Spatial Reasoning and Multiscale Segmentation for Object Recognition in HR Optical Remote Sensing Images

Jordi Inglada† and Julien Michel*
†CNES, 18 avenue E. Belin, BPI 1219, 31401 Toulouse Cedex 9, France
*CS, Parc de la Plaine, Rue Brindejonc des Moulinais, BP 5872 31506 Toulouse Cedex 5, France
jordi.inglada@cnes.fr, julien.michel@c-s.fr

Abstract— **High resolution remote sensing images allow us to access new kinds of information. Classical techniques for image analysis, such as pixel-based classifications or region-based segmentations, do not allow to fully exploit the richness of this kind of images. Indeed, for many applications, we are interested in complex objects which can only be identified and analysed by studying the relationships between the elementary objects which compose them. In this paper, the use of a spatial reasoning technique called Region Connection Calculus for the analysis of high resolution remote sensing images will be presented.**

I. Introduction

Because of the diversity of sensors and the increase of their spatial resolution and repetitivity, the automatic analysis of images is a crucial issue in the remote sensing field. It is therefore needed to design and implement new image analysis techniques which are able to perform complex processing in an efficient manner. In this context, much work has been done for the automatic information extraction from remote sensing images aiming to provide a set of features and descriptors which are compact and parsimonious in order to feed them into learning systems. This type of features come often from low-level processing and are not able to capture the richness and complexity of HR images. Indeed, the improvement of the spatial resolution makes that the objects of interest are not any more limited to several pixels, but they are represented by large areas containing sub-objects. This kind of objects can not be efficiently described by textures, edges, etc. It is therefore useful to use techniques which are able to deal with higher levels of abstraction for the representation and the manipulation of the information contained in the images. In this work, we will use spatial reasoning techniques in order to describe complex objects. The spatial arrangement of the objects in the images gives a crucial information for the recognition and interpretation tasks. This is particularly important when the objects are located in complex environments, which is often the case for high resolution images (meter and sub-meter resolutions). Spatial reasoning can be defined as the set of techniques which allow to manipulate the descriptions of physical objects by taking into account their mutual interactions using the shapes, sizes, positions and orientations. We have chosen to use the well known Region Connection Calculus, RCC-8 [1] in order to describe the relationships between the regions of an image.

The regions used for the spatial reasoning are obtained with a multiscale segmentation approach based on the morphological pyramid [2]. The mathematical morphology approach has been selected because, in contrast to linear multiscale approaches, it allows to select objects in terms of their size. These has been shown to be interesting in meter and sub-meter resolution images for detecting vehicles, buildings, etc. Our algorithm is able to produce regions superimposing across the scales. This allows to exploit the full extent of the RCC-8 relationships set. The regions and their relationships are represented by means of an Attributed Relational Graph, ARG, where the nodes represent the regions and the arcs represent the RCC-8 relationships. Finally, object detection and recognition can be seen as a graph matching problem for which efficient algorithms can be implemented. In these context, an algorithm which uses a greedy search combined with a graph metric which is able to use all the information contained in the ARG has been implemented. In this paper, the theoretical basis of the approach and examples of application to real images will be presented.

II. Region Connection Calculus

Region Connection Calculus [1], is based on the notion of connection between pairs of regions in space. This connection is described by one among a set of possible relationships. One can therefore derive different RCC systems depending on the number and

1-4244-1212-9/07/$25.00 ©2007 IEEE.

the nature of the relationships accounted for. One of these systems, RCC8, is especially interesting, since it is made of exhaustive and disjoint relationships. These properties simplify the reasoning techniques. As shown in figure 1, the 8 RCC8 relationships go from disconnection, DC, to equivalence, EQ, through all intermediate possibilities by taking into account tangecy (EC: external connection, PO: partial overlap, TPP: tangential proper part, NTPP: non-tangential proper part). It is important to note that all these relationships are symmetrical except for TPP and NTPP which are anti-symmetrical. This explains the additional relationships TPPi and NTPPi and their number (8 instead of 6). If one does not take into account tangency, the RCC5 system is obtained.

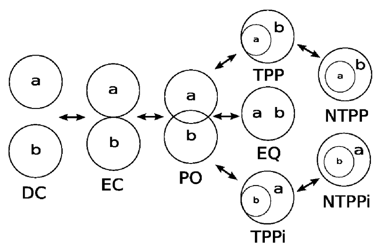

Fig. 1. Transitions between RCC8 relationships

After defining the RCC8 system, one can wonder how to compose relationships: knowing that region (a) is linked to region (b) by relationship R_1 and that region (b) is linked to region (c) by relationship R_2, which is the relationship between regions (a) and (c) ? Even if in some cases one can find a unique answer to this question, most of the cases do not allow this. For instance, if the two known relationships are DC, no information can be inferred about the composition.This is a limitation of the RCC8 system. However the composition table of RCC8 (table I) will help to speed up computations.

