From Images to Bodies: Modelling and Exploiting Spatial Occlusion and Motion Parallax

David Randell, Mark Witkowski and Murray Shanahan
Dept. of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
Exhibition Road
London SW7 2BT
United Kingdom
{d.randell, m.witkowski, m.shanahan}@ic.ac.uk

Abstract

This paper describes the Region Occlusion Calculus (*ROC-20*), that can be used to model spatial occlusion and the effects of motion parallax of arbitrary shaped objects. *ROC-20* assumes the region based ontology of *RCC-8* and extends Galton's *Lines of Sight* Calculus by allowing concave shaped objects into the modelled domain. This extension is used to describe the effects of mutually occluding bodies. The inclusion of van Benthem's axiomatisation of comparative nearness facilitates reasoning about relative distances between occluding bodies. Further, an envisionment table is developed to model sequences of occlusion events enabling reasoning about objects and their images formed in a changing visual field.

1 Introduction

Spatial occlusion (or interposition) arises when one object obscures the view of another. Spatial occlusion is one of several visual cues we exploit to build up our awareness of three-dimensional form and distance. Another is motion parallax, whereby a change in viewpoint causes relative displacements of objects at different distances in the visual field [Braddick and Atkinson, 1982]. Occlusion events help us determine where an object's boundary lies, or infer why an object cannot be seen, and what we need to do in order to render it visible.

For example, consider two objects A and B in an agent's visual field. Suppose the agent moves to its left, while keeping these objects in sight. If object A passes across B, or, when moving toward A, B becomes completely obscured, the agent can infer that A is in front of B. Similarly, if, when moving to the right, no relative change arises, the agent may infer that A and B are far away, or close by and possibly moving in the same direction as itself. Conversely, if A, when visible, always appears to be subtended by B, the agent may infer that A and B are physically connected. In each case, occlusion events and motion parallax are being used to derive an objective model of the world from a naturally restricted viewpoint (Figure 1).

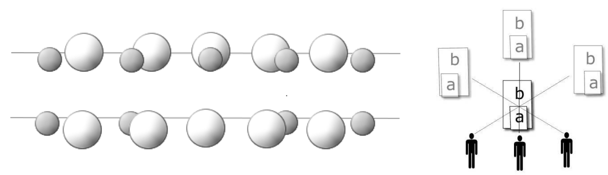


Figure 1: Spatial occlusion at work. Assuming a fixed viewpoint, in the two sequences shown on the left, the smaller ball passes in front of the larger one (top sequence) and behind it (bottom sequence). On the right, occlusion events arise with a change in viewpoint.

While visual occlusion remains a topic of some interest in the machine vision literature [e.g., Plantinga and Dyer, 1990; Geiger, *et al*., 1995], an opportunity arises to investigate occlusion within the Qualitative Spatial Reasoning (QSR) domain. For example, Galton's [1994] *Lines of Sight* calculus outlines a theory of occlusion for modelling convex bodies using a discrete set of 14 occlusion relations. It is natural to take a topological approach to modelling occlusion, since occlusion events are very general, and apply to all objects irrespective of their size, shape and function. Petrov and Kuzmin [1996] provide an axiomatisation of spatial occlusion founded on a point-based ontology.

Randell, *et al.* [1992] develop a mereo-topological theory, *RCC-8*, used to describe spatial relationships between regions based on the primitive relation of connection. Cui, *et al.* [1992] use *RCC-8* to develop a qualitative simulation program to model physical processes by specifying direct topological transitions between these relations over time. Their work is one example of using qualitative spatial representations to model continuous change [Cohn, 1997]. *ROC-20* extends *RCC-8* to reason about relative distances between bodies from occlusion events, and transitions between occlusion events to model the effects of motion parallax from both object motion and changing viewpoints.

2 The Formal Theory

Our universe of discourse includes bodies, regions and points, all forming pairwise disjoint sets. In terms of interpretation, bodies denote physical objects, while regions split into two further disjoint sets that denote either three-dimensional volumes (typically the spaces occupied by bodies) or two-dimensional regions (typically projected *images* of bodies as seen from some viewpoint).

For the purposes of this paper, a set of sorts and a sorted logic are assumed. Within the sorted logic, possible values of variables in formulae are derived implicitly from the specified sort of the argument position in which it appears, allowing *ad hoc* polymorphic functions and predicates to be handled.

The notation and conventions used throughout this paper is as follows: ***type*** $a(\tau 1,.., \tau n): \tau n+1$ means function symbol a is well sorted when its argument sorts are $\tau 1,.., \tau n$ with $\tau n+1$ as the result sort, and ***type*** $a(\tau 1,.., \tau n)$ means predicate a is well sorted when defined on argument sorts $\tau 1,.., \tau n$. Axioms, definitions and theorems are respectively indicated in the text as follows: *(A1,..,An), (D1,..,Dn), and (T1,..,Tn).* Where axiom/definitional schema are used, the numbering in the parentheses reflects the number of object-level axioms and definitions generated, e.g. *(A10-A15)* would indicate that six axioms are defined.

