges of an airport's terminal and the parking zone (see Figure 2.36(b)).

(a)

(b)

Figure 2.36: Structures that can be described using the relation "enters". (a) Artificial construction "entering" the sea. (b) Boarding bridges "entering" the parking zone.

Our proposed meanings are very close to those proposed by [Shariff et al., 1998]. To be consistent with the work performed for the relation "go across", this definition is based on the functions defined in Section 2.4.2. The second proposed meaning coincides better with the relation "go into" rather than with the relation "enter". Therefore, in the following paragraphs we will refer to the relation "go deep into" for this second meaning.

The "enters" model

The first meaning needs to satisfy the two following topological constraints: one of the extremities of the linear object should not intersect the interior of the region, and the interior of the linear object should intersect the interior of the object. These conditions are more general that the ones proposed by [Shariff et al., 1998], since it does not require that the line ends inside the region. This relation is formalized as:

$$\mu_{enter}(L, R) = t(\mu_{int}(L, R^\circ), T\left(\mu_{\neg int}(L_a, R^\circ), \mu_{\neg int}(L_b, R^\circ)\right)) \tag{2.62}$$

As for the relation "go through", the interior of the region can be computed as a fuzzy region. This definition is well adapted to fuzzy regions.

The "go deep into" model

We consider two possible cases for this model: when the region and the linear object have comparable dimensions, and when the region has an infinite size compared to the linear object. To determine whether the region R has an infinite size compared to the line L, we compute the ratio between the length of the linear object and the length of the minimal bounding box of R. If this ratio tends to zero, then R has an infinite size with respect to L, otherwise, we would say that they have comparable sizes.

For the first situation, the relation "go into" the region has the same sense as going deeply into the region for the model (ii) of the relation go across (see Section 2.4.2). Therefore it is directly computed as:

$$\mu_{go\_deep\_into}(L, R) = t(\mu_{enter}(L, R), \mu_{deep}(L, R)) \tag{2.63}$$

This definition is composed of two parts, one is the satisfaction of the "enter" relation and the second one measures how deep into the region does the line reach.

For the second situation the degree to which a linear object L "goes into" R is related to how far away from the boundary is the extremity of L with respect to where it can possibly be. Let DM be the distance map defined in the interior of the region, where for each $p \in R$, $DM(p) = dist(p, R^c)$. Let D_{max} be the maximum distance from the boundary of R reached by L in the interior of R, that is $D_{max} = \max_{x \in L}\{DM(x)\}$. We define X_{deep} as the set of points of L which reach D_{max}:

$$X_{deep} = \{x \in L | DM(x) = D_{max}\}. \tag{2.64}$$

Then the degree of satisfaction of the relation "L goes into R" is obtained by measuring how straight is the trajectory from the entering point to the points in X_{deep}:

$$\mu_{gointo}(L, R) = \min_{x \in X_{deep}} h\left(\frac{l(x, p_0)}{D_{max}}\right), \tag{2.65}$$

where $l(a, b)$ for $a, b \in L$ measures the length of the line between the points a and b, and p_0 is the entering point. If there are several points satisfying this condition, then Equation 2.65 is tested with the point that is the closest (according to the curvilinear axis) to the end point of the line that is in the interior of R. h is a real increasing function such that $h : \mathbb{R} \to [0, 1]$ such that $h(0) = 0$ and $h(1) = 1$. The objective of function h is to measure how straight the trajectory of the line is. Figure 2.37 shows an illustration of the measures involved in this definition.

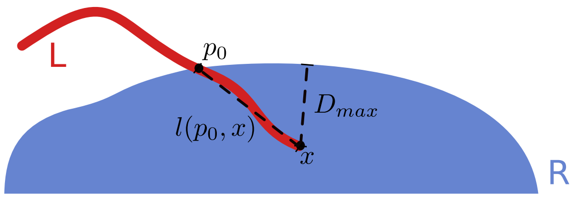
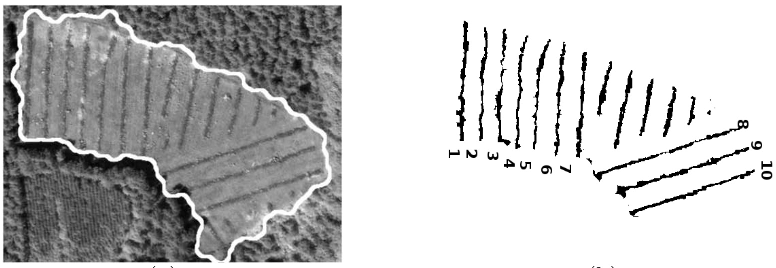


Figure 2.37: Definition of going deep into when the region has an infinite size with respect to the linear object. x represents the point in X_{deep} and p_0 is the entering point.

2.4.4 Illustrative example

The above presented relations were evaluated in two cases, one for ten different paths and a region (Figure 2.38) and the other one for 6 paths and a region (Figure 2.39, where paths are named using their starting and ending points). For the function f and g of Equations 2.56 and 2.58, we used the same functions as in Equations 2.60 and 2.61.

The results are presented in Table 2.5 and in Table 2.6, respectively. In both cases, we observe that the obtained results agree with the perception of the relations. The relation "to go through" verifies the intersection with the interior and the boundaries, and the greater degrees were obtained for the paths that started and ended at the border of the region. Higher values were obtained for the degree of satisfaction of the relation "to go deep into" as the paths went deeply into the region (paths $4-8$ in Figure 2.39 and all the paths in Figure 2.39 except path BC). The notion of opposite sides fits with the intuition. Points situated on "opposite sides" obtained high membership values. The results reflect the need of using two different definitions for the relation "to go across" since in ambiguous cases (for example paths 1 and 10 in Figure 2.38 or path AB in Figure 2.39) it is not possible to reach a consensus.



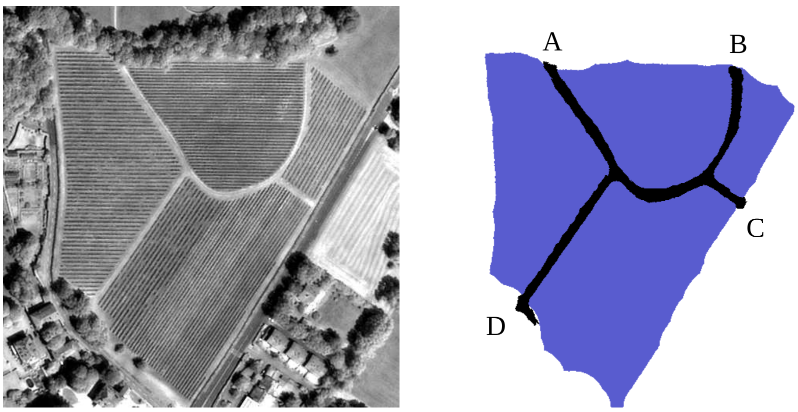
(a) (b)

Figure 2.38: Image for evaluating the “go across”, “go into” and “enter” relations. (a) Original image and region. (b) Ten quasi-linear paths.

Path	μ_{enter}	$\mu_{go\_deep\_into}$	$\mu_{go\_through}$	$\mu_{go\_across1}$	$\mu_{go\_across2}$
1	1.00	0.14	1.00	0.92	0.14
2	1.00	0.53	1.00	1.00	0.53
3	1.00	0.96	1.00	1.00	0.96
4	1.00	1.00	0.00	0.00	0.00
5	1.00	1.00	1.00	1.00	1.00
6	1.00	1.00	1.00	1.00	1.00
7	1.00	1.00	1.00	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00
9	1.00	0.97	1.00	0.97	0.97
10	1.00	0.50	1.00	0.97	0.50

Table 2.5: Results of the relations’ models evaluated on the paths and region of Figure 2.38.

From Table 2.5 we can observe that for all the paths the relation “enter” is satisfied, since they intersect at least one point of the boundary of the region. The paths that are closer to the borders of the region (paths 1, 2 and 10) have lower degrees of satisfaction for the relation “go into”, and therefore have a low value for the relation “go across 2”. All the paths that “go through” the region have a high degree of satisfaction for the relation “go across 1”, since all the paths go from one side to the other.



(a) (b)

Figure 2.39: (a) Original image. (b) Paths and region.

Path	μ_{enter}	$\mu_{go\_deep\_into}$	$\mu_{go\_through}$	$\mu_{go\_across1}$	$\mu_{go\_across2}$
Path AB	1.00	1.00	1.00	0.00	1.00
Path AC	1.00	1.00	1.00	0.86	1.00
Path AD	1.00	1.00	1.00	0.95	1.00
Path BC	1.00	0.53	1.00	0.86	0.53
Path BD	1.00	1.00	1.00	0.95	1.00
Path CD	1.00	1.00	1.00	1.00	1.00

Table 2.6: Results of the relations' models evaluated on the paths and region of Figure 2.39.

2.4.5 Comments about "along" and "go around"

The relation "along" was studied in [Takemura et al., 2005] where the target and the reference regions can have any shape, in particular a line. One particular case of the relation along is the parallel relation, that was studied in Section 2.3.
For the case where L and R are crisp and L is not rectilinear, we use the definition proposed by [Takemura et al., 2005]. To compute the degree to which L goes along R, we first compute the region:

$$\begin{aligned}\beta_t = \bigcup \{[a,b], a \in L, b \in R, D_{LR}(a) < t, D_{LR}(b) < t \\ \text{and }]a,b[\text{ is admissible with respect to } L \text{ and } R\ \}\end{aligned} \tag{2.66}$$

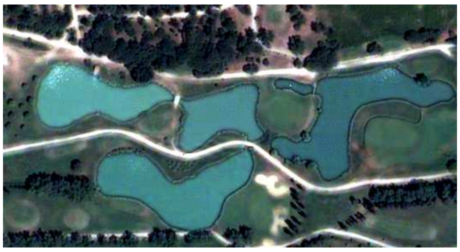
where $D_{LR}(x) = d(x,R) + d(x,L)$, and $d(x,R)$ is the Euclidean distance between the point x and R. To definition of admissible segments is given by Equation 1.8 . The region β_t corresponds to the region between L and R and the segments of L and R which are adjacent to the region between them, these segments are called admissible arcs. To construct the region β_t only the points which are at a distance inferior to t from L and R. To determine the degree to which L is "along" R, we have to measure the degree of elongation of β_t. Several measures of elongation are proposed in [Takemura et al., 2005]. Among the proposed measures, there is one measure that considers that the region should be elongated in the direction where L and R are adjacent, that is:

$$\mu_{along}(L,R) = f_a\left(\frac{l^2(\beta_t)}{S(\beta_t)}\right), \tag{2.67}$$

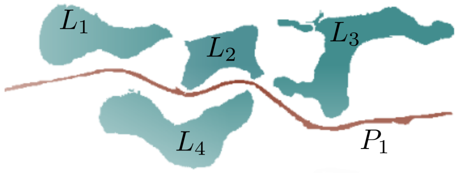
where $l(\beta_t)$ corresponds to the length of the admissible arcs of β_t, $S(\beta_t)$ corresponds to the area of β_t, and f_a is an increasing function, which tends to one when β_t is elongated. In [Takemura et al., 2005] a sigmoid function is proposed for f_a: $f_a(x) = \frac{1-\exp(-ax)}{1+\exp(-ax)}$. This measure can be used to evaluate the relation when only certain parts of L are along R. Table 2.7 shows the evaluation results of the relation "along" between the path labeled P_1 and the lakes L_1, L_2, L_3 and L_4 of Figure 2.40 when using $t = 30$ pixels $\simeq 21m$ and a sigmoid function as f_a with $a = 0.075$. Figures 2.40(c) and 2.40(d) show the regions β_t for the four situations. From the table we can see that in the four cases there is a high satisfiability value for the relation "along", which correctly represents the four situations.

See [Takemura et al., 2005] for an extension of this definition when L or R are fuzzy.

The "go around" relation is a particular case of the surround relation when the reference object is a linear object. Therefore, for the "go around" relation we can use the definitions



(a) Original image.



(b) Segmented objects.

(c) Regions β_t in gray and black, used for the evaluation of $\mu_{along}(P_1, L_1)$ and $\mu_{along}(P_1, L_3)$, respectevely.

(d) Regions β_t in gray, and black, used for the evaluation of $\mu_{along}(P_1, L_2)$ and $\mu_{along}(P_1, L_4)$, respectevely.

Figure 2.40: Illustration of the definition of the relation "along".

	L_1	L_2	L_3	L_4
$\mu_{along}(P_1, x)$	0.86	0.98	0.84	0.97

Table 2.7: Satisfaction values for μ_{along}.

introduced in Section 2.1. To determine if a line "goes around" a region, we must compute the region which is surrounded by L and the degree is obtained by comparing how well R matches the region defined by $\mu_{surround}(L)$. Depending on the size of the region with respect to the curves of L we can use either Equation 2.13 or Equation 2.10 to construct the fuzzy landscape.

2.4.6 Discussion

We have presented several models for the relations "go through", "go across", "enter" and "go deep into". The models of "enter" and "go through" just consider the topological constraints of the relations, while the models for the relations "go across" and "go into" take into account the shape of the object, and in the case of "go into" it gives different alternatives according to the dimensions of the line with respect to the region.

Our definitions satisfy the relations of subsumption between the relations "go across", "go through", "enter" and "go into" which can be organized in a tree as shown in Figure 2.41.

This hierarchical organization can be useful to reason with the relations, and it can be easily introduced in the hierarchy proposed in Figure 1.16.

The models presented in Sections 2.4.2 and 2.4.3 have topological constraints which are two-point based relations. However, when dealing with regions which cannot intersect, these relations can still be evaluated. Let R be the region composed of disjoint parts, and let L be the linear object with approximate width w. Therefore, it is possible to perform a

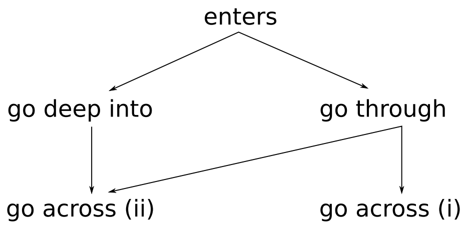
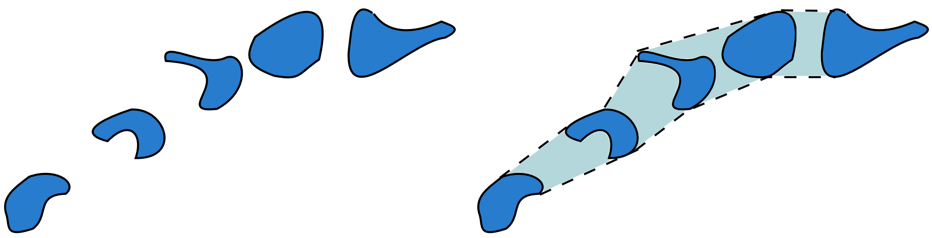


Figure 2.41: Organization of the line region relations.

morphological closing using an structuring element of size $\frac{w}{2}$ in order to construct a region $\tilde{R}$ which can intersect L. Then the evaluation can be performed using $\tilde{R}$ instead of R.

In the Section 2.2 we studied the alignment relation, followed by the parallel relation. When we studied the parallelism which involved a group of aligned objects, we were basically considering the group of aligned objects as if they were a line. This was highlighted by [Herskovits, 1997], "an alignment is an idealization of a line". Moreover, we would like to notice that sometimes the line-region relations that we presented can take place with a group of objects, for instance "the trees along the road", in that case the trees are seen as a group which can have the same behavior as line. In particular, if the road is straight then the group of trees forms an alignment which is parallel to the road. Other examples are "the houses that go around the lake", then the group of houses are organized in such a way that they form a ribbon, which goes around the lake. This property of being able to assimilate a group of objects as a line was defined as polymorphism in [Mathet, 2000], and it depends on the configuration of the group. Figure 2.42 shows an example of polymorphism of a group of points seen as a linear object. Therefore it is possible to evaluate the relations presented in this section with a group of aligned objects. To evaluate the relation between a group $\mathcal{A}$ and a region, we evaluate the relation between $\mathcal{A} \cup \beta_A$ as it was the case for the parallel relation (see Section 2.3).



(a) Group of objects. (b) Linear object.

Figure 2.42: Representation of polymorphism between a group of objects and a linear object.

2.5 Conclusion

In this chapter we have proposed original representations of the spatial relations: surround, alignment, parallelism and line region relations. As it was highlighted at the beginning of each section, each of these relations is frequently found on satellite images, and their modeling is interesting for satellite image interpretation.

For the definitions of all the relations we proposed fuzzy models which allow us to take into account the imprecision in their semantics. Moreover, for some of these relations we proposed or discussed how they can be extended to the case of fuzzy objects which allows considering the imprecision inherent to images and to the segmentation process.

The examples on real objects extracted from satellite images have shown the usefulness and power of the proposed models for scene understanding. We also highlighted how these relations can be used as intermediary steps for extracting objects in images.

In Chapter 1 and in this chapter, we have concentrated on the representations of the relations presented in Figure 1.16. Now that we have a model for each of these relations we can reason about them. In the next chapter we present some of the reasoning techniques and some examples using these relations.

Chapter 3

Spatial reasoning

In the previous two chapters we have presented a collection of spatial relations. In this chapter we discuss how to reason with these relations in the context of object recognition in images. Spatial reasoning deals with the processes used to combine, verify or infer new spatial information. For instance, let O be a set of objects, and R a set of binary spatial relations. To denote that $x \in O$ is the reference object of a relation $r \in R$, and $y \in$ its target object we write $r(x, y)$. For $a, b, c \in O$ and $r_1, r_2, r_3 \in R$, some of the questions which spatial reasoning aims to answer are:

(i) If $r_1(a, b)$ and $r_2(b, c)$, then what is the relation satisfied between a and c ?

(ii) Which are the objects $x \in O$ which satisfy $r_1(x, a)$ and $r_2(b, x)$ and $r_3(c, x)$. In the case where the relations are modeled as fuzzy ones, we can ask to what degree does the conjunction of the three relations is satisfied, or to find the object in O which has the highest degree of satisfaction for the three relations.

(iii) Are there objects $x, y, z \in O$ which satisfy $r_1(x, y)$ and $r_2(y, z)$ and $r_3(z, x)$? In the case where the relations are modeled as fuzzy ones, we can ask to what degree does these objects satisfy these relations.

(iv) When the spatial relations are modeled as fuzzy landscapes, we can ask which is the (possibly fuzzy) region of space which defines the area where r_1 is satisfied with a as reference object, and r_2 with b as reference object and r_3 with c as its reference object.

For simplicity, we have chosen only binary spatial relations and a combination of a maximum of three relations, in the above discussion. Nevertheless, spatial reasoning aims also at answering these questions for relations of a higher arity, and we can combine more than three relations. In the context of using spatial reasoning for object recognition these four questions can be used in different ways. The first question is useful for computational purposes, since by computing two relations we could have the information about a third relation, and therefore more information to help recognizing the objects of interest. The second and third questions are asked to detect inconsistencies or to complete incomplete information. The forth question is particularly useful when extracting objects in an image since it allows us to delimit a region of interest where it is possible to find an object.

According to the type of relations, the representation used and the question to be answered, there are different formalisms to perform spatial reasoning. We give a quick insight into the methods used for qualitative spatial reasoning in Section 3.1. These methods are briefly discussed since qualitative spatial relations are not included in our catalog of relations. In Section 3.2 we present the formalisms that have been used to reason with fuzzy spatial relations. Finally, in Section 3.3 we present two illustrative examples in object recognition reasoning with fuzzy spatial relations.

3.1 Qualitative spatial reasoning (QSR)

Qualitative spatial reasoning addresses the reasoning with qualitative relations. The reasoning is performed in a symbolic way without numerical computations. Qualitative spatial relations are ususally expressed in a crisp way often using logical formalisms. One of the most common examples of qualitative spatial relations are the RCC8 topological relations introduced in Section 1.3.1.1. In this section we only concentrate on the reasoning mechanisms introduced for the RCC8 relations. Nonetheless, other types of qualitative spatial reasoning have been developed for other types of spatial relations, for instance, [Freksa, 1992] for directional relations between objects, [R. Moratz and Freksa, 2003, Moratz and Ragni, 2008, Liu, 1998] for directional relations between points, [Guesgen and Albrecht, 2000] to reason with relative distance, [Clementini et al., 1997] for directional and distance relations. Qualitative spatial reasoning is used to answer questions (i), (ii) and (iii) discussed in the introduction.

A very common reasoning mechanism is to encode spatial relations using a logical formalism, and use the reasoning mechanism of the logical formalism. The RCC8 relations were constructed as an axiomatized first order logic theory, based on a single primitive: Connection (see Section 1.3.1.1). However, first order logic is undecidable, thus there is no effective method to answer questions (i), (ii) and (iii). Nevertheless, [Bennett, 1996, Cohn, 1993] encoded the RCC8 relations in propositional and modal logics which are decidable.

Representing RCC8 as a modal logic allows automatically constructing the composition table of RCC8 relations [Cohn et al., 1997]. Composition tables aim at solving the inference problem proposed by question (i). The RCC8 composition table is shown in Figure 3.1. Composition tables can also be obtained by representing the RCC8 as a lattice of relations [Randell et al., 1992], since this structure allows making inferences about the composition of relations. From the composition table, we can see that the composition of two relations is sometimes unknown, for instance from $DC(a,b)$ and $DC(b,c)$, then it is not possible to know what is the relation between a and c. For other relations there can be several possibilities, for instance if $EC(a,b)$ and $TPP^{-1}(b,c)$ then we can have that $DC(a,c)$ or $EC(a,c)$. Therefore, it is not always possible to answer question (i), nonetheless we can use composition tables to help reducing the possibilities for determining the relations between a and c.

To answer questions (ii) and (iii) the RCC8 relations are represented as relational algebra [Bennett et al., 1997] and QSR can be seen as a qualitative Constraint Satisfaction Problem (CSP) [Davis et al., 1999, Condotta and Würbel, 2007]. The local consistency is tested by using a path consistency algorithm based on the RCC8 composition table.

Other representations of the RCC8 relations is by using a Galois lattice [Le Ber et al., 2001]. This type of representation is further discussed in Section 4.2.3.3.

$\circ$	DC	EC	PO	TPP	NTPP	TPP^{-1}	NTPP^{-1}	EQ
DC	*	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC	DC	DC
EC	DC,EC PO,TPP^{-1} NTPP^{-1}	DC,EC PO,TPP TPP^{-1},EQ	DC,EC PO,TPP NTPP	EC,PO TPP NTPP	PO TPP NTPP	DC,EC	DC	EC
PO	DC,EC PO,TPP^{-1} NTPP^{-1}	DC,EC PO,TPP^{-1} NTPP^{-1}	*	PO TPP NTPP	PO TPP NTPP	DC,EC PO,TPP^{-1} NTPP^{-1}	DC,EC PO,TPP^{-1} NTPP^{-1}	PO
TPP	DC	DC,EC	DC,EC PO,TPP NTPP	TPP NTPP	NTPP	DC,EC PO,TPP TPP^{-1},EQ	DC,EC PO,TPP^{-1} NTPP^{-1}	TPP
NTPP	DC	DC	DC,EC PO,TPP NTPP	NTPP	NTPP	DC,EC PO,TPP NTPP	*	NTPP
TPP^{-1}	DC,EC PO,TPP^{-1} NTPP^{-1}	EC,PO TPP^{-1} NTPP^{-1}	PO TPP^{-1} NTPP^{-1}	PO,EQ TPP TPP^{-1}	PO TPP NTPP	TPP^{-1} NTPP^{-1}	NTPP^{-1}	TPP^{-1}
NTPP^{-1}	DC,EC PO,TPP^{-1} NTPP^{-1}	PO TPP^{-1} NTPP^{-1}	PO TPP^{-1} NTPP^{-1}	PO TPP^{-1} NTPP^{-1}	PO,TPP^{-1} TPP,NTPP NTPP^{-1},EQ	NTPP^{-1}	NTPP^{-1}	NTPP^{-1}
EQ	DC	EC	PO	TPP	NTPP	TPP^{-1}	NTPP^{-1}	EQ

Figure 3.1: Composition table of RCC8 relations. The symbol * stands for no information. Taken from [Renz, 2002]. (cf Figure 1.3 for definition of the relations.)

3.2 Reasoning with fuzzy spatial relations based on fusion operators

We can reason over fuzzy spatial relations to answer questions (ii), (iii) and (iv) discussed in the introduction. Reasoning with spatial relations in images addresses the problem of combining spatial information coming from several sources, which can be formulated as a fusion problem.

In spatial reasoning we use spatial reasoning to combine spatial information. For instance, if we want to answer question (iv) it is necessary to combine the fuzzy landscapes defining each of the regions which represent the regions where the relations are satisfied. Let $\mathfrak{I}$ be the image space. Suppose that the fuzzy landscape that defines the relation r_1 with reference object a is given by the membership function μ_{r_1} over $\mathfrak{I}$, and that μ_{r_2} is the membership function over $\mathfrak{I}$ defining the region of space which satisfies relation r_2 when b is its reference object. Then to determine the region of space where both relations are satisfied we can use a fusion operator F, such that the region of space satisfying both relations noted as μ_r is defined as $\mu_r(x) = F(\mu_{r_1}(x), \mu_{r_2}(x))$.

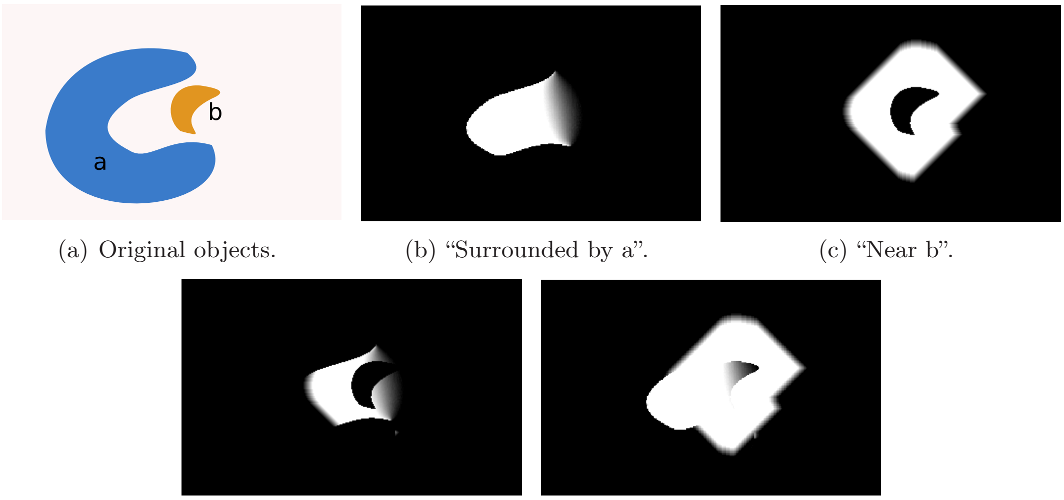
There are several possibilities for choosing F. The fusion operators can be classified into three classes according to their behavior [Bloch, 1996]:

Conjunctive operators: they satisfy $F(\alpha, \beta) \leq \min(\alpha, \beta)$ which correspond to a severe behavior.

Disjunctive operators: they satisfy $F(\alpha, \beta) \geq \max(\alpha, \beta)$ which correspond to an indulgent behavior.

Compromise operators: they satisfy $\min(\alpha, \beta) \leq F(\alpha, \beta) \leq \max(\alpha, \beta)$ which represent a cautious behavior.

In fuzzy logic, conjunctive operators are expressed as t-norms, disjunctive operators as t-conorms and there is a variety of choices for the compromise operators, for instance the arithmetic mean (see [Bloch, 1996] for a complete review of fusion operators).Other operators have a variable behavior, depending on the values of α and β. Figure 3.2 shows an example of fusioning the two spatial relations "Surrounded by a" and "Near b" using a conjunctive and a disjunctive fusion operators. The region of space with satisfies "Surrounded by a" and "Near b" is shown in Figure 3.2(d), which corresponds to the fusion using a conjunctive operator. The region which satisfies "Surrounded by a" or "Near b" is shown in Figure 3.2(d), which corresponds to the fusion using a disjunctive operator.



(a) Original objects. (b) "Surrounded by a". (c) "Near b".

(d) Fusion using min operator. (e) Fusion using max operator.

Figure 3.2: Illustration of fusion between spatial relations. (d) and (e) correspond to the fusion of the relations represented by (b) and (c), using a conjunctive and a disjunctive operator, respectively.

In images, fusion can be performed at object level or at pixel level. Figure 3.2 shows an example of fusion at pixel level. Given a set of objects represented by regions, we can perform fusion at object level. To determine for an object c the degree to which it satisfies a combination of two relations, we need to perform fusion between the degrees of satisfaction of each relation. For instance, the degree of satisfaction to which an object c satisfies a relation r_1 when a is the reference object, or a relation r_2 when b is the reference object, is given by:

$$\mu_{r_1^a \vee r_2^b}(c) = T\left(\mu_{r_1}(a, c), \mu_{r_2}(b, c)\right)$$

where T is a t-conorm representing a disjunctive operator. Two examples of fusion at object level are illustrated in Section 3.3.

Thus we can conclude, that fusion of spatial relations can be used to address questions (ii), (iii) and (iv) discussed at the beginning of the chapter. There are other tools to reason with fuzzy spatial relations, these tools also make use on fusion. For instance in [Bloch, 2006] the fuzzy spatial relations which are defined through mathematical morphology are encoded in modal logics using erosion and dilation as modal operators, this formalism allows to reason on the spatial relations. In [Nempont, 2009] reasoning is performed as a CSP, also involving fusion of spatial relations.

3.3 Illustrative examples: Spatial reasoning for object recognition

In this section we present two examples of how we can combine spatial relations in order to extract patterns on images. This combination is done through fusion of spatial relations, as described in Section 3.2.

The first example deals with urban patterns, in particularly, the detection of residential areas composed of organized houses. In these areas the following spatial relations are satisfied:

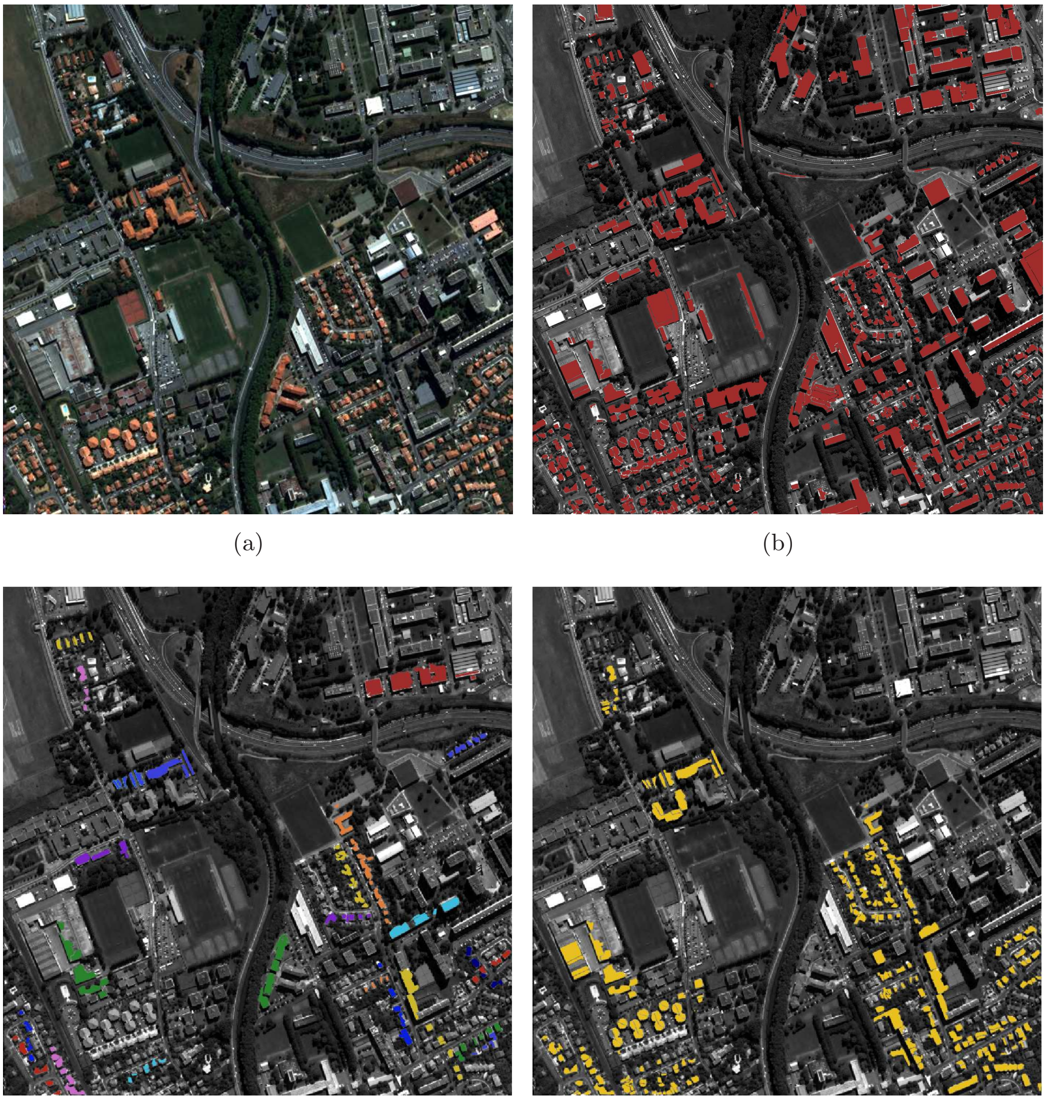
- Houses belong to aligned groups of houses.
- The aligned group of houses is parallel and near another group of aligned houses.

Therefore, to determine the residential areas we propose to determine the groups of aligned houses using the algorithm presented in Section 2.2, then for each group we evaluate the degree to which it is "parallel to" and "near" another group by using a t-norm to fusion the satisfaction degrees of both spatial relations.

The method for determining the *globally* aligned groups of objects was applied on the segmented buildings of Figure 3.3. The buildings were obtained by using the method described in [Poulain et al., 2009]. For the extraction we used a $\beta = 0.85$ and a Voronoi neighborhood constrained by a distance of 30 pixels equivalent to approximately 21 m, since we are interested in residential areas, where houses are usually close to each other. Some of the *globally* aligned groups of houses are shown in Figure 3.3(c). It is not possible to show all the aligned groups found by the algorithm since there are objects which belong to more than one group. The obtained groups contain few elements due to the small neighborhood used to extract them.

From the obtained *globally* aligned groups of houses we extracted the groups which are "parallel to" and "near" another group or which have a group "parallel to" and "near" it with a degree greater than or equal to 0.8. We modeled the "near" relation by constructing a fuzzy landscape using Equation 1.24 using $d_1 = 0$ and $d_2 = 50$ pixels $\simeq 35m$.
The groups of houses that are aligned, which are "parallel to" and "near" another group are shown in Figure 3.3(d). This example demonstrates that by making a conjunctive fusion of three spatial relations it is possible to find the pattern that we searched on the image.

The second example deals with the elimination of false detections of roads in urban areas obtained from a road extractor [Poulain et al., 2010]. Figure 3.4(a) shows the result of a road detection algorithm. Some of the obtained roads are false detections. To eliminate them, we use the following spatial knowledge: in residential areas, aligned groups of houses are "near" and "parallel to" roads. Thus, we first extract the groups of *globally* aligned houses and for each road we evaluate the degree to which a group of aligned houses is "near" and "parallel to" the road. However, some of the roads extracted in Figure 3.4(a)



(a) (b)

(c) (d)

Figure 3.3: (a) Original image. (b) Segmented buildings. (c) Some of the *globally* aligned subsets of houses found by the algorithm with a degree of alignment greater than 0.85. (d) Clusters of houses belonging to *globally* aligned groups which are "parallel to" and "near" other groups with a degree greater than or equal to 0.8.

correspond to roads segments, thus we are also interested in finding the roads which are "near" and "parallel to" an aligned group of houses. Therefore combine in a disjunctive way the degree to which a road is parallel and near an aligned group of houses and the degree to which a road has an aligned group of houses "parallel to" and "near" it. Notice that the two conditions are not the same due to the non symmetry of the "parallel to" relation.

Considering that the groups of aligned buildings which are parallel to roads do not have to satisfy the constraint of buildings being "near" each other, we extracted the groups of aligned buildings using $\beta = 0.85$ and a Voronoi neighborhood constrained by a distance of 70 pixels equivalent to approximately 49 m. Some of the obtained groups of aligned buildings are shown in Figure 3.4(b). If we compare the results of Figure 3.4(b) and of

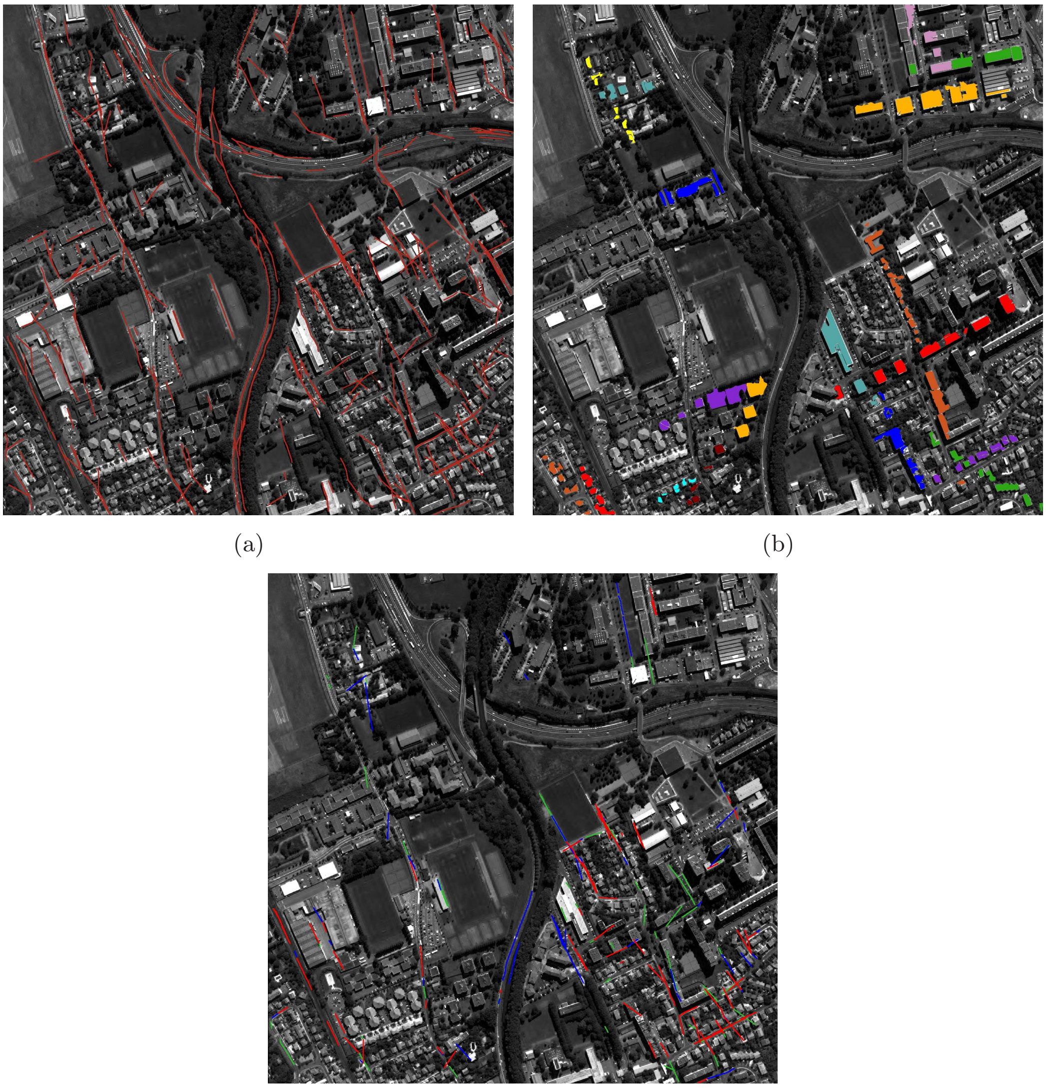
Figure 3.3(c) we can see that we obtain longer groups, and some of the groups of Figure 3.3(c) are included in Figure 3.4(b). However, when allowing a larger distance between the members of an aligned group, we are more permissive and therefore we can obtain groups such as the green group on the bottom right part of the image, which is an aligned group made of distant objects, and does not represent a meaningful alignment for the description of the scene.