III. Implementation

A. Elementary computation

1) Principles: The first step for assessing the usefulness of RCC8 for image analysis, was to implement a technique for the efficient computation of the relationship between two regions of space. The first tool we will use is the 9-intersections matrix, which is a 3×3 matrix of 0's (representing the empty set) and 1's (representing non-empty sets) representing the intersections between the exterior, the boundary and the interior of region A with the interior, exterior and boundary of region B.

$$\mathbf{M} = \begin{pmatrix} card(\breve{A},\breve{B}) & card(\breve{A},\dot{B}) & card(\breve{A},\bar{B}) \\ card(\dot{A},\breve{B}) & card(\dot{A},\dot{B}) & card(\dot{A},\bar{B}) \\ card(\bar{A},\breve{B}) & card(\bar{A},\dot{B}) & card(\bar{A},\bar{B}) \end{pmatrix}$$

Where $card$ is a function which is equal to 1 if the set is not empty and 0 otherwise, $\breve{A}$ is the interior of region A, $\dot{A}$ its boundary and $\bar{A}$ its exterior. Among the 2^9 possible matrices, only 8 are physically plausible and they correspond to the 8 RCC8 relationships. Being able to compute the 9-intersections matrix for a pair of regions of space means knowing the RCC8 relationship between them.

The implementation is as follows. For each region, we have a binary mask coming from the segmentation step. The exterior of the region is obtained by inverting the mask. The interior of the region is obtained by applying a morphological erosion of size 1 pixel. The boundary is obtained by subtracting the interior to the mask. Once these 3 regions are obtained for each one of the regions of interest, we obtain the intersections by applying a binary addition of those.

2) Optimisation: This method has a main drawback: its complexity. Indeed, morphological operations are computing expensive. This complexity is also dependent on the image size. This is particularly annoying, since remote sensing images are very large, while the regions of interest may be small. Therefore, using masks which have the same size as the image can very much slow the process. In order to reduce the computation time, 3 optimisations have been implemented. The first optimisation concerns the cardinal computation. Since what interests us is the fact whether this cardinal of the intersections is zero or not, we implement a lazy operator, which goes though the image and stops when it finds the first non zero pixel.

The following optimisation concentrates on the computation of the elements of the 9-intersections matrix. We have seen that only 8 among the 2^9 possible matrices are physically plausible. This means that the components of the matrix are redundant. Therefore, we have built a binary decision tree where we examine only one component at each step. This tree allows us to find the RCC8 relationship by examining at most 4 components of the matrix.

Despite of these two optimisations, the computation of the relationships is still very time consuming. Since the most common relationship between the regions of the image is disconnection, a particular optimisation for this case has been implemented. We exploit the fact that most of the disconnected regions are also far away from each other. That means that their bounding boxes will also be disconnected. The bounding box

$R_1 \setminus R_2$	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ
DC	no info	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, TPPi, NTPPi	DC, EC, PO, TPP, TPPi, EQ	DC, EC, PO, TPP, NTPP	EC, PO, TPP, NTPP	PO, TPP, NTPP	DC, EC	DC	EC
PO	DC, EC, PO, TPPi, NTPPi	DC, EC, PO, TPPi, NTPPi	no info	PO, TPP, NTPP	PO, TPP, NTPP	DC, EC, PO, TPPi, NTPPi	DC, EC, PO, TPPi, NTPPi	PO
TPP	DC	DC, EC	DC, EC, PO, TPP, NTPP	TPP, NTPP	NTPP	DC, EC, PO, TPP, TPPi, EQ	DC, EC, PO, TPPi, NTPPi	TPP
NTPP	DC	DC	DC, EC, PO, TPP, NTPP	NTPP	NTPP	DC, EC, PO, TPP, NTPP	no info	NTPP
TPPi	DC, EC, PO, TPPi, NTPPi	EC, PO, TPPi, NTPPi	PO, TPPi, NTPPi	PO, TPP, TPPi, EQ	PO, TPP, NTPP	TPPi, NTPPi	NTPPi	TPPi
NTPPi	DC, EC, PO, TPPi, NTPPi	PO, TPPi, NTPPi	PO, TPPi, NTPPi	PO, TPPi, NTPPi	PO, TPP, NTPP, TPPi, NTPPi, EQ	NTPPi	NTPPi	NTPPi
EQ	DC	EC	PO	TPP	NTPP	TPPi	NTPPi	EQ

TABLE I

RCC8 RELATIONSHIP COMPOSITION TABLE

computation is much faster than the morphological operations, so this way of evaluating the DC relationship, which is the first one in the decision tree, will allow an important acceleration of the evaluation of the relationships.