2.1 RCC-8

The mereo-topological theory *RCC-8* [Randell, *et al.*, 1992] is embedded into *ROC-20*. As with *RCC-8*, the same primitive dyadic relation *C/2* is used: '*C(x,y)*' is read as "*x* is connected with *y*" and is interpreted to mean that the topological closures of regions *x* and *y* share a point in common. All the relations defined in *RCC-8* are used, and all carry their usual readings: *DC/2* (disconnected), *P/2* (part), *EQ/2* (equal), *O/2* (overlaps), *DR/2* (discrete) *PO/2* (partial overlap), *EC/2* (external connection), *PP/2* (proper part), *TPP/2* (tangential proper part), *NTPP/2* (non-tangential proper part). *PI/2*, *PPI/2*, *TPPI/2* and *NTPI/2* are the inverse relations for *P/2*, *PP/2*, *TPP/2* and *NTPP/2*, respectively. Of these relations, eight are provably *Jointly Exhaustive and Pairwise Disjoint* (JEPD) and can be singled out for reasoning about state-state topological changes [Cui, *et al.*, 1992]. For brevity this set of relations is referred to as $JEPD^{RCC\text{-}8}$.

Axioms for *C/2* and definitions for the dyadic relations of *RCC-8* are as follows:

(A1) $\forall x\, C(x,x)$
(A2) $\forall x \forall y\, [C(x,y) \rightarrow C(y,x)]$

(D1) $DC(x,y) \equiv_{def.} \neg C(x,y)$
(D2) $P(x,y) \equiv_{def.} \forall z[C(z,x) \rightarrow C(z,y)]$
(D3) $EQ(x,y) \equiv_{def.} P(x,y) \,\&\, P(y,x)$
(D4) $O(x,y) \equiv_{def.} \exists z[P(z,x) \,\&\, P(z,y)]$
(D5) $DR(x,y) \equiv_{def.} \neg O(x,y)$
(D6) $PO(x,y) \equiv_{def.} O(x,y) \,\&\, \neg P(x,y) \,\&\, \neg P(y,x)$
(D7) $EC(x,y) \equiv_{def.} C(x,y) \,\&\, \neg O(x,y)$
(D8) $PP(x,y) \equiv_{def.} P(x,y) \,\&\, \neg P(y,x)$
(D9) $TPP(x,y) \equiv_{def.} PP(x,y) \,\&\, \exists z[EC(z,x) \,\&\, EC(z,y)]$
(D10) $NTPP(x,y) \equiv_{def.} PP(x,y) \,\&\, \neg\exists z[EC(z,x)) \,\&\, EC(z,y)]$
(D11) $PI(x,y) \equiv_{def.} P(y,x)$
(D12) $PPI(x,y) \equiv_{def.} PP(y,x)$
(D13) $TPPI(x,y) \equiv_{def.} TPP(y,x)$
(D14) $NTPPI(x,y) \equiv_{def.} NTPP(y,x)$

type *Φ(Region,Region); where* $\Phi \in \{C,DC,P,EQ,DR,PO,EC,PP,TPP,NTPP,PI,PPI,TPPI,NTPPI\}$

Not reproduced here, but assumed, is an axiom in *RCC-8* that guarantees every region has a nontangential proper part *(A3)*, and a set of axioms *(A4-A9)* introducing Boolean functions for the sum, complement, product, difference of regions, and the universal spatial region, and an axiom that introduces the sort *Null* enabling partial functions to be handled – see [Randell, *et al.*, 1992].

2.2 Mapping Functions and Axioms

ROC-20 uses the set of dyadic relations from *RCC-8* to model the spatial relationship between bodies, volumes, and images. The distinction between bodies and regions is maintained by introducing two functions: '*region(x)*' read as "the region occupied by *x*" and '*image(x,v)*' read as "the image of *x* with respect to viewpoint *v*". The function: *region/1*, maps a body to the volume of space its occupies, and *image/2* maps a body and a viewpoint to its image; i.e. the region defined by the set of projected half-lines originating at the viewpoint and intersecting the body, so forming part of the surface of a sphere of infinite radius centred on the viewpoint. A set of axioms incorporating these functions are defined by the following axiom schema[1]:

(A10-A15) $\forall x \forall y\, [\Phi(region(x),region(y)) \rightarrow \forall v\, [\Phi(image(x,v),image(y,v))]]$

type *region(Body):Region*[2]
type *image(Body,Point):Region*
type *Φ(Region,Region) where* $\Phi \in \{C,O,P,PP,NTPP,EQ\}$

Not all of the defined *RCC-8* relations are shown. For example, given *DC(region(a),region(b))* all image relationships between the *a* and *b* are *possible* depending on the shape of the objects and the viewpoint assumed. This shows that these axioms function as a set of *spatial constraints* between bodies, a given viewpoint, and their corresponding images. This point is re-visited in section 6

[1] Although not developed here, the distinction made between bodies and regions enables one to define the notion of free space and model spatial occupancy – see [Shanahan, 1996].