The resulting roads which are "near" and "parallel to" an aligned group of objects and which are parallel to the group or which have a group parallel to them are shown in Figure 3.4(c)(we call the conjunction of these two conditions the constraint of parallelism). We are interested only in the roads on residential areas, since the hypothesis of the constraint of parallelism is only valid for these areas. Figure 3.5 shows a subregion of the image showing roads which satisfy the constraint of parallelism. We note that most of the roads which have a low degree of satisfaction of this constraint are the roads which can be classified as false detections. However on the bottom of the image we see three false detections that continue to be detected with a high degree. This is due to the fact that there exist groups such as the green aligned group in Figure 3.3(c) for which these roads are parallel. Although there are still some false detections, we can observe that their number has been significantly reduced. Determining the roads which satisfy the constraint of parallelism can be seen as an intermediary step for a road and building extraction method. We can further think of combining the parallel and the alignment relations, with the relation "between" to determine the region between two parallel groups of aligned buildings where it is possible to find a road.

Both examples demonstrated that by fusing spatial relations it is possible to extract high level concepts or improve the results of an algorithm. Moreover, they show the usefulness of the "alignment" and "parallel" relations.

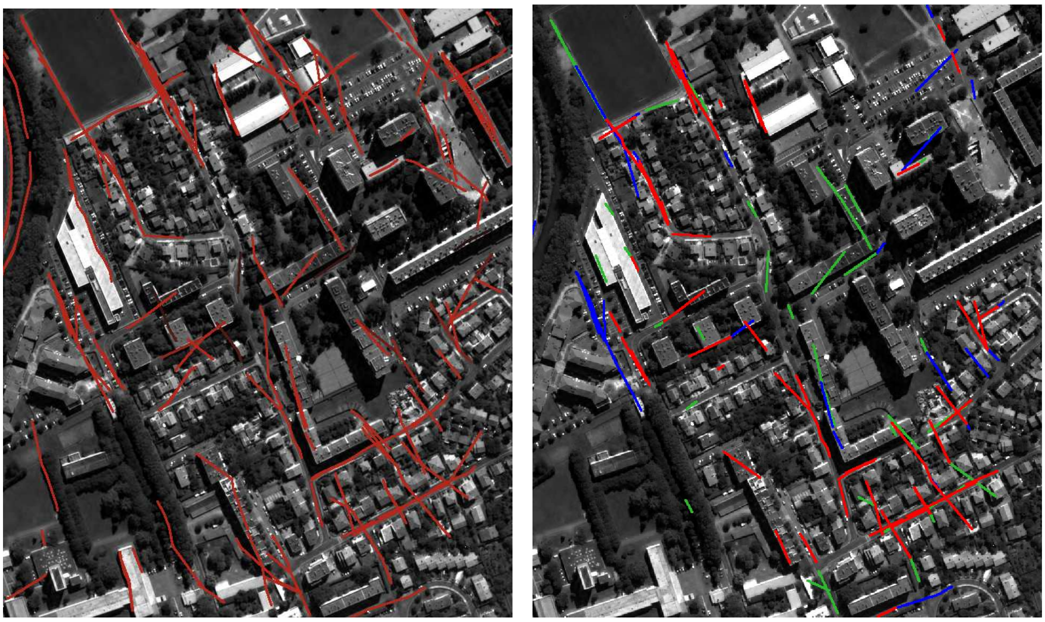
3.4 Conclusion

In this chapter we have presented some of the tools used to perform spatial reasoning. We mainly focus on using fusion to combine different spatial relations. We presented two examples of how it is possible to combine spatial relations by using simple rules such as "in residential areas houses belong to aligned groups of houses which are parallel to and near each other". However, when we have to determine the collection of objects which need to satisfy a combination of several spatial relations it is necessary to structure that knowledge, in order to combine all the relations in an efficient manner. In the next part we focus on knowledge based systems used for image interpretation. These systems rely on a structured representation of the spatial relations that the objects in the image should satisfy, and allow us to reason about that knowledge.



(a) (b) (c)

Figure 3.4: (a) Original roads. (b) Some of the *globally* aligned subsets of houses found by the algorithm with a degree of alignment greater than 0.85. (c) Obtained roads, after eliminating the roads which were not parallel to a group or did not have a group parallel to them. The green roads represent the roads which satisfy the constraint of parallelism with a degree between 0.3 and 0.5, the blue roads between a degree 0.5 and 0.8 and the red roads between a degree of 0.8 and 1.0.



(a) (b)

Figure 3.5: (a) Original roads of a subregion of Figure 3.4(a). (c) Obtained roads, after eliminating the roads which were not parallel to a group or did not have a group parallel to them. The green roads represent the roads which satisfy the constraint of parallelism with a degree between 0.3 and 0.5, the blue roads between a degree 0.5 and 0.8 and the red roads between a degree of 0.8 and 1.0.

Part II

Satellite image interpretation

Chapter 4

Knowledge based image interpretation: an overview

One of the main problems studied in computer vision is the description of the content of images. This problem has been addressed for different types of images: natural, medical and Earth observation images, among others. In this chapter we introduce what is image interpretation within the context of this thesis (Section 4.1), focusing particularly on the main related issues encountered in the interpretation of Earth observation images. Finally we present some of the methods proposed in the literature for image interpretation.

4.1 Introduction

Image interpretation can be defined as the extraction of semantics from an image. This means recognizing the different parts which compose the scene, understanding their spatial organization and constructing a description of it[1]. The description of the scene does not necessarily comprise all the objects nor all the spatial relations that appear in the image, but only the ones which carry a meaningful information about it. Deciding whether a piece of information is meaningful or not depends on the objective of the description.

Image interpretation is easily performed by humans, nonetheless the details and the level of the description depend on several factors which are extrinsic to the image. We explain some of these factors using the image in Figure 4.1, as an example:

Conceptual levels of knowledge: The knowledge of the person allows him to differentiate objects at different detail. For instance, a person who has knowledge in geography can describe the image as "a satellite image containing an industrial and urban area". While a person who does not have this knowledge can describe it as "a satellite image containing buildings and green zones".

Contextual information: It refers to the information which it is not observable in the image but which is relevant for its interpretation [Neumann and Moller, 2008]. For example, if we know that the image shows a part of London, then we can describe it as "An image of London and part of the Heathrow airport".

The vocabulary: Using a different vocabulary to describe the image can result in a different interpretation. For instance, if we have a very basic vocabulary we can describe

[1] http://www.irit.fr/ACTIVITES/EQ_TCI/EQUIPE/dalle/interpretation.html

Figure 4.1: Example of different interpretations. Image taken from Google Earth.

it as “an image containing buildings, roads, green zones and a part of an airport”.

The objective of the description The vocabulary and the level of detail used to describe an image will depend on its objective. For instance if the objective of the description is to show the effects of the airport's expansion in its surrounding areas, then the description can be “an image showing the runway of Heathrow's airport close to a residential area”.

All these factors demonstrate that on image can have innumerable interpretations. The difficulty of image interpretation lies in all the imprescindible information beyond the image which is necessary to perform this task. Thus, even for humans who are able to recognize objects in an image, the task of image interpretation can be subjective.

To automate the image interpretation task in a system, it is necessary to represent all the knowledge that is involved when this task is performed by humans. Due to vast number of factors involved in the image interpretation task, the automation of this process is limited to a specific domain, and even to particular applications. The knowledge that is represented in the system should include the contextual information as well as the different sources of knowledge used to answer the following questions: which are the objects of interest? how to identify these objects or their parts? and how are these objects or parts related?
The acquisition, representation and application of these sources of knowledge are what makes the automatic image interpretation problem very complex, and have led to the use of knowledge based systems as explained in Section 4.2.

In addition to the difficulties encountered in image interpretation, each domain of application has its particular characteristics that should be taken into account in the interpretation. So let us now concentrate on the specific issues encountered on Earth observation images.

Characteristics of earth observation images

Earth observation images contain a large amount of information, which is the outcome of the combination of many different intensities that can represent natural concepts such as vegetation, geomorphological and hydrological concepts, objects constructed by humans such as buildings and roads, and artifacts caused by variations in illumination of the terrain by the Sun, such as shadows [Sowmya and Trinder, 2000]. Due to the diversity of objects found in Earth observation images, there are different conceptual levels to describe an image [Guo et al., 2009]:

- individual objects, for instance a house, a tree, a road segment,
- land cover type, for instance water, bare land, vegetation,
- complex or composite objects, which consist of several spatially related individual objects which form a new semantic concept, for instance an airport, or a harbor.

The conceptual level used to describe the image depends on the objective of the description as well as the resolution of the image. Moreover, Earth observation images contain objects of different sizes. Figure 4.2 illustrates the variety of objects which are found in a satellite image, according to their semantic level and size. The variable size of concepts in an Earth observation makes it unfeasible to analyze all the concepts of the image at a same scale. For instance, a building can be identified at a very high resolution using its shape, a city is better identified at a lower resolution as a texture. Therefore, according to the level of concept we are interested in, we should choose a better scale of observation. By making a multi-resolution analysis of the image it is possible to observe different concepts at different scales, we are interested in the interpretation of very high resolution remote sensing images. These types of images allow us to distinguish more details than in lower resolution images.

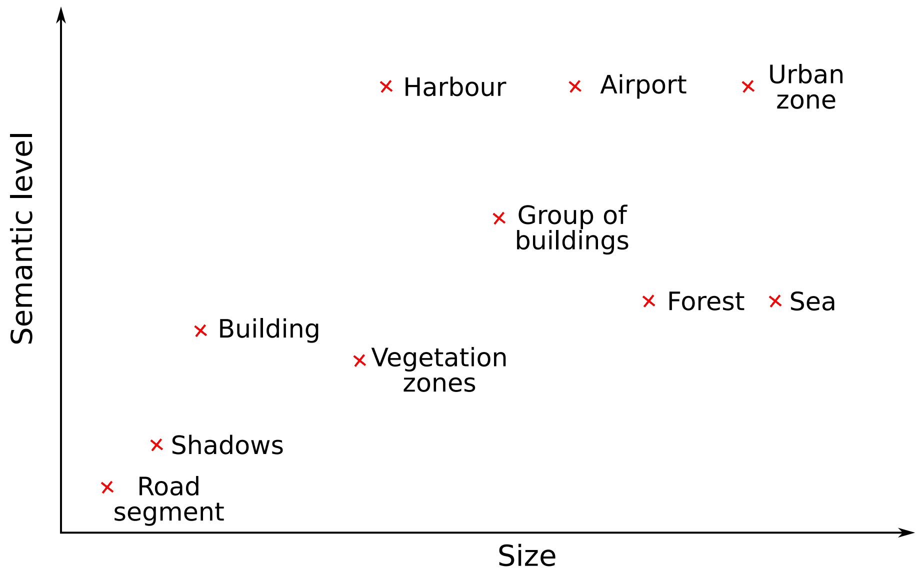


Figure 4.2: Illustration of the different concepts found in satellite images. Image inspired from [Bordes, 2009].

For instance, high resolution remote sensing images allow us to discriminate individual objects that make part of complex concepts. However, determining important details to extract these objects is a difficult task. Traditional methods using spectral or texture features in the images are not enough to extract the individual objects. Therefore, these types of images require to use spatial relations among the parts of the complex concepts, in order to distinguish the individual parts of complex objects, and eventually the complex objects themselves.

4.2 Knowledge based systems (KBS)

The important role played by knowledge in image interpretation explains the large development of knowledge based systems (KBS) in this domain. Some examples of KBS in image interpretation are given Section 4.2.3. A review of these types of systems is given in [Crevier and Lepage, 1997, Sowmya and Trinder, 2000, Le Ber et al., 2006].

KBS are inspired by human reasoning, and consist of representing and modeling the knowledge relative to a domain. The objective of a KBS is to reason on this knowledge in order to solve a problem. Some of the problems which are solved by KBS are identification, recognition, classification, diagnosis, configuration and planification problems, among others [Le Ber et al., 2006]. These systems are usually composed of three parts:

Knowledge base: It is a substitute to represent a particular domain. The knowledge of the domain is represented, as well as the conclusions which can be drawn. The knowledge represented in the knowledge base consists of lexical or ontological knowledge which includes: the symbols allowed in the representation, the constraints representing the structural arrangement of symbols and how they are connected, and rules which represent implicit or general knowledge. There are several knowledge representation schemes (see Section 4.2.1) which can be used to represent these types of knowledge.

Observation base: It contains the data that characterizes the actual problem. It can be represented using a knowledge representation scheme.

Reasoning components: Process the information of the observation base using the knowledge base to solve a problem.

In image interpretation we distiguish three types of knowledge that have to be represented in the system [Matsuyama and Hwang, 1990]:

The image processing knowledge: is used to extract the low level features from the image and their numerical description, so that they can help identifying the objects of interest in the image.

The domain knowledge: concerns knowledge about the semantics of the domain of the image.

Knowledge about the mapping between image features and concepts: it is the knowledge used to make the correspondence or mapping between the low level features into high level concepts of a domain of interest. It is the link between the two previous types of knowledge. The mapping problem is also known as the *semantic gap* [Hudelot et al., 2005].

The advantage of using knowledge based systems is that it is possible to separate all these types of knowledge or subproblems as it is done in [Matsuyama, 1988, Garnesson et al., 1989, Benz et al., 2004, Hudelot, 2005, Maillot, 2005], allowing an application independent system. It contains knowledge which can be reused to solve other problems. It allows to have different specialists involved to specify different types of knowledge used for the interpretation. For instance, the specialist of the domain of application does not need to have knowledge about the image processing tools used to extract the objects of the scene.

Other systems such as [Draper et al., 1989] combine these types of knowledge, where the representation of the objects of a domain also contains knowledge about their extraction. By combining these types of knowledge we obtain a less complex system. However, this system is dedicated to only one type of problem. Moreover, it is complicated to add new concepts to the system, since it is necessary to know the relation between the new concept and the concepts that already belong to the knowledge base, and the way to extract this object. Therefore, the systems which combine the different types of knowledge are not easily expandable to include more information.

In the following we explain in more detail some knowledge representation schemes.

4.2.1 Knowledge representation schemes

Semantic networks and attributed relational graphs: Semantic networks were introduced by [Quillian, 1985]. They consist of a diagram representing connections (binary relations) between concepts which are represented as nodes. The most common links are “Is-A” and “A-Kind-Of” between two concepts and “property” link which assigns a property to a concept. The inference is done by following the links, which makes it intractable, since sometimes in order to answer a query it is necessary to consider each node of the semantic network.

Attributed relational graphs are graphs where each edge has a binary attribute and a value attached to it. The attribute and the value represent the relation between the objects represented by the nodes and the value of the relation. These types of graphs have been frequently used to represent spatial knowledge in an image, for example in [Aksoy, 2006, Petrakis et al., 2002].

Frames: Frames were introduced by [Minsky, 1974]. The objective of frames is to group all the information concerning a concept. These approaches are closely connected to object oriented programming. Each frame is a collection of slots, where each slot contains a value of a characteristic of the object or a pointer to another frame, for example to specify relations or subparts. The slots can even contain a procedure. For instance in [Clément and Thonnat, 1990] a network of frames is used to represent the image processing knowledge.

Semantic networks and frames are not based on logic. This was a criticism to both representation schemes [Woods, 1978] because of the lack of precise semantics which leads to several interpretations that can cause misunderstandings.

Description logics (DL) and Ontologies: Description logics evolved from the semantic networks, frames and predicate logics. They have two components: the TBox, and the ABox. The TBox concerns the terminological knowledge, that is, the classes and relations of the domain represented as concepts and roles (equivalent to the knowledge base). The ABox contains the assertional knowledge about the individuals in terms of the vocabulary of the TBox, where the individuals correspond to an instance of a concept (equivalent to the fact base). Let us consider an example of the type of representations which are possible using DL, presented in [Baader, 2009]: in a conference domain, we may have classes (concepts) like Person, Speaker, Author, Talk, Participant, PhD_student, Workshop, Tutorial; relations (roles) like gives, attends, attended by, likes; and objects (individuals) like Richard, Frank and DL_Tutorial. We can define a speaker as a person that gives a talk, that is:

$$Speaker \equiv Person \sqcap \exists gives.Talk,$$

we can say that Frank is a speaker and attends the DL tutorial using the assertions:

$$Speaker(FRANK), attends(FRANK, DL\_TUTORIAL), Tutorial(DL\_TUTORIAL),$$

and state the constraints that tutorials are only attended by PhD students:

$$Tutorial \sqsubseteq \forall attended\_by.PhD\_student$$

and that the relation "attended by" is the inverse of the relation "attends":

$$attended\_by \equiv attends^{-1}$$

DL is a very rich representational language that allows defining the classes, relations and objects of the domain using concepts, roles, and individuals. Moreover, it permits to state constraints using roles and concepts, and to deduce consequences such as subclass and instance relationships from the definitions and constraints. These characteristics have led to the popularity of DL in many applications, for instance the semantic web [Fensel et al., 2005], or image interpretation [Neumann and Moller, 2008]. In particular the specification of ontologies was developed in several domains, for instance ontologies on medical data [Horrocks et al., 1996, Rosse et al., 2003], on spatial relations [Hudelot et al., 2008, Schulz and Hahn, 2001], and in multimedia concepts [Naphade et al., 2006], among others.
Ontologies are defined as a specification of a domain knowledge [Gruber et al., 1995]. The role of an ontology is two-fold [Bateman and Farrar, 2004]: (i) to set a consistent and well-specified general modeling of the domain, (ii) to support problem solving and inference within the domain of concern. These objectives are achieved by organizing the sets containing all the concepts and relations of the domain using a partial ordering on the concepts and in the relations. Thus, the concept and relation sets are organized in hierarchies, which allow us to infer knowledge from classes to the subclasses.

Conceptual graphs: Conceptual graphs were proposed by [Sowa, 1984] and later their expressiveness was enriched by [Chein and Mugnier, 2008]. Here we will give a quick overview of conceptual graphs as they are introduced in more detail in Chapter 5. Conceptual graphs are built over a vocabulary $\mathcal{V} = (T_C, T_R, I)$, where T_C and T_R correspond to the ontologies representing the set of relations and concepts in the domain, respectively. The set I which corresponds to a set of names, called individual markers, is used for denoting specific objects or entities.
A conceptual graph is defined as a bipartite graph $G = (\mathcal{N}_C, \mathcal{N}_R, \mathcal{E}, l)$ where $\mathcal{N}_C$ denotes the concept nodes, $\mathcal{N}_R$ denotes the relation nodes, and $\mathcal{E}$ is the set of edges. Relation and concept nodes are labeled by types from T_C and T_R using the function

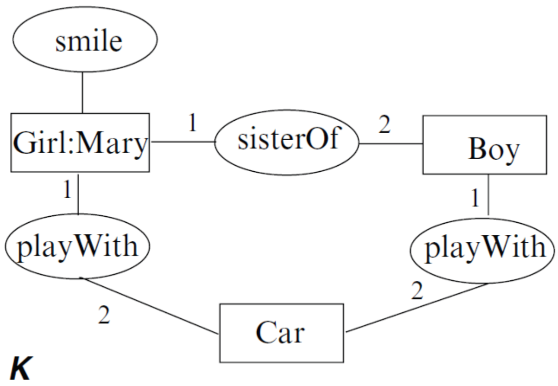


Figure 4.3: A conceptual graph. Image taken from [Chein and Mugnier, 2008].

l. One can indicate that a concept node refers to a specific entity by adding an individual marker to its label.

An example of a conceptual graph is given in Figure 4.3 (the vocabulary is not shown). This example was taken from [Chein and Mugnier, 2008]. The concepts of this graph (represented as squares) are Girl, Boy and Car; the relations (represented as ovals) are smile, playWith, and sisterOf, and there is an individual Mary which is represented inside the square after ":". The graph asserts that: Mary (who is a girl) and her brother are playing with a car, and that Mary is smiling. Notice that in a conceptual graph we can use entities which do not have a specific name, as for example Car and Boy. Additionally, we are allowed to represent relations of any arity.

The core reasoning operator of conceptual graphs is the subsumption relation between graphs. It is based on the notion of graph homomorphism. A graph G subsumes a graph H if there exists a mapping from the nodes of G to the nodes of H, that preserves the relationships between entities of G, and specializes the labels of entities and relationships. The notion of graph homomorphism is defined in Chapter 5.

Moreover, the semantics of the conceptual graphs can be mapped into that of First Order Logic (FOL) [Chein and Mugnier, 2008]. We can describe the conceptual graph of Figure 4.3 using the following formula:

$$\begin{aligned}\Phi(G) = &\exists x \exists y (Girl(Mary) \wedge Boy(x) \wedge Car(y) \wedge smile(Mary) \\ &\wedge sisterOf(Mary, x) \wedge playWith(Mary, y) \wedge playWith(x, y))\end{aligned}$$

Conceptual graphs correspond to a fragment of FOL without functions. Furthermore, [Chein and Mugnier, 1992, Sowa, 1984] demonstrated that homomorphism is sound and complete with respect to FOL semantics. However, the decision problem in this fragment of FOL is NP-complete, and therefore deciding whether a graph G subsumes a graph H is NP-complete [Chein and Mugnier, 2008]. Nevertheless, it has been shown that this problem is polynomially equivalent to other problems such as constraint satisfaction problem [Chein and Mugnier, 2008]. Hence, it is possible to use algorithms of exponential complexity to determine graph subsumption in practical applications.

Conceptual graphs have the great advantage of having a graphical representation which makes them user friendly. Additionally they are logically founded, and this gives them an intrinsic reasoning mechanism.

Both DL and conceptual graphs have formal semantics which make them not only appropriate for representing knowledge but also to reason about it. In DL we can determine whether a concept is subsumed by another concept or check whether the knowledge represented in the ABox is consistent with the one of the TBox. In conceptual graphs we can reason to decide whether a graph G is subsumed by a graph H or whether a graph G is valid. Other types of knowledge are the rules.

Rules: They are of the form *IF-THEN* where the hypothesis *IF* part is represented as an aggregation of statements, and *THEN* represents the consequence. Rule-based knowledge representation systems are especially suitable for reasoning about concrete instance data. Complex sets of rules can efficiently derive implicit facts from explicitly given ones.

We would like to highlight that although we presented the ontologies and the description logics together, we can use other representation schemes to represent the ontological knowledge. For instance the hierarchies used in the vocabulary of the conceptual graphs describe ontological knowledge and they are not represented in a DL formalism. In [Dupin de Saint-Cyr and Prade, 2008] ontologies are presented as graph formulas of proposition logic. In [Bordes, 2009] the ontological knowledge is represented as a semantic network.

We have presented several representation schemes. Most of them have fuzzy extensions, which allows representing imprecise knowledge. The fuzzy attributed relational graphs [Chan and Cheung, 2002] are an extension of attributed relational graphs. In this graph the value of an attribute can be represented as the membership to a fuzzy set.
For DL, the fuzzy DL extension [Straccia, 2006] allows defining fuzzy/vague/imprecise concepts, and reasoners have been developed, which are able to cope with this imprecision [Konstantopoulos and Apostolikas, 2007, Bobillo and Straccia, 2008, Lukasiewicz and Straccia, 2008].
Conceptual graphs are extended to fuzzy conceptual graphs [Thomopoulos et al., 2003a,b] that allow representing fuzzy concepts and markers. Notions such as fuzzy specialization are used to define the homomorphism between two fuzzy conceptual graphs.

4.2.2 Sources and types of information imperfections

In the context of scene description guided by a model, the different types of knowledge are subject to information imperfections. We enumerate some of these information imperfections which are present in image interpretation problems. We only concentrate on the cases where a model is used to represent the objects and the spatial relations that should appear in the scene.

Uncertainty with respect to the model: We are not certain that all the relations and the objects in the model are always valid. Therefore, there can be missing objects in the image or objects in the model which do not belong to a real situation. This situation can be very frequent in models designed to interpret satellite images or natural images, since there is a variability in the relations and in the objects contained in a complex scene. Moreover, even if all the objects in the model appear in the image, we are not certain with respect to the number of instantiations of the objects in the image. For instance, suppose that we want to find the instantiations of the model in Figure 4.4(a) in Figure 4.4(b). The model describes the situation of a house adjacent to a road, however due to the shadows we are not certain that the relation of adjacency is satisfied between the house and road.

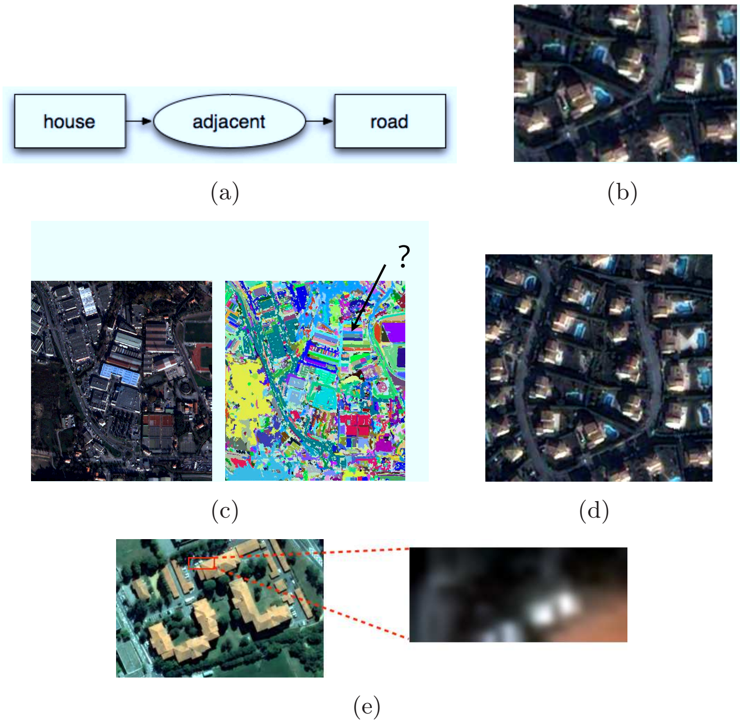


Figure 4.4: Examples of sources of information imperfection: (b)Situation where we are not certain that the objects in the model of Figure (a) appear in the model. (c) Example of uncertainty with labeling the objects in the image after a segmentation. (d) Example of uncertainty with respect to the number of instantiations, the situation described by the model of Figure (a) appears several times in the image. (e) Example of imprecision of objects.

This type of situation is also encountered when the knowledge base contains several models and there is uncertainty with respect to which model better corresponds to the image.

Uncertainty with labeling objects in the image: When labeling the objects in the image after a segmentation, there can be uncertainty with respect to the label given to a region of the image. For example, after we perform a segmentation as in Figure 4.4(c) we can be uncertain on the label that is given to a region.

Imprecision of spatial relations: As discussed in part I, many spatial relations can be imprecise by nature. Their satisfaction can depend on the context or even on the size of the objects.

Imprecision of objects in the image: We can find imprecision of the objects on the images due to the discretization of space (passing form a continuous scene to a digital image), or due to the processing levels (segmentation). For example, in Figure 4.4(e) a zoom in the boundary of the building is performed showing that the boundary of the building is not clearly defined, the pixels bordering the object contain partial volume.

Unknown number of instantiations When we try to instantiate a model in a satellite image it is very common that the pattern described by the model appears an unknown number of times on the image. For example, the situation described by the model of Figure 4.4(a) appears several times in Figure 4.4(d).

4.2.3 Reasoning under imperfections

Usually the reasoning strategy depends on the knowledge representation scheme. However, in the context of scene description guided by a model which describes the spatial structure of the scene, spatial reasoning should also be considered. Moreover, the reasoning strategies have to be adapted to deal with the imperfections described above. In the following, we present some methods for reasoning under uncertainty in the spatial domain. The reasoning strategies and how the imperfections are represented are discussed. Some of the presented works can deal with more than one type of imperfection. This is not an exhaustive survey of the knowledge base image interpretation systems, but the aim here is to give a few relevant examples.

4.2.3.1 Mapping regions into the concepts of the model

First, we discuss the works that address the problem of mapping regions extracted from a segmentation into the concepts of the models. This type of approaches should consider uncertainty with respect to the label given to a region of the image.
[Saathoff and Staab, 2008] developed a system to label regions obtained from a segmentation of a natural image based on several models. Each model is composed of binary spatial relations between two concepts representing a typical arrangement of objects in a natural image, for instance “the sky is above the sea”. The problem of representing the uncertainty of giving the correct label to a region issued from the segmentation is addressed by defining a membership function over the set of regions for each concept in the model. The membership value of each region is set according to the score of a classification. The labeling problem is formulated as a fuzzy constraint satisfaction problem where the spatial relations

represent the constraints, the concepts, the variables, and the domain of the variables is constituted by the image regions. The reasoning is performed by propagation of constraints (representing the spatial relations) which progressively eliminate the set of possible instantiations for each concept. The spatial relations are represented as crisp spatial relations. This approach assumes that the initial segmentation is correct. However, using a generic segmentation is very restrictive since we are considering that it is possible to extract the objects of interest by just using a generic partition criterion such as homogeneity, regularity or coherence in their texture. Moreover, these generic criteria do not allow discriminating between objects belonging to different semantic concepts. For instance, the roof of a building is more homogeneous than a forest when observed in a satellite image. Thus the object of interest can be divided into several regions or two objects can be included in the same region. Therefore, it is unfeasible to extract objects having different visual attributes, as is the case of objects belonging to semantic concepts, by just considering a generic criteria.
To overcome the problems yielded by a generic segmentation it is possible to: (i) include high level knowledge in the segmentation, (ii) use a multi-scale segmentation or (iii) consider an over-segmentation of the image where a concept of the model is represented by several regions. The first possibility is discussed later when we address the systems which address the problem of image segmentation and mapping.
Using a multi-scale segmentation allows to have several segmentations of the same image, where each one considers different values of the partition criteria. For example, this allows extracting a stack of regions corresponding to different homogeneity criteria. Thus the regions obtained correspond to regions satisfying different visual attributes, and therefore they can correspond to objects belonging to different semantic concepts.
The system *eCognition* [Benz et al., 2004] has adopted this approach. An initial multi-scale segmentation through region merging is performed, where the merging criterion is based on homogeneity, shape regularity and size. A fuzzy classification is used to label the regions. This classification is performed by the application of fuzzy rules, for instance "if feature x is low then the region should be assigned to landcover type W". Where "low" is defined by a fuzzy set over the domain of the feature x. The feature x can be a spectral index or a shape feature. Once all the classification rules have been applied, each region is assigned with the label for which it has the highest membership value. Then it is possible to apply rules containing crisp spatial relations to refine the classification results. For example in [Liu et al., 2008b] the system is used to identify cars on a road, so first the road and the cars are identified by using their shape, and then, rules like "cars are adjacent to the road" or "cars are surrounded by road" were used to refine the classification. This approach considers the inconvenients of using a generic segmentation and performs a multi-scale segmentation. The uncertainty of giving the correct label to the regions of the image is also considered. However, decisions about the classification are made very quickly, since the uncertainty of the objects is not considered when performing the spatial reasoning.
The other alternative to get around the generic segmentation problems is to perform an over-segmentation of the image. By performing an over-segmentation, the correspondence between the model and the regions of the image is not one to one, but a group of regions can represent an instantiation of a concept. This approach was studied in [Deruyver and Hodé, 1997, Deruyver et al., 2009, Perchant, 2000]. [Perchant, 2000] proposes to use a fuzzy homomorphism between a graph model represented as an ARG and the graph representing the regions obtained from the over-segmentation. The regions of the image are represented by a fuzzy attributed relational graph which includes information about the satisfiability of the spatial relations represented as fuzzy relations between the regions and

the membership of belonging to a class of objects. The fuzzy homomorphism allows a node in the graph model to correspond to several regions in the image. Different optimization methods were proposed to obtain the homomorphism, such as genetic algorithms [Perchant et al., 1999], estimation of distribution algorithms [Bengoetxea et al., 2002] or tree search [Cesar et al., 2005].
[Deruyver and Hodé, 1997, Deruyver et al., 2009] proposes to represent the knowledge about the model and the regions obtained from the over-segmentation of the image as conceptual graphs. The spatial relations are represented as crisp ones. The problem of labeling the regions in the case where there is complete certainty about the concepts and relations which appear in the model is discussed in [Deruyver and Hodé, 1997]. The mapping between the concept nodes of the model and the regions is an homomorphism which is determined by formulating the problem as a CSP. However, when modeling the problem as a CSP it is not possible to instantiate a concept node of the conceptual graph as a group of regions. Therefore, they introduced the bi-level constraints in which one level contains the classical constraints between the nodes of the graph that should be satisfied by the objects, and the other level contains the intra-node constraints describing the relations between the subparts of the objects. An extension for solving CSPs with a bilevel approach is proposed in [Deruyver and Hodé, 1997]. In [Deruyver et al., 2009] an interesting representation of uncertainty with respect to the model is proposed. This uncertainty is studied in three cases:

1. It is known that there are data (objects or relations) that are missing in the image.

2. It is known that there are additional data in the image.

3. It is not known whether there are missing or additional data in the image.

To deal with these situations two weak arc consistency notions are introduced: quasi-arc consistency and indirect arc consistency. When there is knowledge about missing data in the image, a constraint relaxation is proposed leading to the quasi-arc consistency definition. This relaxation is introduced by adding a function $Relax$ defined over the concept nodes of the model graph. For every conceptual node i in the graph, the value $Relax(i)$ represents the number of allowed relaxations of constraints associated to the node i, that is, the number of constraints which are not satisfied. The second case is handled under the hypothesis that the insertion of new data has a minimum effect on the initial graph, and that the new relations are linked to at least one of the initial conceptual nodes. This leads to the notion of indirect-arc consistency. In the algorithm used to solve this type of constraints it is observed whether the regions or groups of regions representing two concepts i and j which are related in the conceptual graph are directly related or indirectly related through an object k. Two nodes i and j are indirectly related by a node k if i is related to k and j is related to k. Finally, the third case is handled by increasing the $Relax$ value of each concept node iteratively until a solution is obtained. The uncertainty with respect to the model is considered in the reasoning by the introduction of the $Relax$ function and the indirect arc-consistency. The representations proposed in this work allow to make the decision even when there are objects missing in the image, or when there is an unexpected object in the image. In the case where there is an unexpected object in the image it is not necessary to specify any information about the object.

4.2.3.2 Image processing and mapping problems

In [Bloch et al., 2003, Colliot et al., 2006, Nempont, 2009] the segmentation and interpretation problems are addressed simultaneously, and the information from the spatial relations is directly used to help the segmentation process. The imprecision of the spatial relations is considered by representing them as fuzzy landscapes (see Chapter 1). In [Bloch et al., 2003, Colliot et al., 2006], the knowledge is represented as a hierarchical ARG and the interpretation/segmentation are performed by a sequential search. Starting from a structure which is relatively easy to identify, the method searches another structure using its geometric properties and the spatial relations with the previously recognized objects. This method is performed iteratively. Unfortunately, the order of recognition in this method is done empirically depending on the difficulty of recognizing a structure, and it is possible that the information to recognize a structure is not sufficient. An extension of this approach consists of learning the order in a graph reasoning scheme [Fouquier et al., 2008].
In [Nempont, 2009], the model is also given as an ARG and the imprecision on objects is taken into account. The problem is expressed as a Constraint Satisfaction Network, and it is assumed that all the objects that appear in the model also appear in the image. Every object is represented by means of two regions in the image, the upper region which represents the possibility and the lower set which represents the necessity. The set interval formed by the upper and lower sets is reduced using local consistency techniques by taking into account the fuzziness of the spatial relations. By expressing the problem as a Constraint Satisfaction Network not only the knowledge about the spatial disposition about the objects in the image and their properties are taken into account to solve the problem, but also the structure of the whole network.
These approaches are innovative in the sense that they use the spatial relations as another source of information to guide the segmentation and not only to verify at the end if the segmentation is correct. Moreover, they consider the imprecision with respect to spatial relations as well as the imprecision of objects in the image.

4.2.3.3 Categorization problems

In this section we discuss some of the classifications where there are several models in the knowledge base and the objective is to determine which is the model that better describes the situation.

