A QUALITATIVE APPROACH TO LEARNING OF SPATIAL AND TEMPORAL RELATIONS FOR COGNITIVE COMPUTER VISION SYSTEMS

DONA DULHANI MENAKA RANASINGHE

A thesis submitted to The Open University of Sri Lanka in candidature for the degree of Doctor of Philosophy

Department of Mathematics & Computer Science
Faculty of Natural Sciences
The Open University of Sri Lanka
July 2009

Abstract

Cognitive Computer Vision Systems perceive the environment through vision and are able to perform many cognitive tasks such as learning, adapting, developing new strategies for analysis & interpretations and generalizing for new context. The literature review reveals that current approaches to cognitive vision systems are not adequately robust to be able to learn invariant features of spatial and temporal relations. However, it is a well known fact that humans provide a living example for robust visual learning and deliberating on human features such as the ability of qualitative learning, representing a scene in terms of objects and events by exploiting spatial and temporal relations and generating rules for the learned relations provides a rich set of avenues for the development of cognitive computer vision systems. As inspired by this view, this thesis focuses on design and development of an approach to learn rules of a visual scene by formulating the hypothesis that, real world visual scenes can be learned by qualitative realization of spatial and temporal relations.

In the proposed qualitative approach spatial relations are learned through qualitative spatial relation descriptions which mainly consider the relative size of objects (*RSizeRep)*, relative orientation of objects (*ROrientRep)* and topological connections between objects (*TopRep)*. In dynamic scenes time elapsed for change in topological connections are captured by *HybridRep*. Rules for these qualitative spatial relation descriptions are learned inductively, which depict the model of the visual scene. The model developing process is integrated into the

system, QCVision. The design of the QCVision system comprises of several modules. The qualitative relation learning module and the inductive logic programming module are the main modules of the system. The qualitative spatial relation descriptions are learned in the qualitative relation learning module and the rules for these relations are learned in the inductive logic programming module. The inductive logic programming module has been implemented using the ILP language CProgol4.4 and the qualitative relation learning module has been implemented using SWI Prolog. The generated rules are stored in the relation base. The QCVision system is able to develop a model by integrating these learned rules. The action generator module provides the user interface where new symbolic data sets can be input to the system and the developed models are output as graphical scenes.

The evaluation of the QCVision system is carried out to test whether the spatial and temporal relationships of a scene have been qualitatively learned by the QCVision system as per the main objective of this thesis. This evaluation is formed as a comparison of the QCVision system against a set of humans. Ten human subjects were involved in the evaluation. Initially the QCVision system was presented a set of scenes to learn the rules through qualitative realization of the scenes. Then scenes with different number of objects were presented to capture the ability of learning more rules. Later the QCVision system was given a new set of data and was asked to generate scenes based on the already learned rules. The generated scenes were assessed by humans to decide whether the scenes are acceptable to the human. The QCVision system exhibits a higher accuracy level with high example sizes compared to human. Though human

visual learning is about twenty percent higher on average than QCVision system learning, this difference diminishes with increasing example sizes and objects despite higher number of relations to learn by the QCVision system.