Motion Analysis using OCS-14 Transitions

CS-365 Course Project Report

Ashudeep Singh	Abhishek Maheshwari	Amitabha Mukerjee
ashudeep@iitk.ac.in	abhimw@iitk.ac.in	amit@cse.iitk.ac.in
IIT Kanpur	IIT Kanpur	IIT Kanpur

Abstract

Occlusions are a significant phenomenon in the motion analysis in multi-object computer vision. The relative depth information conveyed by occlusions make their study important. But the past studies on the formalization of occlusions-LOS-14, ROC-20, OCC-8 [2, 3, 4] failed to address certain crucial characteristics of occlusions, until OCS-14 [1], a representationally complete formalization, was proposed. Objects in a real world scene transit from one OCS-14 state to the other under a visual activity. The OCS-14 state transitions mark the signatures to some of these visual events. Mining out these states and detecting transitions is an important source of information to study object behaviors and can help gain abstraction about a scene [1]. In this project we propose a transition diagram for these states and discuss the ways in which the transition graph makes the OCS-14 formalization more robust and how these states can be used to extract information from a scene.

1 Introduction

Occlusion is the phenomenon that objects that are spatially separated in 3-d space may make interfering projections on a 2-d image plane. However, studying occlusions is important because occlusions carry information about relative depth ordering of objects, which is important for multi-object tracking[6], activity modeling[7] and studying concepts like object persistence, containment and support amongst infants in Human Cognition[8].

In spatial reasoning literature, there have been formal analyses to study occlusions-LOS-14, ROC-20, OCC-8 etc., but have failed to address certain issues, due to which they have not been widely used in vision applications.

These formalizations had the following drawbacks:

- All of these formalizations ignore these 2 crucial criteria:
  - Whether the visible parts are connected or fragmented.
  - Whether the occluder is dynamic or static

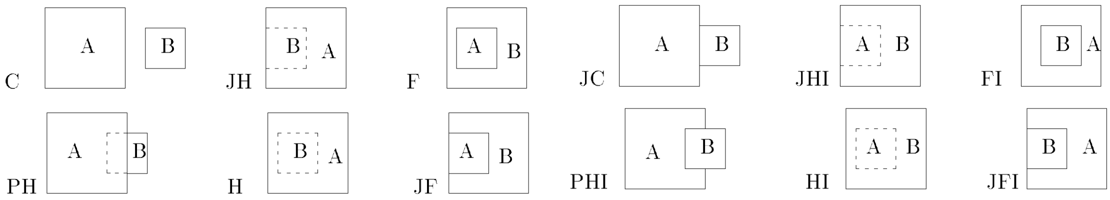


Figure 1: LOS-14 (Line of Sight)

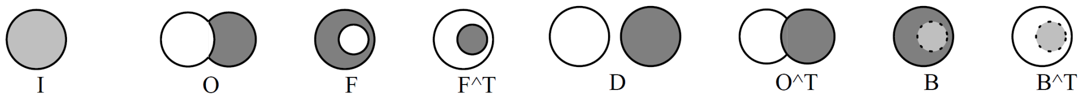


Figure 2: OCC-8 (Occlusion Calculus)

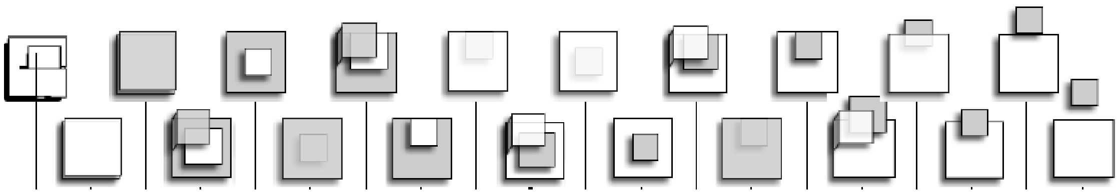


Figure 3: ROC-20 (Region Occlusion Calculus)

- Many unnecessary states are present, primarily in ROC-20, most of which can either not be distinguished nor be easily detected, e.g. precise tangency conditions.
- All these formalizations are based on relational algebra. We have to maintain relations between every pair of two objects. For multi-object tracking systems, the number of such pairs becomes quite high and hence, these are quite expensive.