In [Hudelot et al., 2005, Maillot and Thonnat, 2008] the problem of image interpretation is represented by using the following three levels: the image processing problem, the mapping, and the semantic interpretation problem. In [Maillot and Thonnat, 2008] the problem of isolated objects categorization is treated. For this, a visual concept ontology is proposed. This ontology contains visual descriptors such as geometric properties of objects, texture, color, and spatial relations. This ontology allows describing the domain concepts in terms of visual descriptors. A learning stage is performed to learn these visual descriptors in the image, using machine learning techniques. This makes a link between low level descriptors and high level ones. The categorization of a new image is done by segmenting the image and trying to classify a subpart of the object. An initial hypothesis about the subpart of the object is made and tested by evaluating the attributes corresponding to the visual descriptors, and a probability of matching a subpart is obtained. Then, a global matching is performed by considering the probability of matching the subpart. If the global matching fails, then the hypothesis is dropped and a new hypothesis is proposed, otherwise it tries to identify the other subparts of the object. In this approach the spatial relations are used to verify the hypothesis of the objects. However they do not guide the recognition procedure. In [Hudelot et al., 2005] a similar system is proposed to perform scene analysis. In this system two ontologies are used to communicate between the image processing level and the semantic interpretation level. One of the ontologies is the visual concept ontology of [Maillot and Thonnat, 2008] and the other one is an image processing ontology. The link between the visual concepts and the image low level features is done by using fuzzy sets describing the linguistic variables which evaluate the visual concept. This allows to consider the imprecision of describing the visual attributes of a concept. However, the spatial relations are still modeled as crisp ones. The interpretation is performed by making a fuzzy matching between a hypothesis and the model, and finally by verifying the spatial relations. Using a fuzzy matching between the visual concepts and the objects allows considering the imperfections with respect to the model. These two approaches have the originality of introducing the visual concept ontology to reduce the semantic gap. Both approaches use the spatial relations for verification in the last steps.

[Le Ber and Napoli, 2002] presents a system which recognizes and analyzes spatial structures on satellite images, by performing a match between the labeled image regions and the models of landscapes stored in its knowledge base. The knowledge base contains the vocabulary of a spatial relation hierarchy and a concept object hierarchy organized using the subsumption relation, and the models of landscapes described in an extension of DL, which allows relation quantification. The models can represent information like: the objects of class x satisfy a relation r with exactly one object of the class y, or it satisfies a relation with at least one object of the class y, or with all the objects of the class y. This type of knowledge representation is appropriate for describing objects in satellite images because sometimes the number of times a relation can take place is unknown. The RCC-8 relations (see Section 1.3.1.1) are used to describe the spatial structures. The matching between the models and the regions is performed using the reasoning mechanisms as in DL, however in their framework, there are three types of concept specialization which take into account the subsumption of spatial relations:

- Adding a relation: the class X is specialized into $X - EC - Y$ or $X - DC - Y$.
- Specializing a relation: the class $X - PP - Y$ specialized into $X - TPP - Y$.
- Specializing the range of a relation $X - PP - Y$ specialized into $X - PP - Y - EC - Z$

where EC, DC, PP, TPP correspond to instantiations of RCC-8 relations. The novelty of this approach is the reification of the spatial relations to reason with them. A similar approach was used in [Hudelot et al., 2008] when creating a spatial relation ontology, where spatial relations are considered as concepts. Theoriginal contribution of this work is that the ontology is enriched with fuzzy representations of concepts, where the connection between the fuzzy representation of the spatial relations and the linguistic concepts is explicitly made, allowing to define the semantics of the relations, and therefore to the reduction of the semantic gap.

4.3 Discussion

We have presented different approaches which address the problem of image interpretation considering the spatial relations between the objects and taking into account different types of information imperfections.

The approaches which perform the tasks of segmentation/interpretation simultaneously are not directly confronted with the problem of correctly labeling a region. Moreover, these approaches obtain better results than the ones which separate the problem into a segmentation followed by an interpretation, because all the spatial information as well as the geometry and intensity are considered for making the decision, i.e. segment and recognize a structure. These types of approaches have been applied to the interpretation of brain and thorax images, which are strongly structured scenes and for which is is possible to have in advance knowledge about the types of objects that appear in it. Hence, it is possible to construct models containing all the structures that appear in the scene (in the case of pathologies can be included in the model too [Atif et al., 2007]), as well as the relations between them, and therefore allow a very constrained formulation of the problem.

However, in Earth observation images there are several objects which appear in the scenes and which cannot be predicted and therefore cannot be easily considered in the model. For instance, if we want to construct a model of an airport, there are airports which have satellite terminals, other which do not, an airplane can be near a terminal or on a runaway, according to the size of the airport there can or cannot be an oil storage in the airport; thus all these variations make the models of structures in Earth observation images less constrained. Moreover, the number of instantiations of an object in an image is unknown, for example the number of terminal buildings in an airport is unknown. Thus if we try to formulate the problem of interpretation using a similar approach as in [Bloch et al., 2003, Colliot et al., 2006, Nempont, 2009], it is possible that for a given object in the model, the region of space which represents the conjunction of the relations that should be satisfied by this object, according to the model, is not sufficiently restricted to contribute to the segmentation of the object.

An interesting possibility could be to combine the knowledge of extracting several structures, for instance using an approach as the one proposed in [Hudelot et al., 2005, Maillot and Thonnat, 2008], and use a segmentation/interpretation approach to extract the other structures. Although this proposal is a promising approach, knwoledge acquisition and performing a knwoledge based segmentation is out of the scope of this thesis.

The approaches which perform an over-segmentation of the image, and use spatial reasoning to bring together regions to form objects, also need a very constrained formulation of the model and furthermore they have a high computational cost.

Thus, an intermediary solution is to perform a multi-scale segmentation which allows to have regions which can be candidates for objects in the scene. It is then not necessary to have a lot of knwoledge with respect to their extraction procedure. This solution is subject to the uncertainty of properly labeling the objects in the image, which can be overcome by using membership functions over the set of regions for each concept in the model, as in [Saathoff and Staab, 2008]. The membership function can be constructed using information from a classification procedure or knwoledge about the radiometry of the concepts. In this thesis, we adopted this type of approach, which is further presented in the next chapter.

4.4 Conclusion

In this chapter we have presented the notion of image interpretation in the context of this thesis. Our objective is, giving a model, to determine which are the instantiations of the model in the image. We highlighted the importance of knowledge in the task of image interpretation, and the particular difficulties which are encountered in satellite images.

We presented the main components on a KBS, and focused particularly on the knowledge representation schemes, and we will refer to these schemes in the following Chapter when we describe our choice of representation scheme.

We have reviewed some systems which consider spatial organization for the interpretation, and presented how they represent and use this information and the information imperfections that can take place in the problem. In the following chapter we propose an approach which takes into account the uncertainty of correctly labeling the regions in the image, the imprecision linked to the spatial relations and we discuss how to deal with the uncertainties in the model.

Chapter 5

A new mapping approach for satellite image interpretation using a structural model

A complex scene, as observed in satellite images, is defined as a set of spatially related objects which have a structural arrangement, and which together form a high level object. Some examples of complex scenes include functional complex objects, which impose a given spatial arrangement, such as airports, harbors, train stations, nuclear power plants, toll gates, stadiums, etc. Other examples concern urban morphologies which can be generally described as repetitive patterns. Thus, it is possible to describe the spatial arrangement of these complex objects using a model, which can be used to identify and interpret these complex objects and their parts.

To interpret a complex scene it is not sufficient to recognize the individual objects that belong to the scene. The spatial relations are also of prime importance. Furthermore, some objects in a complex scene cannot always be recognized individually, and often require the recognition of other objects having a spatial relation with them, and then use the spatial relation to identify them. The use of spatial relations allows for instance disambiguating the objects [Rosenfeld et al., 1976]. Moreover, there exist several pattern recognition techniques which allow us to identify certain classes of objects in the images, for instance vegetation, water, shadows, buildings, etc. Then, it is interesting to use the spatial relations which are found in satellite images with respect to these classes (see Part I) to identify the objects belonging to classes which cannot be easily detected. For instance in a harbor scene, the sea can be easily identified and we can use this knowledge to identify the boats.

Our objective in this chapter is to use a model which represents the spatial arrangement of objects in a complex scene in order to detect the scene in the image, and in the case when it exists, be able to detect its components and identify them using semantical concepts.

In this chapter we focus on satellite image interpretation using such a model. As mentioned in the previous chapter, the problem of image interpretation can be separated into several subproblems. Here we only concentrate on the mapping problem and on the knowledge representation problem. In order to demonstrate that the spatial relations that we introduced are of interest for solving the problem of satellite image interpretation, we developed an interpretation system that is presented in the following sections.

In Section 5.1 we introduce the main sources of knowledge of the system and its representation. Section 5.2 deals with the mapping problem, in particular we address the problem of reasoning in the presence of information imperfections.

5.1 Structural model representation

The scene that we want to find in the image is going to be described through a structural model containing the spatial arrangement of objects. The representation scheme used for the structural model should be flexible enough to allow us to represent relations of any arity as well as groups of aligned objects. For the representation of groups of aligned objects the following considerations should be taken into account:

- The number of objects in the alignment is usually unknown.
- The group can satisfy spatial relations with other objects, for example a group of aligned trees parallel to a road.
- Each element of the group can have a relation with other elements. For instance in the case of a group of aligned houses, each house has a spatial relation with its shadow.
- There can be spatial relations among the objects of the group. According to the alignment definition presented in Section 2.2 the consecutive members of the group have to be "near" each other, thus they are always linked by a distance relation.

Due to these specificities, taking into account groups of aligned objects in the proposed model will call for specific developments.

5.1.1 Choice of the representation framework: conceptual graphs vs. description logics

Among the knowledge representation schemes presented in Section 4.2.1, both conceptual graphs and DL allow us to represent spatial relations of any arity (for the case of DL, we have to consider spatial relations as a type of concept, as in [Hudelot et al., 2008]). In the following we present two possible extensions to allow the representation of aligned groups of objects.

Extending the formalism proposed in [Hudelot et al., 2008] it is possible to represent the structural model using DL. The main concepts used in [Hudelot et al., 2008] are illustrated in Figure 5.1. There are two main concepts: SpatialObject and SpatialRelation. For every object that we want to represent in the model, there exists a corresponding concept which is a type of SpatialObject, and all the spatial relations that appear in the model are a type of SpatialRelation. To introduce the notion of groups of aligned objects we first add the concept SpatialObjectGroup to represent the groups of objects. Then we add the concept SpatialAlignedObjectGroup which is a type of SpatialObjectGroup. Then to specify the type of members of the group we introduce the role *hasMember*. So, for instance, if we suppose that the concept House has already been defined, then we could define a group of aligned houses as:

$$SpatialAlignedHouses \sqsubseteq SpatialAlignedObjectGroup \sqcap \exists hasMember.House$$

It is possible to add more concepts and roles to this representation to be able to represent all the considerations that should be taken into account when representing all the possible

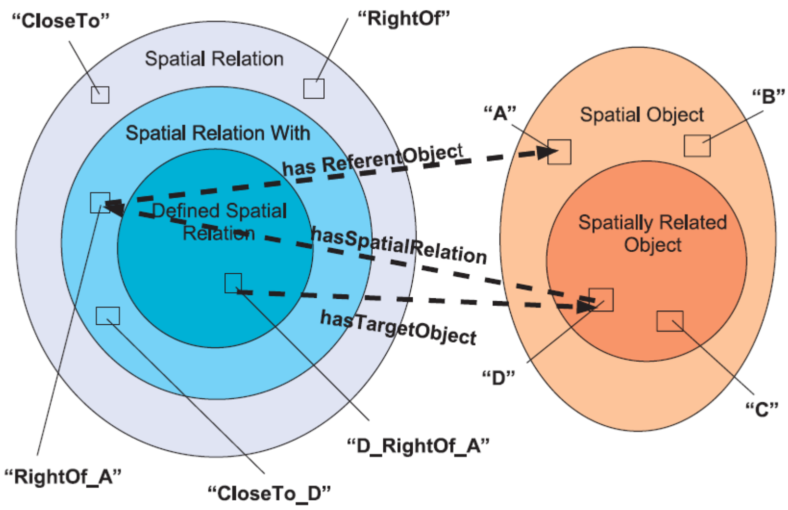


Figure 5.1: Representation of the main concepts of the spatial relation ontology proposed in [Hudelot et al., 2008] as a Venn diagram. This diagram illustrates the different concepts and how they are related. "A" is a SpatialObject, it is the ReferenceObject of the Spatial-RelationWith concept "RightOf_A". "D" is a SpatialObject which has the property of has SpatialRelation which is here the relation "RightOf_A". Image reproduced from [Hudelot et al., 2008].

situations which can involve a group of aligned objects. However, this construction can be time-consuming.

On the other hand, conceptual graphs make a clear distinction between factual and ontological knowledge. The vocabulary of the CG represents the ontological knowledge, while the sets of graphs represent the facts [Chein and Mugnier, 2008]. Therefore we can construct a hierarchy of concepts specifying the "is-a" relation between the concepts and a hierarchy of spatial relations specifying the "is-a" relation between the spatial relations. Hence, for every structural model that we want to search in an image, we can just create a graph which represents its structure, using the concepts and spatial relations of the hierarchies. This practical aspect of the conceptual graphs is what made us use them as the representation scheme for the structural models. Moreover, with respect to representation issues, DL does not adequately represent objects which are related in a non-tree model, and therefore are not appropriate for representing objects containing an embedded structure [Motik et al., 2008], while conceptual graphs do not have any limitation in the shape of the relational structure of the model. For a complete review about the relations between conceptual and DL, one can refer to [Chein and Mugnier, 2008, Baader et al., 1999].

5.1.2 Nested conceptual graphs

In Section 4.2.1 we discussed the definition of conceptual graphs. Let us give a proper definition of conceptual graphs. Conceptual graphs are built over a vocabulary:

Definition 5.1 (Vocabulary [Chein and Mugnier, 2008]). *A vocabulary is a triplet* (T_C, T_R, I) *where:*

- T_C *and* T_R *are pairwise disjoint sets, corresponding to a concept and relation hierarchies;*

- *T_C and T_R are partially ordered by a KindOf relation and each one has a greatest element denoted by $\top$;*

- *in T_R we can find relations with an arity greater than or equal to one. Any two relations with different arity are not comparable;*

- *I is the set of individual markers, which is disjoint from T_C and T_R. The symbol $*$ denotes the generic marker.*

Definition 5.2 (Conceptual Graph (CG)[Chein and Mugnier, 2008])**.** *A conceptual graph is a bipartite graph denoted by $G = \{\mathcal{N}_C, \mathcal{N}_R, \mathcal{E}, l\}$ where:*

- *$\mathcal{N}_C$ and $\mathcal{N}_R$ are the concept node and relation node sets, respectively. The set of nodes of G is equal to $\mathcal{N}_C \cup \mathcal{N}_R$,*

- *$\mathcal{E}$ is the family of edges,*

- *l is a labeling function of the nodes and edges of G which satisfies:*
  - *A concept node $c \in \mathcal{N}_C$ is labeled by $l(c) = (type(c), marker(c))$, where $type(c) \in T_C$ and $marker(c) \in I \cup \{*\}$.*
  - *A relation node $r \in \mathcal{N}_R$ is labeled by $l(r) \in T_R$. $l(r)$ is also called the type of r and is denoted by $type(r)$.*
  - *The degree of a relation node r is equal to the arity of $type(r)$.*
  - *Edges incident to a relation node r are totally ordered and they are labeled from 1 to $arity(type(r))$.*

This definition of conceptual graphs allows us to represent relations of any arity between the nodes representing the concepts, which is useful to represent ternary relations such as "between". The set I is used to represent specific instantiations of the concepts. However, if we do not want to specify a particular instantiation then it is possible to use the generic marker.

These types of graphs are appropriate to represent the spatial relations between objects. However, they cannot represent hierarchically structured knowledge. In [Sowa, 1984] conceptual graphs are extended to *nested conceptual graphs* which allow representing this type of knowledge. In this type of graphs, it is possible to represent internal and external information, zooming, partial description of an entity, or specific contexts. In a nested concept graph, the concept nodes can have a conceptual graph contained in them. So, for instance, if we want to represent a house that is adjacent to a road, and we also want to consider that the house has a shadow located in a direction of 30° with respect to the horizontal axis, then there are two cases:

1. the shadow can be falling onto the road, and maybe it would not be possible to distinguish that the house and the road are adjacent. Nevertheless, it is possible to distinguish that the shadow is adjacent to the road;

2. the shadow does not fall onto road, and in then is possible to distinguish that the house and the road are adjacent.

Therefore, this knowledge can be represented as a node which represents the group containing the house and its shadow with their respective relation, and a road which is adjacent

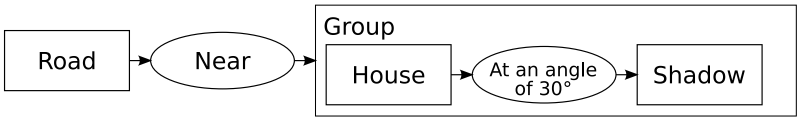


Figure 5.2: Example of nested graph.

to the group as in the conceptual graph of Figure 5.2. The group is drawn as a box. This representation covers the two situations, since a group is adjacent to an object if either of its elements is adjacent to the object.

We will refer to nodes which contain a conceptual graph inside them as complex concept nodes. To specify that a node is a complex concept node, a third field is added to each conceptual node, called *description*. Concept nodes which are not complex will have an empty *description* field noted by **.

Definition 5.3 (Nested Conceptual Graph (NCG))**.** *A nested conceptual graph is a bipartite graph denoted by* $G = \{\mathcal{N}_C, \mathcal{N}_R, \mathcal{E}, l\}$ *where:*

- $\mathcal{N}_C$ *and* $\mathcal{N}_R$ *are the concept node and relation node sets, respectively. The set of nodes of* G *is equal to* $\mathcal{N}_C \cup \mathcal{N}_R$*,*
- $\mathcal{E}$ *is the family of edges,*
- l *is a labeling function of the nodes and edges of* G *which satisfies:*
  - *A concept node* $c \in \mathcal{N}_C$ *is labeled by* $l(c) = (type(c), marker(c), description(c))$*, where* $type(c) \in T_C$*,* $marker(c) \in I \cup \{*\}$ *and* $decription(c) \in \{**\} \cup Desc$*. Desc is a set containing the labels of the descriptions.*
  - *A relation node* $r \in \mathcal{N}_R$ *is labeled by* $l(r) \in T_R$*.*
  - *Edges incident to a relation node* r *are totally ordered and they are labeled from* 1 *to* $arity(type(r))$*.*

The set of complex concept nodes of a conceptual graph is denoted by $D(G)$. The nodes inside a complex concept nodes are called child nodes. A nested conceptual graph can be recursively defined from a basic conceptual graph (Definition 5.2) by adding the field of *description* to the labeling of the concept nodes. Another representation for nested conceptual graphs is a tree of basic conceptual graphs (Refer to [Chein and Mugnier, 2008] for more information about these representations). The label of a simple node c is written as $type(c) : marker(c)$, and when the marker of a node is the generic marker $*$ then, for simplicity the node is labeled as $type(c)$.

To represent the relations between objects inside a complex concept node and concept nodes outside it we can use a coreference concept. Coreference concepts represent two concepts which are equivalent and represent the same entity, and they are joined by a coreference link. It is necessary to use coreference concepts since the knowledge inside the complex node is contextualized by the hierarchical structure representing the group. For instance, in Figure 5.3 we added to the representation of Figure 5.2 that the house is between a green zone and a parking area.

Therefore, we can use nested conceptual graphs to represent groups of objects as complex concept nodes, as well as the relations between them and with objects outside the group.

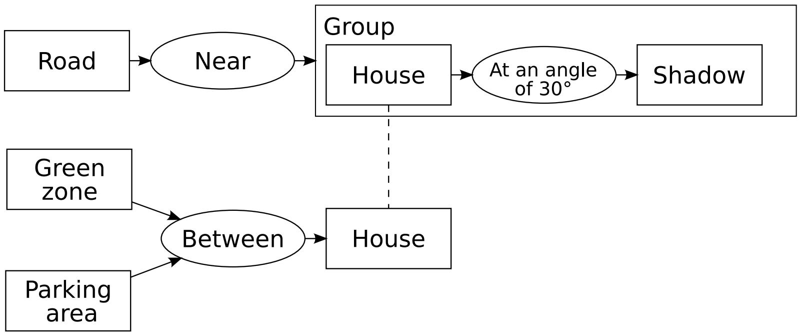


Figure 5.3: Example of nested graph with coreference links, represented as dotted lines.

The connections between a concept node and a relation node are always represented by an arrow even when the relation is symmetric.

Representation of groups of aligned objects in nested conceptual graphs

We propose to represent a group of aligned objects as a complex concept node with the description *AlignedGroup*. In a nested conceptual graph, the number of concepts nodes inside a complex concept node is known. However, when representing a group of aligned objects we face the difficulty that the number of objects in an alignment is unknown. Therefore we propose to represent an aligned group as a complex concept node containing three distinct elements which are related by a distance relation. For instance, Figure 5.4 represents a group of aligned trees parallel to a road. To represent that they are different elements we use a different marker to label each element. We use only three elements to represent a group because every group of aligned objects is composed of subgroups of three aligned objects. Moreover, the reasoning mechanism of conceptual graphs relies on graph homomorphisms (see Definition 5.4). Therefore a group of three aligned objects can be mapped to any subgroup of an aligned group. The distance relation that joins two elements of the group corresponds to the distance relation used in the alignment definition (see Section 2.2).

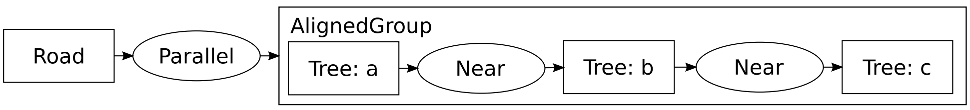


Figure 5.4: Example of a nested conceptual graph representing a group of aligned trees which is parallel to a road.

To represent a relation between the elements of the group and other objects we use coreference links. There are two possible types of relation between an object inside a group and another object outside the complex concept node:

(i) Each element has a relation with another object, for example each tree in Figure 5.4 can have a shadow in a direction of 40° with the horizontal, therefore for each tree there is a shadow.

(ii) All the elements of the group can have a relation with another element, for instance all trees are topologically surrounded by a region of soil.

Both cases are shown in Figure 5.5. The first case is represented by using a different concept node to represent the shadow of each object. For the second case all the objects are related to one node.

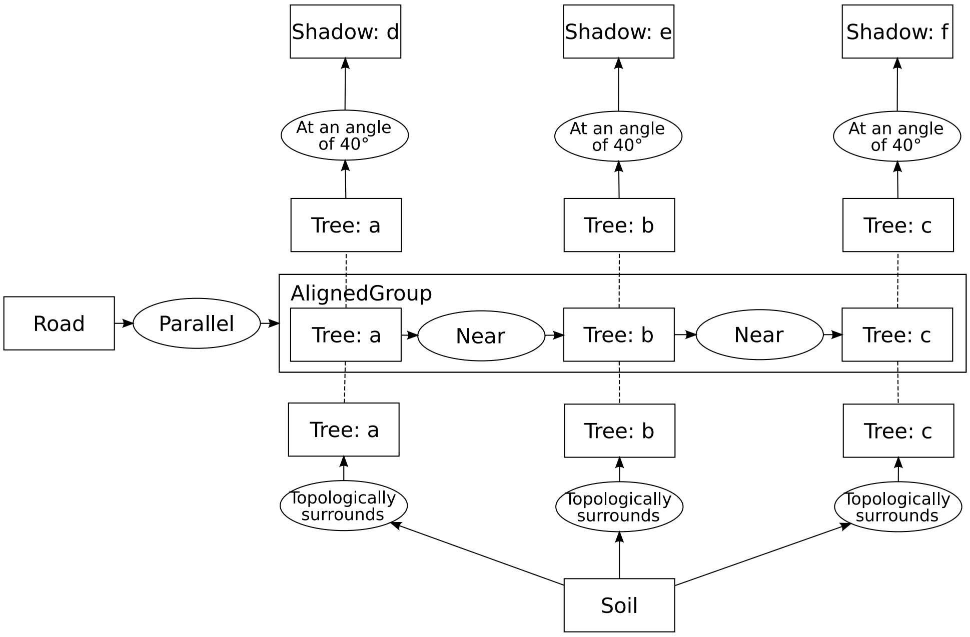


Figure 5.5: Example of a nested conceptual graph representing a group of aligned objects where the members of the group are in relation with other objects.

In conclusion, the nested conceptual graphs are well adapted to our needs. Moreover, they allow us to represent groups of objects (which are not aligned) as the one in Figure 5.3. These groups can be also used in the structural models. As we saw in Definition 5.3 the nested conceptual graphs are built over a vocabulary. In the following section we specify the vocabulary.

5.1.3 Vocabulary

The vocabulary is composed of three parts: concept, relation hierarchies and individual marker sets.
For our application the concept hierarchy consists of a hierarchy containing the objects that are found in the scene and which we want to recognize. This hierarchy depends on the scene that we want to interpret, and will be specified for each scene.

The relation hierarchy consists of a hierarchy of spatial relations (See Figure 5.6). This hierarchy is based on the hierarchy of spatial relations used in the ontology presented in [Hudelot et al., 2008], and extended to include the spatial relations that were introduced in Part I. It does not include the grouping relations, since we proposed to represent alignment as a group property, and therefore alignment is not considered as a possible label for a relation node.

Although the group of aligned objects can be related to other objects or groups of objects, it cannot be the reference object of the "topological surround" and "surround" relations because it is considered as having a linear shape without concavities. These two relations require that the reference object has concavities in order to have a satisfaction degree greater than zero. Even though the aligned group of objects can be the target object of a topological line-region relation, we do not consider these relations for these groups since we have only discussed their extension, but it has not been properly defined yet. In a similar way, we do not consider topological surround as a relation in which a group of objects is its target object.

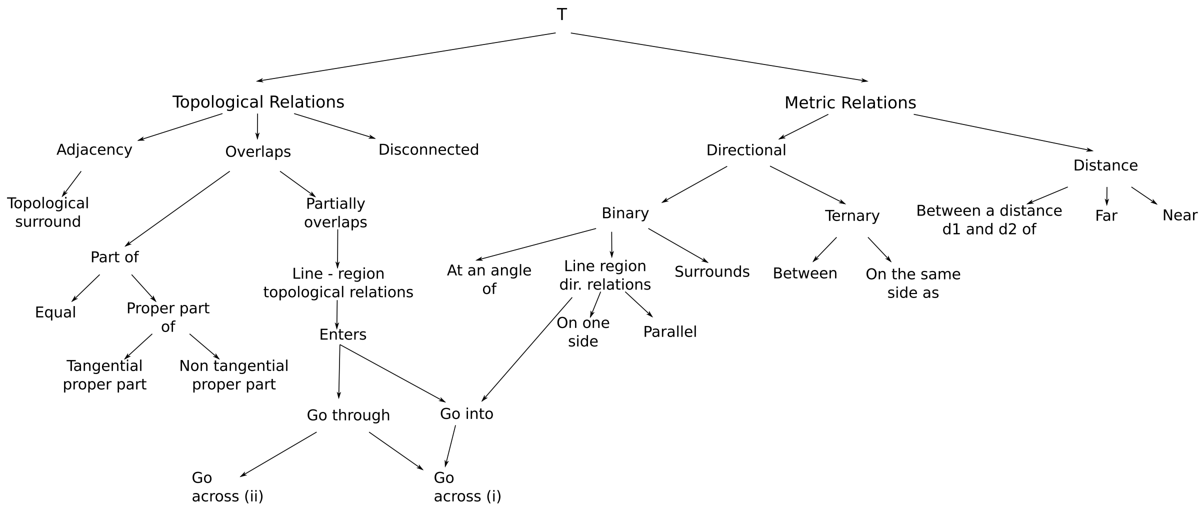


Figure 5.6: Spatial relation hierarchy.

The marker set includes the element $*$ and other markers to represent different objects. Now that we have presented the representation scheme, we address, in the next section, the problem of how to map a conceptual graph to an image.

5.2 Mapping a conceptual graph to an image: general principle

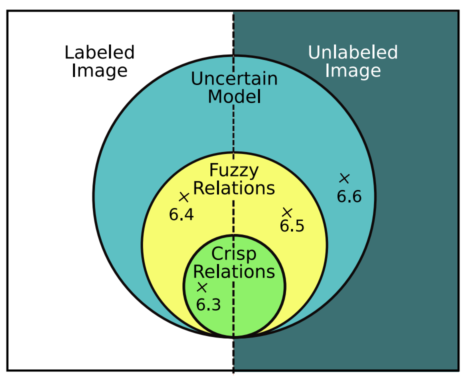


Figure 5.7: Hypotheses used in the following sections for the mapping problem. The disks represent the hypotheses for the model and the spatial relations between the objects of the image. In the left handside we show the situations where we assume that we have a labeled image and in the right hand side we have an unlabeled image. The methods used for the unlabeled image case are also valid for the labeled image case.

To explain the reasoning module of our system we first introduce a solution for a simple problem with very strong hypotheses and then introduce a more complex problem by considering more realistic hypotheses. Figure 5.7 shows a schema of the development of this section, showing different hypotheses that are considered in each subsection. In Section 5.3 we assume that we have a model, for which we are certain that all the relations and objects of the model appear together. Additionally, we have a labeled image and the scene is described as a basic conceptual graph (no complex concept nodes) where the relations between the objects of the image are evaluated as crisp binary or ternary relations. This section is illustrative, and its objective is to introduce the tools used to solve the mapping problem. In Section 5.4 we change the hypothesis of crisp relations to relations which are represented using fuzzy models. Also, the structural model of the scene is a nested conceptual graph with complex concept nodes. These hypotheses are more realistic and consider the imprecision attached to the semantics of the relations. In Section 5.5 we consider an even more realistic situation where all the above hypotheses about the spatial relations and the model continue to be valid, and additionally we suppose that we have an unlabeled image. When having an unlabeled image one of the difficulties lies in properly segmenting the image and correctly labeling each region. Therefore, once we overpass this difficulty the problem is similar to the one when we have a labeled image. Finally, in Section 5.6 we consider the uncertainties with respect to the model. This is of prime importance as discussed in Section 4.2.2. Figure 5.7 illustrates the fact that these problems are nested: the methods used to solve the mapping problem for an image with unlabeled objects are also valid for labeled ones, similarly the ones used for an uncertain model can be used for a model where there is certainty, and the ones for the case where only fuzzy relations take place can be used for the crisp ones, since the crisp case is a particular case of the fuzzy one.

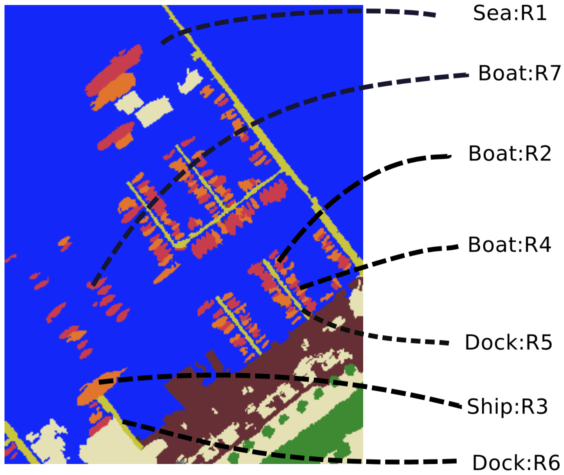
5.3 Simple case

Suppose that we have a basic conceptual graph $G_M = (\mathcal{N}_{C_M}, \mathcal{N}_{R_M}, \mathcal{E}_M, l_M)$, composed of simple concept nodes over the vocabulary $\mathcal{V} = (T_R, T_C, \{*\})$ and a segmented image I_L. Let $\mathcal{P}_L$ be the set of regions on I_L. Suppose that every region $R_i \in \mathcal{P}_L$ is labeled with a concept type from the hierarchy T_C. Our objective is to determine which regions of $\mathcal{P}_L$ satisfy the spatial constraints imposed by G_M. Using the models of spatial relations it is possible to represent the spatial knowledge of I_L as a conceptual graph $G_L = (\mathcal{N}_{C_L}, \mathcal{N}_{R_L}, \mathcal{E}_L, l_L)$ constructed over the vocabulary $\mathcal{V} = (T_R, T_C, \mathcal{P}_L)$. For the sake of simplicity we assume that the relations between the objects of the image are represented by crisp models[1].

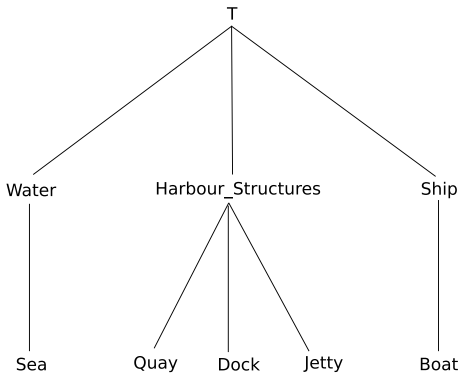
Let Figure 5.8(a) be the image we want to interpret, and Figure 5.8(b) its labeled image L. Let Figure 5.8(c) be the concept hierarchy T_C of the vocabulary $\mathcal{V}$ used to construct the graph G_L shown in Figure 5.8(d) which represents the spatial arrangement of some regions of L. The regions considered in this excerpt of G_L are marked in Figure 5.8(b). Notice that each concept node c of G_L is labeled as $type(c) : marker(c)$ where $type(c) \in T_C$ and $marker(c) \in \mathcal{P}_L$. In the example, the elements of $\mathcal{P}_L$ are labeled as R_i for $i = 0, \ldots, |\mathcal{P}_L|$. The mapping problem can be seen as finding the interpretations of G_M in G_L. Therefore using G_L we can find the instantiations of G_M in I_L. In the conceptual graph literature, this is called finding a graph homomorphism or projection. A graph homomorphism is defined in the following way:

[1] This can be done by only considering the relations which are satisfied above a threshold for instance

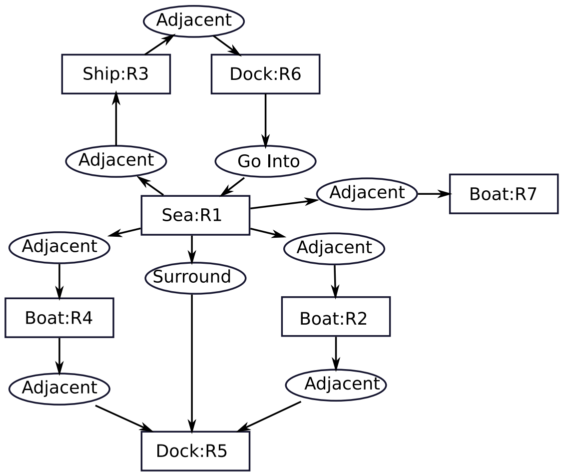
(a) Example image.



(b) Labeled image: The blue regions represent the sea, the red and orange represent ships or boats and the yellow regions represent the docks.



(c) Concept hierarchy T_C in the context of harbors.



(d) Conceptual graph representing the spatial organization of some elements of Figure 5.8(b).

Figure 5.8: Image, concept hierarchy and conceptual graph used in the interpretation example.

Definition 5.4 (Graph homomorphism [Chein and Mugnier, 2008]). *Let* $G_T = (\mathcal{N}_{C_T}, \mathcal{N}_{R_T}, \mathcal{E}_T, l_T)$ *and* $G_H = (\mathcal{N}_{C_H}, \mathcal{N}_{R_H}, \mathcal{E}_H, l_H)$ *be two conceptual graphs defined over the same vocabulary* $\mathcal{V} = (T_R, T_C, I)$*. An homomorphism* π *from* G_T *to* G_H *is a mapping from* $\mathcal{N}_{C_T} \cup \mathcal{N}_{R_T}$ *to* $\mathcal{N}_{C_H} \cup \mathcal{N}_{R_H}$*, which satisfies:*

- $\forall (r, i, c) \in G_T,\ (\pi(r), i, \pi(c)) \in G_H,$
- $\forall e \in \mathcal{N}_{C_T} \cup \mathcal{N}_{R_L}, l_H(\pi(e)) \leq l_T(e).$

where $(r, i, c) \in G$ *represents the edge labeled* i *between a relation* r *and a concept* c*, which means that* c *is the* i*-th argument of* r*.*

The first condition of Definition 5.4 ensures that if a node c is a reference element of a relation r, then its image denoted by $\pi(r)$, is the reference of the relation $\pi(r)$, which is

the image of the relation r. The same happens for the target objects. This also guarantees that relation nodes are mapped to relation nodes and concept nodes are mapped to concept nodes.
The second condition establishes that the label of a concept of G_T should be greater than or equal to the label of its corresponding concept in G_H, the order used in this comparison being the order of the hierarchy T_C. This condition must be also satisfied by the relation nodes with respect to the order of T_R. The homomorphism emphasizes the use of a vocabulary, the second condition depends on the order used over T_C and T_R. To apply a graph homomorphism in our case, it requires that both conceptual graphs are built over the same vocabulary, that is $\mathcal{V} = (T_R, T_C, \{*\} \cup \mathcal{P}_L)$.

Figure 5.9 shows a graph model G_M (on the left) and the conceptual graph of Figure 5.8(d) (on the right). This figure illustrates that there are three possible graph homomorphisms.

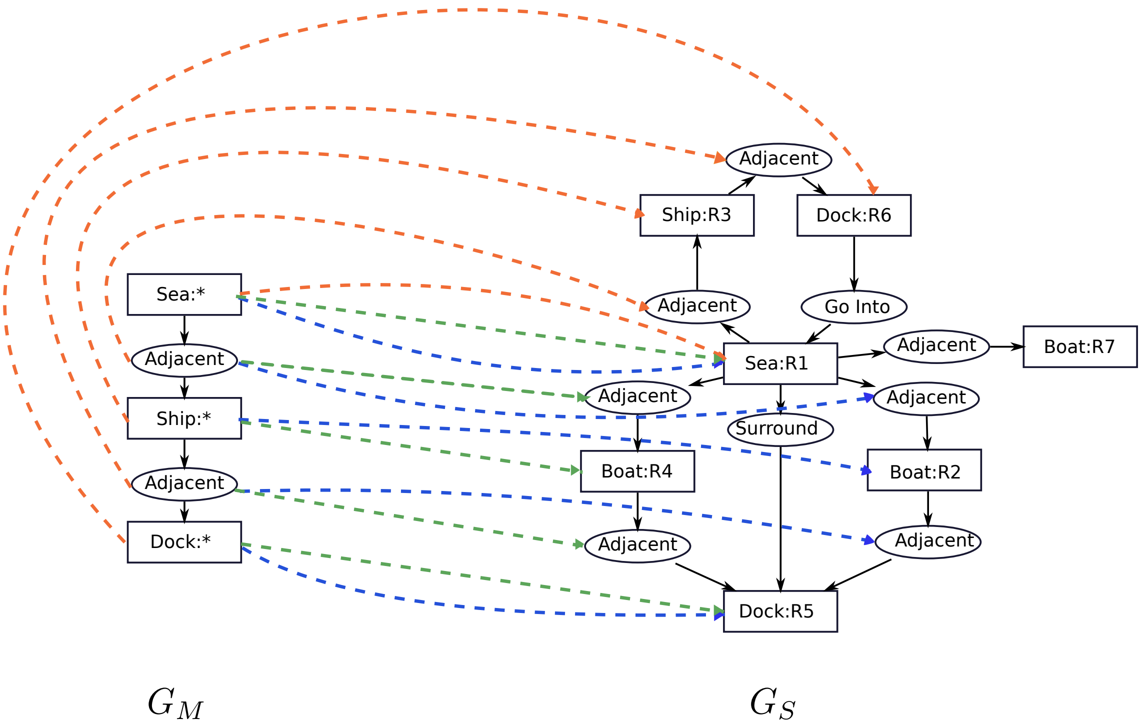


Figure 5.9: Illustration of graph homomorphism. Each color represents a graph homomorphism from G_M to G_L.

