# A METHODOLOGY FOR REAL TIME SCENE ANALYSIS

David C. Hogg*
The University of Western Ontario,
London, Canada.

## Introduction

This short paper outlines a methodology for generating reasonable descriptions, in real time, of simple visual situations involving one or two moving objects (Hogg 1977). An implementation, based upon this methodology, was constructed by the author at the University of Western Ontario using a Spatial Data camera system driven by an Interdate 7/32 mini-computer communicating with a DECsystem-10. The capabilities of this working system are illustrated by the following unmodified extract from the computer generated description of an actual scenario set in the laboratory.

> "An object has appeared at left of scene, call it object A. Object A has begun moving. Object A looks like a person, call it Fred. An object has appeared in middle of scene, call it object B. Fred placed object B in the scene".

Provided the level of overall illumination remains constant, the implementation generates correct descriptions of a variety of scenarios similar to that described above.

## The Methodology

A straightforward picture differencing algorithm is employed to locate moving and, therefore, interesting objects in a scene, perceived by a television camera and digitized into an internal image composed of a matrix of light intensities. Such images are generated repeatedly (every 0.3 seconds in the implementation) and differenced, in real time, with the image of an initial instance of the scene against which distinct regions of change are identified. In general, an object will give rise to several changed regions in a single image, through partial occlusion or insufficient pronouncement of the object against its background. By relating the regions generated from individual and successive images, the appearance of objects in the scene can be postulated and these objects subsequently tracked and possibly identified. To assist in this process a simple relational theory was developed, defining size, shape and positional differences between regions. Tracking is performed by relocating objects through successive images, at each stage guided by object models derived from elementary image features and possibly improved by high level processes. In particular, by identifying movement and shape characteristics, the implementation is able to identify a human being and consequently to provide high level guidance for the tracking process with a subsequent increase in performance, especially when dealing with noisy data or partially occluded objects.

*Supported by National Research Council of Canada grant A-9272.

The methodology is based upon a set of procedures through which control cyles during operation. Each procedure is responsible for recognising instances of a particular concept, characterising such instances by a set of assertions and perhaps producing a descriptive output sentence. Thus, (i) low level feature extraction, including image generation, and (ii) higher level recognition processes, may be closely integrated within a uniform procedural representation. Previous contributions in related fields have tended to concentrate in one of these areas and assumed the other, thereby avoiding many of the problems but as a consequence unable to realise the benefits of a joint consideration (Potter 1974, Badler 1975). During operation, the combined sets of assertions maintained by each procedure constitute an evolving representation, any instance of which provides a detailed description of the current state of the world including object locations, motion primitives and the higher level situations in which these objects are presently involved. This representation serves both as a medium for communication between procedures, enabling the representation itself to evolve, and as a source of immediate high level interpretation for an observer.

## Summary

The working system demonstrates the feasibility of the methodology and, in particular, that high level guidance of tracking increases both reliability and efficiency and can be implemented easily when both high and low level processes are represented uniformly as procedures. An important consequence of the proposed methodology is the support it lends to the supposition that certain significant concepts involving motion can be recognised in a scene without the need for an extensive analysis of individual instances of the scene.

In conclusion, the conceptual framework and procedural representation, only outlined in this short paper, are essentially independent of the basic picture operators and might, therefore, provide a useful foundation for future research, when more sophisticated picture operators will undoubtedly be employed.

## References

Badler, N.I., Temporal Scene Analysis: Conceptual Descriptions of Object Movements. T.R. - 80, Department of Computer Science, University of Toronto, 1975.

Hogg, D.C., Visual Perception in a Changing Environment: A Paradigm for Real Time Scene Analysis. T.R. - 29, Department of Computer Science, University of Western Ontario, 1977.

Potter, J., The Extraction and Utilisation of Motion in Scene Description. Ph.D. Thesis, University of Winsconsin, 1974.