1.1 OCS-14 states

OCS-14 is a state based formalization, in which we just maintain the states of each of the object in a scene. The following three characteristics defining a state, that make this representation complete, are:

- *Nature of Occluder*- Static or Dynamic
- *Visibility of Object*- Fully visible, Partially visible, Fragmented, Fully occluded.
- *Isolation or Grouping* with/from other dynamic objects.

Any representation must attempt to preserve those aspects of the problem that are relevant to the task[6]. These three characteristics define important distinctions relevant to occlusions.

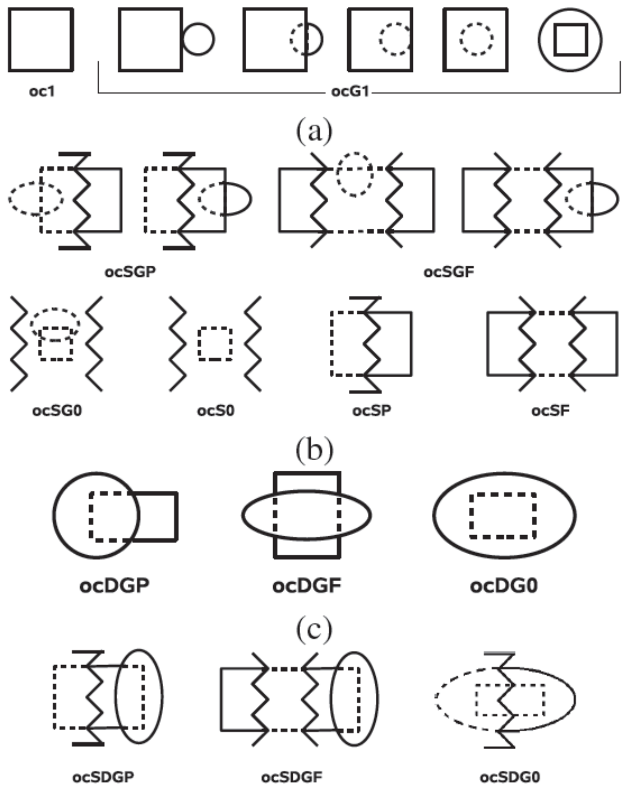


Figure 4: OCS-14 : Occlusion states. The rectangular object is the object of concern, the oval/circle is another dynamic object and the saw-shaped objects are the static occluders.

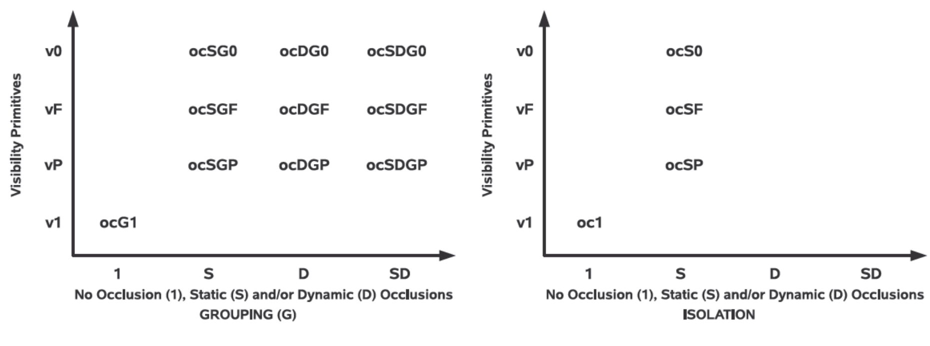


Figure 5: The two halves show ungrouped and grouped states respectively.