Using homomorphisms to map the two graphs permits one graph model to be mapped to several instantiations of the image. This is very useful in the case of satellite images, since most of the time the number of instantiations of a concept in an image is unknown. For instance, in the harbor scene, we do not know in advance the number of docks which are present in the image. Thus, it allows us to represent the spatial structural knowledge without the need of specifying the number of instantiations.

Notice that in this approach we use two types of reasoning. The first one is the spatial reasoning used over the image to determine and represent its spatial knowledge. The second one is over the hierarchies T_C and T_R by using the relation *IS-A-KIND-OF*. This second type of reasoning allows to replace a concept or a relation by its descendants in the hierarchy. This type of reasoning is used when reasoning with ontologies or hierarchies as in [Dupin de Saint-Cyr and Prade, 2008, Metzger et al., 2003].

To find a graph homomorphism there are several algorithmic possibilities (see [Chein and Mugnier, 2008] for a review). One possibility is to express the problem as a Constraint Satisfaction Problem (CSP), where the relations represent the constraints and the concept nodes represent the variables of the CSP, and the values of the CSP are the regions of the image. By representing the graph homomorphism problem as a CSP we can take advantage of all the work that has been developed to find a solution of a CSP in an optimal way. In Section 5.3.1 the CSP and their solving strategies are introduced.

5.3.1 Constraint Satisfaction Problems (CSP)

Constraint Satisfaction Problems are a generic framework for expressing and solving problems whose aim is to find one or all solutions to a set of constraints. A constraint represents a relation, and a constraint satisfaction problem states which relations should hold among a given set of decision variables. A solution of a CSP is an assignment of values to all the variables that satisfy all the constraints. The CSP framework has been used to represent real world problems in different application areas, such as artificial intelligence, operations research, scheduling, supply chain management, graph algorithms, computer vision and computational linguistics, among others. For instance, in computer vision [Waltz, 1975, Rosenfeld et al., 1976] the problem of annotating geometrical figures according to a graph model is implicitly formulated as a CSP, where the checking of local consistencies led to the removal of inconsistent annotations. Other examples are [Tenenbaum and Barrow, 1977, Deruyver et al., 2009, Saathoff, 2006, Dasiopoulou et al., 2008, Saathoff and Staab, 2008] where a CSP is used to label the regions of an image according to a model. [Nempont, 2009] uses a CSP formalism to determine the region of an image where it is possible to find a region representing an object of a model.
A CSP is defined by a triplet $\mathcal{P} = \langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ [Bessiere, 2006] where:

- $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ is a set of n variables. Each variable represents a characteristic of the objects of the real problem;
- $\mathcal{D} = \{D_1, D_2, \ldots, D_n\}$ is a set of n domains. Each domain D_i is associated with the variable x_i and represents the set of values or states which can be assigned to x_i. The size of the largest domain of D is denoted by d;
- $\mathcal{C} = \{C_1, C_2, \ldots, C_t\}$ is a set of t constraints. A constraint C_k is defined through a pair $\langle R_k, S_k \rangle$, where R_k is a subset of the Cartesian product of the domain of the variables in $S_k \subseteq \mathcal{X}$. The arity of a constraint C_k is equal to the number of variables involved in it. We denote by r the maximum arity of the constraints of $\mathcal{P}$.

The CSP with $r = 2$ are called binary CSP, and are frequently represented as a graph, where the edges of the graph represent the constraints and the vertices the variables. We call $A = \{a_1, a_2, \ldots, a_n\}$ a solution of a CSP $\mathcal{P}$ if every $a_i \in D_i$ and for each $C_j \in \mathcal{C}$ its corresponding relation R_j holds on the projection of A onto S_j. We denote the projection of A onto a variable i by $A \downarrow_i$ and onto a set of variables S by $A \downarrow_S$. The set of all solutions of a $\mathcal{P}$ is denoted by $Sol_{\mathcal{P}}$. We say that $\mathcal{P}$ is consistent if $Sol_{\mathcal{P}} \neq \emptyset$.

Different problems can be associated to a CSP $\mathcal{P} = \langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$, some of them are:

(i) Determine whether $\mathcal{P}$ is consistent.

(ii) Search a solution to $\mathcal{P}$, that is search $A \in Sol_{\mathcal{P}}$.

(iii) Find the number of solutions.

(iv) Find the set of solutions $Sol_{\mathcal{P}}$.

Determining whether a CSP is consistent is an NP-complete decision problem, and finding a solution is NP-hard. The size of the research space of a CSP is equal to $\Pi_{1 \leq i \leq n} |D_i|$ [Condotta and Würbel, 2007]. The problem of consistency is usually relaxed by defining local consistencies, and this methodology is used in the filtering strategies, as we will see in the following paragraphs.

To find a solution to a CSP there are two basic types of procedures:

Search procedures: They consist of exploring one by one every combination of the domain of each variable and rejecting those combinations which do not satisfy one of the constraints. They are usually solved using backtracking and/or branch and bound algorithms. The conjunctive nature of the CSP ensures that this type of procedure always finds a solution when there exists one.

Inference or filtering: They correspond to algorithms used to simplify the CSP and to reduce the number of unnecessary explorations in the search procedures. They consist of reducing the domain of the variables by applying local consistency algorithms.

Usually a filtering procedure is applied followed by a search procedure.
In the following we describe in more detail the different types of local consistencies and some of the algorithms that have been proposed to solve them.

Consider a CSP $\mathcal{P} = \langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$ as defined above. The different types of local consistencies are:

Node consistency: Let C_k denote the unary constraint for the variable x_i. Then $\mathcal{P}$ is said to be node consistent if $D_i \subseteq R_k$. A CSP can be made node consistent by replacing, for each variable x_i, the domain D_i by $D_i \cap R_k$.

Arc consistency: This is the most common type of consistency. This type of consistency was first defined for binary CSP. The name "arc" refers to the directed edge in the constraint graph. Let C_k denote a binary constraint between the variables x_i and x_j. We say that the arc (i, j) is consistent if for every $a_i \in D_i$, there is a corresponding $b_j \in D_j$ such that $(a_i, b_j) \in R_k$. A CSP network is arc-consistent if all its arcs are arc-consistent. The notion of arc consistency can also be applied to n-ary networks, by defining arc-consistency on the domains of the variables [Chein and Mugnier, 2008]. Given a n-ary constraint C_k, and a variable $x_i \in S_k$, we say that the domain D_i of the variable x_i is arc-consistent relative to the constraint C_k, if $D_i \neq \emptyset$ and for every $v \in D_i$, there exists a tuple $A \in R_k$ such that $A \downarrow_i = v$. A domain D_i is arc-consistent if it is arc-consistent relative to all the constraints in $\mathcal{C}$. $\mathcal{P}$ is arc-consistent if all its domains are arc-consistent.
A CSP $\mathcal{P}$ can be made arc consistent by deleting for each variable x_i all the elements $v \in D_i$ for which there does not exist a corresponding tuple $A \in R_k$ such that $A \downarrow_i = v$. There are several algorithms for obtaining an arc-consistent CSP; the most simple one is AC-1 [Mackworth, 1977] which checks all the domains. The AC-2 algorithm [Mackworth, 1977] consists in removing arc inconsistencies which can never be part of any global solution. When those inconsistencies are removed they may propagate in neighboring arcs that were previously consistent. Those inconsistencies are in turn removed until the algorithm converges. The AC-3 algorithm [Mackworth, 1977] is

a generalization and simplification of AC-2 algorithm. The AC-4 algorithm [Mohr and Henderson, 1986] is again a generalization of AC-3 algorithm, it has a better worst case complexity than AC-3 algorithm, however it always reaches the worst case situation [Rossi et al., 2006], that is why the AC-3 algorithm continues to be one of the most used algorithms. Additionally, the AC-3 algorithm has the advantage of being independent of the data structure, and can be adapted to constraints of any arity. Algorithm 2 shows the AC-3 algorithm for n-ary constraints proposed by [Chein and Mugnier, 2008]. The algorithm keeps record of the constraints which have not yet been checked or must be checked again because the domain of at least one of their variables has been modified. These constraints are stored in the $ToCheck$ list, the algorithm ends when $ToCheck$ is empty. The sub-algorithm Revise (Algorithm 3) makes the domain of the variables in a constraint C_k arc-consistent relative to this constraint and marks the variables whose domain has changed so that they are considered for the consistency check of the other constraints.

k-consistency: This is the more general type of consistency. A CSP is k-consistent if for every tuple of k variables $(x_{i_1}, \ldots, x_{i_k})$ and for every instantiation $A = (a_1, \ldots, a_{k-1})$ of $(k-1)$ variables $(x_{i_1}, \ldots, x_{i_{k-1}})$. There exists $v \in D_k$, such that $(a_1, \ldots, a_{k-1}, v)$ is consistent. The arc-consistency is equivalent to 2-consistency.

Checking for arc-consistency to reduce the domain of variables is a good compromise between the reduction of the domain and the computational complexity. Only checking for node consistency does not make significant reduction on the domain, since no interaction between the variables is considered. Searching for k-consistency with $k > 3$ can considerably reduce the domain, while on the other hand it has a huge computational complexity.

As in [Chein and Mugnier, 2008] the graph homomorphism problem is solved by representing the problem as a CSP and using arc-consistency algorithm to reduce the domain of the variables, followed by a backtracking search.

Backtracking Once we have performed the reduction of the domains we can perform a backtrack algorithm to find the solutions to the CSP. The backtrack algorithm basically consists of extending a partial assignment of a solution by adding new values to the uninstantiated variables. If the partial solution cannot be extended because it has reached an inconsistency, then we reject the value of the last instantiated variable and try another value for this variable. This process is often represented as a search tree, where each node (below the root) represents a choice of a value for a variable, and each branch represents a candidate partial solution [Rossi et al., 2006].

5.3.2 Illustration

Consider the two conceptual graphs G_M and G_L of Figure 5.9. Suppose that we want to find the graph homomorphism from G_M to G_L using the CSP framework. The relation nodes of G_M correspond to the constraints, the concept nodes of G_M to the variables, and concept nodes of G_L to the domain of those variables. In our example there are two constraints representing the two relation nodes of graph G_M. One constraint corresponds to the adjacency relation between the node labeled Sea and the node labeled Ship, which we call Adjacency1. The other constraint is the adjacency relation between the node labeled

Input: A constraint network $\mathcal{P} = (\mathcal{X}, \mathcal{D}, \mathcal{C})$
Output: Computes the arc-consistent closure of $\mathcal{P}$ if it exists, otherwise returns Failure

$ToCheck \longleftarrow \mathcal{C}$;
while $ToCheck \neq \emptyset$ **do**
    Select C_k from $ToCheck$;
    foreach $x_{k_i} \in S_k$ **do**
        `// Marked as false in the vector. Change all the variables which belong to the domain of` C_k
        $Changed[k_i] \longleftarrow false$
    end
    $result \longleftarrow$ `Revise`(C_k) ; `// see Algorithm 3`
    if $result = EmptyDomain$ **then**
        return Failure ;
    if $result = Changed$ **then**
        foreach $C_k \neq C_l$ *such that there is* $x_j \in S_k \cap S_l$ *and* $Changed[j] = true$ **do**
            $ToCheck \longleftarrow ToCheck \cup \{C_l\}$;
        end

end

Algorithm 2: Basic algorithm used for arc-consistency checking.

as Ship and the one labeled as Dock, which we denote as Adjacency2. The set of variables is $\mathcal{X} = \{Sea : *, Ship : *, Dock : *\}$ and the initial domains are:

Domain Ship
Boat: R2
Ship: R3
Boat: R4
Boat: R7

Domain Sea
Sea: R1

Domain Dock
Dock: R5
Dock: R6

The constraints of the problem are:

Constraint Adjacency1	
Sea	Boat
Sea: R1	Boat: R2
Sea: R1	Ship: R3
Sea: R1	Boat: R4
Sea: R1	Boat: R7

Constraint Adjacency2	
Boat	Dock
Boat: R2	Dock: R5
Ship: R3	Dock: R6
Boat: R4	Dock: R5

This CSP problem is not arc-consistent, since there is no corresponding value in the domain of $Dock : *$ which is related to $Boat : R7$. To make the CSP arc consistent we apply the algorithm AC-3. When we revise the constraint Adjacency2, we remove $Boat : R7$ from the $Ship : *$ domain, and then the CSP becomes arc-consistent.

Now that the problem is arc-consistent, we can proceed to find the set of solutions. This is done by a backtracking procedure. Figure 5.10 illustrates the backtracking search tree for this example. First, we initialize a partial assignment of a solution by selecting the region $R2$ as a possible instantiation of the variable $Ship$. Then we add $R5$ for the instantiation of the variable $Dock$. The partial solution continues to be consistent, since the

Input: A constraint C_k
Output: Makes C_k arc-consistent if possible and marks variables whose domain has changed; returns EmptyDomain if a domain has been emptied, NoChange if no domain has been modified, otherwise Changed

`// Remove from` R_k `values which have been suppressed from domains`
foreach *variable* $x_{k_i} \in S_k$ **do**
    remove from R_k every tuple A such that $A \downarrow_{k_i} \notin D_{k_i}$
end
`// Make domains arc-consistent with respect to` C_k
$result \longleftarrow NoChange$
foreach *variable* $x_{k_i} \in S_k$ **do**
    foreach $v \in D_{k_i}$ **do**
        if $\nexists A \in R_k$ *such that* $v = A \downarrow_{k_i}$ **then** `//` v `has no support in` R_k
            remove v from D_{k_i} ;
            if $D_{k_i} = \emptyset$ **then**
                return EmptyDomain;
            else
                $result \longleftarrow Changed$
                $Changed[k_i] \longleftarrow true$
    end
end
return result;

Algorithm 3: Revise algorithm, sub-part of algorithm 2.

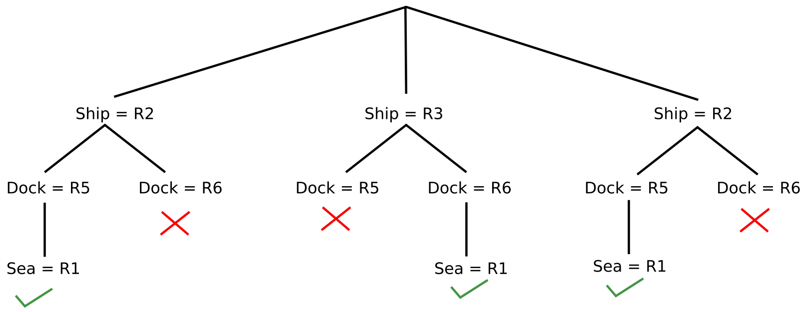


Figure 5.10: Backtracking search tree for example CSP.

relation Adjacency2 between $R2$ and $R5$ is satisfied. Finally we add $R1$ as the instantiation of the Sea variable. This assignment is complete and is consistent, therefore it is one of the solutions. Now we start another partial assignment by backtracking from the last solution and keeping all the elements equal to the last obtained solution except the last assigned variable, which is assigned with another value of its domain. However, since the Sea domain only consists of one value, there are no other possibilities, so we backtrack to the second to last assigned value and replace the instantiation of $Dock$ by R6. This partial assignment is inconsistent since it does not satisfy the Adjacency2 constraint, so the algorithm backtracks and tries another value for $Dock$. Since all the domain of dock has already been explored in this branch, then we backtrack and try another partial solution using $R3$ as an instantiation of $Ship$. This procedure is continued until the backtracking tree has been revised.

Finally, we obtain that $Sol(\mathcal{P}) = \{\{Sea : R1, Boat : R2, Dock : R5\}, \{Sea : R1, Boat : R4, Dock : R5\}, \{Sea : R1, Ship : R3, Dock : R6\}\}$. These three sets of solutions correspond to the configurations of the model. Since the model represented the spatial arrangement of a habor, we can conclude that the image contains a habor.

5.3.3 Discussion

In this section we presented a methodology to find the instantiations of a graph G_M in an image I_L based on the work of [Chein and Mugnier, 2008]. This problem was represented as a CSP, where the set of variables corresponds to the set of concept nodes in G_M, the domain of variables are the regions of the image and the relation nodes are the constraints. The relation nodes are the spatial relations that we want to verify in the image. For every variable x_i, corresponding to a concept node c, the domain D_i is composed of all the regions which have a label type less than the type of c. These regions are the candidates for being an instantiation of c.

Due to the CSP formalism it is not necessary to compute the graph of the labeled image I_L, when we want to verify if a constraint C_k, corresponding to a relation node of G_M, is satisfied in the image. We only need to observe if the relation is satisfied between the domains of the variables involved in the constraint, and these domains are sets of regions of the image. Thus, it is not necessary to compute all the possible relations between the regions but only those relations which appear in the graph model G_M. For instance in the situation of Figure 5.9, the relations $Surround$ or $GoInto$ that appear in the graph G_L of our example are not evaluated, since they do not provide any information to determine the instantiations of G_M in I_L. Therefore, the CSP formalism allows us to find the instantiations of the concepts directly in the image, and it is not necessary to perform the step of the graph creation representing all the possible relations that appear between the regions of the image. Computing only the relations which appear in the graph model is more efficient than computing all the possible relations.

5.4 Including fuzzy spatial relations

As in the previous section, suppose that we have a model G_M and a labeled image I_L. The spatial relations that appear among the objects in the image are now represented through the fuzzy models of Chapters 1 and 2. Also the graph model is a nested conceptual graph G_M, where we can have complex concept nodes representing groups of objects. When the spatial relations are evaluated as a degree, then determining whether a set of regions of I_L satisfies G_M becomes a matter of degree. Therefore, the notions presented in the previous section need to be revised.

If we continue to represent our problem as a constraint satisfaction problem, we need a formalism where it is possible to express fuzzy constraints. [Dubois et al., 1996] introduced the Fuzzy CSP (FCSP) to deal with flexible constraints. Flexible constraints include: fuzzy relations, soft constraints which express preferences between relations, and prioritized constraints which express the constraints which can be violated in the case where there exists a conflict.

The FCSP formalism is well adapted to our problem since it allows us to represent fuzzy relations. However, the representation of complex concept nodes needs an adaptation of this formalism that is presented in Sections 5.4.2 and 5.4.3.

5.4.1 Fuzzy constraint satisfaction problems (FCSP)

A FCSP is defined as $\mathcal{P} = \{\mathcal{X}, \mathcal{D}, \mathcal{C}\}$, where $\mathcal{X}$ and $\mathcal{D}$ correspond to a set of variables and the set of domains, as in a CSP. $\mathcal{C} = \{C_1, \ldots, C_t\}$ is a set of t flexible constraints. A flexible constraint C_k is defined through a pair $\langle R_k, S_k \rangle$, where $S_k \subset \mathcal{X}$ is the set of variables which involve C_k, and R_k is a fuzzy relation over the Cartesian product of the domains $D_{k_1} \times \ldots \times D_{k_n}$ of the variables in S_k. R_k is defined through its membership function $\mu_{R_k} : D_{k_1} \times \ldots \times D_{k_n} \to [0, 1]$ [Dubois et al., 1996]. Moreover, let $\{v_{k_1}, \ldots, v_{k_n}\} \in D_{k_1} \times \ldots \times D_{k_n}$, then:

$$\begin{aligned}
\mu_{R_k}(v_{k_1}, \ldots, v_{k_n}) = 1 &\text{ means } (v_{k_1}, \ldots, v_{k_n}) \text{ totally satisfies } C_k. \\
\mu_{R_k}(v_{k_1}, \ldots, v_{k_n}) = 0 &\text{ means } (v_{k_1}, \ldots, v_{k_n}) \text{ totally violates } C_k. \\
0 < \mu_{R_k}(v_{k_1}, \ldots, v_{k_n}) < 1 &\text{ means } (v_{k_1}, \ldots, v_{k_n}) \text{ partially satisfies } C_k.
\end{aligned}$$

The consistency of a FCSP is also a matter of degree. Given an instantiation of $\{v_1, \ldots, v_n\} \in D_1 \times \ldots \times D_n$ the degree to which $\{v_1, \ldots, v_n\}$ satisfies $\mathcal{P}$ is given as a conjunction of the satisfaction of each of its constraints [Dubois et al., 1996]:

$$Cons(v_1, \ldots, v_n) = \min_{\tilde{C}_k \in \mathcal{C}} \mu_{R_k}((v_1, \ldots, v_n) \downarrow_{S_k}) \tag{5.1}$$

where $(v_1, \ldots, v_n) \downarrow_{S_k}$ represents the projection of $(v_1, \ldots, v_n)$ onto the set of variables S_k. The consistency of an instantiation in this context is further discussed in Section 5.4.6.

With the introduction of fuzzy relations in the FCSP model, determining whether a value $v \in D_i$ is suitable for representing a variable x_i becomes a matter of degree. Then we can represent this information as a fuzzy set μ_{x_i} over D_i:

$$\begin{aligned} \mu_{x_i} : \quad D_i &\longrightarrow [0,1] \\ v &\longmapsto \mu_{x_i}(v) \end{aligned}$$

where μ_{x_i} accounts for the imprecise or incomplete knowledge about x_i. The membership value of $v \in D_i$ depends on the information that we have about its satisfaction of the constraints in which it is involved and on its reliability. Initially when we have not verified which are the constraints that are verified or violated by v, we can consider that $\mu_{x_i}(v) = 1$ since we know that v belongs to D_i. Subsequently when we have more knowledge about the degree of satisfaction of the relations which include x_i, then the membership function is affected by this knowledge. We will refer to μ_{x_i} as the membership function of the variables x_i.

Most of the definitions of CSP can be extended to FCSP (see [Dubois et al., 1996]). [Dubois et al., 1996] extend the definition of arc-consistency for binary networks. Given a FCSP $\mathcal{P} = \{\mathcal{X}, \mathcal{D}, \mathcal{C}\}$ of binary fuzzy constraints, we say that $\mathcal{P}$ is arc-consistent, if for every constraint C_k involving the variables x_i and x_j, we have that every $u \in D_i$ satisfies:

$$\mu_{x_i}(u) \leq \sup_{(u,v)\in D_i \times D_j} \min[\mu_{R_k}(u,v), \mu_{x_j}(v)] \tag{5.2}$$

This means that the fuzzy set μ_{x_i}, representing the possible values of x_i, should be included in the projection on x_i of the conjunction of the fuzzy set μ_{R_k} with the cylindrical extension of the fuzzy set μ_{x_j}. Alternatively, this can be seen as a translation into fuzzy expressions of the arc-consistency expression in the crisp case:

$$\forall u \in D_i \ \exists v \in D_j \text{ such that } R(u,v)$$

Moreover, we can also define arc-consistency over the domains, as it was done for CSP in [Chein and Mugnier, 2008] (see Section 5.3.1). The advantage of defining arc-consistency over the domains is that it is not only restricted to binary constraints. Let C_k be a n-ary constraint and x_i a variable belonging to S_k. In the crisp case, we say that the domain D_i of the variable x_i is arc-consistent relative to the constraint C_k, if $D_i \neq \emptyset$ and for every $v \in D_i$, there exists a tuple A which satisfies R_k and for which all its members belong to their respective domains. So we propose to translate this to the fuzzy case as:

$$\mu_{x_i}(v) \leq \sup_{A=(a_{k_1},\ldots,a_{k_n}):A\downarrow_i = v} \min[\mu_{R_k}(a_{k_1}, \ldots, a_{k_n}), \min_{j=k_1,\ldots,k_n} \mu_{x_j}(a_{k_j})]. \tag{5.3}$$

The AC-3 algorithm used to compute the arc-consistent closure of a CSP was extended to the FCSP framework by [Dubois et al., 1996]. However, the algorithm in [Dubois et al., 1996] only deals with binary constraints. In Algorithms 4 and 5 we propose an extension of the AC-3 algorithm to deal with fuzzy constraints of any arity. In the following we refer to this algorithm as FAC-3. This algorithm is based on the algorithms presented in [Dubois et al., 1996] for FCSP and the algorithm of [Chein and Mugnier, 2008] for CSP containing constraints of any arity (shown in Algorithm 3). The main algorithm is identical to the one of [Dubois et al., 1996]. In this algorithm we keep record of the constraints which have not been revised to ensure that its domains are arc-consistent. When the membership function of a variable changes, all the constraints related to that variable are added to the check list so that they can be revised for arc-consistency. The variable *ConsSup* saves the maximum

consistency value for a solution of the FCSP $\mathcal{P}$.
We propose a RevisedFuzzyConstraint function which checks constraints of any arity as in [Chein and Mugnier, 2008] for the crisp case. When calling RevisedFuzzyConstraint(C_k) the domain of the variables in S_k are made arc-consistent by modifying their membership functions. For every value $v \in D_i$ of a variable $x_i \in S_k$, the membership value of $\mu_{x_i}(v)$ is replaced to the maximum of Equation 5.3. If $\mu_{x_i}(v)$ is equal to zero then v is removed from the domain of D_i. If μ_{x_i} changes, then the variable x_i is marked as changed in order to consider the arc-consistency check with respect to other constraints.

For each variable x_i the initial membership function of μ_{x_i} is a constant function over D_i equal to one. However, as we have more knowledge about the degree of satisfaction of a relation which involves x_i for the regions in the domain D_i, the membership μ_{x_i} is modified to incorporate this new knowledge.

When we evaluate a relation which can be represented as a fuzzy landscape (see Part I) we will use a mean measure (Equation 1.16) to evaluate the satisfiability of the spatial relation.

Input: A FCSP $\mathcal{P} = (\mathcal{X}, \mathcal{D}, \mathcal{C})$
Output: Computes the arc-consistent closure of $\mathcal{P}$ if it exists, otherwise returns Failure

$ConsSup = 1$
$ToCheck \longleftarrow \mathcal{C}$
while $ToCheck \neq \emptyset$ **do**
  Select C_k from $ToCheck$
  foreach $x_{k_i} \in S_k$ **do**
    $Changed[k_i] \longleftarrow false$
  end
  $result \longleftarrow$ `ReviseFuzzyConstraint`(C_k) ; `// see Algorithm 5`
  if $result = EmptyDomain$ **then**
    return Failure ;
  if $result = Changed$ **then**
    foreach $C_k \neq C_l$ *such that there is* $x_j \in S_k \cap S_l$ *and* $Changed[j] = true$ **do**
      $ToCheck \longleftarrow ToCheck \cup \{C_l\}$
    end
end
return $ConsSup$;

Algorithm 4: FAC-3 algorithm used for determining the arc-consistency of a FCSP [Dubois et al., 1996].

The extension of the AC-3 algorithm to fuzzy constraints increases the complexity by a factor t_α, where t_α is the number of satisfaction levels used to discretized the interval [0,1] to represent the satisfaction of the relations and of the membership functions. Therefore the complexity of FAC-3 is $O(t_\alpha|\mathcal{C}|cd^{c+2})$ where d is the maximum length of a domain and c is the maximum arity of a constraint.

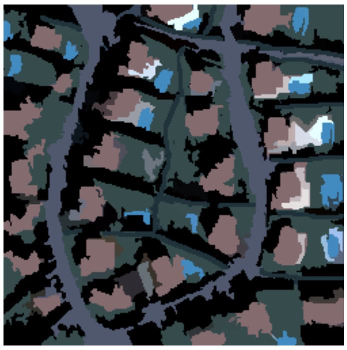
Input: $C_k, Changed, ConsSup$

Output: Makes C_k arc-consistent if possible and marks variables whose domain has changed in the $Changed$ vector

foreach *variable* $x_{k_i} \in S_k$ **do**
  remove from R_k every tuple A such that $A \downarrow_{k_i} \notin D_{k_i}$
end
$Height \longleftarrow 0$
$result \longleftarrow NoChange$
foreach *variable* $x_{k_i} \in S_k$ **do**
  foreach $v \in D_{k_i}$ **do**
    $newDegree \longleftarrow 0$
    foreach *tuple* $A = (a_{k_1}, \ldots, a_{k_n})$ *in the domain of* R_k *such that* $v = A \downarrow_{k_i}$ **do**
      $eval \longleftarrow \min[\mu_{R_k}(a_1, \ldots, a_{k_n}), \min_{j=k_1,\ldots,k_n} \mu_{x_j}(a_{k_j})]$
      $height \longleftarrow \max(eval, height)$
      $newDegree \longleftarrow \max(eval, newDegree)$
    end
  end
  if $newDegree = 0$ **then**
    Delete v from D_{k_i}
    if $D_{k_i} = \emptyset$ **then**
      return $EmptyDomain$
  if $newDegree \neq \mu_{x_{k_i}}(v)$ **then**
    $Change[k_i] \longleftarrow true$
    $\mu_{x_{k_i}}(v) \longleftarrow newDegree$
    $result \longleftarrow Change$
end
$ConsSup \longleftarrow \min(ConsSup, Height)$
return $result$

Algorithm 5: ReviseFuzzyConstraint method.

(a) Example image.

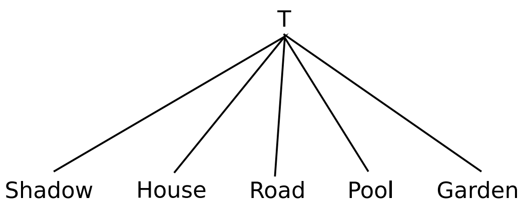


(b) Segmented and manually labeled image.

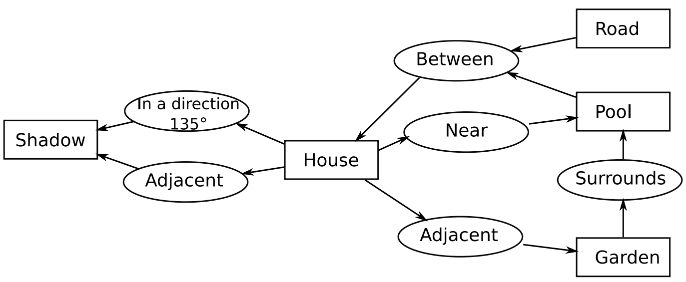
Figure 5.11: Example image and manually labeled image. The labels in Figure (b) are: green represent the gardens, orange represents the houses, black represents the shadows, blue the pools, and gray represents the roads.

Illustration of the algorithm

To illustrate how the FAC-3 algorithm works, we are going to search in Figure 5.11(a), the houses having a pool which is located in the garden at the "back" of the house and the house has a shadow. There are several ways to describe the spatial organization that is satisfied between the houses, their shadow, the pool, the garden and the road. One possible way is given by the conceptual graph shown in Figure 5.12(b), which is built over the vocabulary shown in Figure 5.12(a). We represented that the pool is at the "back" of the house through the relation "between", by saying that the house is "between" the road and the pool.



(a) Concept hierarchy.



(b) Conceptual graph.

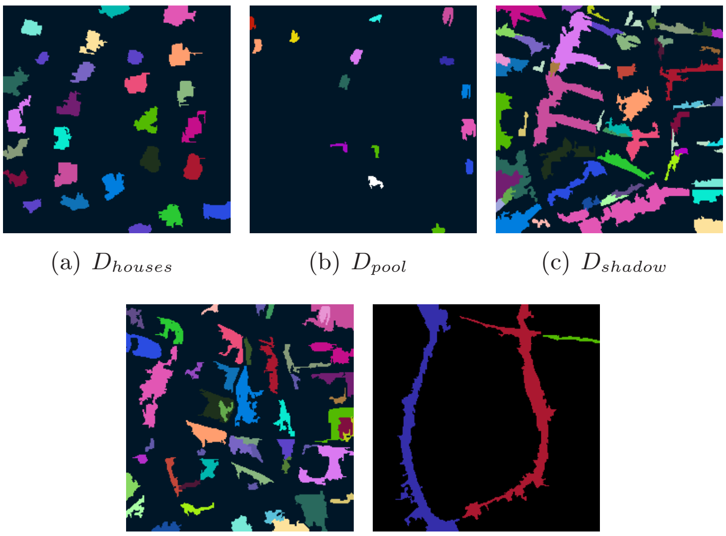
Figure 5.12: Concept hierarchy and conceptual graph describing "the house has a pool in its garden, the house is between the pool and the road, and the house has a shadow".

To search instantiations of the conceptual graph in the image, we search for the instantiations of the graph on the image shown in Figure 5.11(b) which corresponds to a segmentation of Figure 5.11(a) which has been manually labeled. The corresponding CSP problem is formulated as:

- $\mathcal{X} = \{x_{shadow}, x_{house}, x_{road}, x_{pool}, x_{garden}\}$ containing the variables representing the concept nodes of the conceptual graph.

- $\mathcal{C} = \{C_{direction}, C_{adjacent_1}, C_{between}, C_{near}, C_{surrounds}, C_{adjacent_2}\}$, where $C_{direction} = \langle \mu^{135^\circ}_{direction}, \{x_{shadow}, x_{house}\}\rangle$, $C_{adjacent_1} = \langle \mu_{adjacent}, \{x_{shadow}, x_{house}\}\rangle$, $C_{between} = \langle \mu_{between}, \{x_{pool}, x_{road}, x_{house}\}\rangle$, $C_{near} = \langle \mu_{near}, \{x_{house}, x_{pool}\}\rangle$, $C_{surrounds} = \langle \mu_{surrounds}, \{x_{garden}, x_{pool}\}\rangle$ and $C_{adjacent_2} = \langle \mu_{adjacent}, \{x_{house}, x_{garden}\}\rangle$. Notice that there are two constraints representing the spatial relation of adjacency, which we named $C_{adjacent_2}$ and $C_{adjacent_1}$ to differentiate them.

- $\mathcal{D} = \{D_{shadow}, D_{house}, D_{road}, D_{pool}, D_{garden}\}$ where D_i represents the possible candidates for x_i. The regions in each D_i are shown in Figure 5.13.



(a) D_{houses} (b) D_{pool} (c) D_{shadow}

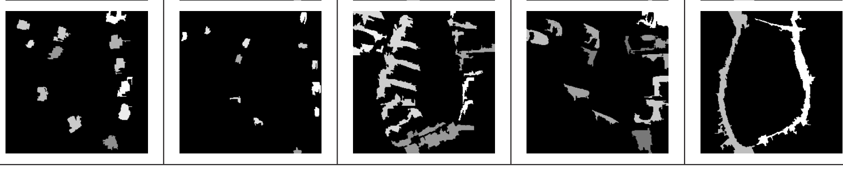
(d) D_{garden} (e) D_{road}

Figure 5.13: Regions in the domains of each variable of the CSP problem representing the instantiations of the conceptual graph in Figure 5.12 in the image shown in Figure 5.11(b). (manually labeled)

Table 5.1 shows the evolution of each μ_{x_i} over D_i for $i = \{house, pool, shadow, garden, road\}$ at different iterations after calling ReviseFuzzyConstraint. The membership value of a region $v \in D_i$ is represented by the gray label of the region. Lighter colors represent a higher membership value.
In the first iteration $C_{between}$ is checked, mainly modifying the membership values of $\mu_{x_{houses}}$ and eliminating the regions in D_{house} which do not satisfy this constraint. At iteration 2, the constraint $C_{direction}$ is checked, which clearly modifies μ_{shadow}. Between iterations 2 and 7 the constraints $C_{between}$ and $C_{direction}$ are rechecked until the membership values representing the domains of the variables in $S_{between}$ and $S_{direction}$ are arc-consistent with respect to these two constraints. Then in iteration 7 the constraint $C_{surrounds}$ is revised which causes the elimination of some elements of D_{pool} and D_{garden}. The algorithm continues revising all the constraints, until it converges to a solution at

iteration 32, illustrated in the last row of Table 5.1. μ_{houses} has a high membership value for those houses which have a pool in the garden at the "back" of the house.

Table 5.1: Evolution of the fuzzy sets μ_{x_i} over D_i for $i = \{houses, pool, shadow, garden, road\}$ when making each D_i arc-consistent.

Iteration	μ_{houses}	μ_{pool}	μ_{shadow}	μ_{garden}	μ_{road}
0					
1					
2					
7					
10					
13					
17					
24					
32					

Algorithm 4 is well adapted to reduce the domains of the variables when we only consider fuzzy relations over simple concept nodes. Nevertheless, when we deal with complex concept nodes representing groups of regions satisfying certain relations between them, it is necessary to adapt the FCSP.

The idea of incorporating a formalism to a CSP to deal with groups of objects has also been discussed in [Deruyver and Hodé, 1997], where the authors propose bi-level constraints to solve the problem of labelling the regions in an image using a conceptual graph (only containing simple nodes). In [Deruyver and Hodé, 1997] groups are defined using an equivalence relation over the set of regions, then each group is a subset of the members of an equivalence class and each group represents a concept of the conceptual graph. All the members of a group belong to a same domain. A bi-level constraint consists of one constraint representing the relations between the groups and a second constraint representing the equivalence relation between the members of a group. Although the formalism proposed in [Deruyver and Hodé, 1997] is appropriate for interpreting an image using a conceptual graph only having simple concept nodes, it cannot be adapted to our context, because we consider groups of objects which can have several relations between them and which belong to different domains. Additionally, in [Deruyver and Hodé, 1997], due to the spatial relations used in their vocabulary, they have the property that a relation is satisfied between two groups if there exists a member of each group for which the relation is satisfied. In our case we do not have this property since we are dealing with relations such as parallelism for which it is necessary to consider the group as a whole, and not only its members. Therefore, it is necessary to introduce another formalism which is more appropriate for our context.

We can see the complex concept nodes as having a duality, since they can be seen as a constraint or as a variable. They are seen as a variable when considered as objects that can satisfy spatial relations, and they are seen as a constraint when we evaluate the relations or spatial properties that should be satisfied by the group of objects. Therefore, we propose to define this constraint/variable by a relation representing the conjunction of all the conditions that should be satisfied inside the complex concept node, and a membership degree to the domain of groups which depends on the satisfaction of the conditions inside the nested node, as well as the satisfaction of the relations of the group with other objects. In the following, we explain in more detail this conception of the complex concept nodes, which is an important feature of our contribution. We differentiate between the complex concept nodes which represent an alignment and those which do not. We refer to the first ones as alignments and to the latter ones as groups.

