

## Reasoning about Space and Images

Paulo E Santos\*

Lab of Artificial Intelligence Applied to Automation, University center of FEI, Brazil

### Introduction

Much work in computer vision in the 70's and 80's aimed at the development of high-level vision, whereby the numerical processes feed a symbolic level of knowledge with which an agent is capable of interpreting the world. These early attempts were frustrated by the non-existence at the time of efficient algorithms for dealing with uncertainty, of tractable knowledge representation formalisms and also by the rudimentary stage of image-processing algorithms. Since then, important advances in Artificial Intelligence (AI) suggest that we may be at the stage of bridging the gap between AI and Computer Vision. One possible way of bridging this gap is the development of Qualitative Spatial Reasoning (QSR) methods based on sensor data. This idea is intrinsically connected to the tradition on logic-based image interpretation. This paper presents a brief introduction to these fields and discusses a possible research agenda for the future.

### Qualitative Spatial Reasoning (QSR)

The goal of Qualitative Spatial Reasoning (QSR) in Artificial Intelligence is to formalize spatial knowledge using elementary entities such as spatial regions, directions, line segments among others [1]. Traditionally, however, QSR formalisms are independent from an observer's viewpoint, which makes them not applicable to computer vision or robotic problems. There is, however, a growing interest in the development of dynamic formalisms about space in which qualitative changes observed by a mobile robot are the building blocks of the system [2-4].

The development of viewpoint-based QSR formalisms not only has a theoretical interest, but it is also an essential step towards equipping robots with the capability of interpreting its sensor data using high-level knowledge, since the notion of space is ubiquitous in our knowledge of the external world. In dos Santos et al. [5] the formalism developed in Souchanski [4] was applied to the task of image understanding from the point of view of the robot, where the notions of approximation and separation of visual objects were used to describe the scenes observed. Following a similar line, the perception of cast shadows, to make sense of a robot's environment, was recently formalized into a spatial reasoning system [6].

This sub-field of QSR, whose aim is the investigation of viewpoint dependent Qualitative Spatial Reasoning, is related to the long tradition of logic-based image interpretation, as we shall see in the next section.

### Logic-based Scene Interpretation

The modern research on scene interpretation is largely concerned with the development of probabilistic methods motivated by the need to deal with sensor noise and image uncertainty [7]. Probabilistic methods, however, are propositional, imposing restrictions in their capability to represent general domain knowledge and their applicability on problems containing a possibly unbounded number of objects Tran and Davis, 2008. Logic-based image interpretation, on the other hand, tackles exactly the problem of the effective representation of general facts about the domain, as well as the generalization of these

facts to problems with infinite variables [8]. Thus, research on logic-based image interpretation does not preclude the use of probabilistic methods, but complements them by making explicit the knowledge content of a domain.

The first framework for a logic-based scene interpretation system was proposed in Reiter and Mackworth [9] where three sets of axioms were defined to constrain the number of possible interpretations of the scenes observed. Therefore, the scene interpretation process is reduced to a constraint satisfaction problem. The SIGMA system [10] successfully deploys the ideas proposed by Reiter and Mackworth on the field of aerial image interpretation. Some properties of the formalism used by the SIGMA system were further developed in Schroeder and Neumann [11] and recently revisited and incorporated into a description logic setting [12].

In this context, there is a branch of research that follows the work presented in Shanahan [13] where a logical formalism is developed to rigorously define the information obtained from a robot's sensors in terms of symbols hypothesizing the existence, location and shape of the observed objects. Following these ideas Santos and Shanahan [14], Santos [3], dos Santos et al. [5] presented a theory aiming at the automatic scene understanding from a robot's viewpoint. In particular, Santos [3] presents formalism capable to interpret events such as approaching, receding, or coalescing from pairs of subsequent images obtained by a mobile robot's stereo pair. In order to further interpret these image-related events, an abductive procedure was developed for hypothesizing on the possible changes that might have occurred with the domain objects that could explain the image events.

In another work dos Santos et al. [5], a framework capable of interpreting events based on an arbitrary long image sequence was proposed. In this case, events such as the rotation of an object around a reference point; one object following another; and, one object trespassing another were formally defined within a dynamic logic and further used in a scene interpretation procedure involving perceptually indistinguishable objects. This system, however, could not handle uncertainty in the scenes. In Fenelon et al. [15], Santos et al. [16] a spatial reasoning system was developed within probabilistic logics that was applied on the task of interpreting images from traffic scenes from the viewpoint of a camera at the driver's position [17]. This work, along with others [18,19], constitute the initial attempts at the development of computer vision methods that are both: robust to uncertainties, and capable of executing automated high-level reasoning. However, much work remains to be done on this front.