2 Logic Formulations

- ${}^w x$ - World Points
- ${}^w B$ - Static background points
- ${}^w S_i$ - Surface of a connected set.
- $isProjection(x, {}^w x)$ - A 2d point x is a projection of ${}^w x$
- $pointOrderLOS({}^w x, {}^w y)$ - Both and ${}^w x$ and ${}^w y$ lie on the same line of sight and ${}^w x$ is nearer to the viewpoint.
- $S_i = support({}^w S_i) = \{x : (\exists {}^w x \in {}^w S_i) isProjection(x, {}^w x)\}$.
- $occludes({}^w x, {}^w y) = pointOrderLOS({}^w x, {}^w y)$ and $visible({}^w x) = \neg\exists\, {}^w z\; occludes({}^w z, {}^w x)$.
- An object is isolated from other objects if its support (S_i) doesn't overlap with any other: $\forall k \neq i (S_i \wedge S_k \neq \phi)$
- Visual Support of an object: $V_i = \{x : (\exists {}^w x \in {}^w S_i) isProjection(x, {}^w x) \wedge visible(^w_x)\}$; $V \subseteq S_i$.
- **Visibility**
  - Full visibility $\Rightarrow V_i = S_i$
  - Partially Visible $\Rightarrow V_i \subset S_i$
  - Invisible $\Rightarrow V_i = \phi$
- **Static and Dynamic Occlusions:**
  - Set of points under static occlusions :
    ${}^w S_{stat}(i) = \{{}^w x \in {}^w S_{exp}(i) : (\exists {}^w z \in {}^w B) \wedge occludes({}^w z, {}^w x) \wedge visible({}^w z)\}$
  - Set of points under dynamic occlusions :
    ${}^w S_{dyn}(i) = \{{}^w x \in {}^w S_{exp}(i) : (\exists {}^w z \in {}^w S_j) \wedge (i \neq j) \wedge occludes({}^w z, {}^w x) \wedge visible({}^w z)\}$

- ${}^{W}S_{vis}(i) ={}^{w} S_{exp}(i) -{}^{w} S_{stat}(i) \cup{}^{w} S_{dyn}(i) = \{{}^{w}x : ({}^{w}x \in{}^{w} S_i) \wedge visible({}^{w}x)\}$

- So, the following situations can occur depending on the sizes of the above sets ${}^{w}S_{stat}$ and ${}^{w}S_{dyn}$.
  - Only Static Occlusions (S): $([{}^{w}S_{stat}(i) \neq \phi] \wedge [{}^{w}S_{dyn}(i) = \phi])$
  - Only Dynamic Occlusions (D): $([{}^{w}S_{stat}(i) = \phi] \wedge [{}^{w}S_{dyn}(i) \neq \phi])$
  - Both Static and Dynamic Occlusions (SD) : $([{}^{w}S_{stat}(i) \neq \phi] \wedge [{}^{w}S_{dyn}(i) \neq \phi])$
  - No Occlusions (1) : $([{}^{w}S_{stat}(i) = \phi] \wedge [{}^{w}S_{dyn}(i) = \phi])$

  The above 4 conditions are important to distinguish between feasible and unfeasible transitions amongst S, D, SD and 1 states.

3 OCS-14 Transitions

Out of the total possible 182 transitions, there are only a few feasible in real world visual scenes. An example of unfeasible transitions is shown below.