B. *Analysis for a set of regions*

The application of the computation of the relationship between a pair of regions to all the pairs of regions in the image may be very time consuming. In order to reduce the number of computed relationships we will take into account the symmetry properties of the relationships. This allows to reduce by a factor of 2 the number of computed relationships. We also use the fact that the multiscale segmentation gives disconnected regions for a given level of the pyramid. Using the composition table I, further optimisations can be implemented. Indeed, this table allows us to unambiguously infer the relationship in 27 cases out of 64. In 16 other cases, the composition table allows to jump the first level of the binary decision tree. In other 7 cases, the composition table allows to decide the third level of the tree. In order to fully exploit this information, before computing the relationship between a region (a) and a region (c), we will look for an intermediate region (b) for which the relationships which link it to (a) and (c) have already been computed. If such a region exists, 2 cases may appear. First, the knowledge of the already computed relationships allows to jump the first or the third level of the decision tree: this information is stored and the examination of the intermediate regions found goes on. Second, the knowledge of the already computed relationships allows to unambiguously determine the new relationship: the computation is finished. If at the end of the examination of the intermediate regions it

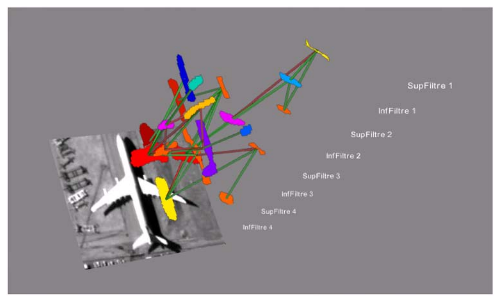

Fig. 2. RCC-8 graph of a plane

was impossible to determine the needed relationship, a computation using the optimisations of the previous section is done by using all the information stored during this pre-computation step. An example of the type of RCC8 graph which can be obtained with this computation is presented in figure 2 for the case of a plane in a sub-meter resolution image.

IV. GRAPH MATCHING

In order to measure a purely structural similarity between two graphs, many measures have been proposed. Most of these measures are based on the research of the largest common sub-graph between the two compared graphs. Sorlin and Solnon [3] defined a generic measure which unifies several approaches. It is based on an evaluation of a function which depends on the two compared graphs. This function depends on the common characteristics of the two graphs (nodes, arcs and labels associated to both of them) and it may be tuned using two functions f and g which determine the nature of the measure. The following definitions are useful for the remainder of the presentation:

Fig. 3. Object detection by RCC-8 graph matching

Definition 1 (Graph match): Let G and G' be two graphs defined respectively over V and V'. A **graph match** is a subsample of $m \subseteq V \times V'$ containing the set of pairs (v, v') so that the nodes v and v' are matched.

Definition 2 (Common descriptors): We call **common descriptors** of a graph match m the set $descr(G) \sqcap_m descr(G')$, where the operator $\sqcap_m$ stands for the set of all nodes, node's labels, arcs and arc's labels which are common to G and G' within a match m.

Definition 3 (Split nodes): We call **Split nodes** or **splits(m)** of a match m of two graphs G and G' the set of nodes of G matched to several nodes of G' and the nodes of G' matched to several nodes of G.

These definitions allow us to introduce the following generic similarity measure:

Definition 4 (Similarity with respect to a match): The similarity between two graphs G and G' with respect to a match m is defined by:

$$sim_m(G, G') = \frac{f(descr(G) \sqcap_m descr(G')) - g(splits(m))}{f(descr(G) \cup descr(G'))} \tag{1}$$

where f and g are two application dependent functions.

Definition 5 (Similarity between two graphs): The similarity between two graphs G and G' is defined by:

$$sim_m(G, G') = \max_{m \subseteq V \times V'} \frac{f(descr(G) \sqcap_m descr(G')) - g(splits(m))}{f(descr(G) \cup descr(G'))} \tag{2}$$

where f and g are two application dependent functions.

In our case, f and g will simply be the cardinals of the sets, which correspond to the isomorphism of the graphs.

V. Results and discussion

Using the multiscale segmentation presented in [2] followed by the RCC8 graph computation presented in section III and the graph matching technique introduced in section IV, an object recognition system can be implemented. The first step will consist in building a template RCC8 graph to match. This can be done by computing the RCC8 graph for an image patch containing an example of the type of object to be recognized, as it is shown in figure 2. The following step is to compute the graph of the scene to be analysed. Finally, the computation of the score function (eq. 2) between the template graph and the sub-graphs extracted from the analysed scene, allows to detect likely matches. An example of the kind of result which can be obtained is presented in figure 3.

The system presented in this paper has been implemented within the ORFEO Toolbox, a free software library for remote sensing image analysis developed by CNES. It is available for download at `http://smsc.cnes.fr/PLEIADES/lien3_vm.htm`.

References

[1] N. M. Gotts, J. M. Gooday, and A. G. Cohn, "A connection based approach to common-sense topological description and reasoning," *Topology for philosophers*, vol. 79, no. 1, pp. 51–75, 1996.

[2] F. Laporterie, G. Flouzat, and O. Amram, "The morphological pyramid and its applications to remote sensing : multiresolution data analysis and features extraction," *Image Anal Stereol*, vol. 21, pp. 49–53, 2002.

[3] S. Sorlin and C. Solnon, "Reactive Tabu Search for Measuring Graph Similarity," in *Graph-Based Representations in Pattern Recognition. Lecture Notes in Computer Science*, pp. 172–182, Springer Berlin / Heidelberg, 2005.

1-4244-1212-9/07/$25.00 ©2007 IEEE.