Task-based (Cognitive) Control for ActIPret Project

Hilary Buxton
School of Cognitive and Computing Sciences,
University of Sussex, Falmer, Brighton BN1 9QH, UK

December 18, 2001

Abstract

In this position paper we propose task-control techniques to be used in advanced vision applications. Dynamic Decision Networks (DDNs) have been developed for task-based selection, scheduling and communication with central system 'attention' control policies. These focus the processing for particular tasks using a graphical model approach with probabilistic reasoning, which interprets the evidence gathered by selected 'operators' (visual routines or modules). This approach allows the most likely explanation of the evidence to be demanded at any particular moment as well as cutting off processing when criteria for confirmation or rejection of a particular behaviour hypothesis have been fulfilled. We propose extensions these methods for: 1) compiling the task policy from a declarative 'Activity Plan', 2) development of evaluation criteria to meet real-time constraints for compatibility with RT CORBA, 3) learning to refine the DDN task control. Also, with partners we need to develop behaviour (user-camera) coordination in an integrated system.

1 Introduction

The general problem of selecting programs and setting tuneable parameters to get the required information from images for a particular application has now given rise to its own subfield of 'intelligent image processing'. For example, in Japan research has focussed on this problem [7], in the US [3] and in Europe general tools have been developed [1, 2]. There exist real-time knowledge-based (cognitive) program supervision approaches, e.g. PLANETE which has been used in driving assistance [8, 9]. The advantage of this approach is an explicit, declarative description of the relationships between program modules, their parameters and operating conditions etc. at the conceptual level. We also have a requirement for a symbolic description, our 'Activity Plan' in ActIPret, mainly for the purpose of explanation and repair with indexing into this conceptual description in the 'tutor mode'. However, this Activity Plan can also serve as a top-level description of the task but we know that adaptive (reactive) processing, learning and really fast execution can be problematic using explicit, knowledge-based techniques.

2 Proposal

The way we propose to tackle the very real-time constraints in task-based control for ActIPret, therefore, is to use our prior research on Dynamic Decision Network (DDN) approach [4, 6], which uses graph-based probabilistic reasoning models with reactive planning. The advantages here are: 1) they naturally give an evolving probabilistic interpretation that can be demanded at any time in the processing, 2) it is possible to build an evaluation control loop based on utility/priority of the results with respect to the task. This is then a way of mapping conceptual temporal constraints onto hard RT constraints to best meet task demands in the system. The full information integration offered by this type of approach is what we originally proposed for ActIPret. However, following the kick-off meeting and further reflection, there are three crucial areas of further work beyond the functionality in our HIVIS-WATCHER system [5, 6].

- First, the automatic compiling of the DDN mechanism and task policies from the Activity Plan is important as we then have, for the ideal case demonstrated by the expert, a general method of mapping into the on-line task-control and interpretation 'engine'.

- Second, the development of the evaluation loop to meet hard real-time constraints so that it can map into RT CORBA. The main constraints we considered in previous research were qualitative orderings of operations and quantitative thresholds that set criteria for confirming of rejecting behaviour hypotheses. The system did in fact operate in real-time (see [4] for timings of different control policies). However, for general control on an open, distributed RT platform, further work is required to guarantee a result, under optimal scheduling of program modules and resources, for any possible instantiation of processing within the system. THIS IS PROBABLY THE MOST URGENT EXTENSION.

- Third, the development of refinement techniques for the DDNs and task policies themselves will be useful. Much as HMMs and BBNs are known to benefit from incremental learning, we can improve the task-based reactive policies, derived from our Activity Plans, for use in the on-line system. In particular, if our experts initially demonstrate presegmented sequences, there can be improvement for a better synchronised continuous version with smoother transitions.