5.4.2 Dealing with alignments

For the sake of simplicity, the aligned groups of objects are only considered for objects belonging to the same concept type. Let x_i be the variable representing the type of the objects involved in the alignment, and let D_i be its domain. Let x_g denote the variable which represents the aligned group of objects when seen as a variable, and let C_w represent the constraint of alignment of the group.

When the group is seen as a variable x_g, the domain D_g is a subset of the power set of D_i. For a group $V = \{v_1, \dots, v_p\} \in D_g$ the membership value $\mu_{x_g}(V)$ depends on three factors:

1. the degree of alignment of V,
2. the degree of satisfaction of the spatial constraints (spatial relations) that are supposed to be satisfied by x_g, and
3. the membership value of its members to D_i.

When the groups are extracted we do not have any information about the satisfaction of the spatial relations which involve V. Therefore, its initial degree of satisfaction is equal to the conjunction of the condition of being aligned and of the degree of satisfaction that each $v_j \in V$ is an instantiation of the concept represented by the variable x_i, that is:

$$\mu_{x_g}(V) = \min[\mu_{ALIG}(V), \min_{j=1,\dots,p} \mu_{x_i}(v_j)] \tag{5.4}$$

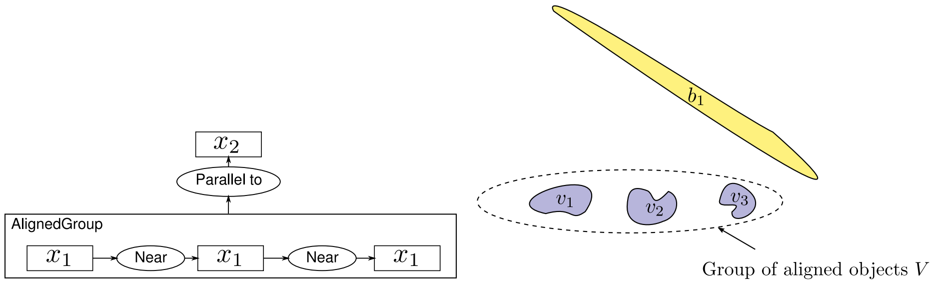
where μ_{ALIG} is the degree of global alignment given by Equation 2.30. The membership value $\mu_{x_g}(V)$ changes as we acquire more information about the membership value of its members to D_i and about the satisfaction degree for V of the relations in which the instantiations of the variable x_g should be involved.

When the group is seen as a constraint C_w, it evaluates the property of alignment of a group. Remember that every constraint C_k is defined through a pair $\langle R_k, S_k \rangle$, where R_k is the relation that represents the constraint and S_k is set containing the variables involved in C_k. Usually a relation representing a constraint is defined as a subset of a Cartesian product of the domains of its variables, and the number of domains and variables is fixed. However, this is not the case for the constraint C_w, since each group can have a different number of elements, which is a priori unknown. Therefore, to properly define the relation representing this constraint, it would be necessary to specify for each $n \geq 3$, a relation with arity n, to define the groups of aligned objects in D_i^n. Due to the lack of knowledge about the number of elements in a group, we use as a simplification the same notation to define the relations for each possible arity. Then the degree of satisfaction of the relation of a tuple $(v_1, \dots, v_n) \in D_i^n$ is equal to the degree of alignment of the set $V = \{v_1, \dots, v_n\}$, and of the conjunction of the degrees μ_{x_i} of its elements:

$$\mu_{R_w}(v_1, \dots, v_n) = \min[\mu_{ALIG}(V), \min_{j=1,\dots,p} \mu_{x_i}(v_j)] \tag{5.5}$$

where $S_w = \{x_i\}$, where x_i represents the variable of the objects that satisfy the relation since the relation is only defined over one type of objects. At the beginning both μ_{R_w} and μ_{x_g} are identical. However, as we make D_g arc-consistent with respect to other constraints, the values of μ_{R_w} and μ_{x_g} start to differ. For a group $V \in D_g$, the value $\mu_{x_g}(V)$ includes the information regarding the satisfaction of the constraints involving x_g, while μ_{R_w} only represents the information concerning the satisfaction of the alignment relation and of the degree of membership of each element $v \in V$ to D_i. For example Figure 5.14 shows an example where $\mu_{R_W} = \mu_{x_g}$ are equal when we do not have any information about the parallel relation. After we have calculated the degree of parallelism between V and b_1, the value of μ_{x_g} changes, while μ_{R_W} remains the same.

When having a complex concept node representing an aligned group of objects, it is necessary to consider both degrees μ_{R_w} and μ_{x_g} because they represent different types of information. The membership μ_{R_W} remains independent of the relations that the group satisfies, while the membership value μ_{x_g} varies according to the interaction of the group with other elements.



(a) Conceptual graph. (b) Domains.

Figure 5.14: Illustration of the dual characteristic of the group V of aligned objects. Let $D_1 = \{v_1, v_2, v_3\}$, $D_2 = \{b_1\}$. Suppose that for every $v_j \in V$ we have $\mu_{x_1}(v_j) = 1.0$. Before we have any information about the satisfaction of the parallel relation we have $\mu_{x_3}(V) = \mu_{R_{aligned}}(V) = 1$, where x_3 is the variable that represents the group, and $C_{aligned} = \langle R_{aligned}, S_{aligned} \rangle$ is the alignment constraint of the group. Suppose that we find that $\mu_{R_{parallel}}(V, b) = 0.3$, then the membership value of V belonging to the domain of x_3 becomes $\mu_{x_3}(V) = 0.3$, while $\mu_{R_{aligned}}(V) = 1$.

The dual characteristic of this constraint obliges us to be careful with its evaluation and it is necessary to evaluate the constraint before it is evaluated as a variable. This precaution should be considered when searching the arc-consistency closure. Additionally, there are several considerations that should be taken into account when making D_g arc-consistent. In the following we discuss these considerations.

Considerations for making the domain of an aligned group arc-consistent

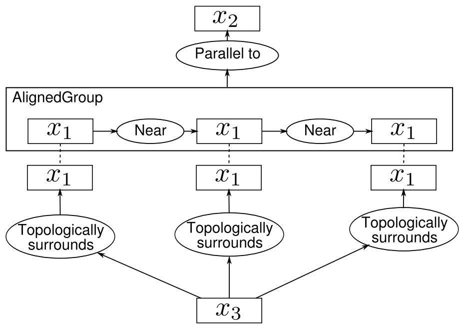
The algorithm presented in Section 2.2 to extract groups of aligned objects allows us to extract long groups of objects. When we have to determine a group of objects that satisfies a constraint, we are interested in searching for the longest group satisfying that constraint. Therefore we propose to find the groups of aligned objects, and, when it is necessary, to prune the elements of the group, and then find the subgroups which are aligned until one of these subgroups satisfies the constraint or disappears. If we have a group of aligned objects V such that $\mu_{ALIG}(V) = \alpha$, and suppose we have to eliminate an element $v \in V$, then an aligned subgroup of $V \setminus \{v_j\}$ is an aligned group $U \subseteq V \setminus \{v_j\}$ such that $\mu_{ALIG}(V) \leq \mu_{ALIG}(U)$. We can assume that the orientations of U and V are very similar.

When considering only some groups of aligned objects to verify the constraints, there are certain precautions that should be taken into account. First, we will discuss the considerations dealing with the changes in the membership function μ_{x_i}, and then, those referring to the situation when the group is seen as an object.

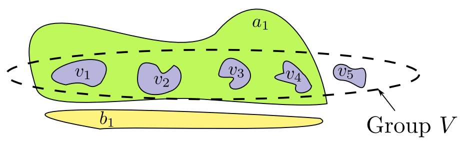
A member $v_j \in V$ no longer belongs to D_i

Suppose that there is $v_j \in V$ which has a zero degree with respect to the membership function of the variable x_i, i.e: $\mu_{x_i}(v_j) = 0$. By Equation 5.4 we know that this implies that $\mu_{x_g}(V) = 0$. However, it is possible that there is an aligned subgroup $U \subseteq V \setminus \{v_j\}$ which is part of a consistent solution. For example, suppose that we want to find the instantiations of the conceptual graph of Figure 5.15(a) on the image of Figure 5.15(b). The corresponding CSP problem is formulated as:

- $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$, where x_1, x_2 and x_3 correspond to the concept nodes of Figure 5.15(a) and x_4 corresponds to the alignment variable.
- $\mathcal{D} = \{D_1, D_2, D_3, D_4\}$, where $D_1 = \{v_1, v_2, v_3, v_4, v_5\}$, $D_2 = \{b_1\}$ $D_3 = \{a_1\}$ and $D_4 = \{V\}$ for $V = \{v_1, v_2, v_3, v_4, v_5\}$.
- $\mathcal{C} = \{C_{topologically\_surround}, C_{parallel\_to}, C_{aligned}\}$ where $C_{topologically\_surround} = \langle \mu_{topologically\_surround}, \{x_3, x_1\}\rangle$, $C_{parallel\_to} = \langle \mu_{||}, \{x_4, x_2\}\rangle$ and $C_{aligned} = \langle \mu_{ALIG}, \{x_1\}\rangle$.



(a) Conceptual graph



(b) Domains

Figure 5.15: The element v_5 belonging to the group V, formed by $V = \{v_1, v_2, v_3, v_4, v_5\}$, does not satisfy the relation "topologically surrounds" which should be satisfied by all the members of the group. However, there are subgroups of V which satisfy all the constraints.

Suppose that we follow the same procedure as in the previous section to make the domains of $\mathcal{D}$ arc-consistent. When we check the arc-consistency condition for the domains involved in $C_{topologically\_surround}$ we obtain $\mu_{x_1}(v_5) = 0$, and therefore a modification of the fuzzy set μ_{x_1} over D_1. Thus, the constraint $C_{aligned}$ has to be checked, because $x_1 \in S_{aligned}$. When we revise $C_{aligned}$ the membership function μ_{x_4} is updated, and $\mu_{x_4}(V) = 0$. Hence, D_4 is empty and no solution is found. Nevertheless, the group $U = \{v_1, v_2, v_3, v_4\} \subseteq V$ satisfies the constraints $C_{aligned}$ and $C_{parallel}$ and all its members satisfy $C_{topologically\_surround}$, thus it makes part of a consistent instantiation. Therefore, eliminating V without considering if any of its subgroups satisfied the relations was a hurried decision. Therefore, we propose to add the aligned subgroups of $V \setminus \{v_5\}$ to the domain D_4 and reconsider the new domain for the constraint checking. In this case we obtain a solution.

By starting with a large group V of N elements and reconsidering the aligned subgroups of V, we have a worst case scenario of $O(Nlog(N))$ to make D_g arc-consistent. Another strategy is to consider initially all the aligned subgroups of V for the domain of D_g, then in a worst case scenario, the size of the domain of the constraint is $\sum_{k=3}^{N} \binom{N}{k}$. If we had followed this second strategy, the result for the example would have been 4 groups of 3 elements and 1 group of 4 elements. But, we are interested in the longest group which satisfies the constraint, so the result is reduced to the group of 4 elements which is the same answer as the one we obtain by using the first strategy. The low complexity obtained when using one group and then prunning it, compared to the strategy of using all the possible subsets, justifies why it is more efficient to use this first strategy.

The membership function μ_{x_i} has changed

Let $v_j \in V$ be a member of a group. Suppose that the degree of membership of v_j

to its domain D_i has decreased after we have checked the arc-consistency condition. The decrease of $\mu_{x_i}(v_j)$ means that it is less possible that x_i takes the value v_j than when we extracted the aligned subsets of D_i. To integrate this new information into the membership function of V i.e. $\mu_{x_g}(V)$, we propose to continue consider V as a possible candidate of D_g. However, we propose to add to D_g the groups of aligned objects belonging of $V \setminus \{v_j\}$, and update μ_{x_g}. By adding these new groups to D_g, the relations that involve x_g have to be reevaluated, since D_g has changed.

The group V, when seen as a variable, does not satisfy a constraint

Due to the fact that we are considering only certain groups, we need to consider what happens to the aligned subgroups of a group V when $\mu_{x_g}(V) = 0$, that is, when V does not satisfy a constraint that should be satisfied by the members of D_g.

Let us first remember that when a spatial relation is represented as a fuzzy landscape (see Part I) we are using a mean measure (Equation 1.16) to evaluate the satisfiability of the spatial relation with respect to a target object. For instance, suppose that A and B are two regions of the image and we want to evaluate to what degree A is "near" B. Let γ_{near} be the membership function representing the fuzzy landscape defining the region "near" B, then the degree of satisfaction of the relation A is "near" B is given by:

$$\mu_{near}(b,a) = \frac{\sum_{p \in A} \gamma_{near}(p)}{|A|}$$

Now let us consider the situation when x_g is the target object of the relation.

Proposition 5.1. *Let $A = \{a_1, \ldots, a_n\} \in D_g$ be an aligned group and $C_t = \langle \mu_{R_t}, S_t \rangle$ be a binary constraint, such that $S_t = \{x_g, x_u\}$. Let $m \in D_u$, suppose that x_g is the target object of the relation represented by C_t and that $\mu_{R_t}(m, A) = 0$. Then for every aligned subgroup $B \subseteq A$ we have $\mu_{R_t}(m, B) = 0$.*

Proof. We consider each of the possible relations in T_R in which an alignment can be involved.

(a) Let R_t be a relation which produces a fuzzy landscape denoted by γ_R.
If $\mu_{R_t}(m, A) = 0$, then it means that for every point $p \in A$ we have $\gamma_R(p) = 0$. This holds in particular for the points $p \in B \subset A$, therefore $\mu_{R_t}(m, B) = 0$.

(b) If R_t corresponds to the adjacency relation, and $\mu_{adj}(m, A) = 0$, then it is straightforward to see that $\mu_{adj}(m, B) = 0$. The same is true when x_g is the reference object of the relation because of the symmetry of the relation.

(c) If R_t corresponds to the parallel relation, and $\mu_{parallel}(m, A) = 0$, then it is either because the visibility condition is zero or because the orientations are not similar (see Section 2.3). If the visibility condition is zero, then this is also true for B for the same reason as in (a). If the orientations are not similar then the orientation of B is not similar to the one of m, since the orientation of B can be considered equal to the orientation of A (see Part I), and therefore $\mu_{parallel}(m, B) = 0$

□

Now, let us consider the case when x_g is the reference object of the relation:

Proposition 5.2. *Let* $A = \{a_1, \ldots, a_n\} \in D_g$ *be an aligned group and* $C_t = \langle \mu_{R_t}, S_t \rangle$ *be a binary constraint, such that* $S_t = \{x_g, x_u\}$*. Let* $m \in D_u$*, suppose that* x_g *is the reference object of the relation represented by* C_t *and that* $\mu_{R_t}(A, m) = 0$*. Then for every aligned subgroup* $B \subseteq A$ *we have that* $\mu_{R_t}(B, m) = 0$*.*

Proof. (a) If R_t is a relation which produces a fuzzy landscape, then R_t should be one of the following metric relations: distance relation, "on one side" or "in a direction of angle α". All of these relations produce a fuzzy landscape which is increasing with respect to the reference object. Therefore, the landscape produced by B is contained by the one produced by A. So, $\mu_{R_t}(B, m) = 0$.

(b) If R_t is the parallel relation, and $\mu_{parallel}(m, A) = 0$, then it means that either the visibility condition is not satisfied or the orientation condition is not satisfied. In the case when the visibility condition is not satisfied, then as in previous case it is not satisfied by B. If the orientation condition is not satisfied then it is not satisfied by B either.

□

Using similar arguments we can extend the previous proposition to deal with the ternary relation "between":

Proposition 5.3. *Let* $A = \{a_1, \ldots, a_n\} \in D_g$ *be an aligned group and* C_t *be a constraint representing the "between" relation, such that* $S_t = \{x_g, x_u, x_v\}$*. Let* $m \in D_u$*,* $q \in D_v$*, and suppose that* A *is one of the reference objects of the relation and that* $\mu_{R_t}(A, m, q) = 0$*. Then for every aligned subgroup* $B \subseteq A$ *we have* $\mu_{R_t}(B, m, q) = 0$*.*

Proposition 5.4. *Let* $A = \{a_1, \ldots, a_n\} \in D_g$ *be an aligned group and* C_t *be a constraint representing the "between" relation, such that* $S_t = \{x_g, x_u, x_v\}$*. Let* $m \in D_u$*,* $q \in D_v$*, and suppose that* A *is the target object of the relation and that* $\mu_{R_t}(m, q, A) = 0$*. Then for every aligned subgroup* $B \subseteq A$ *we have* $\mu_{R_t}(m, q, B) = 0$*.*

We can conclude that when a group V does not satisfy a relation, then none of its aligned subgroups $U \subseteq V$ satisfies the constraint. Therefore, we can remove V from D_g without worrying about the subgroups. Notice that the propositions are not guaranteed to be true if we use other measures to evaluate the relations which are represented as fuzzy landscapes. For instance, that would be the case if we used the necessity (Equation 1.7).

Using a degree μ_{x_g} to represent the group as an object, and a degree μ_{R_w} to represent it as the spatial property of the group, allows us to identify the above cases and to handle them in an appropriate way.

5.4.3 Dealing with groups (not necessarily aligned)

The other types of complex concept nodes are the groups which have a conceptual graph not representing an alignment. Suppose that we want to represent a complex concept node with child concept nodes represented by the set of variables $\{x_{k_1}, \ldots, x_{k_n}\}$ and child relations $\{R_{t_1}, \ldots, R_{t_p}\}$.

When it is observed as a variable x_g its domain is $D_g = D_{k_1} \times \ldots \times D_{k_n}$. Let $V = (v_{k_1}, \ldots, v_{k_n}) \in D_{k_1} \times \ldots \times D_{k_n}$, its membership value $\mu_{x_g}(V)$ depends on:

- the degree of satisfaction of the child relations,

- the degree of satisfaction of the spatial constraints (spatial relations) that are supposed to be satisfied by x_g, and
- the membership value of its members to their respective domains.

Therefore, for the initial membership value $\mu_{x_g}(V)$ we do not have any information about the degree of satisfaction of V with other regions, so its initial degree is equal to the conjunction of the degree of satisfaction of all the child relations with the conjunction of the respective membership values of all of its members, that is:

$$\mu_{x_g}(V) = \min \left[\bigwedge_{j=t_1}^{t_p} \mu_{R_j}(V \downarrow_{S_j}), \min_{h=k_1,\ldots,k_n} \mu_{x_h}(v_h) \right] \tag{5.6}$$

When it is seen as a constraint C_w, it evaluates that all the conditions inside the complex concept node are satisfied. Unlike the alignment case, we know in advance which are the members inside the complex concept node. Therefore the set of variables in C_w is $S_w = \{x_{k_1}, \ldots, x_{k_n}\}$, which is the union of the sets of variables of its child relations: $S_w = \cup_{j=t_1}^{t_p} S_j$. For a tuple $V = (v_{w_1}, \ldots, v_{w_n}) \in D_{w_1} \times \ldots \times D_{w_n}$ the degree of satisfaction of the relation is:

$$\mu_{R_w}(v_{w_1}, \ldots, v_{w_n}) = \min \left[\bigwedge_{j=t_1}^{t_p} \mu_{R_j}(V \downarrow_{S_j}), \min_{h=w_1,\ldots,w_n} \mu_{x_h}(v_h) \right] \tag{5.7}$$

As for the aligned groups, the degrees $\mu_{R_w}(V)$ and $\mu_{x_g}(V)$ are equal at the beginning but as more information about the satisfaction of the relations between V and the other region is acquired, the degrees $\mu_{R_w}(V)$ and $\mu_{x_g}(V)$ become different.

Behavior of the members of a group, when the group is the target object of a metric relation

Let V be a group of regions (aligned or not) and B a region. Suppose that A and V are the reference and target objects of a metric relation R_t which is represented as a fuzzy landscape. Then we know that the degree of satisfaction of the relation is greater than zero if and only if at least one member of V satisfies the relation with B. Moreover, if the members of the group have a high satisfaction degree of R_t, then the group has a high degree of satisfaction of R_t.

Therefore, assume that the model conceptual graph has a relation node representing a metric relation R_t where the target objects is a complex concept node representing a group (aligned or not) and the reference object is another concept node c_{ref} (which can be complex or not), and R_t is modeled as a fuzzy landscape, then we can add a relation node having the same type as R_t between each concept node inside the complex concept node and the node c_{ref}, as reference. This would help us find an instantiation of the group with high satisfaction of the relation R_t.

When the model has a complex concept node representing a group of aligned objects, which is the target object of a relation of type "parallel to" with a concept node c_{ref}, then we can add the relation node of type "on one side", between each member of the group and the node c_{ref}, because when a group V is "parallel to" a object B. Then all the members of V area "on one side" of B.

Other relations such as adjacency have a different behavior, and for a group to be adjacent to an object it is only necessary that one of its members is adjacent to the object.

5.4.4 Proposed algorithm

In this section, we propose a new algorithm based on the FAC-3 which takes into account the nested constraints. The set of constraints $\mathcal{C}$ is made of the constraints representing the relation nodes and the complex concept nodes. The representation of the constraints representing complex concept nodes of the conceptual graph is explained in Sections 5.4.2 and 5.4.3.

Basic algorithm used for arc-consistency checking.

The main algorithm presented in Algorithm 6 has the same structure as the one used in the original FAC-3 (see Algorithm 4). For each constraint we revise the arc-consistency condition for each of its domains. If during the checking, the domain of one of the variables has changed, then all the constraints involving that domain are reconsidered for the arc-consistency checking procedure. The only difference with respect to the FAC-3 algorithm is the separation of the revise relation according to the type of the constraint. For the constraints representing complex concept nodes, we use the functions ReviseGroupConstraint and ReviseAlignmentConstraint, and for the other constraints ReviseSimpleConstraint. We refer to these three functions as Revise methods. When updating the $ToCheck$ list (lines 25 to 33 of Algorithm 6) we take into account the duality of the complex concept nodes. If the domain of a variable which is also a constraint has changed, then the corresponding constraint should be added to the $ToCheck$ list.

To take into account that the variables which represent a complex concept node are only instantiated when the respective constraint is evaluated, we use the vector $Instantiated$ which indicates whether the domain of the variable has been already instantiated or not. In the Revise methods, we first verify that all the variables involved in the relation have been instantiated. In the case where they have not been instantiated, the arc-consistency check is not performed.

ReviseSimpleConstraint

The function ReviseSimpleConstraint in Algorithm 7 is composed of two parts. In the first part, if the constraint has not been evaluated, then it is evaluated for the first time and the domain of the relation is created. To create the domain we evaluate the relations presented in Chapters 1 and 2. If the relation is modeled in such a way that it creates a fuzzy landscape, we compute the landscape for each of the elements in the reference object's domain, and evaluate the relation with all the elements in the domain of the target objects. Otherwise, we find the satisfaction degree for each possible tuple in the relation domain. To update the domain we remove the tuples which contain an element which does not belong to its respective domain. The second part of the function deals with the updating of the membership function μ_{x_i} of each of the variables involved in the constraint. This part is identical to the Revise function of the FAC-3 algorithm (see Algorithm 5).

ReviseGroupConstraint

The function ReviseGroupConstraint presented in Algorithm 8 is also composed of two parts. In the first part, we create or update the relation's domain (Algorithm 9). The domain of the relation is created by making an exhaustive search of its domain and only adding the tuples for which the satisfaction degrees of all the child relations of the nested node are greater than zero. The updating of the relation's domain is made by eliminating those tuples for which there is at least one variable which does not belong to the domain.

Input: A constraint network $\mathcal{P} = (\mathcal{X}, \mathcal{D}, \mathcal{C})$
Output: Computes the arc-consistent closure of $\mathcal{P}$ if it exists, otherwise returns Failure

begin
    $ToCheck = \mathcal{C}$
    `// Initialize the` *Instantiated* `vector, mark as true all the variables which do not represent a group`
    foreach *variable* $x_i \in \mathcal{X}$ **do**
        if x_i *does not represent a group* **then** $Instantiated[i] = true$ **else** $Instantiated[i] = false$
    end
    foreach $C_k \in \mathcal{C}$ **do**
        FirstEvaluation[k]=false
    end
    while $ToCheck \neq$ **do**
        Select C_k from $ToCheck$
        foreach *variable* $x_{k_i} \in S_k$ **do**
            $Changed[k_i] = false$
        end
        if C_k *represents a group* **then**
            $result =$ `ReviseGroupConstraint`(C_k) ; `// see Algorithm 8`
        else
            if C_k *represents an alignment* **then**
                $result =$ `ReviseAlignmentConstraint`(C_k) ; `// see Algorithm 10`
            else
                $result =$ `ReviseSimpleConstraint`(C_k) ; `// see Algorithm 7`

        if $result = EmptyDomain$ **then**
            return Failure
        if $result = Changed$ **then**
            foreach $C_l \neq C_k$ *such that* $x_i \in S_l$ *and* $Changed[i] = true$ **do**
                $ToCheck = ToCheck \cup \{C_l\}$
                $Change[i] \longleftarrow false$
            end
            if C_k *is inside a nested constraint* C_l *or is related to a variable that represents a nested node seen as a constraint* C_k **then**
                $ToCheck = ToCheck \cup \{C_l\}$

        end
    end
end

Algorithm 6: Basic algorithm used for arc-consistency checking in a nested constraint network with complex concept nodes.

Input: a constraint C_k

Output: Makes C_k arc-consistent if possible and marks variables whose domain has changed; returns EmptyDomain if a domain has been emptied, NoChange if no domain has been modified, otherwise Changed

```
begin
    // Verify that all the variables in $S_k$ are instantiated
    if Exists $x_{k_i} \in S_k$ such that $Instantiated[k_i] = false$ then
        return $NoChange$

    if $FirstEvaluation[k]$ then // Create the domain for $R_k$
        Let $S_k = \{x_{k_1}, \ldots, x_{k_n}\}$
        foreach $A = (a_1, \ldots, a_{k_n}) \in D_{k_1} \times \ldots \times D_{k_n}$ do
            if $\mu_{R_k}(a_1, \ldots, a_{k_n}) > 0$ then Add $A$ to the domain of $R_k$
        end
        $FirstEvaluation[k] \longleftarrow false$
    else
        foreach variable $x_{k_i} \in S_k$ do
            remove from the domain of $R_k$ every tuple $A$ such $A \downarrow_{k_i} \notin D_{k_i}$
        end

    // Make the domains of the variables in $S_k$ arc-consistent with
       respect to $C_k$
    $result \longleftarrow NoChange$
    $height \longleftarrow 0$
    foreach variable $x_{k_i} \in S_k$ do
        foreach $v \in D_{k_i}$ do
            $NewDegree = 0$
            foreach $A = (a_{k_1}, \ldots, a_{k_n})$ in the domain of $R_k$ such that $A \downarrow_{k_i} = v$ do
                $eval \longleftarrow \min[\mu_{R_k}(a_{k_1}, \ldots, a_{k_n}), \min_{s=k_1,\ldots,k_n} \mu_{x_s}(a_s)]$
                $height \longleftarrow \max(eval, height)$
                $NewDegree \longleftarrow \max(NewDegree, eval)$
            end
            if $NewDegree = 0$ then
                Delete $v$ from $D_{k_i}$
                if $D_{k_i} = \emptyset$ then return $EmptyDomain$
            if $NewDegree \neq \mu_{x_{k_i}}(v)$ then
                $\mu_{x_{k_i}}(v) \longleftarrow NewDegree$
                $result \longleftarrow Changed$
                $Changed[k_i] \longleftarrow true$

        end
    end
    return $result$
end
```

Algorithm 7: ReviseSimpleConstraint: Revise algorithm for constraints which do not represent alignments nor groups, sub-part of algorithm. 6

When we update the relation's domain, the domain of its respective variable is also modified and is marked as changed in the *Changed* vector. The second part of Algorithm 8 corresponds to the updating of the domain of the variables representing the child nodes. We update the value of the membership function of x_i by considering the value of the corresponding constraint, since the membership function μ_{x_g} considers the interactions between the members of the group as well as the satisfaction of the relations between the group and other variables, which certainly affect the satisfaction of the membership function of the child nodes.

ReviseAlignmentConstraint

In the first part of ReviseAlignmentConstraint presented in Algorithm 10, we construct or update the domain of the constraint and its respective variable in function CreateOrUpdateAlignmentDomain (Algorithm 11). To create the domain we search for globally aligned groups using several satisfaction degrees α. For each α we only consider the subset of D_{k_i} which has a membership value greater than α. Due to the conjunctive condition of Equation 5.4, the membership degree of the group μ_{x_g} is always bounded by the membership degrees of its members, therefore even if the group has a higher value for μ_{ALIG}, its membership value μ_{x_g} would be equal to the membership value μ_{x_i} of its members.
When we update the domain D_g we observe for each element a_j of each group A if its membership value $\mu_{x_i}(a_j)$ is equal to zero. In that case we find the aligned subgroups $B_t \subset A \setminus a_j$ and add them to the domain D_g.

The second part of ReviseAlignmentConstraint deals with making the domain D_i arc-consistent with respect to C_k. For each value $a_j \in D_i$ we verify that there exists a group in D_g which contains it. To do this we replace the membership degree of $\mu_{x_i}(a_j)$ by:

$$\mu_{x_i}(a_j) = \max_{\{V \in D_g | a_j \in V\}} \mu_{x_g}(V)$$

Input: a constraint C_k
Output: Makes C_k arc-consistent if possible and marks variables whose domain has changed; returns EmptyDomain if a domain has been emptied, NoChange if no domain has been modified, otherwise Changed

Let x_g be the variable representing the nested group and D_g its domain
begin
  if *Exists* $x_{k_i} \in S_k$ *such that* $Instantiated[k_i] = false$ **then**
    return $NoChange$
  $height \longleftarrow 0$
  $result \longleftarrow$ `CreateOrUpdateGroupDomain` $(C_k,\ D_g)$; `// See Algorithm 9`
  if $D_g = \emptyset$ **then**
    return $EmptyDomain$

  `// Make the domains of the variables in` S_k `arc-consistent with respect to` C_k
  foreach *variable* $x_{k_i} \in S_k$ **do**
    foreach $v \in D_{k_i}$ **do**
      $NewDegree = 0$
      foreach $A = (a_{k_1}, \ldots, a_{k_n})$ *in the domain of* R_k *such that* $A\downarrow_{k_i} = v$ **do**
        $height \longleftarrow \max(\mu_{x_g}(A), height)$
        $NewDegree \longleftarrow \max(NewDegree, \mu_{x_g}(A))$
      end
      if $NewDegree = 0$ **then**
        Delete v from D_{k_i}
        if $D_{k_i} = \emptyset$ **then**
          return $EmptyDomain$

      if $NewDegree \neq \mu_{x_{k_i}}(v)$ **then** `// Update` $Changed$ `and the membership value of` v
        $\mu_{x_{k_i}}(v) \longleftarrow NewDegree$
        $result \longleftarrow Changed$
        $Changed[k_i] \longleftarrow true$
        $Changed[g] \longleftarrow true$

    end
  end
  return $result$
end

Algorithm 8: ReviseGroupConstraint:Revise algorithm for the constraint which represents a group, sub-part of algorithm 6.

Input: A constraint C_k representing a nested group, D_g domain of the corresponding variable

Output: Creates or updates the domain of C_k and of x_g

Let $S_k = \{x_{k_1}, \ldots, x_{k_n}\}$ be the child variables, $T_k = \{C_{k_1}, \ldots, C_{k_p}\}$ be the constraints representing the child relations.

begin
  if $FirstEvaluation[k]$ **then** `// Create the domain for` R_k `and` x_g
    foreach $A = (a_1, \ldots, a_{k_n}) \in D_{k_1} \times \ldots \times D_{k_n}$ **do**
      $\mu_{R_k}(a_1, \ldots, a_{k_n}) \longleftarrow \wedge_{j=k_1}^{k_p} \mu_{R_j}(a_1, \ldots, a_{k_n})$
      if $\mu_{R_k}(a_1, \ldots, a_{k_n}) > 0$ **then**
        Add A to the domain of R_k and to D_g
        $\mu_{x_g}(A) \longleftarrow \min\left[\min_{s=k_1,\ldots,k_n} \mu_{x_s}(a_s), \mu_{R_k}(a_1, \ldots, a_{k_n})\right]$
    end
    $FirstEvaluation[k] \longleftarrow false$, $Instantiated[g] \longleftarrow true$, $Changed[g] \longleftarrow true$, $result \longleftarrow Changed$
  else
    foreach *variable* $x_{k_i} \in S_k$ **do**
      remove from the domain of R_k every tuple A such $A \downarrow_{k_i} \notin D_{k_i}$
    end
    foreach $A = (a_1, \ldots, a_{k_n}) \in D_g$ **do**
      //Update the membership value
      $NewDegree \longleftarrow \min\left[\min_{s=k_1,\ldots,k_n} \mu_{x_s}(a_s), \mu_{R_k}(a_1, \ldots, a_{k_n}), \mu_{x_g}(A)\right]$
      if $NewDegree = 0$ **then**
        Delete A from R_k and from D_g
      if $NewDegree \neq \mu_{x_g}(A)$ **then** `// Update` $Changed$ `and the membership value of` A
        $\mu_{x_g}(A) \longleftarrow NewDegree$
        $Changed[g] \longleftarrow true$
        $result \longleftarrow Changed$
    end
  return $result$
end

Algorithm 9: CreateOrUpdateGroupDomain sub-part of algorithm 8.

Input: a constraint C_k
Output: Makes C_k arc-consistent if possible and marks variables whose domain has changed; returns EmptyDomain if a domain has been emptied, NoChange if no domain has been modified, otherwise Changed

Let x_g be the variable representing the nested group and D_g its domain
begin
    if *Exists* $x_{k_i} \in S_k$ *such that* $Instantiated[k_i] = false$ **then**
        return $NoChange$
    $height \longleftarrow 0$
    $result \longleftarrow$ `CreateOrUpdateAlignmentDomain` $(C_k,\, D_g)$; `// See Algorithm 11`
    if $D_g = \emptyset$ **then**
        return $EmptyDomain$

    `// Make the domains of the variables in` S_k `arc-consistent with respect to` C_k
    foreach $v \in D_{k_i}$ **do**
        $NewDegree = 0$ **foreach** $A \in D_g$ *such that* $v \in A$ **do**
            $height \longleftarrow \max(\mu_{x_g}(A), height)$
            $NewDegree \longleftarrow \max(NewDegree, \mu_{x_g}(A))$
        end
        if $NewDegree = 0$ **then**
            Delete v from D_{k_i} ;
            if $D_{k_i} = \emptyset$ **then**
                return $EmptyDomain$

        if $NewDegree \neq \mu_{x_{k_i}}(v)$ **then** `// Update` $Changed$ `and the membership value of` v
            $\mu_{x_{k_i}}(v) \longleftarrow NewDegree$
            $result \longleftarrow Changed$
            $Changed[k_i] \longleftarrow true$
            $Changed[g] \longleftarrow true$

    end
    return $result$
end

Algorithm 10: ReviseAlignmentConstraint:Revise algorithm for the constraint which represent an alignment, sub-part of algorithm 6.

Input: A constraint C_k representing an aligned group, D_g domain of the corresponding variable
Output: Creates or updates the domain of C_k and of x_g, returns EmptyDomain if a domain has been emptied, NoChange if no domain has been modified, otherwise Changed

Let $S_k = \{x_{k_1}\}$ be the variable involved in the alignment.
begin
    $domainStatus \longleftarrow NoChange$
    if $FirstEvaluation[k]$ **then** `// Create the domain for` R_k `and` x_g
        `// Find the alignments for several satisfaction degrees`
        for $\alpha \in]0,1]$ **do**
            Find $\mathfrak{G}$, the group of globally of aligned groups to a degree α of objects belonging to the set $\{a \in D_{k_1} : \mu_{x_{k_1}}(a) > \alpha\}$
            foreach $A = \{a_1, \ldots, a_{k_n}\} \in \mathfrak{G}$ **do**
                $\mu_{R_k}(a_1, \ldots, a_{k_n}) \longleftarrow \mu_{ALIG}(A)$; `// See Equation 2.30`
                Add A to D_g and add the tuple $(a_1, \ldots, a_{k_n})$ to the domain of R_k
                $\mu_{x_g}(A) \longleftarrow \min[\min_{s=k_1,\ldots,k_n} \mu_{x_s}(a_s), \mu_{ALIG}(A)]$
            end
        end
        $FirstEvaluation[k] \longleftarrow false$; $Instantiated[g] \longleftarrow true$
        if $D_k \neq \emptyset$ **then**
            $Changed[g] \longleftarrow true$; $domainStatus \longleftarrow Changed$
        else return $EmptyDomain$
        return $domainStatus$
    else
        `// Revise the alignment condition on each group`
        foreach $A = \{a_1, \ldots, a_{k_n}\} \in D_g$ **do**
            if *There is an element* $a_i \in A$ *which no longer belongs to* D_k **then**
                Remove A from D_g and R_k
                for $\alpha \in]0,1]$ **do**
                    Find $\mathfrak{G}_A$, the group of globally of aligned groups to a degree α of objects belonging to the $A \cap D_{k_1}$
                    foreach $B = \{b_1, \ldots, b_{k_m}\} \in \mathfrak{G}_A$ **do**
                        $\mu_{R_k}(b_1, \ldots, b_{k_m}) \longleftarrow \mu_{ALIG}(B)$ Add the set B to D_g
                        Add the tuple $(b_1, \ldots, b_{k_m})$ to the domain of R_k
                        $\mu_{x_g}(B) \longleftarrow \min[\min_{s=k_1,\ldots,k_n} \mu_{x_s}(b_s), \mu_{ALIG}(B)]$
                        $Changed[g] \longleftarrow true$
                        $domainStatus \longleftarrow Changed$
                        Set to $true$ the value of $FirstEvaluation$ for all the constraints involving x_g
                    end
                end
        end
end

Algorithm 11: CreateOrUpdateAlignmentDomain sub-part of algorithm 8.

5.4.5 Complexity and improvements

In this section we study the complexity of the functions of each part of the algorithm independently. Let d be the size of the largest domain in $\mathcal{D}$, n_o the maximum number of points in a region, and N the number of points in the image.

ReviseSimpleConstraint For a constraint C_k the worst case time complexity for calling ReviseSimpleConstraint(C_k) corresponds to the case where the domain has to be created and the relation corresponds to the between relation, which is the one with the largest complexity (see Table 5.2). Creating the domain has a complexity of $O(d(d-1)(NN_\nu + n_o^2) + dn_o)$, and N_ν corresponds to the size of the structuring elements used in Equation 1.19. When the function ReviseSimpleConstraint(C_k) is called and it only has to verify for arc-consistency without creating the domain of the relation, then we have to compare every value in the domains belonging to S_k to each tuple satisfying C_k. In the worst case there are d^3 tuples and d values in the domain, therefore the worst case complexity in this case is $O(d^4)$. In the worst case, the function ReviseSimpleConstraint(C_k) is called t_α times for every value in its domain. We have $|S_k| = 3$ and so it can be called $3t_\alpha d$ times, where t_α is the number of satisfaction levels in which the interval $[0, 1]$ is discretized.