[2] Sortal declarations given here are not as restricted as they could be, for example we could declare: ***type*** *region(Body):3DRegion*, and ***type*** *image(Body,Point):2DRegion***,** where *2DRegion* and *3DRegion* are (disjoint) subsorts of the sort *Region.*

below, where a change in viewpoint, or a change in the relative positions of bodies with respect to a viewpoint, is discussed.

2.3 Occlusion Defined

A second primitive relation: '*TotallyOccludes(x,y,v)*', read as "*x* totally occludes *y* with respect to viewpoint *v*", is now added, and is axiomatised to be irreflexive and transitive (and is, by implication, asymmetric):

$$(A16)\quad \forall x \forall v\ \neg TotallyOccludes(x,x,v)$$

$$(A17)\quad \forall x \forall y \forall z \forall v\ [[TotallyOccludes(x,y,v)\ \&\ TotallyOccludes(y,z,v)] \rightarrow TotallyOccludes(x,z,v)]$$

type *TotallyOccludes(Body,Body,Point)*

The intended *geometric* meaning of total occlusion is as follows. Let *line(p1,p2,p3)* mean that points *p1*, *p2* and *p3* fall on a straight line with *p2* strictly between *p1* and *p3*. Then, *x* totally occludes *y* from *v iff* for every point *p* in y, there exists a point *q* in *x* such that *line(v,q,p)*, and there are no points *p′* in *y*, and *q′* in *x*, such that *line(v,p′,q′)*. Given the transitivity of total occlusion, an object *x* can totally occlude an object *y* even if *x* itself is totally occluded by another object.

Several axioms are now introduced to embed *RCC-8* into this theory:

$$(A18)\quad \forall x \forall y \forall z \forall v\ [[TotallyOccludes(x,y,v)\ \&\ P(region(z),region(y))] \rightarrow TotallyOccludes(x,z,v)]$$

i.e. if *x* totally occludes *y*, *x* totally occludes any part of *y*.

$$(A19)\quad \forall x \forall y \forall v\ [TotallyOccludes(x,y,v) \rightarrow \forall z[P(region(z),region(y))] \rightarrow \neg TotallyOccludes(z,x,v)]$$

i.e. if *x* totally occludes *y* no part of *y* totally occludes *x*.

$$(A20)\quad \forall x \forall y \forall v\ [TotallyOccludes(x,y,v) \rightarrow \forall z[[P(region(z),region(x))\ \&\ P(region(u),region(y))] \rightarrow \neg TotallyOccludes(u,z,v)]]$$

i.e. if *x* totally occludes *y* no part of *y* totally occludes part of *x*.

This latter axiom excludes cases where the occluding body has parts that wrap 'behind' the occluding object. That is to say, while some nested bodies satisfy this relation, not all do, as in the case where, for example, a body is totally enveloped by another. This particular model is an example of *mutual occlusion*, which is defined below in definition (*D17*).

$$(A21)\quad \forall x \forall v \exists y \exists z[P(region(y),region(x))\ \&\ P(region(z),region(x))\ \&\ TotallyOccludes(y,z,v)]$$

i.e. every *x* has a part that totally occludes another part of *x*. This axiom guarantees that bodies have 'depth'.

$$(A22)\quad \forall x \forall y \forall v\ [TotallyOccludes(x,y,v) \rightarrow P(image(y,v),image(x,v))]$$

i.e. if *x* totally occludes *y*, the image of *x* subtends the image of *y*. Note that (A22) is not a biconditional because the *P/2* relation does not take account of relative distance, a topic to be considered shortly.

By separating out volumes and images, two non-identical bodies having identical images (as in the case where one body exactly occludes another) can be modelled without inconsistency. Spatial identity in terms of co-location still applies, but is restricted to the dimensionality of the regions being modelled.

Next, the relation of occlusion is weakened to include, for example, partial occlusion: '*Occludes(x,y,v)*' is read as "*x* occludes *y* from viewpoint *v*":

$$(D15)\quad Occludes(x,y,v) \equiv_{def.} \exists z \exists u[P(region(z),region(x))\ \&\ P(region(u),region(y))\ \&\ TotallyOccludes(z,u,v)]$$

type *Occludes(Body,Body,Point)*

i.e. *x* occludes *y* if a part of *x* totally occludes a part of *y*.

Total occlusion between two objects implies occlusion, which in turn implies region overlap between their corresponding images:

$$(T1)\quad \forall x \forall y \forall v[TotallyOccludes(x,y,v) \rightarrow Occludes(x,y,v)]$$

$$(T2)\quad \forall x \forall y \forall v\ [Occludes(x,y,v) \rightarrow O(image(x,v),image(y,v))]$$

Occludes/3 is *non-symmetrical*. By contrast, the *O/2* relation in *RCC-8* is symmetrical, which renders it unsuitable for modelling occlusion relationships. Hence the need to augment *RCC-8* with an additional primitive relation.