The DDN control itself must be organised in 2 PARTS as I suggested in the kick-off meeting. There should be a set of global, simple operators related to the task e.g. 'mutual proximity' (of hand and CD candidates), which attach 'preconceptual' software markers (indexes) to the relevant moving and static objects. This then allows selection and scheduling of 'attentional' processing, i.e. further local and possibly complex operators related to the task can obtain detailed evidence to confirm or reject the expected sequence of activities. The set of attentive operators are structured by what we called 'typical object behaviour models' which allow reactive planning of the visual data and attach 'postconceptual' or functional software markers (indexes) to the objects. For example, the confirmed hand could be a 'picker-upper' and the confirmed CD could be a 'picked-object'. Further semantic consequences can be attached to these functional roles but what is essential is to compute the deictic description of their relationship. This spatio-temporal relationship is the evidence required for confirming or rejecting hypotheses about 'approach' and 'pick-up' activities using a Bayes net. Note that this is what gives the system active (covert attention) control as it restricts the processing performed by the system to that required by the task rather than all that can be computed in the dynamic scene. For active camera (overt attention) control, we can add panning and zooming in on the behaviour interaction.

3 Conclusion

I hope this short paper clarifies aspects of the approach, but not the detail, for the Cognitive Vision Framework WP1 task-based control in ActIPret. Please comment urgently as it is crucial for the planned work of Kingsley Sage in the COGS team. We propose to schedule work on DDN development and mapping onto RT CORBA with PROFACTOR asap in the New Year. We propose to collaborate with INRIA team ORION on the compiler/engine subtask in the first half of next year. Further work on the refinement subtask remains for years 2/3.

In addition, it is clear that there is much that is affected in the system by this proposed approach, especially for the activity and behaviour interpretation work. In particular, we have agreed work with INFA on the computation of deictic descriptions under task-control. We have mentioned the link to attentive behaviour with moving camera viewpoints with PROFACTOR. Here we can further note that the proposed software marker (index) mechanism is an important component for combining information across viewpoints. It is also clear that there are organisational issues for the lower-level vision processing in recognition and tracking of objects. In particular, I suggest that the modules requested in the kick-off meeting are mainly of the 'attentive' processing kind (3D trajectories FORTH and 3D fully recognised objects CMP) so they need to be decomposed and controllable in real-time. Simpler modules that only compute 2D ground-plane trajectories and likely candidates for the objects are also likely to be useful in the completely specified system as they can cue when this more complex interpretation processing is required. This will save processing time in the fully integrated system.

References

[1] S.A. Chien and H.B. Mortensen. "Automating image processing for scientific data analysis of a large image database". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:854–859, 1996.

[2] V. Clement and M. Thonnat. "Integration of image processing procedures: OCAPI, a knowledge based approach". *Computer Vision, Graphics and Image Processing: Image Understanding*, 57:166–184, 1993.

[3] B. Draper, A Hanson, and E. Riseman. "Knowledge-directed vision: Control, learning, and integration". *IEEE Transactions on Signals and Symbols*, 84(11):1625–1637, 1996.

[4] R. Howarth. Interpreting a dynamic and uncertain world: Task-based control". *Artificial Intelligence*, 100:5–85, 1998.

[5] R. Howarth and H. Buxton. "Visual surveillance monitoring and watching". In *European Conference on Computer Vision*, pages 321–334, Cambridge, UK, 1996.

[6] R. Howarth and H. Buxton. "Conceptual descriptions from monitoring and watching image sequences". *Image and Vision Computing*, 18:105–135, 2000.

[7] T. Matsuyama. "Expert systems for image processing: Knowledge based composition". *Computer Vision, Graphics and Image Processing: Image Understanding*, 48:22–49, 1989.

[8] S. Moisan, C. Sheklar, and M. Thonnat. "Real-time perception program supervision for vehicle driving assistance". In *International Conference on Recent Advances in Mechatronics*, pages 173–179, Istanbul, Turkey, 1995.

[9] M. Thonnat, S. Moisan, and M. Crubezy. "Experience in integrating image processing programs". In *International Conference on Vision Systems*, pages 200–215, Gran Canaria, Spain, 1999.