ReviseGroupConstraint For a constraint C_k representing a group, the function Create-OrUpdateGroupDomain has a worst case complexity of $O(d^{|S_k|})$ when the domain is created. The revision of the arc-consistency condition for this function has a worst case complexity of $O(d^{|S_k|+1})$ corresponding to the case where for every value in the domain of S_k we have to check its existence in every tuple in the domains of the relation, where the maximum number of tuples is $d^{|S_k|}$. Using the same arguments as for the ReviseSimpleConstraint function, this function can be called $|S_k|t_\alpha d$ times.

ReviseAlignedConstraint Let C_k be a constraint representing an alignment. The complexity of creating the domain is given by $O(d^2 + N_A^2 n_0^2 + N)$ where N_A is the number of local alignments (see Section 2.2.5). Let N_g be the number of aligned groups that are obtained, N_g is bounded by $\sum_{j=3}^{d} \binom{d}{k}$, and let N_o be the maximum number of elements in a group, which is bounded by d. Then the function ReviseAlignedConstraint can be called a maximum of $O(t_\alpha N_o log(N_o))$ times.

So the complexity of the whole algorithm is $O(|\mathcal{C}|ct_\alpha d^{c+2} + N_A^2 n_0^2 + N + d^2(NN_\nu + n_o^2))$, where $c = \max_k |S_k|$. We can see that the computational complexity of the algorithm can be very high. For the AC-3 algorithm several works [Wallace and Freuder, 1992, Boussemart et al., 2004, Schulte and Stuckey, 2004] have proposed strategies to sort the constraints to be revised. We propose to consider the computational time of each constraint to determine the order, which is one of the strategies also proposed in [Schulte and Stuckey, 2004]. First we revise the constraints which do not represent a complex concept node. The order in which these constraints are chosen depends on their computational time (see Table 5.2), so we select the constraint in the *ToCheck* list which has the lowest computational time. When there are no more constraints representing relation nodes in the *ToCheck* list we can proceed to evaluate the constraints representing the complex concept nodes.

It can be interesting to adapt other criteria that have been used to improve the efficiency of the AC-3 algorithm. For instance, [Wallace and Freuder, 1992, Boussemart et al., 2004, Schulte and Stuckey, 2004] have proposed to choose the constraint to be checked according

Spatial Relation	Complexity	Commentaries
Adjacency	$O(2N_BN_\nu)$	N_ν number of points in the definition of neighborhood, usually 8 or 26.
Topological surround	$O(2N_BN_\nu)$	N_ν number of points in the definition of neighborhood, usually 8 or 26.
Distance	$O(2N)$	When using crisp objects, the fuzzy landscape representing a distance can be constructed by computing a distance map and then applying the distance function on each point of the distance map. We can construct a distance map using the Danielsson map, which has a linear execution time [Danielsson, 1980].
At an angle of	$O(N(1+2N_\nu))$	This complexity is used by using the propagation method (see [Bloch, 1999]), where N_ν number of points of the considered in the neighborhood, usually 8 or 26.
On one side of	$O(2N(1+2N_\nu))$	This complexity is used by using the propagation method (see [Bloch, 1999]), where N_ν number of points of the considered in the neighborhood, usually 8 or 26.
Between	$O(NN_\nu+2N_B)$	N_ν is the cardinality of the support of the structuring elements used for the computation of the region.
On the same side	$O(NN_\nu+N_B)$	N_ν cardinality of the support of the structuring elements used for the computation of the region..
Surrounds	$O(N_Blog(N_B) + N_dN)$	N_d is the number of points used in the near function.
Go through	$O(N_BN_\nu+N_T)$	N_ν: number of points in the definition of neighborhood.
Go across 1	$O(N_BN_\nu + N_T + N_Blog(N_B))$	N_ν: number of points in the definition of neighborhood, usually 8 or 26.
Go across 2	$O(2N_B+N_B(2N_\nu+log(N_B))$	N_ν: number of points in the definition of neighborhood, usually 8 or 26.
Go into	$O(2N_B+N_R(2N_\nu+log(N_R))$	N_ν: number of points in the definition of neighborhood, usually 8 or 26.
Parallel	$O(N(1+2N_\nu))$	N_ν: number of points in the definition of neighborhood, usually 8 or 26.
Alignment	$O(N_O^2+N_A^2N_B^2+N)$	N_O: number of objects used to find the local alignments. N_A number of locally aligned groups obtained.

Table 5.2: Computational complexity for each spatial relation. N_B refers to the maximum number of points in objects. For the relations represented as a fuzzy landscape, we assume that the region where the landscape is computed is a square, with length $\sqrt{N}$ and that N is the number of points in this region.

to its capability of making a more significant reduction of the domains of the variables, and therefore reducing the calls of the Revise function. Different heuristics are proposed to determine this.

5.4.6 Finding a solution

Once we have identified all the regions in the image which can be an instantiation of a concept node of the model, we have to determine which are the "best" solutions. For this we need a criterion to determine how to choose a set of instantiations over another.

A first approach is to organize the solutions according to the degree to which they satisfy all the constraints, i.e. according to the consistency value of each solution:

$$Cons(V) = \min_{\tilde{C}_k \in \mathcal{C}} \mu_{R_k}(V \downarrow_{S_k}) \tag{5.8}$$

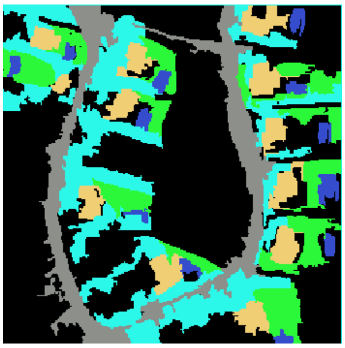
This approach is very severe, and does not let us discriminate between two solutions which have the same minimum value but different values for other constraints [Dubois et al., 1996]. For instance, suppose that the we have two instantiations V_1 and V_2 and that the satisfaction degrees of the constraints is given by $Sat_1 = \{0.4, 0.55, 0.42, 0.65\}$ and $Sat_2 = \{0.4, 0.8, 0.97, 0.8\}$, respectively. Therefore, we would say that V_1 and V_2 are equally good, since $Cons(V_1) = Cons(V_2)$, although V_2 has a greater than or equal satisfaction value for all the constraints. Thus, saying that V_1 and V_2 are equally good is not very adequate.

A more convenient approach is to aim at maximizing the number of satisfied constraints. For instance, the solutions can be ordered according to the leximin order as in [Fargier, 2006, Möller and Näth, 2008, Saathoff and Staab, 2008], that is, for every solution $V_i \in Sol(\mathcal{P})$ construct a vector Sat_i containing the satisfaction degrees for each constraint of $\mathcal{C}$ organized from minimum to maximum. The leximin ordering is given by applying the lexicographic ordering over those vectors. We denote the leximin order by $<_{LEXIMIN}$. Using the same example as in the previous paragraph, we have $Sat_1 = \{0.4, 0.42, 0.55, 0.65\}$ and $Sat_2 = \{0.4, 0.8, 0.8, 0.97\}$, so $Sat_1 <_{LEXIMIN} Sat_2$, and therefore V_2 is better than V_1.

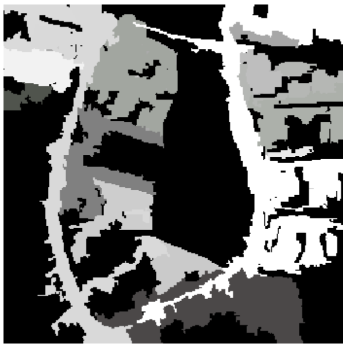
The solutions can be represented as two images, one containing the instantiations and the other containing, for each element, a color representing its position in the order of the solutions. Figure 5.16 shows the solutions of the example of Figure 5.12. Figure 5.16(a) shows all the instantiations, Figure 5.16(b) shows the order image. The best solution are those having a lighter color in the order image.

Illustration In this section we present an example to illustrate the algorithm proposed in Section 5.4.4 .
Consider again the image in Figure 5.11 and suppose that we want to find the houses which satisfy the condition: "the houses having a pool which is located in the garden at the "back" of the house and the houses have a shadow", but in addition, we want the houses that satisfy the condition of being neighbor houses in the same road. We represent "neighboring houses of the same road", as a group of aligned houses which are near each other. The conceptual graph in Figure 5.17 shows a possible way to describe the spatial organization of the scene that we want to find in the image. For simplicity, in the conceptual graph of Figure 5.17 we only draw one of the coreference links to indicate the relations that the houses inside the group have to satisfy.

The corresponding CSP problem is formulated as:



(a) Instantiations



(b) Order

Figure 5.16: Solutions of example of Figure 5.12.

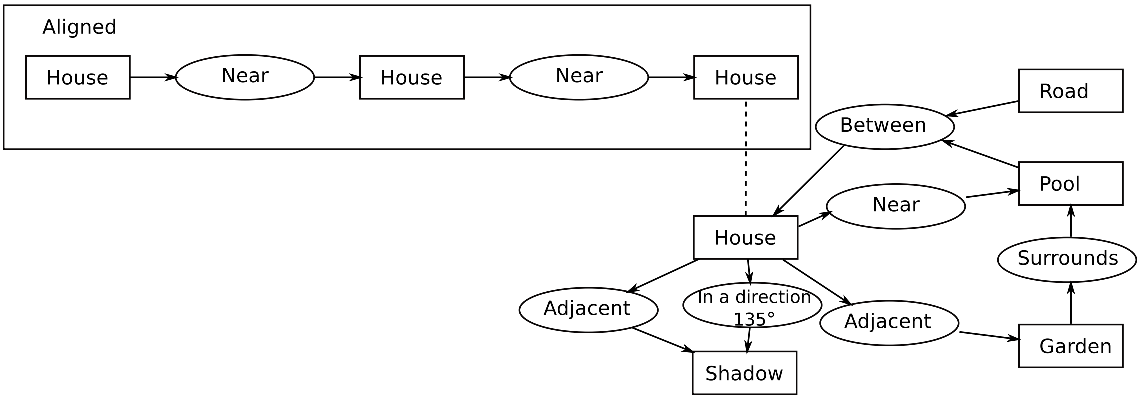


Figure 5.17: Conceptual graph representing "the group of neighboring houses forming an aligned group which have a pool located in the garden at the "back" of the house, and which have a shadow".

- $\mathcal{X} = \{x_{shadow}, x_{house}, x_{road}, x_{pool}, x_{garden}, x_{group\_houses}\}$ containing the variables representing the concept nodes of the conceptual graph, where x_{group_houses} is the variable representing the group of aligned houses.

- $\mathcal{C} = \{C_{direction}, C_{adjacent_1}, C_{between}, C_{near}, C_{surrounds}, C_{adjacent_2}, C_{aligned}\}$, where $C_{direction} = \langle \mu^{135^\circ}_{direction}, \{x_{shadow}, x_{house}\}\rangle$, $C_{adjacent_1} = \langle \mu_{adjacent}, \{x_{shadow}, x_{house}\}\rangle$, $C_{between} = \langle \mu_{between}, \{x_{pool}, x_{road}, x_{house}\}\rangle$, $C_{near} = \langle \mu_{near}, \{x_{house}, x_{pool}\}\rangle$, $C_{surrounds} = \langle \mu_{surrounds}, \{x_{garden}, x_{pool}\}\rangle$, $C_{adjacent_2} = \langle \mu_{adjacent}, \{x_{house}, x_{garden}\}\rangle$ and $C_{aligned} = \langle \mu_{ALIG}, \{x_{house}\}\rangle$.

- $\mathcal{D} = \{D_{shadow}, D_{house}, D_{road}, D_{pool}, D_{garden}, D_{group\_houses}\}$ where D_i represents the possible candidates for x_i. The regions in each D_i for $i \in \{shadow, house, road, pool, garden\}$ are shown in Figure 5.13.

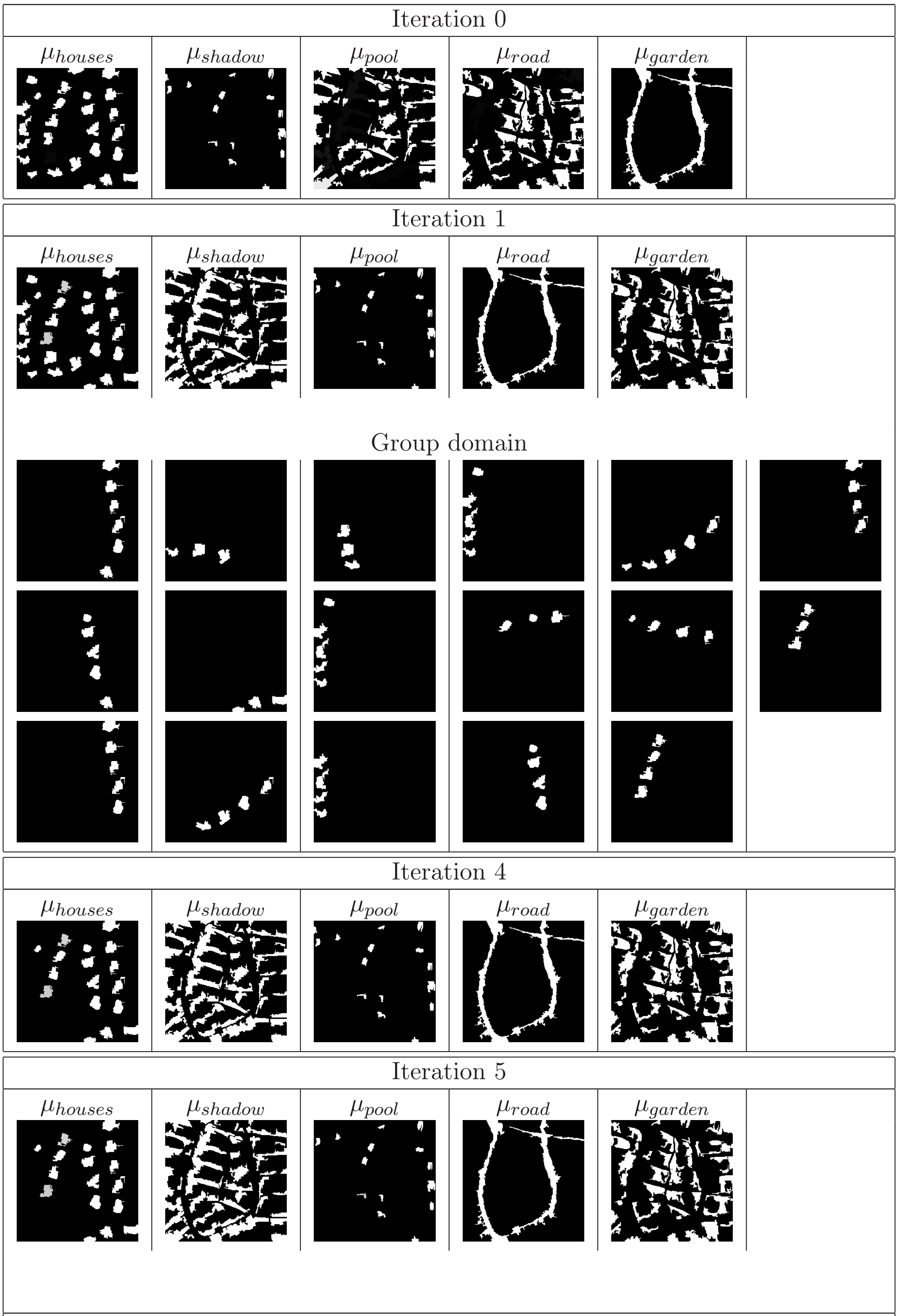
Table 5.3 shows the evolution of μ_{x_i} for $i \in \{shadow, house, road, pool, garden\}$ and the domain of x_{group_houses}, we only show D_{group_houses} when it has changed. The membership degree of a region to μ_{x_i} for $i \in \{shadow, house, road, pool, garden\}$ is represented by its gray value. The images representing the groups belonging to D_{group_houses} only show the elements of D_{group_houses} but the images do not contain information about the membership values. For illustrative purposes we decided to evaluate the alignment constraint in the

first iteration in order to show the evolution of D_{group_houses}.
At the first iteration we check $C_{aligned}$ and obtain 17 groups. In iteration 4 the constraint $C_{between}$ is checked and the domain of x_{houses} is modified. As D_{houses} has been modified the alignment condition is checked again, and several groups are eliminated. In iteration 15 the constraint $C_{adjacent_2}$ is checked, and again the domain of x_{houses} is modified. Hence, in iteration 16 the $C_{alignment}$ is checked and more groups are eliminated. The final results are shown in iteration 38, where two groups are obtained. Since one groups is contained in the other group, this is equivalent as having one only group.

5.4.7 Discussion

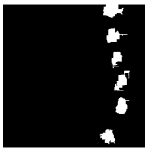
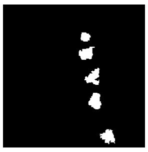
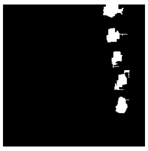
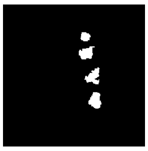
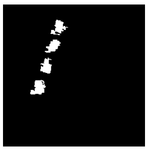
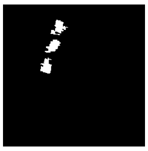
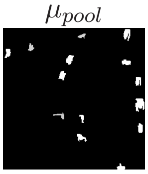
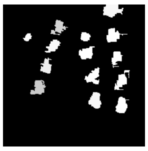
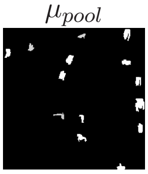
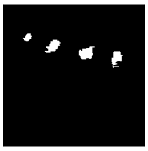
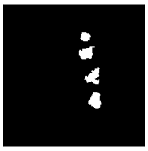
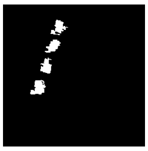
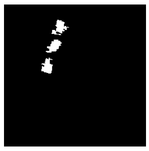
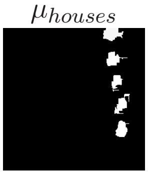
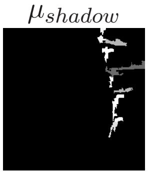
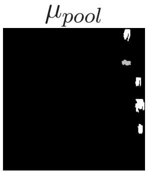
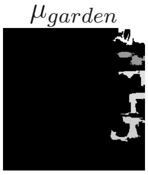
We have extended the definition of arc-consistency for fuzzy constraints with an arity greater than or equal to 2. Using this definition we extended the fuzzy AC-3 algorithm to deal with these constraints. Furthermore, this new algorithm was adapted to handle constraints that represented complex concept nodes, such as alignments or groups. To represent this type of constraints we had to make the algorithm flexible enough to accept constraints where the number of variables inside the constraint is not defined, for the alignment relation. We also explored the double nature of such complex concept nodes: constraint/ variable. Although we adapted the system for handling particular constraints, the proposed algorithms can be used for any constraint with similar characteristics. For instance, the bi-level constraints proposed in [Deruyver and Hodé, 1997] can be represented in the proposed context. The bi-level constraints consist of two levels of constraints: one dealing with the constraints between the variables and the other between the values of the variables. To formalize this type of constraints [Deruyver and Hodé, 1997] propose a binary crisp relation $Cmpi$ which determines whether two values belonging to a same constraint are compatible. Thus, we can represent the same information of a bi-level constraint by using a complex concept node where the number of elements inside the node is not specified and which only needs to satisfy the binary compatibility relation between them. The constraints between two variables in a bi-level constraint can be represented as relations between the complex concept nodes.

Table 5.3: Evolution of the fuzzy sets μ_{x_i} over D_i for $i = \{houses, pool, shadow, garden, road\}$ when making each D_i arc-consistent, and of D_{group_houses}.



Continued on next page

Table 5.3 – continued from previous page

Group domain					
					
					
Iteration 15					
μ_{houses} 	μ_{shadow} 	μ_{pool} 	μ_{road} 	μ_{garden} 	
Iteration 16					
μ_{houses} 	μ_{shadow} 	μ_{pool} 	μ_{road} 	μ_{garden} 	
Group domain					
					
					
Iteration 38					
μ_{houses} 	μ_{shadow} 	μ_{pool} 	μ_{road} 	μ_{garden} 	
Group domain					
					

5.5 Interpreting an unlabeled image

When dealing with an unlabeled image, the difficulty lies in adequately creating the domain of regions which can represent each concept node of the model conceptual graph. This problem can be divided into two steps: the first one concerning the image partition and the second one concerning the labeling of the regions.

5.5.1 Image partition

As presented in Section 4.2.3 there are several strategies for determining the initial regions, two very common ones being:

- Use high level knowledge to guide, tune or improve low level feature extraction and image segmentation, some examples are [Matsuyama, 1988, Garnesson et al., 1989, Clement and Thonnat, 1993], or,
- Perform an unsupervised segmentation and label each region according to its characteristics, as in [Benz et al., 2004, Durand et al., 2007, Saathoff and Staab, 2008].

A knowledge based segmentation can correctly extract regions which correspond to instantiations of the concept nodes of the model. However, in practice these methods are domain specific, and cannot be easily reused. Moreover, the development of a knowledge based segmentation is out of the scope of this work. Therefore, we have decided to use a multi-scale segmentation. Multi-scale segmentation is well adapted to Earth observation images since it allows us to extract objects of different sizes, which are observed in these type of images. Moreover, the multi-scale segmentation has an explicit hierarchical organization of the regions which can be useful for spatial reasoning, as discussed in Section 5.5.3. We used a multi-scale segmentation obtained from a hierarchical Mean Shift [Paris and Durand, 2007, DeMenthon and Megret, 2002].

When performing a multi-scale segmentation we can assume that the regions obtained from the segmentation can be candidates of an object. By considering this hypothesis we do not have to worry about the segmentation problems such as a region containing two objects or an object split into several regions.

5.5.2 Region labeling

After performing the segmentation we label the regions of the image by using two types of information;

- the approximate size of objects represented by the concept nodes of the graph, and
- the knowledge about the extraction of certain types of concepts in T_C.

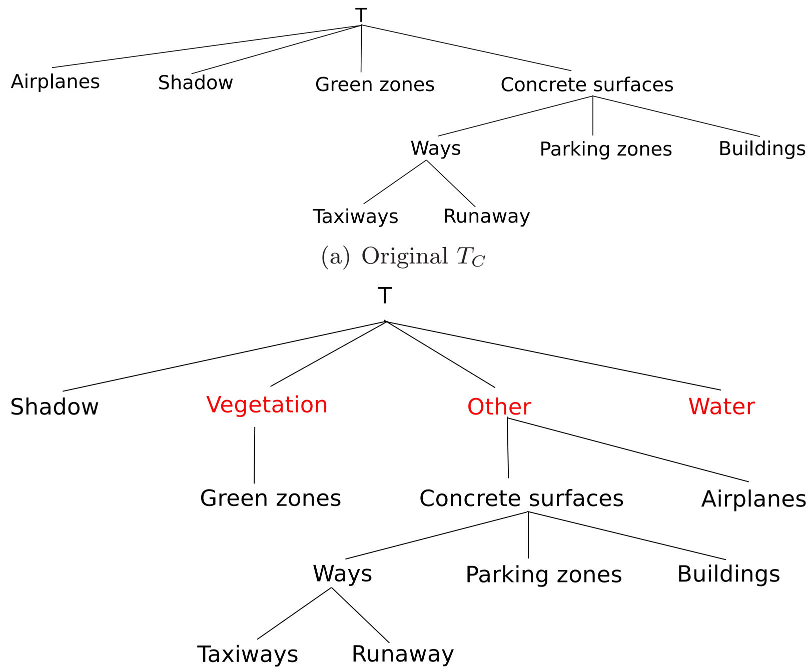
For the first type of information, we are going to assume that it is possible to know the typical sizes of the objects that we are searching, as in [Ciucu et al., 2002]. This information is given as linguistic terms { very small, small, medium, large, very large }. For each of these terms we construct a membership function over $\mathbb{R}$, which defines its semantics. The parameters defining these functions can be learned according to the scene.

For the second type of information, we want to exploit the fact that we know how to extract certain classes of concepts, for instance:

- water (using Normalized Difference Water Index NDWI [McFeeters, 1996]),

- vegetation (using Normalized Difference Vegetation Index NDVI [Goward et al., 1991]),
- shadow (using a hysteresis threshold over the intensity image),
- ...

Let T_C be the concept hierarchy of the vocabulary over which the conceptual graph is built. Let H_C be a set representing the classes of concepts that we know how to extract. We propose to add the concepts of H_C to T_C, and also add a concept "Other" to represent all the concept classes which are not "A_KIND_OF" one of the concepts in H_C. Figure 5.18 illustrates an example of the concepts of H_C added to T_C: Figure 5.18(a) illustrates the initial concept hierarchy, then the modified concept hierarchy of Figure 5.18(b) is obtained by adding the concepts of $H_C \cup$ "$Other$".



(a) Original T_C

(b) Concept hierarchy after adding new concepts

Figure 5.18: Addition of new categories to the concept hierarchy T_C, the new concepts are marked in red.

To label the regions of the images we make use of the inclusion relation between the categories of T_C. First we extract in the image the regions corresponding to the classes of concepts in H_C. Then we compute the "Other" class as the complement of the disjunction of the known classes. Suppose that we can define every $c \in H_C \cup \{"Other"\}$ as a fuzzy set μ_c over the image space. Then for every concept node in the conceptual graph, represented as a variable x_i, we construct its membership function over the regions of the image as:

$$\mu_{x_i}(v) = \left[\bigwedge_{\{c \in H_C | type(x_i) \leq c\}} F(\mu_c, v) \right] \wedge \mu_{size_i}(v) \tag{5.9}$$

where μ_{size_i} represents the membership function corresponding to the size of the objects represented by x_i, and F is a comparison measure which evaluates how well v matches μ_c.

For instance, F can be a mean measure (Equation 1.16), a necessity measure (Equation 1.7), a possibility measure (Equation 1.6). The first term of Equation 5.9 is the conjunction of all the membership values of the classes of H_C for which the type of the concept x_i is a sub category.

Now that we have estimated the initial membership functions for the variables that represent the concept nodes of the model, we can apply Algorithm 6 to find the arc-consistent domains. The proposed method is illustrated in Figure 5.19.

5.5.3 Spatial reasoning over the tree of regions

In this section we mention some practical issues of how to use the tree of regions generated by the multi-scale segmentation for spatial reasoning. The result of the multi-scale segmentation can be represented on a tree structure, where the root node corresponds to the whole image. Figure 5.20 illustrates an example of a tree.
For a region r_i in the image we will denote by the same name the node in the tree which represents it. Let denote by $level(r_i)$ the level in the tree where node r_i is situated, with the level 0 corresponding to the lower level. $child_l(r_i)$ denotes the set of child nodes of node r_i which are l levels below $level(r_i)$. For example in the tree of Figure 5.20, we have $child_1(r_{21}) = \{r_{10}, r_{11}\}$ and $child_2(r_{21}) = \{r_{00}, r_{01}, r_{02}, r_{03}, r_{04}\}$.

We have the following properties which are easily proved.

Proposition 5.5. *Let r_i and r_j be two disjoint regions obtained from a multi-scale segmentation, and let R_L be a binary relation which is represented as a fuzzy landscape. Then for every $l \in \mathbb{N}$ such that $0 < l \leq level(r_i)$ it is true that:*

$$\mu_{R_L}(r_i, r_j) = \frac{\sum_{k: r_k \in child_l(r_j)} |r_k| \mu_{R_L}(r_i, r_k)}{|r_j|}, \tag{5.10}$$

where $\mu_{R_L}(r_i, r_j)$ is the degree of satisfaction (evaluated using a mean) of the relation S_L when r_i is the reference object and r_j is the target object.

Proposition 5.6. *Let r_i and r_j be two disjoint regions obtained from a multi-scale segmentation. Then for every $l \in \mathbb{N}$ such that $0 < l \leq level(r_i)$ it is true that:*

$$\mu_{ADJ}(r_i, r_j) > 0 \text{ if and only if } \exists r_k \in child_l(r_j) \text{such that } \mu_{ADJ}(r_i, r_k) > 0 \tag{5.11}$$

where $\mu_{ADJ}(r_i, r_j)$ is the degree of satisfaction of the adjacency relation between r_i and r_j.

Therefore, it is only necessary to compute the satisfaction degrees of the relations which produce a fuzzy landscape for the regions belonging to the lower level of the tree, and we can obtain the degree of satisfaction of all the other regions by using Proposition 5.5.

5.5.4 Finding a solution

When finding the instantiations of a model in an unlabeled image there can be conflict among the different instantiations. This information has to be considered when finding the final solution. Every instantiation can be seen as a source of information. Let V_1 and V_2 be two instantiations, suppose that according to V_1 a region u is of type c_1, and according to V_2 it is of type c_2. If there is a relation between c_1 and c_2 in the concept hierarchy T_C and $c_1 \wedge c_2 \neq \perp$ then we can combine the information of V_1 and V_2 and label u with the type $c_1 \wedge c_2$. However, if $c_1 \wedge c_2 = \perp$ then V_1 and V_2 are in conflict, and to solve the conflict we eliminate the instantiation with a worst instantiation according to leximin ordering.

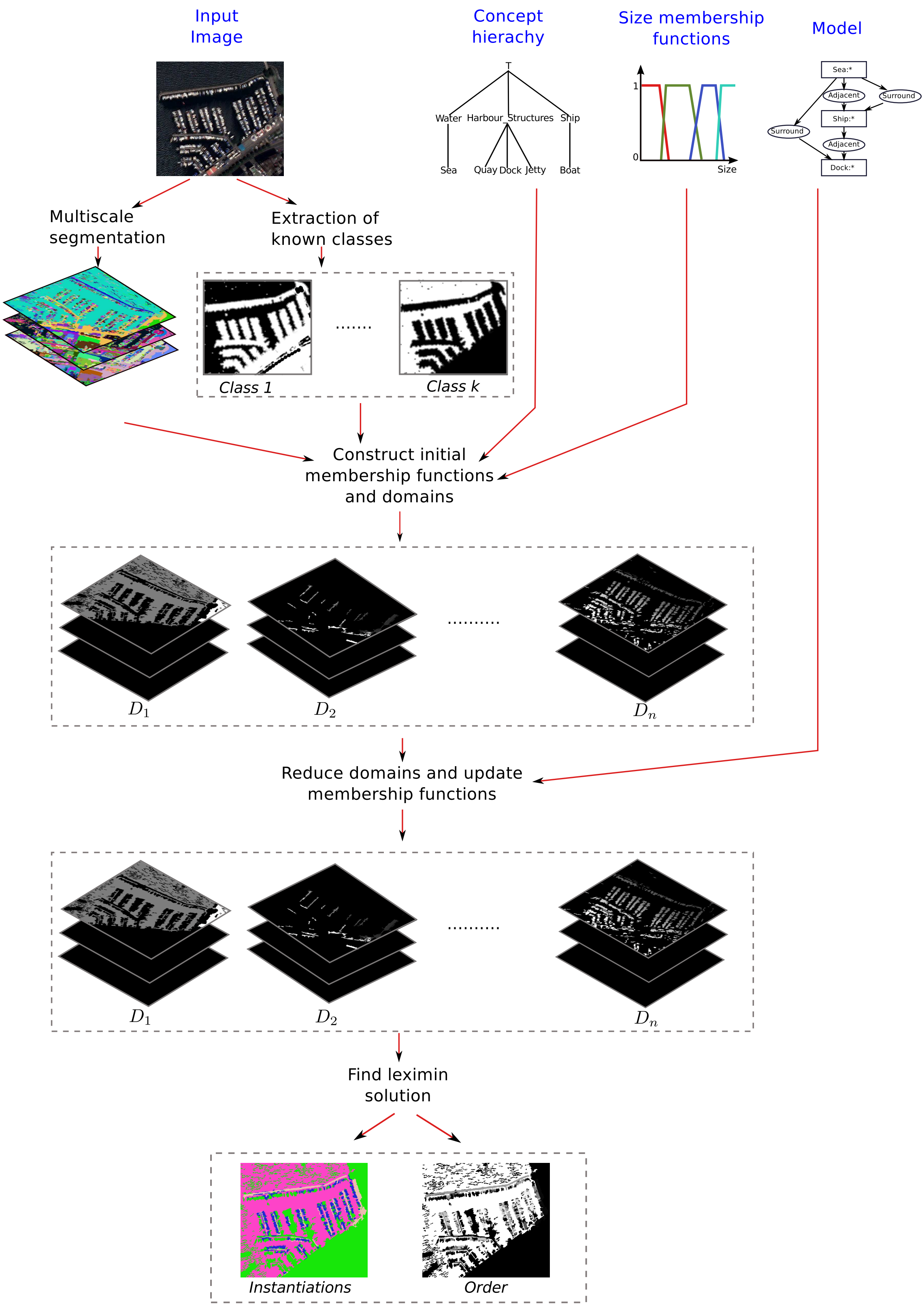


Figure 5.19: Proposed method for determining the model's instantiations

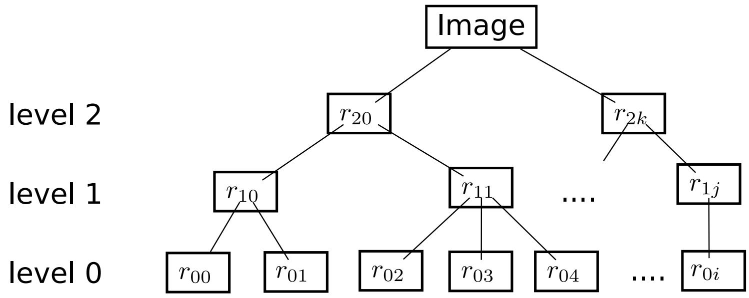


Figure 5.20: Tree representation of multiscale segmentation.

5.5.5 Results

In this section we present some experimental results of the proposed method. We apply the method in two situations: searching for the harbors in an image, and the interpretation of an image containing an airport.

Finding harbors

We applied the method represented in Figure 5.19 to the image shown in Figure 5.21 to extract the harbors in the image. The conceptual graph that we used to represent the structure of a harbor is given in Figure 5.8(d) and the concept hierarchy is the one given in Figure 5.8(c). It was not possible to use a conceptual graph containing complex concept nodes representing the alignment of the boats because the boats were not appropriately segmented; therefore the algorithm extracting the alignments did not had a good performance.

The size of the image shown in Figure 5.21 is of 5901×11801 pixels. To apply the proposed methodology we divided the image into 72 tiles of approximately 985×985 pixels, and applied the methodology in each tile. To extract the initial candidate regions for each domain, we used the set H_C containing the known classes was $\{water, vegetation\}$, and in addition we used the fact that docks have a linear structure[2]. Figure 5.22 shows the concept hierarchy after adding the new concepts.

For instance, let us observe the results for the image tile shown in Figure 5.23. The extraction of the known classes is shown in Figure 5.24 and the initial objects in the domain of each variable are shown in Table 5.4.

[2]To evaluate if a region is linear we computed the ratio of its largest principal moment by the smallest principal moment. And we considered that an object was linear if this ratio was equal or greater than 4. This value was chosen from the experimental results.

Figure 5.21: Image of a lake.

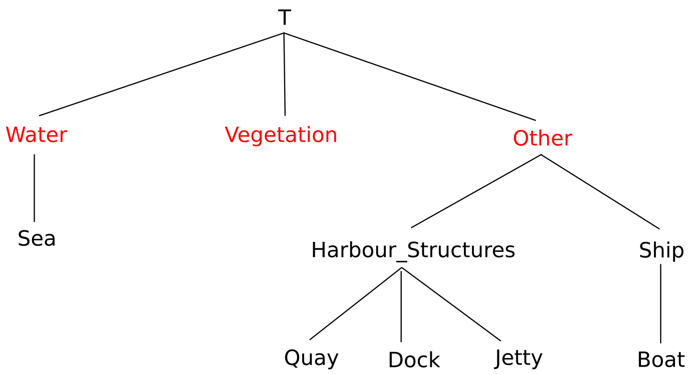
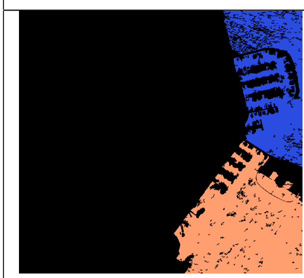
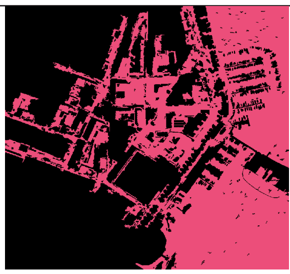


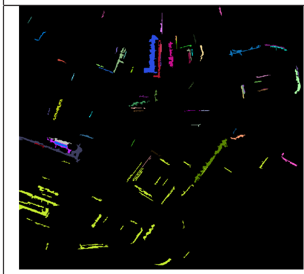
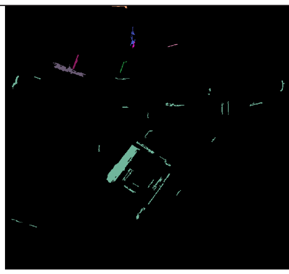
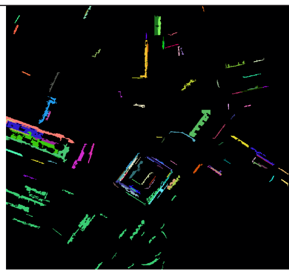
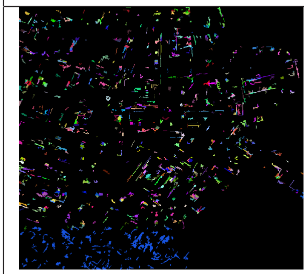
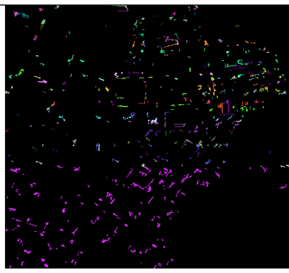
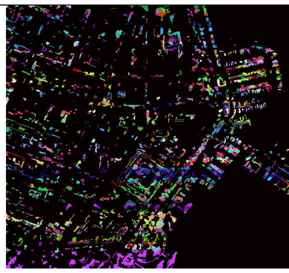
Figure 5.22: Concept hierarchy after adding new concepts. The new concepts are marked in red.

Table 5.4: Initial domains for the concept nodes representing sea, ship and dock.