Other more specific occlusion relations may now be defined: partial, and mutual occlusion. An example of mutual occlusion is two interlinked rings. These relations will then be finessed further by combining them with the set of *RCC-8* relations:

$$(D16)\quad PartiallyOccludes(x,y,v) \equiv_{def.} Occludes(x,y,v)\ \&\ \neg TotallyOccludes(x,y,v)\ \&\ \neg Occludes(y,x,v)$$

type *PartiallyOccludes(Body,Body,Point)*

i.e. *x* occludes (but does not totally occlude) *y*, but *y* does not occlude *x*.

(D17) $MutuallyOccludes(x,y,v) \equiv_{def.}$
$Occludes(x,y,v)\ \&\ Occludes(y,x,v)$

type *MutuallyOccludes(Body,Body,Point)*

i.e. *x* and *y* occlude each other.

For completeness (not listed here) inverse relations for *Occludes/3*, *TotallyOccludes/3* and *PartiallyOccludes/3* are defined (*D18-D20)*; leaving the null case: *NonOccludes/3*, where no occlusion arises:

(D21) $NonOccludes(x,y,v) \equiv_{def.}$
$\neg Occludes(x,y,v)\ \&\ \neg Occludes(y,x,v)$

type *NonOccludes(Body,Body,Point)*

The six relations: *NonOccludes/3*, *MutuallyOccludes/3*; and *TotallyOccludes/3*, *PartiallyOccludes/3*, and their inverses are pairwise disjoint.

Finally, these new occlusion relations must be mapped to their *RCC* analogues:

(A23) $\forall x \forall y \forall v[NonOccludes(x,y,v) \rightarrow DR(image(x,v),image(y,v))]$

(A24) $\forall x \forall y \forall v[PartiallyOccludes(x,y,v) \rightarrow [PO(image(x,v),image(y,v)) \vee PP(image(x,v),image(y,v))]]$

(A25) $\forall x \forall y \forall v[MutuallyOccludes(x,y,v) \rightarrow [PO(image(x,v),image(y,v)) \vee P(image(x,v),image(y,v)) \vee PI(image(x,v),image(y,v))]]$

2.4 Finessing the Occlusion Relations

Although a variety of occlusion relations have now been defined, they are still very general, as no spatial relation stronger than *P/2* from *RCC-8* is used. Total occlusion, for example, covers three cases: (i) where the *image* of the occluded body is a tangential proper part of that of the occluding body, (ii) where it is a nontangential proper part, or (iii) the images are identical because one body exactly occludes the other. By refining the existing set of occlusion relations in this manner, a total set of 20 JEPD relations become definable. These are generated using the following definitional schemas:

(D22-D33) $\Phi\Psi(x,y,v) \equiv_{def.}$
$\Phi(x,y,v)\ \&\ \Psi(image(x,v),image(y,v))$

(D34-D41) $X\Psi^{-1}(x,y,v) \equiv_{def.}$
$X(y,x,v)\ \&\ \Psi(image(y,v),image(x,v))$

type Φ*(Body,Body,Point)*

where if:
Φ = *NonOccludes, then* $\Psi \in \{DC,EC\}$
Φ = *TotallyOccludes, then* $\Psi \in \{EQ,TPPI,NTPPI\}$
Φ = *PartiallyOccludes, then* $\Psi \in \{PO,TPP,NTPP\}$
Φ = *MutuallyOccludes, then* $\Psi \in \{PO,EQ,TPP,NTPP\}$
and where if:
X = TotallyOccludes, then $\Psi \in \{EQ,TPPI,NTPPI\}$
X = PartiallyOccludes, then $\Psi \in \{PO,TPP,NTPP\}$
X = MutuallyOccludes, then $\Psi \in \{TPP,NTPP\}$

e.g. $TotallyOccludesEQ(x,y,v) \equiv_{def.}$
$TotallyOccludes(x,y,v)\ \&$
$EQ(image(x,v),image(y,v))$

$TotallyOccludesEQ^{-1}(x,y,v) \equiv_{def.}$
$TotallyOccludesEQ(y,x,v)\ \&$
$EQ(image(y,v),image(x,v))$

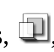
type Φ*(Body,Body,Point) where:* Φ *is an element from the set of all 20 occlusion relations.*

It is this part of the *Region Occlusion Calculus* that is now referred to as *ROC-20*, and the set of 20 JEPD relations as $JEPD^{ROC\text{-}20}$.

3 Theory Comparisons

It is now possible to map out the relationship between *RCC-8*, Galton's [1994] Lines of Sight Calculus (*LOS-14*), and *ROC-20*. Consider the $JEPD^{RCC\text{-}8}$ *overlap* relations first, i.e. *{PO, TPP, NTPP, EQ, TPPI, NTPPI}*. These relations are indifferent to relative distance with respect to a viewpoint, and each conflates a pair of Galton's relations. For example, given only that *x* partially overlaps *y*, it is impossible to say whether *x* is in front of or behind *y*. In both *LOS-14* and *ROC-20*, these two cases are distinguished.