Figure 6: State of the person goes from $ocS1$ to $ocSF$ through $ocSP$. A direct transition: $ocS1 \rightarrow ocSF$ is not feasible in real world scenes

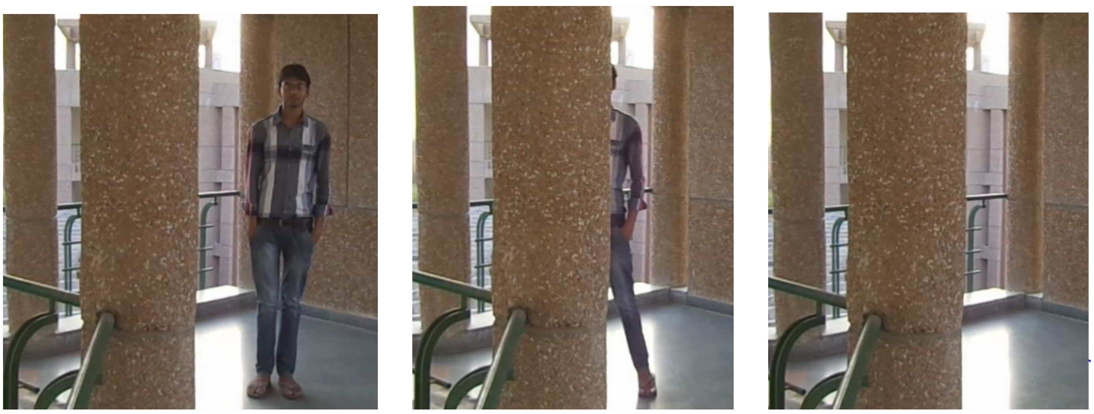
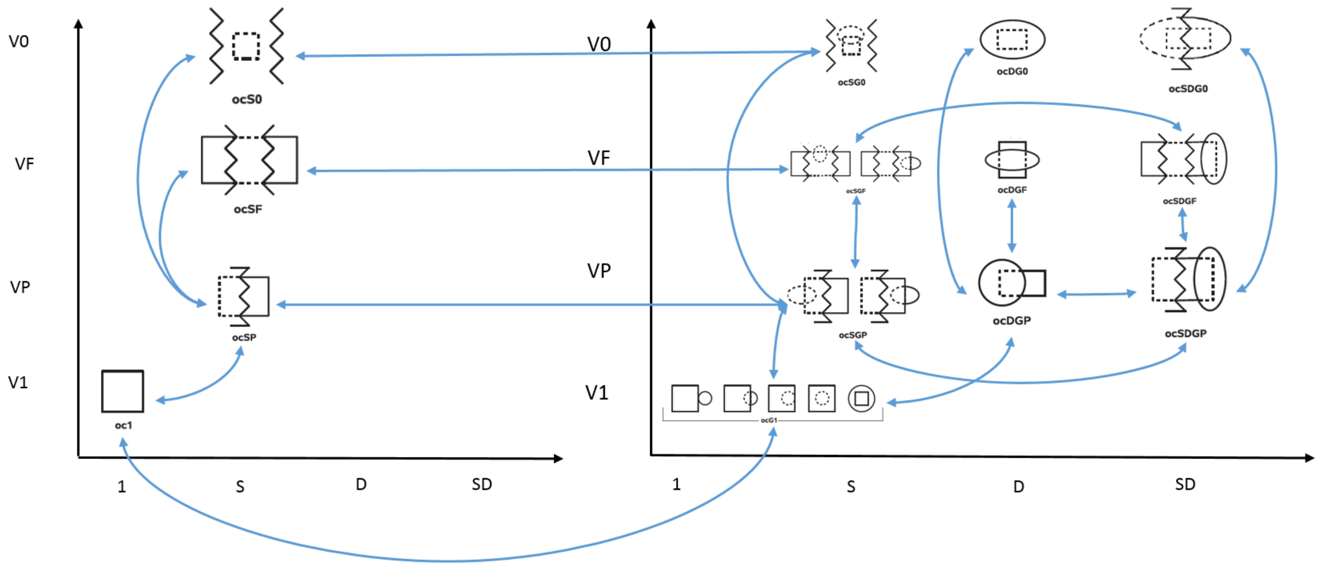
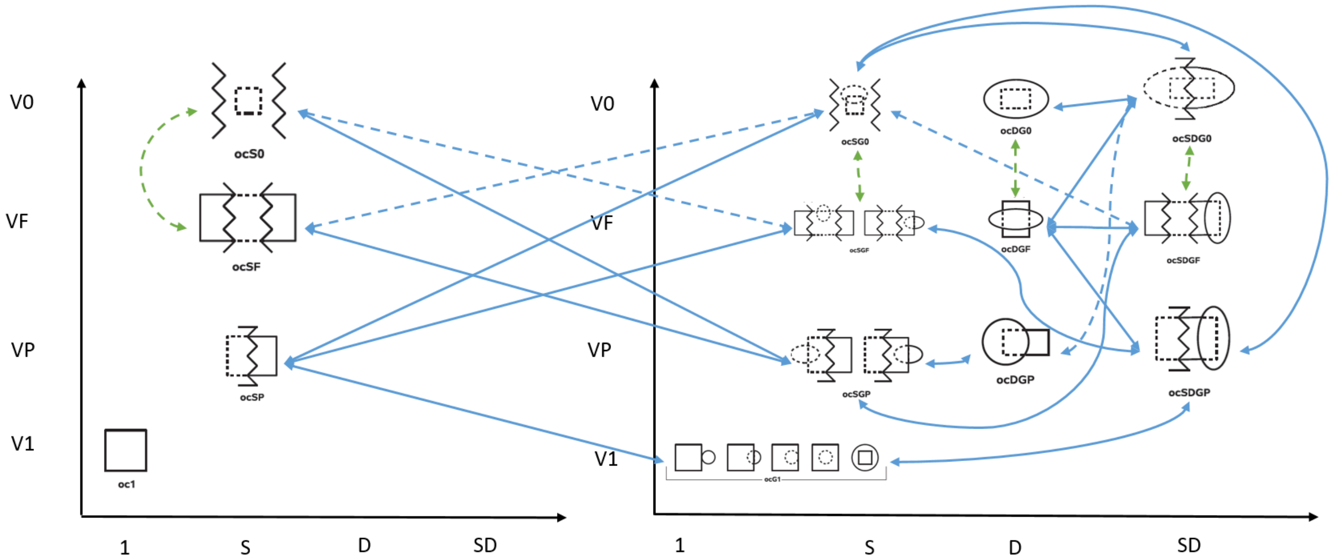