D_{sea}		
		

Continued on next page

Table 5.4 – continued from previous page

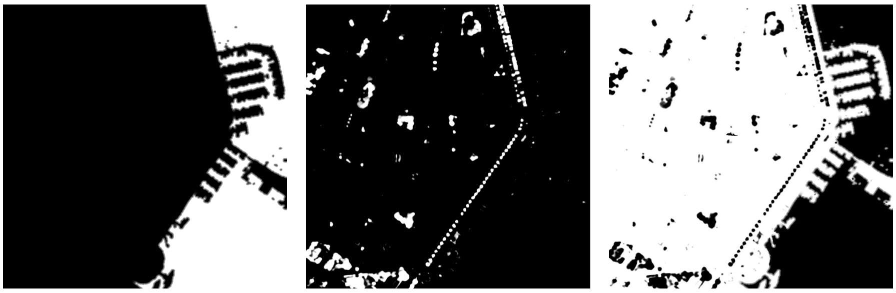
D_{dock}		
		
D_{boat}		
		

For this tile image, the obtained instantiations and ordering are shown in Figure 5.25. We can see that the method obtains satisfactory results, it detects most of the configurations of the model in the image. This example shows how by using the spatial relations it is possible to reduce the domain of the concepts in the conceptual graph to obtain the desired model.

The instantiations of the graph in the whole image are given in Figure 5.26. We can see that although we used a very simple graph it was possible to extract the zones of the image which correspond to the harbors. Table 5.5 shows some examples of the instantiations of the model. We can see that in the Figures (a) to (h) the harbors are correctly detected. Even if the harbors of Figures (b), (e) and (h) were very small, the method finds satisfactory instantiations.

Figure (e), (i), (j), (k) and (l) have some examples of false detections. Most of the correct instantiations have a higher consistency value (Equation 5.8) than the false detections of Figures (e), (i), (j) and (k). The result of Figure (l), as well as the missing instantiations in all of the Figures can be attributed to segmentation errors. This examples brings up the fact that the results of our methodology depend on the quality of the segmentation. However, our methodology can be used as an intermediate step in a segmentation chain. Moreover, the results can be improved by adding more spatial relations which constrained the model, for instance we can add a constrint which says that the port should be adjacent to the land.

Figure 5.23: Image tile example.



(a) Water (b) Vegetation (c) Other

Figure 5.24: Results from the extraction of the known classes.

Table 5.5: Examples of instantiations of the model of Figure 5.8(d) in the image tiles of Figure 5.21.

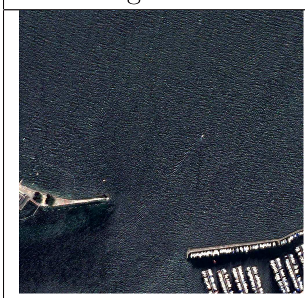
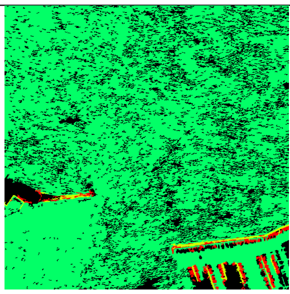
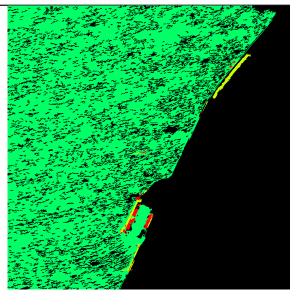
Original tile	Instantiations	Original tile	Instantiations
			
(a)		(b)	
			Continued on next page

Table 5.5 – continued from previous page

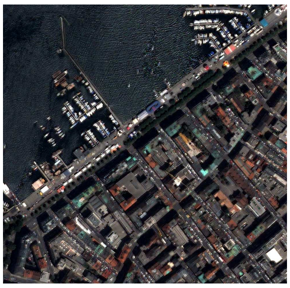
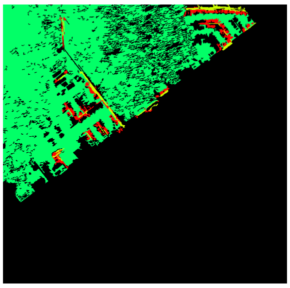
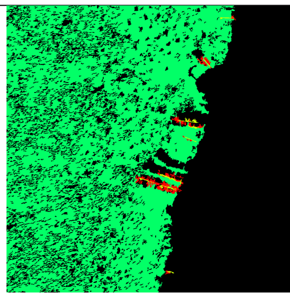
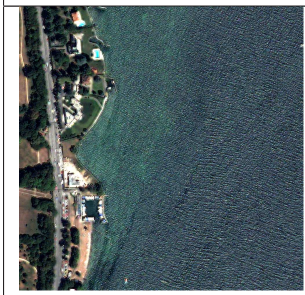
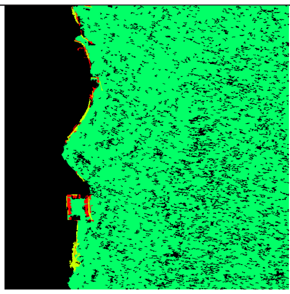
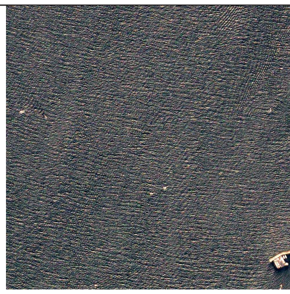
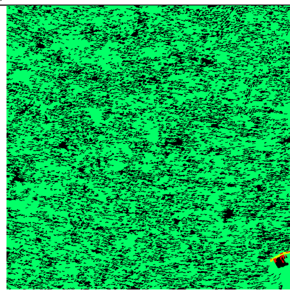
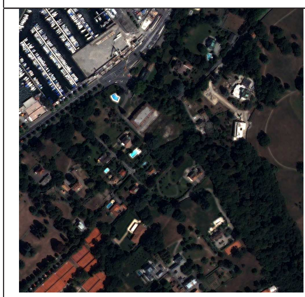
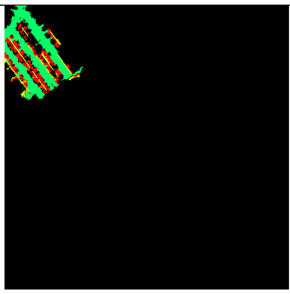
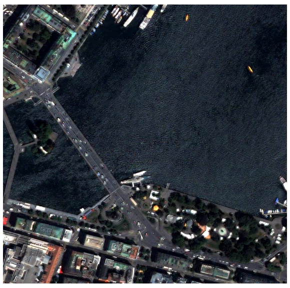
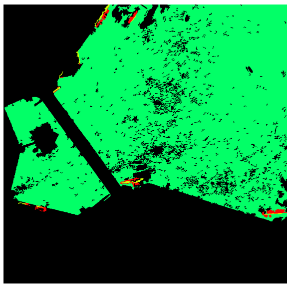
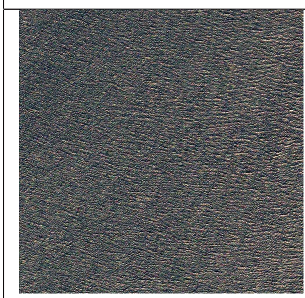
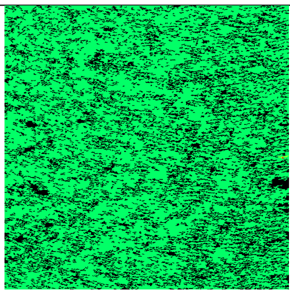
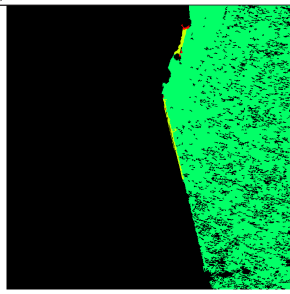
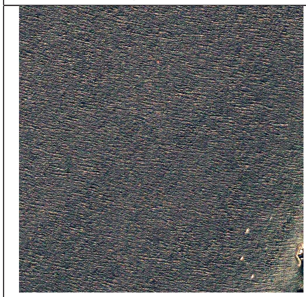
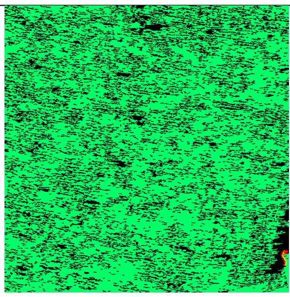
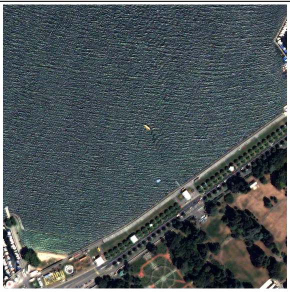
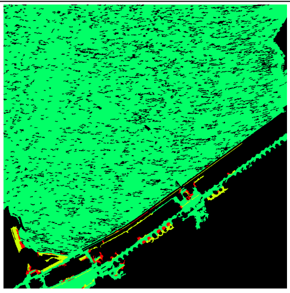
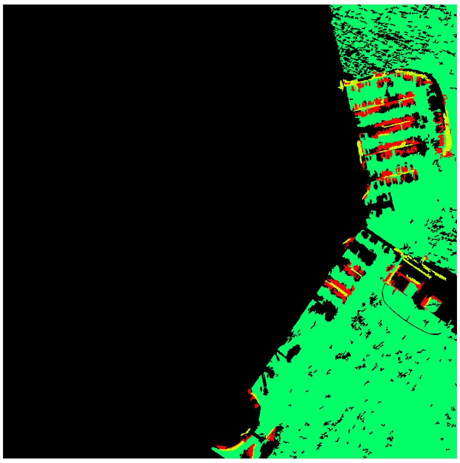
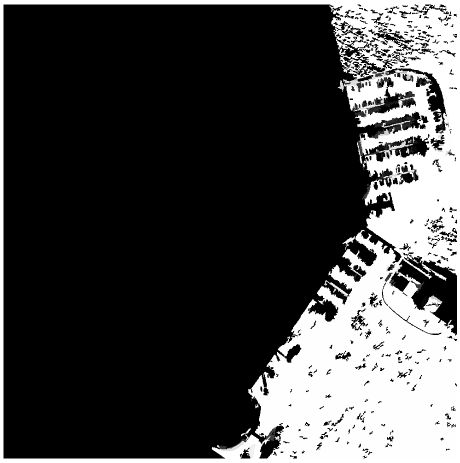
Original tile	Instantiations	Original tile	Instantiations
			
(c)		(d)	
			
(e)		(f)	
			
(g)		(h)	
			
(i)		(j)	
			Continued on next page

Table 5.5 – continued from previous page

Original tile	Instantiations	Original tile	Instantiations
			
(k)		(l)	



(a) Instantiations



(b) Order

Figure 5.25: Solutions obtained for the tile image of Figure 5.23. In green the sea, in red the boats and in yellow the docks.

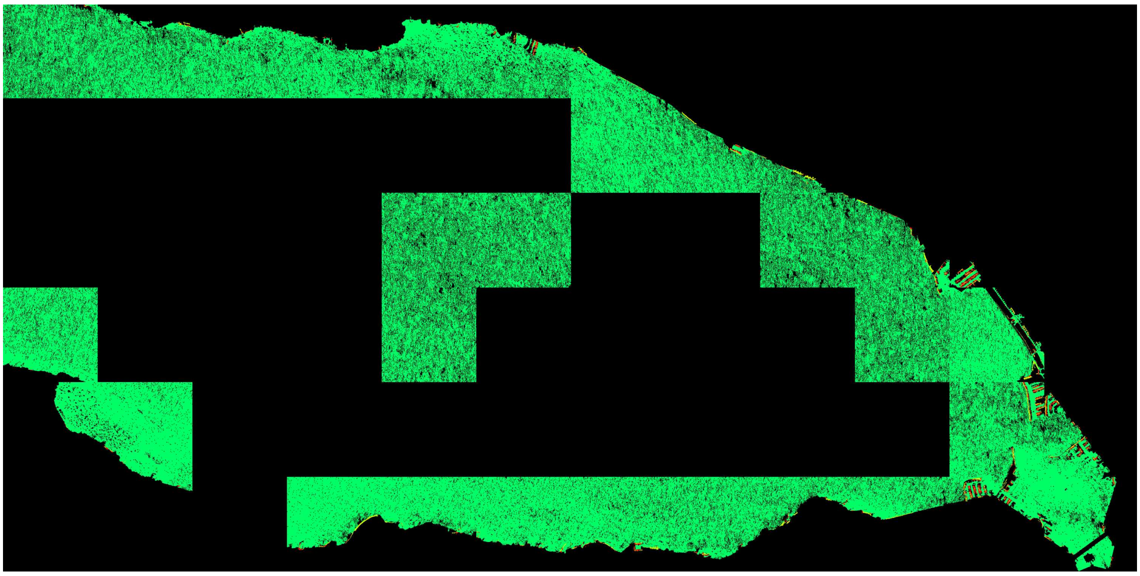


Figure 5.26: Instantiations of the model in Figure 5.8(d) in the image in Figure 5.21. In green the sea, in red the boats and in yellow the docks.

Interpretation of an airport image

The second example addresses the problem of performing the interpretation of the airport image of Figure 5.27. For this example we used the model in Figure 5.28 which is based on the concept hierarchy of Figure 5.18(a), using the same concepts of H_C as in Figure 5.18(b).

The results of the multiscale segmentation are shown in Figure 5.29. And the initial domains for each of the concepts of the conceptual graph are shown in Table 5.6. The initial domain for the green areas inside the alignment and the one outside the alignment are the same and are equivalent to D_{green_areas}.

The instantiations of the model are shown in Figure 5.30. The instantiations coincide

Figure 5.27: Airport image.

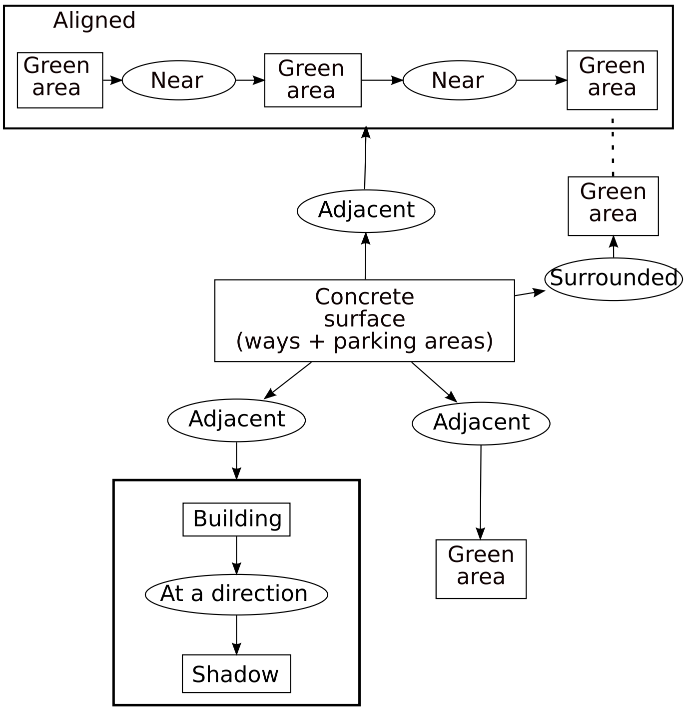


Figure 5.28: Conceptual graph describing the spatial arrangement of objects in an airport. Only one of the coreference links between the concrete surface is shown, however this link is made between every member of the aligned group and the concrete area.

with the airport. There is only one building which was not detected, and this is because it was separated into two regions in the segmentation and one of the regions satisfies the condition of being adjacent to its shadow, but they do not satisfy the condition of being adjacent to the concrete surface.

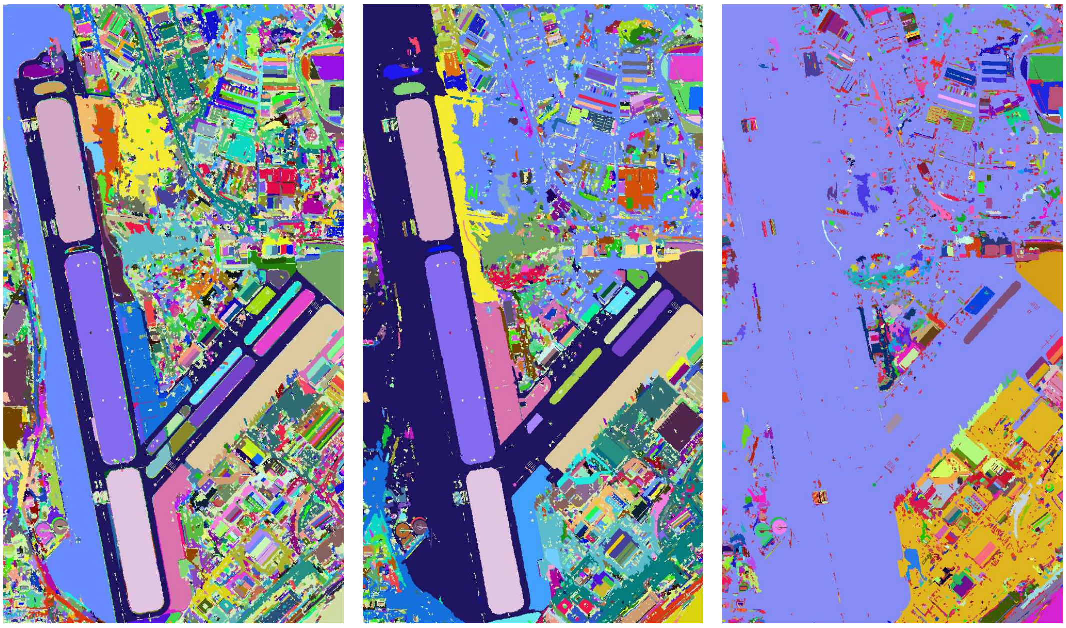


Figure 5.29: Results of multiscale segmentation of image in Figure 5.27. For visualization purposes we rotated the results by an angle of 90 degrees.

Table 5.6: Initial domains for the concept nodes representing green area, building, shadow and concrete surface.

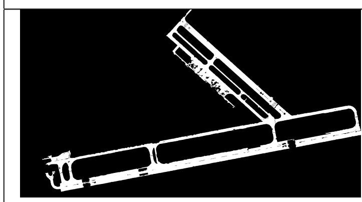
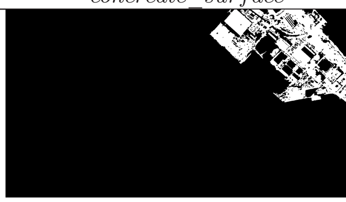
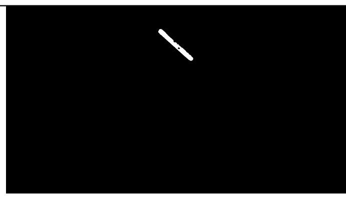
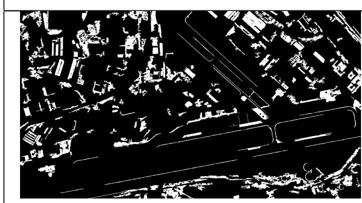
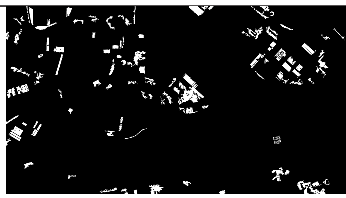
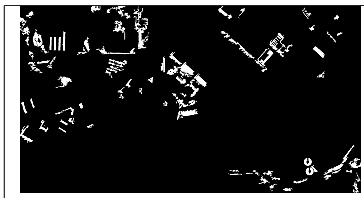
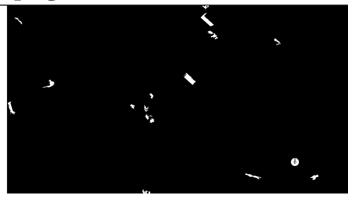
$D_{concreate\_surface}$		
		
D_{green_area}		
		
$D_{building}$		
		
D_{shadow}		
		Continued on next page

Table 5.6 – continued from previous page

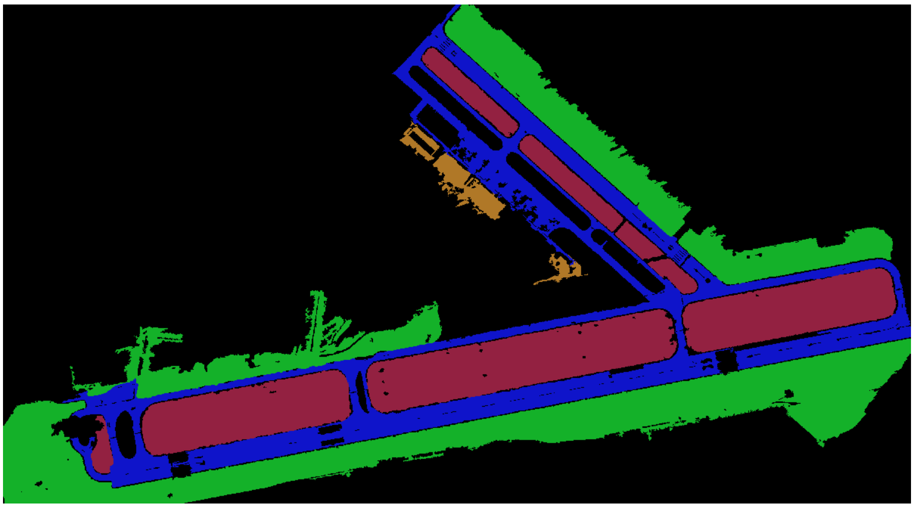


Figure 5.30: Instantiations of the conceptual graph of Figure 5.28 in the image of Figure 5.27. In red the aligned green zones, in green the green zone which does not belong to the aligned group, in blue the concrete area, and in yellow the building and its shadow.

Discussion

Through these two examples we have shown the interest of the proposed approach for obtaining the instantiations of a complex objects in an image. Even if the segmentations in both examples were not perfect the methodology was able to detect the instantiations of the models in the images. Moreover, we were able to extract the instantiations of the complex objects without the need of specifying a method for extracting each object of the scene.

The use of homomorphism between the model conceptual graph and the image allows us to retrieve several instantiations of the objects composing a complex objects, for instance we were able to retrieve the two groups of aligned green zones in the airport, or different boats which are adjacent to a dock. Moreover, using the conceptual graphs for the representation of the desired structure, allow us to represent groups having a relation with other objects and members of the group having relations with objects outside the group. The richness of this representation permits us to describe a complex structure like an airport in a simple way.

In these two examples we used very simple concept hierarchies and only few classes H_C to help us to identify the initial candidates for the instantiations of the concept nodes in the image. However, as H_C has more elements it is possible to have a smaller set for the initial candidates and the number of false detections can be reduced. The development of concept hierarchies (ontologies) is one of the objectives of the ANR project DAFOE, thus integrating the ontologies of DAFOE into our methodology can improve the results.

5.5.5.1 Implementation remarks

The whole methodology was implemented as generic C++ classes for the Orfeo Toolbox library (http://www.orfeo-toolbox.org/otb/).

5.6 Towards the introduction of priorities and uncertainties

As mentioned in Section 4.2.2 there can be uncertainty with respect to the model. In this section we discuss this problem of uncertainty.

When dealing with uncertainty in the spatial relations of the model, the question of how to deal with this uncertainty and how to propagate it in our system is raised. Relying on the possibility theory is adequate for this type of situations, since it allows to model the existence and not existence of an instantiation.

Let $G_M = (\mathcal{N}_C, \mathcal{N}_R, \mathcal{E}_M, l_M\}$ be our model conceptual graph. We can suppose as in the case of FCSP [Dubois et al., 1996] that there exist several subsets of $\mathcal{N}_R$ which define the real situation. Therefore, we could construct a possibility distribution:

$$\pi_R = 2^{\mathcal{N}_R} \rightarrow [0,1], \tag{5.12}$$

Where for $\mathcal{Q} \subset \mathcal{N}_R$, the value $\pi_R(\mathcal{Q})$ represents the possibility that the set of relations $\mathcal{Q}$ appear in the image. To construct π_R we could think of learning π_R from a data base of labeled images by evaluating the occurrence of each relation. Another alternative is to construct π_R from a priority distribution (given by a user) over $\mathcal{N}_R$ as in [Dubois et al., 1996]. Therefore, the user can construct a function:

$$\begin{aligned} Pr_R: \quad \mathcal{N}_R &\rightarrow [0,1] \\ R_i &\rightarrow Pr_R(R_i) \end{aligned}$$

where $Pr_R(R)$ indicates the degree to which it is necessary to satisfy the relation R. For example, if we want to model the relations that take place in a golf field, then the constraint specifying the relation $\mu_{SURROUNDED}(sand, green\_area)$ will have higher priority than the relation $\mu_{NEAR}(sand, water)$ since the first relation is always true, while the second one is not. If given two relations R_1 and R_2, we have that $Pr_R(R_1) > Pr_R(R_2)$ then it means that the satisfaction of R_2 is less necessary. Hence, this can express the uncertainty of the existence of R_2.

To consider the priorities in Algorithm 6, the degree of satisfaction of a prioritized relation R becomes [Dubois et al., 1996]:

$$\mu_{\tilde{R}_k}(v_1, \ldots, v_n) = \max[1 - Pr(R_k), \mu_{R_k}(v_1, \ldots, v_n)] \tag{5.13}$$

where $\mu_{R_k}(v_1, \ldots, v_n)$ is the degree of satisfaction of the relation when applied to the set $v_1, \ldots, v_n$. Finally, the possibility distribution π_R is given by:

$$\pi_R(\mathcal{Q}) = \min_{\{i: R_i \notin \mathcal{Q}\}} c(Pr_R(c(\mu_{R_k}))), \tag{5.14}$$

where c is a complement function, for example, $c(x) = 1 - x$. Thus if for every relation $R \in \mathcal{N}_R$ we use $\mu_{\tilde{R}}$ rather than μ_R in the algorithm of FCSP, then we are able to propagate this uncertainty throughout the system.

If the uncertainty refers to whether an object exists or not in the model then we have to model it over the set $\mathcal{N}_C$ of concepts by constructing a possibility distribution over the concepts. The posssibility distribution over the concepts has to be linked with the one over the relations. Therefore, we can construct a priority distribution Pr_C over the concepts, that will be attached to the priority distribution Pr_R over the relations. For a binary fuzzy relation R over fuzzy variables a and b, we have that for every x in the domain of a, and every y in the domain of b the following inequality is true:

$$\mu_{R(a,b)}(x, y) \leq \min(\mu_a(x), \mu_b(y)).$$

This inequality can be extended to relations of any arity. Therefore, if we construct a priority distribution over the concepts Pr_C then we can modify the priority distribution over the relations by:

$$Pr_R(R) = \min(Pr_R(R), \min_{C_i \in S_R} Pr_C(C_i)), \tag{5.15}$$

where S_R contains the variables involved in the relation R. The priorities over the concepts can represent the key concepts in an scene, for example in a model describing a harbor the concept of water is necessary. The use of priorities on the objects is also used in [Deruyver et al., 2009] when a degree of freedom of each relation is introduced by using the bi-level constraints.

5.7 Conclusion

In this chapter we addressed the problem of incorporating complex spatial relations in the representation of a model which represents a scene that we want to find in an image. For this, we first adapted a representation scheme to introduce this type of information in a model, then we addressed the problem of identifying the model in the image, by simultaneously considering some of the possible information imperfections which are present in this type of problem.
The mapping problem was represented as a graph homomorphism which allows us to have several instantiations of a model in an image. This flexibility is adequate for satellite images since most of the time the number of instantiations of a model in an image is unknown.

The problem of obtaining the graph homomorphism in an image was formulated as a CSP, as in [Chein and Mugnier, 2008]. However, due to the imprecision of the spatial relations, it was necessary to move to a more flexible formalism such as the FCSP. The arc-consistency algorithm proposed by [Dubois et al., 1996] only deals with binary relations, thus we adapted this algorithm to deal with relations having an arity greater than 2. Moreover, we adapted the algorithm to deal with groups of objects which can be aligned or not.

Finally, we proposed a methodology to find the instantiations of a conceptual graph in an unlabeled image. Our method was successfully applied in unlabeled images obtaining adequate results, even if there were segmentation errors. The results demonstrate the interest of using the spatial relations for the interpretation of images.

Chapter 6

Conclusion

6.1 Main contributions

This thesis focused in spatial reasoning applied to very high resolution remote sensing images. The main contributions of this thesis are on one hand the novel definitions of spatial relations within a fuzzy set framework, and on the other hand the proposition of a methodology to perform spatial reasoning using the newly defined relations, among others, applied to the interpretation of Earth observation images using a model. In the following we discuss these contributions.

6.1.1 Novel definitions of spatial relations

In the first part of the thesis we concentrated on spatial relations present in Earth observation images. In Chapter 1 we defined a set containing these relations. From this set of relations we concentrated on 4 types of spatial relations, which had a definition which was not adapted to this type of images. These relations are:

Surround: We proposed a definition for the relation "surround", based on the idea of angular coverage used in the definitions which had been already proposed in the literature [Rosenfeld and Klette, 1985, Miyajima and Ralescu, 1994b]. However, our approach differs from previous work in the sense that we propose to model the relation as a fuzzy landscape. This definition integrates two types of information: the angular coverage by the concavities of the object (and not the whole object [Rosenfeld and Klette, 1985]), and the distance to the target object which is linked to the size of the target object. This second type of information had not been considered previously, and allow us to give a definition which is adapted the cases where the size of the object is very small compared to the concavities of he reference object. However, this second type of information can make the computation of the relation inefficient when we have to evaluate the relation for several target objects having different sizes, because it would be necessary to compute one fuzzy landscape for each object. Thus in such cases we could think of constructing just one fuzzy landscape without considering the distance information, and add the distance information later.

Alignment and parallelism: We introduced two novel definitions for alignment of objects: *local* and *global* alignment. Both definitions are based on the notion of neighborhood of an object and on relative direction. We proposed an original method for

extracting the groups of *locally* and *globally* aligned objects. The proposed definitions and methods allow us to consider different definitions of the neighborhood of an object, making them very flexible. These definitions of alignment consider the object as a whole making them more appropriate for determining the alignment between objects, than the definitions which simplify the object by reducing it to just one point.
We studied the parallel relation mainly in the context of determining when a group of *globally* aligned objects is parallel to another group or to another group of *globally* aligned objects. The combination of parallelism and alignment allows us to describe complex spatial configurations, which is an important contribution with respect to previous work.

Line-region relations: We proposed definitions dedicated to the cases where one of the objects involved in the relation is a linear object. We focused on four relations: "go through", "go across", "go deep into" and "enter". We defined these relations using topological and geometrical components. The definitions were based on results obtained from human-subject tests and provide results in accordance with human perception. However, there are several aspects of the relations that need a further study, for instance, what is the perception of the relation when there are considerable changes in the curvature of the linear object along its curvilinear axis.

All of the proposed definitions were modeled in such a way that they represent the imprecision linked to the semantics of the relations. Moreover, we proposed extensions for the "surround", "alignment" and "parallel" relations to deal with fuzzy objects.

6.1.2 Spatial reasoning and image interpretation guided by a model

The second part of this thesis concerns the integration of fuzzy spatial relations, in particular the "alignment" relation, into a system to perform spatial reasoning applied to the interpretation of images guided by a model. The main contributions concerning this second part are:

Representation of spatial scenes: We proposed conceptual graphs as models for representing the spatial knowledge of a scene. To represent all the spatial relations we adapted the nested conceptual graphs to represent an aligned group of objects.

Extension of fuzzy CSP: We introduced a formal framework, based on existing fuzzy CSP of [Dubois et al., 1996], to address the problems of:

(i) Determining arc-consistency closure of a network containing constraints representing relations of any arity.

(ii) Determining arc-consistency closure of a CSP network representing a conceptual graph containing complex concept nodes, with an unspecified number of elements, and which contains spatial relations or satisfies the spatial property of alignment.

We addressed the first problem by extending the definition of arc-consistency for fuzzy binary constraints proposed by [Dubois et al., 1996], to fuzzy n-ary constraints. For the second problem, we introduced a novel type of constraints which can also been seen as variables. These constraints/variables are modeled using two different membership functions: one dealing with the satisfaction of the constraints in the interior of the complex node, and the other representing the satisfaction of the relations between the concept complex node and the other concept nodes, in addition to the satisfaction of the constraints in the interior of the complex node. This representation enables the integration of the complex spatial information, represented by the alignments, into the fuzzy CSP. Moreover, we proposed an algorithm to determine the arc-consistency closure of a network containing constraints/variables.

We presented the results of the methodology in two examples which demonstrate the interest of using the proposed spatial relations for the interpretation of complex objects. Moreover, we were able to obtain the desired intantiations despite the segmentation errors. The methodology was applied on Earth observation images, nonetheless it can be applied to any type of images.

6.2 Perspectives

6.2.1 Short-term perspectives

Study of the polymorphism between a group of aligned objects and a linear object

In [Mathet, 2000] the property of treating a group of objects as only one object, and in particular as a linear object, was discussed. Therefore, it can be interesting to extend the definitions of line-region relations when observing a group of aligned objects as a linear object. Moreover, we could use other gestalt laws, apart from alignment, to form groups which can be seen as an object.

Introduction of the uncertainty of the model into the interpretation method

As it was noted in Chapter 4, one of the difficulties of obtaining the instantiations of a model in an image is the uncertainty on whether all the objects and relations that appear in the model are present in the image. This type of situation is very common in Earth observation images due to the occlusions, and moreover due to the variety of spatial structures which can form a semantic concept. In Section 5.6 we proposed a way to address this problem. It would be useful to integrate this proposal into the system.

Optimization of the algorithm for finding the arc-consistency closure of a nested constraint network with complex concept nodes

In Algorithm 6 the order in which a constraint is chosen, to check the arc-consistency of its domains, depends on its computation time. We give a higher priority for those constraints having a low complexity value. This approach is an efficient way to construct the domain of each relation in its first evaluation. However, once the domain of the relation has been created this approach does not contribute to reducing the algorithm's computational time. Thus, it would be more convenient to use another strategy. For instance, choosing the constraints involving a variable for which the domain was considerably reduced, as in [Boussemart et al., 2004, Nempont, 2009]. The idea behind this strategy is that if the domain of a variable x_i has been reduced, then the domains of the variables related to x_i are also probably reduced.

Extraction of initial regions and information labeling

In the interpretation of an unlabeled image we used a generic multi-scale segmentation.

However, this type of segmentation does not correctly extract the fine structures of the image. Hence, it would be more appropriate to use another approach which considers the geometry of the regions, for example, the cocons model proposed in [Guigues et al., 2003], or the segmentation proposed in [Bin, 2007]. This method consists in filtering the tree of shapes of the images, which consists of the boundaries of the level-sets of the image, using scale and contrast criteria, where the scale depends on the geometry of the region. These two types of segmentations were conceived for earth observation images, and allow to recognize linear objects. Furthermore, for the initial labeling of the regions it can be interesting to incorporate the geometric information included in the scale proposed in [Bin, 2007].

Using ontologies from DAFOE

The DAFOE (Differential and Formal Ontology Editor) platform[1] allows the creation and management of ontologies. An image ontology and a scene ontology for very high resolution satellite image interpretation were developed in the Competence Center (COC) CNES/DLR/Telecom ParisTech under the supervision of Marine Campedel. The image ontology describes the processes used to extract the low level features, and the scene ontology describes the concepts on the scene. The objective of these two ontologies is to reduce the semantic gap between the image processing experts and the photo interpreters. An annotation tool to assist photo interpreters using these ontologies was proposed.
It would be interesting to use the knowledge represented by these ontologies to obtain the initial candidate regions of our system, and therefore propose an alternative annotation tool, which allows the photo interpreter to specify the spatial relations between the objects of interest.

6.2.2 Long-term perspectives

Integration of the system into a query based architecture with user relevance feedback

When constructing a conceptual graph containing the spatial information of a scene that we want to search in an image, it can be difficult to consider all the possible relations that take place among the objects. Thus, we can think of integrating the methodology that we proposed for the interpretation of images into an architecture which allows the user to give an initial conceptual graph and the information to extract the scene of interest. The system will use this information to give a first result. Then the user can add more relations or concept nodes to the graph in order to obtain the desired result.

Study of the relevance of spatial relations for describing a spatial scene, and automatic creation of conceptual graphs representing the spatial structure of a scene

There can be several ways of describing a same scene. For instance, the situation searched by the conceptual graph of Figure 5.17 can also be described by the conceptual graph of Figure 6.1. These two conceptual graphs can be used to search for the "the group of neighboring houses forming an aligned group, which have a pool located in the garden at the "back" of the house, and which have a shadow". Since both graphs have a different number of spatial relations and the relations are different, the computational time of the arc-consistency would be different. Therefore it can be interesting to study which description

[1] http://dafoe4app.fr/

is more relevant and within a description which relations are more relevant. One possibility for this could be to adapt the relevance measures proposed in [Dessalles, 2008] for natural language descriptions, to our context. Furthermore, after having a criterion for evaluating the relevance of a spatial relation in a model it could be interesting to study learning algorithms which could automatically generate relevant conceptual graphs to describe a class of complex scenes, from a labeled image base.

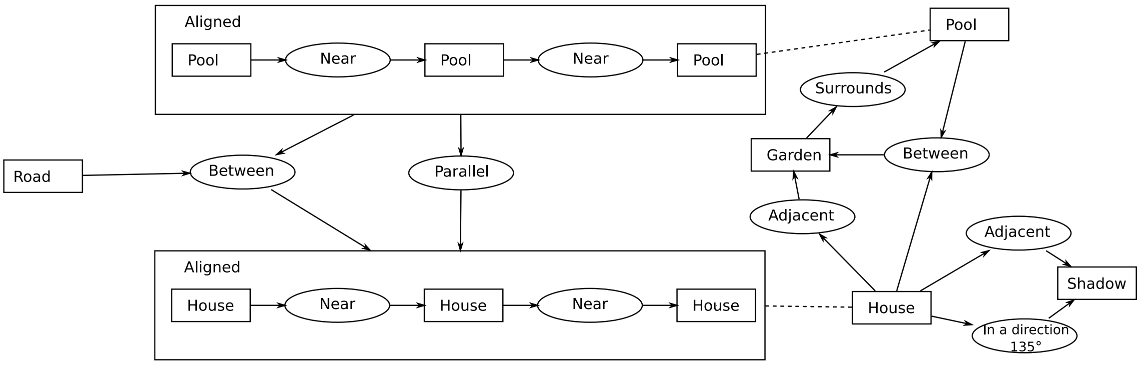


Figure 6.1: Alternative conceptual graph for representing "the group of neighboring houses forming an aligned group, which have a pool located in the garden at the "back" of the house, and which have a shadow".