This leaves two *RCC-8* relations *{DC, EC}*. These map respectively to the *LOS-14* relations *C/2* (clears) and *JC* (just clears), and to the two *ROC-20* relations: *NonOccludesDC/3* and *NonOccludesEC/3*. The six remaining relations of *ROC-20* are precisely the cases where non-convex bodies (ruled out in *LOS-14*) are allowed into the modelled domain. These correspondences are illustrated in table 1.

In table 1 mutually occluding objects are shown thus, , indicating that the lighter coloured 'U'-shaped object both occludes and is occluded by the darker. In the special case of *MutuallyOccludesEQ/3* part of the darker body lies behind the lighter body, and is exactly subtended by it, while a (visible) part of this, extends through a slot in the lighter body and occludes it.

The formal relationship between *LOS-14* and *ROC-20* is also illustrated by the following theorem:

(T3) $\forall x \forall y \forall v\ [\neg MutuallyOccludes(x,y,v) \leftrightarrow [NonOccludes(x,y,v) \vee TotallyOccludes(x,y,v) \vee PartiallyOccludes(x,y,v) \vee TotallyOccludes^{-1}(x,y,v) \vee PartiallyOccludes^{-1}(x,y,v)]]$

where the five disjuncts are provably pairwise disjoint, and where each disjunct in turn respectively splits into 2+3+3+3+3 = 14 of the $JEPD^{ROC\text{-}20}$ 'base' relations.

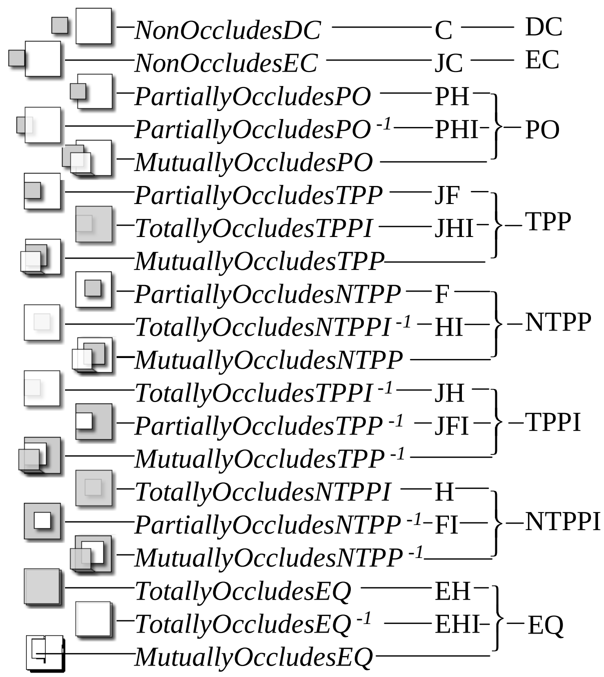


Table 1: The comparison between the *JEPD* relations of *ROC-20*, *LOS-14* and *RCC-8*. In each case the dark and light objects in the model respectively maps to the *x*, *y* variables of each $\Phi(x,y,v)$ relation of *ROC-20*, and to each corresponding $\Phi'(x,y)$ relation of *LOS-14* and *RCC-8*.

4 Comparative Distance and Occlusion

While the notion of relative distance between bodies appears in this theory, it only forms part of the *interpretation* resulting from the model used, and is implicit. Made explicit, a robot, for example, can exploit this information to reason about partial orderings of radial distances between itself and bodies based on their observed or inferred occlusion properties. A reworked subset of comparative distance axioms originally proposed by van Benthem [1982] is embedded into the theory. The primitive relation: '$N(x,y,z)$' used here, is read as "point *x* is nearer to body *y* than *x* is to body *z*", while '$E(x,y,z)$' is read as "body *y* is as near to point *x* as is body *z*":

(A26) $\forall x \forall y \forall z \forall u\ [[N(x,y,z) \,\&\, N(x,z,u)] \rightarrow N(x,y,u)]$
(A27) $\forall x \forall y\ \neg N(x,y,y)$
(A28) $\forall x \forall y \forall z \forall u\ [N(x,y,z) \rightarrow [N(x,y,u) \vee N(x,u,z)]]$

(D42) $E(x,y,z) \equiv_{def.} \neg N(x,y,z) \,\&\, \neg N(x,z,y)$

type $\Phi(Point,Body,Body)$, where: $\Phi \in \{N,E\}$

Comparative distance is related to occlusion, and is embedded into *ROC-20*, with the following axioms:

(A29) $\forall x \forall y \forall v\ [TotallyOccludes(x,y,v) \rightarrow N(v,x,y)]$

i.e. if *x* totally occludes *y* with respect to some viewpoint *v*, then *x* is nearer to *v* than *y* is to *v*.

(A30) $\forall x \forall y \forall v\ [N(v,x,y) \rightarrow$
$\forall z[P(region(z),region(y)) \rightarrow N(v,x,z)]]$

i.e. if *v* is nearer to *x* than *y*, then *v* is nearer to *x* than any part of *y*.