Figure 7: State of the person goes from $ocS1$ to $ocS0$ through $ocSP$. A direct transition: $ocS1 \rightarrow ocS0$ is not feasible in real world scenes

3.1 OCS-14 Transition Diagram





- – (a) represents the transition diagram containing the most probable transitions
  – (b) represents the transition diagram containing the transitions under certain assumptions or constrained motion of multiple objects.

The diagrams

- The left half (4 states) of the state diagram contains states in which occludersare static i.e. parts of the static background.
- The right half (10 states) of the diagram shows the states in which the object is occluded by both static and dynamic occluders.
- Figure (a) shows transitions that take place which requires one or more (dynamic) objects to move without any constraint.
- Figure (b) shows the transitions when certain special movements like shrinking (moving far), expanding (moving close) and special simultaneous motions are required.
- Transitions of figure-(b) have a very low probability of occurrence in real world visual scenes.
  - (Green Dashed arrow) Objects are allowed to shrink (move far), expand (move close) and also the two fragmented portions disappear together.

Figure 8: Transition from $ocSF$ to $ocS0$ is highly improbable as both the fragments disappear simultaneously.

  - (Blue Dashed arrows) Objects are allowed to shrink (move far), expand (move close) while allowing another dynamic object to come in contact with it in the projection (grouping).
  - (Blue Undashed arrows) These transitions occur under the assumption that object move simultaneously so as to reach another state from one state e.g. as soon as an object becomes unoccluded by a static occluder, it gets grouped with a dynamic object.