Notations

	Fuzzy sets, spatial relations, mathematical morphology and possibility theory
$\mathfrak{I}$	discrete domain (i.e $\mathbb{Z}^2$)
$\mu_A : \mathfrak{I} \to [0,1]$	membership function representing a fuzzy set A over $\mathfrak{I}$
$(\mu_A)_\alpha$	α-cut of μ
$\mu_{int}(A,B)$	degree of intersection between two fuzzy sets A and B (Equation 1.2)
$\mu_{\neg int}(A,B)$	degree of non-intersection between two fuzzy sets A and B (Equation 1.4)
c	fuzzy complementation
$\mu_{inclusion}(A,B)$	degree of inclusion of A in B (Equation 1.5)
Π	possibility degree (Equations 1.6)
N	necessity degree (Equation 1.7)
M	mean measure (Equation 1.16)
t	t-norm
T	t-conorm
$D_{V_c}(A)$	morphological dilation of A by the structuring element V_c
$\mu_{adj}(A,B)$	degree of adjacency between A and B (Equation 1.11)
$H^A(B)$	angle histogram from A to B (Equation 1.14)
$\mu_{distance}(A)(x)$	membership function of the landscape representing "at a distance between d_1 and d_2 from A" (Equation 1.24)
$D_{\nu_\theta}(A)$	directional dilation of a fuzzy set A in a direction $\vec{u}_\theta$
$\mu_{surround}(A)(x)$	membership function of the landscape representing "surrounded by A" (Equation 2.10)
$\mu_{surround}(A,\mu_n)(x)$	membership function of the landscape representing "surrounded by A and near" (Equation 2.13)
$O(A,B)$	orientation histogram between A and B
$\mu_{ALIG}(\mathcal{S})$	degree of *global* alignment of a group of objects $\mathcal{S}$ (Definition 2.1)
$\mu_{LA}(\mathcal{S})$	degree of *local* alignment of a group of objects $\mathcal{S}$ (Definition 2.3)
$\mu_{parallel}(A,B)$	degree of satisfaction of the relation "A is parallel to B", where A and B can represent objects (fuzzy or not), or groups of *globally* aligned objects (Section 2.3)
$\mu_{go\_through}(L,R)$	degree to which L(linear object) "goes through" R (Equation 2.53)
$\mu_{go\_across1}(L,R)$	degree of satisfaction of the relation "L goes across R" using the meaning of going from one side to the opposite one (Equation 2.57)

$\mu_{go\_across2}(L,R)$	degree of satisfaction of the relation "L goes across R" using the meaning of going deep into the region (Equation 2.59)
$\mu_{enter}(L,R)$	degree of satisfaction of the relation "L enters R" (Equation 2.62)
$\mu_{go\_deep\_into}(L,R)$	degree of satisfaction of the relation "L goes deep into R" (Equation 2.63)
$\mu_{along}(L,R)$	degree of satisfaction of the relation "L is along R" (Equation 2.63)

	Conceptual graphs and constraint satisfaction problems
$\mathcal{V} = (T_C, T_R, I)$	vocabulary (Definition 5.1) of a conceptual graph, where: T_C concept ontology T_R relation ontology I set of individual markers
$G = \{\mathcal{N}_C, \mathcal{N}_R, \mathcal{E}, l\}$	conceptual graph (Definition 5.2), where: $\mathcal{N}_C$ concept nodes set $\mathcal{N}_R$ relation nodes set $\mathcal{E}$ is the family of edges l is a labeling function
$\mathcal{P} = \langle \mathcal{X}, \mathcal{D}, \mathcal{C} \rangle$	contraint satisfaction problem, where: $\mathcal{X}$ set of variables $\mathcal{D}$ set of domains $\mathcal{C}$ set of constraints
$C_k = \langle \mu_{R_k}, S_k \rangle$	fuzzy constraint C_k , where μ_{R_k} is the fuzzy relation representing C_k and S_k is the set containing the variables involved in C_k

Basic definitions of fuzzy set theory

Let $\mathfrak{I}$ be the image space, in our case $\mathbb{Z}^2$. A fuzzy set A defined over $\mathfrak{I}$, is represented through its membership function $\mu_A : \mathfrak{I} \rightarrow [0,1]$. For any point $x \in \mathfrak{I}$, $\mu_A(x)$ is the degree to which x belongs A. Let us note by $\mathcal{F}$ the set of all fuzzy sets defined on $\mathfrak{I}$.
For $\mu_A \in \mathcal{F}$ we can defined the following crisp sets: its core defined as $Core(\mu_A) = \{x \in \mathfrak{I}, \mu_A(x) = 1\}$, its support $Supp(\mu_A) = \{x \in \mathfrak{I}, \mu_A(x) > 0\}$ and its a-cuts (for $\alpha \in [0,1]$) $(\mu_A)_\alpha(x) = \{x \in \mathfrak{I}, \mu_A(x) \geq \alpha\}$.

Fusion operators [Bloch, 1996]

A t-norm is an operator t from $[0,1] \times [0,1]$ into [0,1] which is commutative, associative, increasing in both variables and that admits 1 as unit element. It represents a conjunction and generalizes intersection and logical "and". Typical examples are $\min(A,B)$, AB, $\max(A+B-1,0)$, the last one being known as the Lukasiewicz t-norm. A t-conorm is an operator T from $[0,1] \times [0,1]$ into $[0,1]$ which is commutative, associative, increasing in both variables and that admits 0 as unit element. It represents a disjunction and generalizes union and logical "or". Typical examples are $\max(A,B)$, $A+B-AB$, $\min(A+B,1)$, the last one being the Lukasiewicz t-conorm.

Appendix A

List of publications

International Conferences with Peer-review:

- M.C. Vanegas, I. Bloch and J. Inglada, "Detection of Aligned Objects for High Resolution Image Understanding". In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing (IGARSS 2010), Honolulu, Hawai, USA, July 2010.
- M.C. Vanegas, I. Bloch and J. Inglada, "Searching aligned groups of objects with fuzzy criteria". In the Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU 2010), Dortmund, Germany, June 2010.
- M.C. Vanegas, I. Bloch and J. Inglada, "Fuzzy Spatial Relations for High Resolution Remote Sensing Image Analysis: The Case of “To Go Across” ". In Proceedings of the IEEE International Conference on Geoscience and Remote Sensing (IGARSS 2009), Cape Town, South Africa, July 2009.
- M.C. Vanegas, I. Bloch, H. Maître and J. Inglada, "Approximate Parallelism Between Fuzzy Objects: Some Definitions", In Proceedings of the International Workshop on Fuzzy Logic and Applications (WILF 2009), Palermo, Italy, June 2009.

Bibliography

R. F. Abler. The national science foundation national center for geographic information and analysis. *International Journal of Geographical Information Systems*, 1:303–326, 1987.

S. Aksoy. Modeling of remote sensing image content using attributed relational graphs. In D.-Y. Yeung, J. Kwok, A. Fred, F. Roli, and D. de Ridder, editors, *Structural, Syntactic, and Statistical Pattern Recognition*, volume 4109 of *Lecture Notes in Computer Science*, pages 475–483. Springer Berlin / Heidelberg, 2006.

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

J. Atif, C. Hudelot, G. Fouquier, I. Bloch, and E. D. Angelini. From generic knowledge to specific reasoning for medical image interpretation using graph based representations. In *International Joint Conference on Artificial Intelligence IJCAI'07*, pages 224–229, 2007.

F. Baader. Description logics. In *Reasoning Web: Semantic Technologies for Information Systems, 5th International Summer School 2009*, volume 5689 of *Lecture Notes in Computer Science*, pages 1–39. Springer–Verlag, 2009.

F. Baader, R. Molitor, and S. Tobies. Tractable and Decidable Fragments of Conceptual Graphs. In *Proceedings of the 7th International Conference on Conceptual Structures: Standards and Practices*, pages 480–493, 1999.

C. Barber, D. Dobkin, and H. Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4):469–483, 1996.

J. A. Bateman and S. Farrar. Spatial ontology baseline. SFB/TR8 internal report I1-[OntoSpace]: D2, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany, 2004.

E. Bengoetxea, P. Larranaga, I. Bloch, A. Perchant, and C. Boeres. Inexact Graph Matching by Means of Estimation of Distribution Algorithms. *Pattern Recognition*, 35:2867–2880, 2002.

B. Bennett. Modal logics for qualitative spatial reasoning. *Logic Journal of the IGPL*, 4 (1):23–45, 1996.

B. Bennett, A. Isli, and A. Cohn. When does a composition table provide a complete and tractable proof procedure for a relational constraint language. In *Proceedings of the IJCAI-97 workshop on Spatial and Temporal Reasoning*, Nagoya, Japan, 1997.

U. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4):239–258, 2004.

C. Bessiere. *Foundations of Artificial Intelligence*, volume 2, chapter Constraint propagation, pages 29–83. Elsevier, 2006.

L. Bin. *Echelle et résolution en imagerie de télédétection*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Dec. 2007.

I. Bloch. Fuzzy spatial relationships for image processing and interpretation: a review. *Image and Vision Computing*, 23:89–110, 2005.

I. Bloch. Spatial reasoning under imprecision using fuzzy set theory, formal logics and mathematical morphology. *International Journal of Approximate Reasoning*, 41:77–95, 2006.

I. Bloch. Information combination operators for data fusion: A comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.

I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32(11):1873–1895, 1999.

I. Bloch. Fuzzy Relative Position between Objects in Image Processing: a Morphological Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7): 657–664, 1999.

I. Bloch and H. Maître. Fuzzy Mathematical Morphologies: A Comparative Study. *Pattern Recognition*, 28(9):1341–1387, 1995.

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

I. Bloch, C. Pellot, F. Sureda, and A. Herment. Fuzzy Modelling and Fuzzy Mathematical Morphology applied to 3D Reconstruction of Blood Vessels by Multi-Modality Data Fusion. In D. D. R. Yager and H. Prade, editors, *Fuzzy Set Methods in Information Engineering: A Guided Tour of Applications*, chapter 5, pages 93–110. John Wiley and Sons, New-York, 1996.

I. Bloch, H. Maître, and M. Anvari. Fuzzy adjacency between image objects. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(6):615–653, 1997.

I. Bloch, T. Géraud, and H. Maître. Representation and Fusion of Heterogeneous Fuzzy Information in the 3D Space for Model-Based Structural Recognition - Application to 3D Brain Imaging. *Artificial Intelligence*, 148(1-2):141–175, Aug. 2003.

I. Bloch, O. Colliot, and R. Cesar. On the Ternary Spatial Relation Between. *IEEE Transactions on Systems, Man, and Cybernetics SMC-B*, 36(2):312–327, Apr. 2006.

F. Bobillo and U. Straccia. fuzzyDL: An expressive fuzzy description logic reasoner. In *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2008.*, pages 923–930. IEEE, 2008.

J. B. Bordes. *Inférence de connaissances sémantiques : application aux images satellitaires.* PhD thesis, Ecole Nationale Supérieure des Télécommunications, Apr. 2009.

B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy Sets and Systems*, 84(2):143–153, 1996.

F. Boussemart, F. Hemery, and C. Lecoutre. Revision ordering heuristics for the Constraint Satisfaction Problem. In *Proceeding of the First international workshop "Constraint propagation and implementation" of the 10th International Conference on Principles and Practice of Constraint Programming (CP'2004)*, pages 29–43, Toronto, Canada, 2004.

R. Cesar, E. Bengoetxea, I. Bloch, and P. Larranaga. Inexact Graph Matching for Model-Based Recognition: Evaluaton and Comparison of Optimization Algorithmss. *Pattern Recognition*, 38:2099–2113, 2005.

K. Chan and Y. Cheung. Fuzzy-attribute graph with application to Chinese character recognition. *IEEE Transactions on Systems, Man and Cybernetics*, 22(1):153–160, 2002.

M. Chein and M.-L. Mugnier. *Graph-based Knowledge Representation and Reasoning — Computational Foundations of Conceptual Graphs*. Advanced Information and Knowledge Processing. Springer, 2008.

M. Chein and M.-L. Mugnier. Conceptual graphs: Fundamental notions. *Revue d'Intelligence Artificielle*, 6(4):365–406, 1992.

S. Christophe and A. Ruas. Detecting building alignments for generalisation purposes. In *Proceedings of 10th International Symposium on Spatial Data Handling*, pages 419–432, 2002.

M. Ciucu, P. Héas, M. Datcu, and J. C. Tilton. Scale space exploration for mining image information content. In *Mining Multimedia and Complex Data, KDD Workshop MDM/KDD 2002, PAKDD Workshop KDMCD 2002, Revised Papers*, pages 118–133. Springer, 2002.

V. Clément and M. Thonnat. Integration of image processing procedures: OCAPI, a knowledge-based approach. Technical report, Institut National de Recherche en Informatique et en Automatique, 1990.

V. Clement and M. Thonnat. A knowledge-based approach to integration of image processing procedures. *CVGIP: Image Understanding*, 57(2):166–184, Mar. 1993.

E. Clementini and P. D. Felice. A spatial model for complex objects with broad boundary supporting queries on uncertain data. *Data and Knowledge Engineering*, 37:285–305, 2001.

E. Clementini, P. Felice, and D. Hernández. Qualitative representation of positional information. *Artificial Intelligence*, 95(2):317–356, 1997.

A. Cohn and N. Gotts. The 'egg-yolk' representation of regions with indeterminate boundaries. *Geographic objects with indeterminate boundaries*, 2:171–187, 1996.

A. Cohn, B. Bennet, J. Gooday, and N. Gotts. Qualitative spatial representation and reasoning with the region connection calculus. *GeoInformatica*, 1:275–316, 1997.

A. G. Cohn. Modal and non-modal qualitative spatial logics. In F. Anger, H.-W. Guesgen, and J. van Benthem, editors, *Proceedings of the Workshop on Spatial and Temporal Reasoning at the 13th International Joint Conference on Artificial Intelligence*, pages 95–100, Aug. 1993.

A. G. Cohn, D. A. Randell, and Z. Cui. Taxonomies of logically defined qualitative spatial relations. *International Journal of Human-Computer Studies*, 43(5/6):831–846, 1995.

O. Colliot, O. Camara, and I. Bloch. Integration of Fuzzy Spatial Relations in Deformable Models - Application to Brain MRI Segmentation. *Pattern Recognition*, 39(8):1401–1414, Aug. 2006.

J. Condotta and E. Würbel. Réseaux de contraintes temporelles et spatiales. In F. Le Ber, G. Ligozat, and O. Papini, editors, *Raisonnements sur l'espace et le temps: des modèles aux applications (Traité IGAT, série géomatique)*, chapter 7, pages 181–224. Lavosier, 2007.

D. Crevier and R. Lepage. Knowledge-Based Image Understanding Systems: A Survey* 1. *Computer Vision and Image Understanding*, 67(2):161–185, 1997.

P. Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14 (3):227–248, 1980.

S. Dasiopoulou, C. Saathoff, P. Mylonas, Y. Avrithis, Y. Kompatsiaris, S. Staab, and M. Strinztis. Introducing context and reasoning in visual content analysis: An ontology-based framework. *Semantic Multimedia and Ontologies*, pages 99–122, 2008.

E. Davis, N. M. Gotts, and A. G. Cohn. Constraint Networks of Topological Relations and Convexity. *Constraints*, 4(3):241–280, 1999.

D. DeMenthon and R. Megret. Spatio-Temporal Segmentation of Video by Hierarchical Mean Shift Analysis. *Language*, 2:1, 2002.

A. Deruyver and Y. Hodé. Constraint satisfaction problem with bilevel constraint: application to interpretation of over-segmented images. *Artificial intelligence*, 93(1-2):321–335, 1997.

A. Deruyver, Y. Hodé, and L. Brun. Image interpretation with a conceptual graph: Labeling over-segmented images and detection of unexpected objects. *Artificial Intelligence*, 173(14):1245–1265, 2009.

A. Desolneux, L. Moisan, and J. Morel. A grouping principle and four applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4):508–513, 2003.

A. Desolneux, L. Moisan, and J. M. Morel. *From Gestalt Theory to Image Analysis: A Probabilistic Approach*, volume 34 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 2008.

J. Dessalles. *La pertinence et ses origines cognitives: nouvelles théories*. Hermès science publications, 2008.

E. Dogrusoz and S. Aksoy. Modeling urban structures using graph-based spatial patterns. In *IEEE International Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007*, pages 4826–4829, 2007.

B. Draper, R. Collins, J. Brolio, A. Hanson, and E. Riseman. The schema system. *International Journal of Computer Vision*, 2(3):209–250, 1989.

D. Dubois and H. Prade. *Fuzzy sets and systems: theory and applications*, volume 144 of *Mathematics in Science and Engineering*. Academic Press, 1980.

D. Dubois, H. Prade, and C. Testemale. Weighted fuzzy pattern matching. *Fuzzy Sets and Systems*, 28(3):313–331, 1988.

D. Dubois, H. Fargier, and H. Prade. Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty. *Applied Intelligence*, 6(4):287–309, 1996.

F. Dupin de Saint-Cyr and H. Prade. Logical handling of uncertain, ontology-based, spatial information. *Fuzzy Sets and Systems*, 159(12):1515–1534, 2008.

N. Durand, S. Derivaux, G. Forestier, C. Wemmert, P. Gançarski, O. Boussaid, and A. Puissant. Ontology-based object recognition for remote sensing image interpretation. In *19th IEEE International Conference on Tools with Artificial Intelligence, ICTAI*, pages 472–479, 2007.

M. Egenhofer and J. Herring. Categorizing Binary Topological Relations Between Regions, Lines, and Points in Geographic Databases. Technical report, Department of Surveying Engineering, University of Maine, 1990.

M. J. Egenhofer and R. Franzosa. Point-set topological spatial relations. *International Journal of Geographical Information Systems*, 5(2):161–174, 1991.

H. Fargier. *Modèles et algorithmes pour l'aide à la décision.* Habilitation à diriger des recherches, Université Paul Sabatier, Toulouse, France, octobre 2006.

D. Fensel, F. Van Harmelen, I. Horrocks, D. McGuinness, and P. Patel-Schneider. OIL: An ontology infrastructure for the semantic web. *IEEE Intelligent Systems*, 16(2):38–45, 2005.

S. Fortune. A sweepline algorithm for Voronoi diagrams. *Algorithmica*, 2(1):153–174, 1987.

G. Fouquier. *Optimisation de séquences de segmentation combinant modèle structurel et focalisation de l'attention visuelle. Application à la reconnaissance de structures cérébrales dans des images IRM 3D.* PhD thesis, Télécom ParisTech, 2010.

G. Fouquier, J. Atif, and I. Bloch. Sequential Spatial Reasoning in Images based on Pre-Attention Mechanisms and Fuzzy Attribute Graphs. In *European Conference on Artificial Intelligence ECAI*, pages 611–615, Patras, Greece, jul 2008.

H. Freeman. On the encoding of arbitrary geometric configurations. *IEEE Transactions on electronic computers*, 10:260–268, June 1961.

H. Freeman. Computer processing of line-drawing images. *ACM Computing Surveys*, 6: 57–97, Mar. 1974.

J. Freeman. The modelling of spatial relations. *Computer Graphics and Image Processing*, 4(2):156–171, June 1975.

C. Freksa. Using orientation information for qualitative spatial reasoning. In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. International ConferenceGIS—From Space to Territory"*, pages 162–178, 1992.

P. Garnesson, G. Giraudon, and P. Montesinos. MESSIE : un systeme multi spécialistes en vision. Application à l'interprétation en imagerie aérienne. Technical Report 1012, INRIA, 1989.

S. Goward, B. Markham, D. Dye, W. Dulaney, and J. Yang. Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. *Remote Sensing of Environment*, 35(2-3):257–277, 1991.

T. Gruber et al. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies*, 43(5):907–928, 1995.

H. W. Guesgen and J. Albrecht. Imprecise reasoning in geographic information systems. *Fuzzy Sets and Systems*, 113(1):121–131, 2000.

L. Guigues, H. L. Men, and J. P. Cocquerez. The hierarchy of the cocoons of a graph and its application to image segmentation. *Pattern Recognition Letters*, 24(8):1059–1066, 2003.

D. Guo, H. Xiong, V. Atluri, and N. Adam. Object discovery in high-resolution remote sensing images: a semantic perspective. *Knowledge and Information Systems*, 19(2): 211–233, 2009.

D. Hernandez, E. Clementini, and P. Di Felice. Qualitative distances. In A. Frank and W. Kuhn, editors, *Spatial Information Theory - Proceedings of COSIT'95*, pages 45–57, 1995.

A. Herskovits. Language, spatial cognition, and vision. In O. Stock, editor, *Spatial and Temporal Reasoning*, pages 155–202. Springer Netherlands, 1997.

I. Horrocks, A. Rector, and C. Goble. A description logic based schema for the classification of medical data. *Knowledge Representation Meets Databases-KRDB*, 96, 1996.

C. Hudelot. *Towards a Cognitive Vision Platform for Semantic Image Interpretation; Application to the Recognition of Biological Organisms.* PhD thesis, University of Nice Sophia Antipolis, 2005.

C. Hudelot, N. Maillot, and M. Thonnat. Symbol grounding for semantic image interpretation: from image data to semantics. In *Tenth IEEE International Conference on Computer Vision Workshops, 2005. ICCVW'05*, Beijing, China, 2005.

C. Hudelot, J. Atif, and I. Bloch. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, 159(15):1929–1951, 2008.

J. Inglada and J. Michel. Qualitative spatial reasoning for high resolution remote sensing image analysis. *Transactions on Geoscience and Remote Sensing*, 47(2):599–612, Feb. 2009.

H. Ip and W. Wong. Detecting perceptually parallel curves: Criteria and force driven optimization. *Computer Vision and Image Understanding*, 68(2):190–208, November 1997.

L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 (11):1254–1259, 2002.

M. Jooyandeh, A. Mohades, and M. Mirzakhah. Uncertain Voronoi diagram. *Information Processing Letters*, 109(13):709 – 712, 2009.

H. Kang and E. Walker. Characterizing and controlling approximation in hierarchical perceptual grouping. *Fuzzy Sets and Systems*, 65(2-3):187–223, 1994.

H.-B. Kang and E. Walker. Perceptual grouping based on fuzzy sets. In *IEEE International Conference on Fuzzy Systems*, pages 651–659, 1992.

J. Keller and X. Wang. A fuzzy rule-based approach to scene description involving spatial relationships. *Computer Vision and Image Understanding*, 80(1):21–41, 2000.

K. Koffka. *Principles of Gestalt Psychology.* Harcourt-Brace, New York, 1935.

S. Konstantopoulos and G. Apostolikas. Fuzzy-DL reasoning over unknown fuzzy degrees. In *Proceedings of the 2007 OTM Confederated international conference on On the move to meaningful internet systems-Volume Part II*, pages 1312–1318. Springer-Verlag, 2007. ISBN 3540768890.

R. Krishnapuram, J. Keller, and Y. Ma. Quantitative analysis of properties and spatial relations of fuzzyimage regions. *IEEE Transactions on Fuzzy Systems*, 1(3):222–233, 1993.

B. Kuipers. Modeling spatial knowledge. *Cognitive Science*, 2(2):129–154, 1978.

B. Kuipers and T. Levitt. Navigation and mapping in large scale space. *AI magazine*, 9 (2):25, 1988.

Y. Kurata and M. Egenhofer. The 9+-Intersection for Topological Relations between a Directed Line Segment and a Region. In *1st International Symposium for Behavioral Monitoring and Interpretation*, pages 62–76, 2007.

B. Landau and R. Jackendoff. What and where in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16(2):217–265, 1993.

F. Le Ber and A. Napoli. Object-based representation and classification of spatial structures and relations. In *14th IEEE International Conference on Tools with Artificial Intelligence ICTAI*, pages 268–275, 2002.

F. Le Ber, L. Mangelinck, and A. Napoli. Design and comparison of lattices of topological relations for spatial representation and reasoning. Technical report, INRIA RR-4321, 2001.

F. Le Ber, J. Lieber, and A. Napoli. Les systèmes à base de connaissances. In J. Akoka and I. Comyn Wattiau, editors, *Encyclopédie de l'informatique et des systèmes d'information*, pages 1197–1208. Vuibert, 2006.

L. Likforman-Sulem and C. Faure. Extracting lines on handwritten documents by perceptual grouping. In C. Faure, P. Keuss, G. Lorette, and A. W. (eds), editors, *Advances in Handwriting and Drawing : a multidisciplinary approach*, pages 21–38, 1994.

J. Liu. A method of spatial reasoning based on qualitative trigonometry. *Artificial Intelligence*, 98(1-2):137–168, 1998.

Y. Liu, Q. H. Guo, and M. Kelly. A framework of region-based spatial relations for non-overlapping features and its application in object based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(4):461 – 475, 2008a.

Y. Liu, Y. Zhanga, and Y. Gaoa. Gnet: A generalized network model and its applications in qualitative spatial reasoning. *Information Sciences*, 178:2163–2175, 2008b.

D. G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31(3):355–395, 1987.

T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):291–308, 2008.

A. Mackworth. Consistency in networks of relations. *Artificial intelligence*, 8(1):99–118, 1977.

N. Maillot. *Ontology Based Object Learning and Recognition*. PhD thesis, University of Nice Sophia Antipolis, 2005.

N. Maillot and M. Thonnat. Ontology based complex object recognition. *Image and Vision Computing*, 26(1):102–113, 2008.

D. Mark and M. Egenhofer. Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing. *Cartography and Geographic Information Systems*, 21(4):195–212, 1994a.

D. Mark and M. Egenhofer. Calibrating the Meanings of Spatial Predicates from Natural Language: Line-Region Relations. In *Proceedings, Spatial Data Handling 1994*, volume 1, pages 538–553, 1994b.

D. Mark and M. Egenhofer. Topology of prototypical spatial relations between lines and regions in English and Spanish. In *Twelfth International Symposium on Computer-Assisted Cartography*, volume 4, pages 245–254, 1995.

D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, New York, 1982.

Y. Mathet. *Etude de l'expression en langue de l'espace et du déplacement : Analyse linguistique, modélisation cognitive et leur expérimentation informatique*. PhD thesis, Université de CAEN, 2000.

P. Matsakis and S. Andréfouet. The fuzzy line between among and surround. In *Proceedings of the IEEE International Conference on Fuzzy Systems, FUZZY-IEEE 2002*, volume 2, pages 1596–1601, 2002.

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

T. Matsuyama. Expert systems for image processing-knowledge-based composition of image analysis processes. In *9th International Conference on Pattern Recognition*, pages 125–133, 1988.

T. Matsuyama and V. Hwang. *SIGMA: A Knowledge-Based Aerial Image Understanding System*. Plenum, 1990.

S. McFeeters. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7):1425–1432, 1996.

J. Metzger, F. Le Ber, and A. Napoli. Modeling and representing structures for analyzing spatial organization in agronomy. In A. de Moor, W. Lex, and B. Ganter, editors, *Conceptual Structures for Knowledge Creation and Communication, 11th International Conference on Conceptual Structures, ICCS 2003*, volume 2746, pages 215–228, Dresden, Germany, july 2003. Springer.

M. Minsky. A Framework for Representing Knowledge. In P. Winston, editor, *The Psychology of Computer Vision*. McGraw-Hill, 1974.

K. Miyajima and A. Ralescu. Spatial organization in 2D segmented images: Representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems*, 65:225–236, 1994a.

K. Miyajima and A. Ralescu. Spatial organization in 2d images. In *Proceedings of the Third IEEE Conference on Fuzzy Systems, FUZZ-IEEE'1994*, pages 100–105, 1994b.

R. Mohan and R. Nevatia. Using perceptual organization to extract 3d structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11):1121–1139, 1989.

R. Mohan and R. Nevatia. Perceptual organization for scene segmentation and description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(6):616–635, 1992.

R. Mohr and T. C. Henderson. Arc and path consistency revisited. *Artificial Intelligence*, 28:225–233, 1986.

R. Möller and T. H. Näth. Implementing probabilistic description logics: An application to image interpretation. In A. G. Cohn, D. C. Hogg, R. Möller, and B. Neumann, editors, *Logic and Probability for Scene Interpretation*, Dagstuhl Seminar Proceedings, pages 1862–4405, Dagstuhl, Germany, 2008. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.

R. Moratz and M. Ragni. Qualitative spatial reasoning about relative position. *Journal of Visual Laguages and Computing*, 19:75–98, 2008.

J. Mota, G. Câmara, M. Escada, O. Bittencourt, L. Fonseca, and L. Vinhas. Case-based reasoning for eliciting the evolution of geospatial objects. In K. S. Hornsby, C. Claramunt, M. Denis, and G. Ligozat, editors, *9th International Conference on Spatial Information Theory, COSIT'09*, volume 5756 of *Lecture Notes in Computer Science*, pages 405–420, Aber Wrac'h, France, Sept 2009. Springer.

B. Motik, B. C. Grau, I. Horrocks, and U. Sattler. Representing structured objects using description graphs. In G. Brewka and J. Lang, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference, KR 2008, Sydney, Australia, September 16-19, 2008*, pages 296–306. AAAI Press, 2008.

M. Naphade, J. Smith, J. Tesic, S. Chang, W. Hsu, L. Kennedy, A. Hauptmann, and J. Curtis. Large-scale concept ontology for multimedia. *IEEE Multimedia*, 13(3):86–91, 2006.

O. Nempont. *Modèles structurels flous et propagation de contraintes pour la segmentation et la reconnaissance d'objets dans les images. Application aux structures normales et pathologiques du cerveau en IRM.* PhD thesis, Ecole Nationale Supérieure des Télécommunications, Mar. 2009.

B. Neumann and R. Moller. On scene interpretation with description logics. *Image and Vision Computing*, 26(1):82–101, Jan. 2008.

M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, 72(2): 107–132, 2007.

S. Paris and F. Durand. A topological approach to hierarchical segmentation using mean shift. In *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR'07.*, pages 1–8. IEEE, 2007. ISBN 1424411807.

A. Perchant. *Morphisme de graphes d'attributs flous pour la reconnaissance structurelle de scènes.* PhD thesis, Ecole Nationale Supérieure des Télécommunications, 2000.

A. Perchant, C. Boeres, I. Bloch, M. Roux, and C. Ribeiro. Model-based Scene Recognition Using Graph Fuzzy Homomorphism Solved by Genetic Algorithm. In *Graph-Based Representations in Pattern Recognition, GbRPR*, pages 61–70, Castle of Haindorf, Austria, 1999.

E. Petrakis, C. Faloutsos, and K. Lin. ImageMap: An image indexing method based on spatial similarity. *IEEE Transactions on Knowledge and Data Engineering*, 14(5): 979–987, 2002.

M. Peura and J. Iivarinen. Efficiency of simple shape descriptors. In *Third International Workshop on Visual Form*, pages 443–451, 1997.

V. Poulain, J. Inglada, M. Spigai, J.-Y. Tourneret, and P. Marthon. Fusion of high resolution optical and SAR images with vector data bases for change detection. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2009*, juillet 2009.

V. Poulain, J. Inglada, M. Spigai, J.-Y. Tourneret, and P. Marthon. High resolution optical and sar image fusion for road database updating. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2010*, Hawaii, juillet 2010.

M. Quillian. Word concepts: a theory and simulation of some basic semantic capabilities. In *Readings in Knowledge Representation*, pages 97–118, Los Altos, CA, 1985.

B. N. R. Moratz and C. Freksa. Qualitative spatial reasoning about relative position : The tradeoff between strong formal properties and succesful reasoning about route graphs. In *Spatial Cognition III*, pages 385–400, Berlin Heidelberg, 2003. Springer-Verlag.

J. Rabin, J. Delon, and Y. Gousseau. Circular earth mover's distance for the comparison of local features. In *International Conference on Pattern Recognition, ICPR'08*, pages 1–4, Tampa, USA, Dec. 2008.

A. L. Ralescu and J. G. Shanahan. Perceptual organization for inferring object boundaries in an image. *Pattern Recognition*, 32(11):1923–1933, Nov. 1999.

D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, pages 165–176, San Mateo, CA, 1992. Morgan Kaufmann.

J. Renz. *Qualitative Spatial Reasoning with Topological Information*, volume 2293 of *Lecture Notes in Computer Science*. Springer, 2002. ISBN 3-540-43346-5.

G. Retz-Schmidt. Various views on spatial prepositions. *AI Magazine, Summer 1988*, pages 95–105, 1988.

A. Rosenfeld. Fuzzy digital topology. *Information and Control*, 40(1):76–87, 1979.

A. Rosenfeld and R. Klette. Degree of adjacency or surroundedness. *Pattern Recognition*, 18(2):169–177, 1985.

A. Rosenfeld, R. Hummel, and S. Zucker. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man and Cybernetics*, 6(6):420–433, 1976.

C. Rosse, J. Mejino, et al. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *Journal of biomedical informatics*, 36(6):478–500, 2003.

F. Rossi, P. Van Beek, and T. Walsh. *Handbook of constraint programming*. Elsevier Science Ltd, 2006.

J. Rouco, M. Penas, M. Penedo, M. Ortega, and C. Alonso-Montes. Artificial Intelligence Techniques and Recognition-Certainty Measure of Pairwise Line Segment Perceptual Relations Using Fuzzy Logic. In *Iberoamerican Congress on Pattern Recognition*, volume 4756, pages 477–486. Springer, 2007.

N. Roussopoulos, C. Faloutsos, and T. Sellis. An efficient pictorial database system for psql. *IEEE Transactions on Software Engineering*, 14(5):639 –650, may. 1988.

C. Saathoff. Constraint reasoning for region-based image labelling. In *IET International Conference on Visual Information Engineering, 2006. VIE 2006*, pages 138–143, 2006.

C. Saathoff and S. Staab. Exploiting spatial context in image region labelling using fuzzy constraint reasoning. In *WIAMIS '08: Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 16–19, Washington, DC, USA, 2008.

S. Schockaert, M. D. Cock, C. Cornelis, and E. E. Kerre. Fuzzy region connection calculus: Representing vague topological information. *International Journal of Approximate Reasoning*, 48:314–331, 2008.

C. Schulte and P. J. Stuckey. Speeding up constraint propagation. In M. Wallace, editor, *Tenth International Conference on Principles and Practice of Constraint Programming*, volume 3258 of *Lecture Notes in Computer Science*, pages 619–633, Toronto, Canada, Sept. 2004. Springer-Verlag.

S. Schulz and U. Hahn. Mereotopological reasoning about parts and (w)holes in bio-ontologies. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, pages 210–221. ACM, 2001.

A. Schwering. Evaluation of a semantic similarity measure for natural language spatial relations. In *Conference on Spatial Information Theory (COSIT07)*, number 4736 in Lecture Notes in Computer Science, pages 116–132, Berlin, 2007. Springer.

A. Shariff, M. Egenhofer, and D. Mark. Natural-Language Spatial Relations Between Linear and Areal Objects: The Topology and Metric of English-Language Terms. *International Journal of Geographical Information Science*, 12(3):215–245, 1998.

D. Shen, W. Wong, and H. Ip. Affine-invariant image retrieval by correspondence matching of shapes. *Image and Vision Computing*, 17(7):489–499, 1999.

J. F. Sowa. *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, MA, 1984.

A. Sowmya and J. Trinder. Modelling and representation issues in automated feature extraction from aerial and satellite images. *ISPRS journal of photogrammetry and remote sensing*, 55(1):34–47, 2000.

S. Steiniger and R. Weibel. Relations among map objects in cartographic generalization. *Cartography and Geographic Information Science*, 34(3):175–197, 2007.

U. Straccia. Answering vague queries in fuzzy DL-Lite. In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems,(IPMU-06)*, pages 2238–2245. Citeseer, 2006.

C. Takemura, R. Cesar, and I. Bloch. Fuzzy modeling and evaluation of the spatial relation along. In *Progress in Pattern Recognition, Image Analysis and Applications*, volume 3773 of *Lecture Notes in Computer Science*, pages 837–848. Springer Berlin / Heidelberg, 2005.

L. Talmy. How language structures space. In *Spatial Orientation: Theory, Research, and Application*, pages 225–282. Plenum Press, New York, 1983.

S. Tellex and D. Roy. Spatial routines for a simulated speech-controlled vehicle. In *Proceedings of the 1st ACM SIGCH/SIGART conference on Human-robot interaction*, pages 156–163. ACM New York, NY, USA, 2006.

J. M. Tenenbaum and H. G. Barrow. Experiments in interpretation-guided segmentation. *Artificial Intelligence*, 8(3):241 – 274, 1977.

R. Thomopoulos, P. Buche, and O. Haemmerlé. Different kinds of comparisons between fuzzy conceptual graphs. *Conceptual Structures for Knowledge Creation and Communication*, pages 54–68, 2003a.

R. Thomopoulos, P. Buche, and O. Haemmerlé. Representation of weakly structured imprecise data for fuzzy querying. *Fuzzy Sets and Systems*, 140(1):111–128, 2003b.

S. Toh. Extracting viewpoint invariance relations using fuzzy sets. In *Proceedings of the Intelligent Vehicles' 92 Symposium*, pages 325–329, 1992.

M.-C. Vanegas, I. Bloch, and J. Inglada. Fuzzy Spatial Relations for High Resolution Rmote Sensing Image Analysis: The Case of "To Go Across". In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2009*, Cape Town, July 2009a.

M.-C. Vanegas, I. Bloch, H. Maître, and J. Inglada. Approximate Parallelism Between Fuzzy Objects: Some Definitions. In *International Workshop on Fuzzy Logic and Applications WILF*, volume 5571, pages 12–19, Palermo, June 2009b.

M.-C. Vanegas, I. Bloch, and J. Inglada. Detection of aligned objects for high resolution image understanding. In *IEEE IGARSS 2010*, Honolulu, Hawai, USA, July 2010a.

M.-C. Vanegas, I. Bloch, and J. Inglada. Searching aligned groups of objects with fuzzy criteria. In *International Conference on Information Processing and management of Uncertainty in Knowledge-Based Systems, IPMU 2010*, pages 605–613, Dortmund, Germany, June 2010b.

R. Wallace and E. Freuder. Ordering heuristics for arc consistency algorithms. In *Proceedings of the Biennial Conference-Canadian Society for Computational Studies of Intelligence*, pages 163–170. Canadian Information Processing Society, 1992.

D. Waltz. Understanding line drawings of scenes with shadows. *The psychology of computer vision*, pages 19–91, 1975.

M. Wertheimer. Laws of organization in perceptual forms. In W. H. Ellis, editor, *Source Book of Gestalt Psychology*. Routledge and Kegan Paul, London, 1938.

W. Woods. Semantics and quantification in natural language question answering. *Advances in computers*, 17(3), 1978.

R. Zwick, E. Carlstein, and D. V. Budescu. Measures of similarity among fuzzy concepts: A comparative analysis. *International Journal of Approximate Reasoning*, 1:221–242, 1987.