Note that the named viewpoint is not necessarily identified with an agent, and intentionally so. For example, if the agent holds and aligns two objects (one in each hand) where the one totally occludes the other, it does not follow the agent is closer to the occluding object, than the one occluded. It is also because of the guiding projective geometry assumed here (and which interprets the *image/2* function) that a viewpoint is identified with a point, and not an extended region in space.

5 Relative Orientation

If one body lies just to the left of another with respect to a line of sight, and is closer to the observer, movement to the right will typically increase the apparent separation between them. The relative left-right hand positions of the bodies will reverse as the line of sight intersects both bodies and passes to the left of that point. In order to be able to model and exploit this example of motion parallax, the ternary primitive relation: '$Left(x,y,v)$', read as "*x* is to the left of *y* from viewpoint *v*", is added and axiomatized. Its dual *(Right/3)* is also defined:[3]

(A31) $\forall x \forall v \neg Left(x,x,v)$
(A32) $\forall x \forall y \forall v\ [Left(x,y,v) \rightarrow \neg Left(y,x,v)]$
(A33) $\forall x \forall y \forall z \forall v\ [[Left(x,y,v) \,\&\, Left(y,z,v)] \rightarrow$
$Left(x,z,v)]$

(D43) $Right(x,y,v) \equiv_{def.} Left(y,x,v)$

type $\Phi(Body,Body,Point)$ where: $\Phi \in \{Left,Right\}$

For completeness, the relation: '$NonLeftRight(x,y,v)$' read as "*x* is neither to the left or right of *y* relative to viewpoint *v*", is added:

(D44) $NonLeftRight(x,y,v) \equiv_{def.}$
$\neg Left(x,y,v) \,\&\, \neg Right(x,y,v)$

type $NonLeftRight(Body,Body,Point)$

Here it is assumed that the observer's horizon is fixed, and that the field of view is restricted. Without these assumptions, the transitivity of *Left/3*, for example, would fail in the intended model. This would be the case if the agent were at the centre of a circular arrangement of objects (Stonehenge,

[3] Other spatial orientation duals with exactly the same properties (irreflexivity, etc.) are easily definable, e.g. forward/rearward, or above/below.

for example), entailing each object could be both to the left and the right of itself.

The primitive relation *Left/3* is embedded into the theory using the following axioms:

$$(A34)\ \forall x \forall y \forall v\ [Left(x,y,v) \rightarrow [\exists z\ [P(region(z),region(x))\ \&\ Left(z,y,v)]\ \&\ \neg\exists u\ [P(region(u),region(x))\ \&\ Left(y,u,v)]]]$$

$$(A35)\ \forall x \forall y \forall v\ [Left(x,y,v) \rightarrow \neg P(image(x,v),image(y,v))]$$

i.e. in the first case (*A34)* if from *v*, *x* is left of *y*, some part of *x* projects to the left of *y*, while no part of *x* projects to the right of *y*; while in the second case *(A35)*, from *v*, if *x* is left of *y* then *x* is not subtended by *y*.

It is now straightforward to see how *ROC-20* can be further developed. For example, where one object lies to the left of another and is disjoint, to the left and in boundary contact, and so on. All the distinct states depicted in figure 1 can then be modelled.

6 Relative Viewpoints

A change of viewpoint always carries the possibility of a change in the apparent spatial relationships holding between bodies in the domain (figure 1). If, for example, two bodies are physically separated, and an agent is allowed to freely move around, several apparent spatial relationships may be seen to apply. However, for two bodies forming a part-whole relation, no change in the viewpoint will coincide with both bodies separating. These and other configuration possibilities form the basis of the set of global *spatial constraints* introduced in section 2.2. There still remains the question of singling out additional *dynamic* spatial constraints, this time arising from instantaneous transitions between temporally ordered sequences of occlusion events.

As with many discrete based QSR theories, the set of $JEPD^{ROC\text{-}20}$ relations can be worked into an *envisionment*, where a set of axioms lay out the dynamic possibilities and constraints of spatial relationships deemed to hold between bodies over consecutive moments in time [Cohn, 1997]. For *ROC-20* this is represented as a table (table 2) where legal/illegal (instantaneous) transitions between spatial relationships are respectively denoted by "y" (yes) or "n" (no) entries mapping to pairs of named occlusion relations. A path formed by linking together pairs of nodes denotes a possible projected sequence of states from an initial state (at time *t*) via successor states (at times *t+1* … *t+n*).

The symmetry about the highlighted diagonal indicates the symmetrical relationship between each pair of named nodes. For example, the relation *NonOccludesEC/*3 has five such legal transitions, as read across the named row or down the named column. This means the relation *NonOccludesEC/3* from time *t* to the next instant *t+1*, now re-worked as an envisionment axiom (assuming a fixed viewpoint and the continued existence of the bodies from time *t* to *t+1*), has the following form:

$$\forall x \forall y \forall v \forall t\ [HoldsAt(NonOccludesEC(x,y,v),t) \rightarrow [HoldsAt(NonOccludesEC(x,y,v),t+1) \vee HoldsAt(NonOccludesDC(x,y,v),t+1) \vee HoldsAt(PartiallyOccludesPO(x,y,v),t+1) \vee HoldsAt(PartiallyOccludesPO^{-1}(x,y,v),t+1) \vee HoldsAt(MutuallyOccludesPO(x,y,v),t+1)]]$$