Important observations

- All the transitions are reversible. It is intuitive, since we can play a video backwards and still the scene remains a real world scene.
- An object cannot go from oc1 state to a fragmented or a fully occluded state without going through a Partial visibility state.
- From states with only static occlusions we have to first reach a state with static occlusions but grouped with another dynamic object, before going to other grouped states.
- The transitions between static, dynamic, static-dynamic states can be seen with respect to set-sizes as mentioned in Section 2(Logical Formulations) of the report.

4 OCS-14 Transitions as Signatures to Visual Activity

Occlusion transitions are an important visual signature of the interaction between objects. If we can mine out these transitions from a scene, we can gain useful abstraction of the scene and recognize the events taking place in the scene. Also, once we have

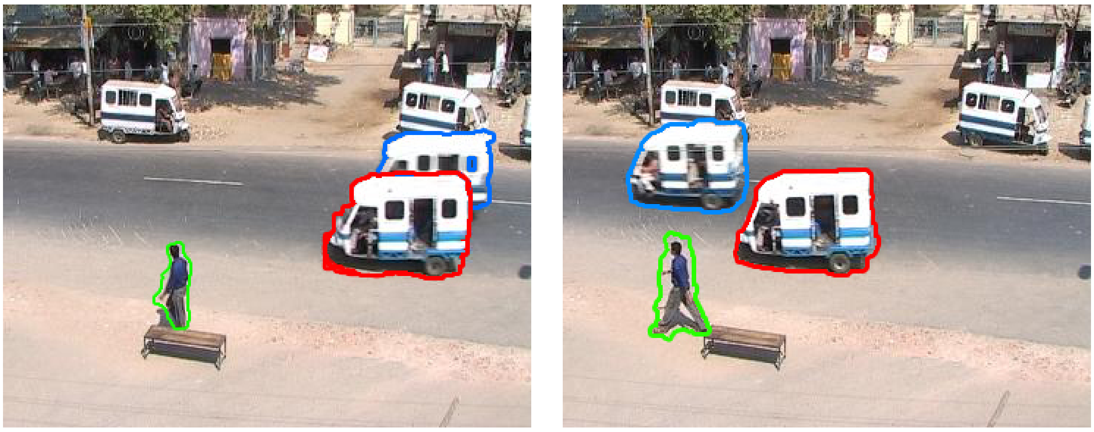


Figure 9: The event auto-rickshaw(blue) overtaking another auto-rickshaw(red) is characterized by transition $ocDGP \rightarrow oc1$ for blue auto-rickshaw and from $ocG1 \rightarrow oc1$ for red auto-rickshaw. If we have a signature set of transition for 'overtaking' event, we can detect this event through occlusion state transitions

a transition graph for the occlusion states available, we can search for the next state only in the reachable states from the current state. However, due to limitations like less frame rate, we may miss out on some transitions (as in the above example).

5 Validating the transition graph

Fig. 10 shows relations between Object A and Object B as a map of OCS-14 states of both objects. Using this map we know all the LOS-14 relations an object in ocG1 can be in. Then referring to Fig. 11, we can find all the relations reachable from current relation and then map this back to OCS-14 state(s) for object A. This is a useful

	ocG1	**ocDGP**	**ocDGF**	**ocDG0**
ocG1	JC	PH, JF, F	PH	H, JH, EH
ocDGP	PHi, JFi, Fi	MuOccPO	MuOccPO	×
ocDGF	PHi	MuOccPO	MuOccPO	×
ocDG0	Hi, JHi, EHi	×	×	×

Figure 10: OCS-14 to LOS-14 Map. Given ocs14 state of 2 objects, we can tell the occlusion relation between them