\*Corresponding author: Paulo E Santos, Lab of Artificial Intelligence Applied to Automation, University center of FEI, Brazil, Email: [psantos@fei.edu.br](mailto:psantos@fei.edu.br)

Received July 24, 2012; Accepted July 26, 2012; Published July 28, 2012

Citation: Santos PE (2012) Reasoning about Space and Images. Adv Robot Autom 1:e108. doi:10.4172/2168-9695.1000e108

Copyright: © 2012 Santos PE. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Conclusion

Spatial reasoning is present in almost all human interactions in the real world, from tying a shoe-lace to executing complicated navigation tasks, or the simple interpretation of visual scenes. However, current computer vision systems have largely overlooked recent developments in the field of Qualitative Spatial Reasoning (QSR) in Artificial Intelligence. In order to foster across-fertilization between QSR and Computer Vision, this article briefly presents the main QSR formalisms for qualitative spatial reasoning whose aim is the automatic interpretation of scenes using high-level concepts. We set as a challenge the further development of QSR methods, grounded on computer vision data (and robust to uncertainty), that are capable of making sense of images using our knowledge of the external world<sup>1</sup>.

## References

- Cohn AG, Renz J (2008) Handbook of Knowledge Representation, chapter Qualitative Spatial Representation and Reasoning 3: 551-596.
- Santos P, Shanahan M (2003) A logic-based algorithm for image sequence interpretation and anchoring. In Proc of IJCAI, Acapulco, Mexico 1408-1410.
- Santos PE (2007) Reasoning about depth and motion from an observer's viewpoint. *Spatial Cognition and Computation* 7: 133-178.
- Souchanski M, Santos P (2008) Reasoning about dynamic depth profiles. In Proc of the 18th European Conference on Artificial Intelligence (ECAI), Amsterdam, The Netherlands, IOS Press 30-34.
- dos Santos MV, de Brito RC, Park HH, Santos P (2009) Logic-based interpretation of geometrically observable changes occurring in dynamic scenes. *Applied Intelligence* 31: 161-179.
- Santos P, Dee HM, Fenelon V (2009) Qualitative robot localization using information from cast shadows. In IEEE International Conference on Robotics and Automation (ICRA) 220-225.
- Lavee G, Rivlin E, Rudzsky M (2009) Understanding video events: A survey of methods for automatic interpretation of semantic occurrences in video. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 39: 489-504.
- Santos PE (2010) Tutoriais do XVIII Congresso Brasileiro de Automatica, volume 1, chapter Raciocinio e percepcao espacial: uma abordagem logica 127-147.
- Reiter R, Mackworth AK (1989) A logical framework for depiction and image interpretation. *Artificial Intelligence* 41: 125-155.
- Matsuyama T, Hwang VS (1990) SIGMA: A Knowledge-Based Image Understanding System. Plenum Press, New York, U.S.
- Schroeder C, Neumann B (1996) On the logics of image interpretation: Model construction in a formal knowledge representation framework. In International Conference on Image Processing 2: 785-788.
- Neumann B, Moller R (2007) On scene interpretation with description logics. *Cognitive Vision System* 3948: 247-275.
- Shanahan M (1996) Robotics and the common sense informatics situation.
- Santos P, Shanahan M (2002) Hypothesising object relations from image transitions. In van Harmelen, F, Proc of ECAI, Lyon, France 292-296.
- Fenelon V, Hummel B, Santos PE, Cozman F (2010) Encoding spatial domains with relational bayesian networks. In Spatio-temporal Dynamics Workshop 49-54.
- Santos P, Hummel B, Fenelon V, Cozman F (2010) Probabilistic encoding of spatial domains. In Proc of the first International Workshop on Uncertainty in Description Logics, 1-6.
- Souza CRC, Santos PE (2011) Probabilistic logic reasoning about traffic scenes. In The 12<sup>th</sup> Conference towards Autonomous Robotic Systems (TAROS) Berlin, Heidelberg Springer-Verlag 6856: 219-230.
- Tran SD, Davis LS (2008) Event modeling and recognition using Markov Logic Networks. *Computer Vision – ECCV* 5303: 610-623.
- Morariu VI, Davis LS (2011) Multi-agent event recognition in structured scenarios. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 3289-3296.