	NonOccludesDC	NonOccludesEC	PartiallyOccludesPO	PartiallyOccludesTPP	PartiallyOccludesNTPP	TotallyOccludesEQ	TotallyOccludesTPPI	TotallyOccludesNTPPI	MutuallyOccludesPO	MutuallyOccludesTPP	MutuallyOccludesNTPP	PartiallyOccludesPO^{-1}	PartiallyOccludesTPP^{-1}	PartiallyOccludesNTPP^{-1}	TotallyOccludesEQ^{-1}	TotallyOccludesTPPI^{-1}	TotallyOccludesNTPPI^{-1}	MutuallyOccludesTPP^{-1}	MutuallyOccludesNTPP^{-1}	MutuallyOccludesEQ
NonOccludesDC	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
NonOccludesEC	y	y	y	n	n	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n
PartiallyOccludesPO	n	y	y	y	n	y	y	n	y	y	n	n	n	n	n	n	n	y	n	y
PartiallyOccludesTPP	n	n	y	y	y	y	n	n	y	y	y	n	n	n	n	n	n	n	n	y
PartiallyOccludesNTPP	n	n	n	y	y	y	n	n	n	y	y	n	n	n	n	n	n	n	n	y
TotallyOccludesEQ	n	n	y	y	y	y	y	y	y	y	y	n	n	n	n	n	n	y	y	y
TotallyOccludesTPPI	n	n	y	n	n	y	y	y	y	n	n	n	n	n	n	n	n	y	y	y
TotallyOccludesNTPPI	n	n	n	n	n	y	y	y	n	n	n	n	n	n	n	n	n	y	y	y
MutuallyOccludesPO	n	y	y	y	n	y	y	n	y	y	n	y	y	n	y	y	n	y	n	y
MutuallyOccludesTPP	n	n	y	y	y	y	n	n	y	y	y	y	n	n	y	y	y	n	n	y
MutuallyOccludesNTPP	n	n	n	y	y	y	n	n	n	y	y	n	n	n	y	y	y	n	n	y
PartiallyOccludesPO^{-1}	n	y	n	n	n	n	n	n	y	y	n	y	y	n	y	y	n	y	n	y
PartiallyOccludesTPP^{-1}	n	n	n	n	n	n	n	n	y	n	n	y	y	y	y	n	n	y	y	y
PartiallyOccludesNTPP^{-1}	n	n	n	n	n	n	n	n	n	n	n	n	y	y	y	n	n	y	y	y
TotallyOccludesEQ^{-1}	n	n	n	n	n	n	n	n	y	y	y	y	y	y	y	y	y	y	y	y
TotallyOccludesTPPI^{-1}	n	n	n	n	n	n	n	n	y	y	y	y	n	n	y	y	y	n	n	y
TotallyOccludesNTPPI^{-1}	n	n	n	n	n	n	n	n	n	y	y	n	n	n	y	y	y	n	n	y
MutuallyOccludesTPP^{-1}	n	n	y	n	n	y	y	y	y	n	n	y	y	y	y	n	n	y	y	y
MutuallyOccludesNTPP^{-1}	n	n	n	n	n	y	y	y	n	n	n	n	y	y	y	n	n	y	y	y
MutuallyOccludesEQ	n	n	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y

Table 2: The envisionment table for *ROC-20*

The envisionment table can be interpreted two ways: either in terms of the viewpoint changing, or where the positions of the bodies change. In the former case, an additional predicate is required: *ChangePos(v1,v2)*, (meaning viewpoint *v1* changes to viewpoint *v2*), which relates *v1* at time *t* in the antecedent of the envisionment axiom to *v2* at time *t+1* in the consequent. These sequences of occlusion events can then be viewed as building the *topology* of motion parallax into the model. Obviously, where orientation information is added the number of relations and nodes increase, as does the overall complexity of the new set of permissible transitions between specified named occlusion relations.

7 Discussion and Conclusions

ROC-20 presents an axiomatisation of spatial occlusion. It assumes the region based ontology of *RCC-8* [Randell, *et al.*, 1992] and extends the work of Galton [1994] by allowing both convex and concave shaped bodies. It is this extension that describes occlusion events of mutually occluding bodies. The inclusion of van Benthem's [1982] notion of comparative nearness facilitates reasoning about relative distance between occluding bodies. An envisionment table models sequences of occlusion events to enable reasoning about objects and the images that may be formed in a visual field.

Several directions for future work are indicated. The axiomatisation of the primitive relation: *TotallyOccludes/3*, currently rules out models where the occluding body has a part that wraps behind the occluded body. In the theory this is a case of *mutual occlusion*. However, we can see potential gains by re-working the current axiomatisation (and relaxing this restriction) so that *any* degree of enclosure of one body by another (from some assumed viewpoint) could be a case of total occlusion.