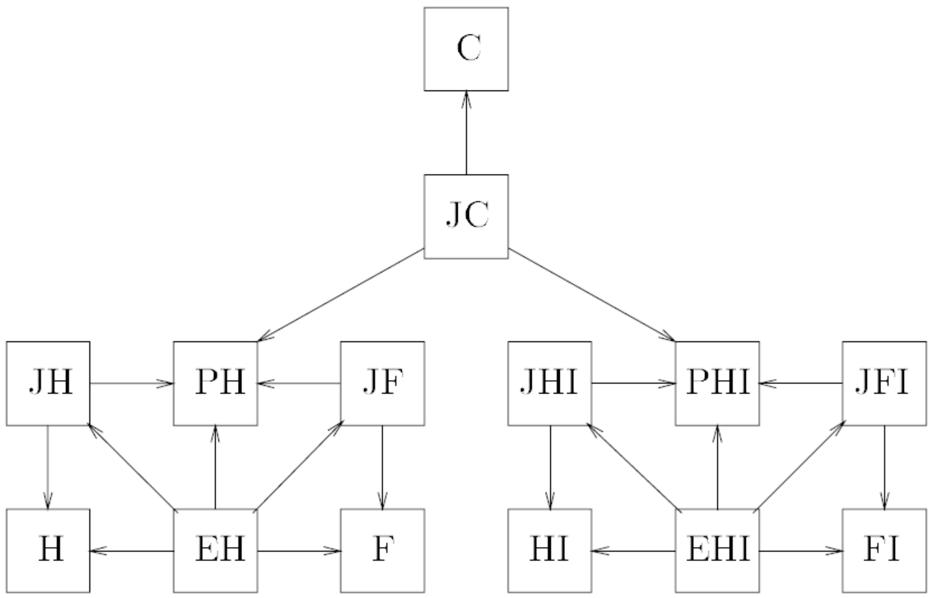


Figure 11: Transitions amongst LOS-14 relations

validation as LOS-14 has been widely accepted by the spatial reasoning community as an important formalization.

6 Conclusion

We discussed above OCS-14, a representationally complete formalization, which was proposed to address the drawbacks of other formalizations in visual applications . Objects in a real world scene transit from one OCS-14 state to the other under a visual activity. The OCS-14 state transitions mark the signatures to some of these visual events. Mining out these states and detecting transitions is an important source of information to study object behaviors and can help gain abstraction about a scene[1]. We have proposed a transition diagram for these states and discussed the ways in which transition graph makes the OCS-14 formalization more robust and how these states can be used to extract information from a scene.

References

[1] *Guha, P., Mukerjee, A., & Venkatesh, K. S. (2011, July).* OCS-14: you can get occluded in fourteen ways. *In Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two (pp. 1665-1670). AAAI Press.*

[2] *Galton, A. (1994, September).* Lines of sight. *In AISB Workshop on Spatial and Spatio-Temporal Reasoning.*

[3] *Randell, D., Witkowski, M., & Shanahan, M. (2001, August).* From images to bodies: Modelling and exploiting spatial occlusion and motion parallax. *In International Joint Conference on Artificial Intelligence (Vol. 17, No. 1, pp. 57-66).*

[4] *Köhler, C. (2002, September).* The occlusion calculus. *In Cognitive Vision Workshop.*

[5] *Mukerjee, A., & Joe, G. (1990).* A qualitative model for space. *Texas A & M University, Computer Science Department.*

[6] *Yilmaz, A., Javed, O., & Shah, M. (2006).* Object tracking: A survey. *Acm Computing Surveys (CSUR), 38(4), 13.*

[7] *Prithwijit Guha, Amitabha Mukerjee, and K.S. Venkatesh (2007).* Pattern Recognition Technologies and Applications: Recent Advances. *Idea Group Inc., 2007.*

[8] *Baillargeon, R., & Wang, S. H. (2002).* Event categorization in infancy. *Trends in cognitive sciences, 6(2), 85-93.*

[9] *Guha, P., Mukerjee, A., & Venkatesh, K. S. (2008).* Occlusion Sequence Mining for Activity Discovery from Surveillance Videos. *Pattern Recognition Technologies and Applications: Recent Advances, 212.*

Acknowledgements

We are very thankful to *Prof. Amitabha Mukerjee* and *Prof. Prithwijit Guha* for their valuable ideas and guidance throughout the course of the project.