TOWARD CONSISTENT DESCRIPTIONS IN VISION SYSTEMS*

Steven W. Zucker
Department of Electrical Engineering
McGill University
3480 University Street
Montreal, Canada H3A 2A7

1. Introduction

Part of the complexity in vision systems derives from the need to associate potentially enormous numbers of final descriptions to original intensity arrays. One solution for coping with this complexity is to organize the system into levels, so that the removal of possibilities at one level implicitly eliminates possibilities at other levels. If, in addition, the various structures and processes within such a vision system can be expressed in terms of a single, unified framework, then questions relating to the necessary and sufficient interactions between processes can be posed. The goal of this paper is to discuss, in a canonical fashion, a number of different kinds of interactions within a low-level vision system and, then, to study these interactions with respect to scheduling strategies, focus of attention mechanisms, and global system stability.

2. Information Flows

If vision systems are to be conceptualized in terms of levels of processing, then two fundamental kinds of information flow are necessary. The first kind can be thought of as propagating from the original image upwards. Its purpose is to establish the initial descriptions at various levels of abstraction. In principle it only needs to flow in one direction and can be implemented by production rules. In the event that the system is operating in a sufficiently restricted domain, i.e., one in which all data patterns are uniquely predefined, then the highest level structure will be the unique interpretation. However, realistic domains are not like this and, in general, the intermediate descriptive structures will contain ambiguities and multiple possibilities. These ambiguities multiply across the levels of abstraction and eventually lead to many possible final descriptions.

To reduce these ambiguities, a second kind of information flow-feedback-is necessary. Conceptually, this allows information to move both horizontally 'across all positions at a single level of abstraction and vertically across all levels of abstraction at a single projected position. Diagonal kinds of feedback, which flow across both position and level of abstraction, can be expressed in terms of these two orthogonal flows.

3. Relaxation Processes for Dot Labeling and Grouping

Feedback can be implemented uniformly in both directions by relaxation labeling processes (RLPs). These are parallel, iterative schemes for using consistency or compatibility relationships to reduce local ambiguities. In order to apply RLPs to dot labeling and grouping, a hierarchical arrangement of two horizontal RLPs and one vertical RLP must be defined. One horizontal process attempts to disambiguate the assignment of functional labels from the set {EDGE, INTERIOR, NOISE} onto individual dots. It iterates toward horizontal (i.e., spatial) consistency between the labels attached to neighboring dots. The second horizontal process performs the grouping operation by labeling the links between dots as either CONNECTED or NOT-CONNECTED. The vertical RLP allows interactions between the link labels and the dot labels. It iterates toward descriptions which are consistent across both levels of abstraction. (For example, a CONNECTED link is consistent with two neighboring INTERIOR dots.)

4. Conclusions

Experiments with this RLP system for labeling and grouping dots have revealed a number of conclusions relevant to the design of more complex systems.

(1) The conceptualization of vision systems in terms of levels of description and processes that operate across and between levels provides a framework in which many different systems can be compared with one another. For example, given this basis, an essential difference can be seen between Shirai's context-sensitive line finder and Marr's primal sketch. In Shirai's system the feedback is across widely separated levels of abstraction while in Marr's system the feedback is more local because it is restricted to more adjacent levels. Since, from a systems-theoretic viewpoint, local feedback is often sufficient for noise reduction, the existence of stable, independent modules becomes a possibility. Furthermore, the lack of generality in systems such as Shirai's might be counterbalanced by introducing more local forms of feedback.

(2) The strategy selected for scheduling individual processes can have a drastic effect on the system's overall performance. In particular, letting the horizontal and vertical processes run simultaneously rather than alternately will, under certain conditions, allow the system to converge rather than oscillate. Since all large systems (e.g. Hearsay-II) contain a scheduling mechanism for coping with finite resources, this mechanism must be designed with regard to total system performance. A unifying conceptual structure could certainly facilitate this design.

(3) The cooperative/competitive nature of the computations described eliminates the need for "best-first" attention-focusing mechanisms. With the many local processes operating in parallel, distinct activity only appears to take place around well-defined positions (or levels); in ambiguous situations the various processes counteract each other's effects. The net effect is that activity is focused implicitly - there is no need to define an explicit (typically heuristic) rule.

---

* This research was supported in part by the National Research Council under grant A4470. For a more extensive discussion of these ideas, see Technical Report 77-4, Dept. of E.E., McGill.