Additional work is required to generate the *composition table* [see Cohn, 1997] for *JEPD* subsets of the defined occlusion relations. Also of note is the question whether there are any decidable and tractable subsystems of *ROC-20*, as has already been shown for RCC-8 [Bennett, 1994; Renz and Nebel, 1998]. Further computational gains may be made by adding information about the relative size of bodies or regions acting as additional constraints when checking for consistency of sets of these relations [c.f. Gerevini and Renz, 1998].

ROC-20 lays the theoretical foundations for further work in Cognitive Robotics, in which the images of objects are used to infer the spatial arrangement of objects in a robot's world - ultimately with map building and route planning in mind. We argue that the modelling of occlusion and motion parallax within a traditional QSR approach offers a uniform framework to achieve this. Galton [1994] has already shown these lines of sight relations can partition an idealized plan view of the embedding space into a set of polygonal regions. For each (view) point in that space exactly one of the JEPD line of sight relations holds. Where objects of varying shapes and sizes exist, many named sight lines that form tangents to objects naturally intersect at points. These correspond in this theory to a conjunction of atomic formulae drawn from the set of *JEPD* relations used. This gives rise to a set of extrinsic *reference points* determined completely by the objective spatial arrangement of the objects in the robot's world. With these points, *localization* becomes possible, while enabling qualitative and metric quantitative information to be combined. Spatial constraints and envisionment axioms now lead into map building and route planning. The robot then acquires the means to plan and execute moves [Levitt and Lawton, 1990; Schlieder, 1993] while constantly monitoring and relating its own direction of movement to the observed change and sequence of occlusion events in its visual field.

Acknowledgements

Work described here has been supported by EPSRC project GR/N13104, "Cognitive Robotics II". We wish to thank Paulo Santos, Brandon Bennett and Antony Galton for fruitful discussions during the development of ideas presented in this paper; and the comments given by the anonymous referees.

References

[Braddick and Atkinson, 1982] Braddick O. J. and Atkinson J., Higher Functions in Vision, in Barlow, H. B. and Mollon, J. D. (eds.) *The Senses*. Cambridge, Cambridge University Press, page 224

[Bennett, 1994] Bennett, B., Spatial Reasoning with Propositional Logic, *Proc. 4th Int. Conf. on Knowledge Representation and Reasoning (KR-94)*, pages 51-62

[Cohn, 1997] Cohn, A.G., Qualitative spatial representation and reasoning techniques, *Proc. KI-97*, *Advances in Artificial Intelligence*, LNAI 1303, Springer, pages 1-30

[Cui, *et al.*, 1992] Cui, Z, Cohn, A.G and Randell D.A., Qualitative simulation based on a logical formalism of space and time, *Proc. AAAI-92*, pages 679-684

[Galton, 1994] Galton, A.P., Lines of Sight, *AISB Workshop on Spatial and Spatio-Temporal Reasoning*.

[Geiger, *et al*., 1995] Geiger, D., Ladendorf, B. and Yuille, A., Occlusions and Binocular Stereo, *Int. J. of Computer Vision*, Vol. 14, pages 211-226

[Gerevini and Renz, 1998] Gerevini, A and Renz, J., Combining Topological and Qualitative Size Constraints for Spatial Reasoning, *Proc. 4th Int. Conf. on Principles and Practice of Constraint Programming (CP-98)*, pages 220-234

[Levitt and Lawton, 1990] Levitt, T.S. and Lawton, D.T., Qualitative Navigation for Mobile Robots, *Artificial Intelligence*, Vol. 44, pages 305-360

[Petrov and Kuzmin, 1996] Petrov, A.P. and Kuzmin, L.V., Visual Space Geometry Derived from Occlusion Axioms, *J. of Mathematical Imaging and Vision*, Vol. 6, pages 291-308

[Plantinga and Dyer, 1990] Plantinga, H. and Dyer, C.R., Visibility, Occlusion, and the Aspect Graph, *Int. J. of Computer Vision*, Vol. 5, pages 137-160

[Randell, *et al.*, 1992] Randell, D. A.; Cui, Z and Cohn A. G., A Spatial Logic Based on Regions and Connection, *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning (KR-92)*, Morgan Kaufmann, San Mateo, pages 165-176

[Renz and Nebel, 1998] Renz, J. and Nebel, B., Efficient Methods for Qualitative Spatial Reasoning, *Proc. 13th Euro. Conf. on Artificial Intelligence (ECAI-98)*, John Wiley & Sons, pages 562-566

[Schlieder, 1993] Schlieder, C., Representing Visible Locations for Qualitative Navigation, in Piera Carrate, N. and Singh, M.G. (eds.) *Qualitative Reasoning and Decision Technologies*, Barcelona: CIMNE, pages 523-532

[Shanahan, 1996] Shanahan, M.P., Robotics and the Common Sense Informatic Situation, *Proc. 12th Euro. Conf. on Artificial Intelligence (ECAI-96)*, pages 684-688

[van Benthem, 1982] van Benthem, J., The Logic of Time, Synthese Library Vol. 56, Reidel, London, Appendix A.