IIZUKA ´94, Fukuoka, Japan
August 1-7, 1994, Invited paper

Connectionist Decision Systems for a Visual Search Problem

Uwe R. Zimmer

University of Kaiserslautern - Computer Science Department - Research Group Prof. E. v. Puttkamer
67663 Kaiserslautern - Germany
Phone: …49 631 205 2624 - Fax: …49 631 205 2803 - Telex: 4 5627 unikl d
e-mail: uzimmer@informatik.uni-kl.de

Visual Search has been investigated by many researchers inspired by the biological fact, that the sensory elements on the mammal retina are not equably distributed. Therefore the focus of attention (the area of the retina with the highest density of sensory elements) has to be directed in a way to efficiently gather data according to certain criteria.

The work discussed in this article concentrates on applying a laser range finder instead of a silicon retina. The laser range finder is maximal focused at any time, but therefore a low-resolution total-scene-image, available with camera-like devices from scratch on, cannot be used here. By adapting a couple of algorithms, the edge-scanning module steering the laser range finder is able to trace a detected edge. Based on the data scanned so far, two questions have to be answered. First: "Should the actual (edge-) scanning be interrupted in order to give another area of interest a chance of being investigated?" and second: "Where to start a new edge-scanning, after being interrupted?".

These two decision-problems might be solved by a range of decision systems. The correctness of the decisions depends widely on the actual environment and the underlying rules may not be well initialized with a-priori knowledge. So we will present a version of a reinforcement decision system together with an overall scheme for efficiently controlling highly focused devices.

1. Efficient Visual Search

Visual Search is a dynamic research field, investigated by researchers from many different disciplines. We will highlight only some works concerning efficient visual search applying connectionist techniques (mainly self-organizing feature maps).

Ahmad and Omohundro have presented a multi-feature-map-system being able to search for and to locate target objects with certain features (namely equilateral triangles, [2]) in static camera images and more general to detect feature correspondence between the camera frame and the model space [1]. Marshall [4] has shown general techniques to handle context-sensitive problems like trajectory tracing, grouping and steering in visual motion perception with the "EXIN"-network.

The visual search system presented here is mainly conceived as a support system for the geometric environment abstractions on an autonomous mobile robot and is a part of the sub-symbolic abstraction project SPIN[1].

2. Concepts & Structure

The main structure is based on the decomposition in three main parts as shown in figure 1:

- An edge-scanner, able to perform 3-d edge-following and to a certain degree autonomous, i.e. beginning at a given starting

1. *see [5], [6] for more information about other parts of the SPIN-project*

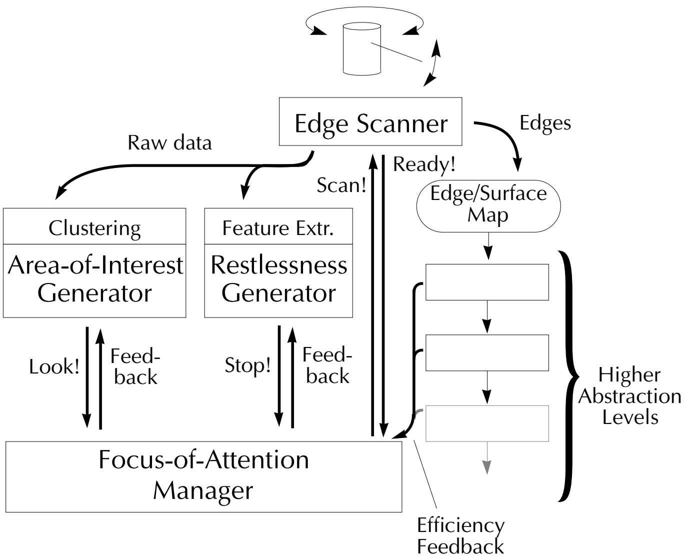

figure 1 : Focus of attention system

point, the edge-scanner finds recursively all edges, which are noticed while following the actual edge. The result is the edge / surface map, which represents discrete orientations and curvatures of all scanned edge-fragments.

- Based on the edge / surface-map a long pipeline of abstraction steps (up to a generalized, completed and classified cluster representation, which will be used by a symbolic level) follows. The "efficiency[1]" of the actual data-flow is measured at each stage.
- Finally the edge-scanner is controlled by the focus-of-attention manager, whose decisions are based on the efficiency feedback from the abstraction pipeline. The control task is divided into two sub-modules: The area-of-interest generator, which should select a region in the actual environment where an efficient knowledge accumulation might be possible, and a "disturbance module" called restlessness generator, which determines when to choose a new target area.

In the rest of the paper we will concentrate the discussion on these two decision sub-modules.

3. Area-of-Interest Generator

The area-of-interest is selected on the basis of statistical informations about the environment. The statistical informations are spacial densities of edges, free-space and unknown regions. In order to collect and represent these densities we use a special, dynamic self-organizing map with 3-dimensional simplexes (tetrahedrons) as the basic structure. The simplexes dividing the space in Voronoi regions, attributed with informations about the amount of sensor data collected in this region as well as densities (edges / free-space). The network is adapted by replacing single cells by new simplexes or by fusing neighbouring simplexes, according to several rules (see [7] for all the insertion / fusion criteria).

This unsupervised process gives the needed input for the area-of-interest selection. Up to now this decision is done by simple heuristics (search strategies), without using the efficiency feedback from the higher abstraction levels. In several tests the number of scanned edges / surfaces is significantly increased (in comparison to recursive edge-following) only by applying the clustering together with simple heuristics, but we can do even more by using the restless-generator, whose decisions are already based on the efficiency feedback (see next section).

This module is actually a framework for search strategies, which are adapted manually to the needs of the abstraction pipeline. Although the results are already encouraging, we are about to *select* the search strategies by a reinforcement system (as applied in the restlessness generator) as well as a neural fuzzy decision system (SPIN-NFDS [5]) for a further improvement of this task.

The tests with several search strategies seem to imply that the key issue of our visual search scenario is an adequate clustering rather than a sophisticated search heuristic.

4. Restlessness Generator

Due to the fact that one edge scan may spend a lot of time gathering information of low usefulness (e.g. scanning walls), a module which interrupts the actual scanning in order to give another (perhaps more productive) area in the environment the chance of being explored. In contrast to camera based systems, which are able to use the whole (low resolution) scene for this decision, the highly focused laser range finder delivers only information of the actually scanned edge. Therefore the features used for

1. a term which has to be determined by each individual abstraction step

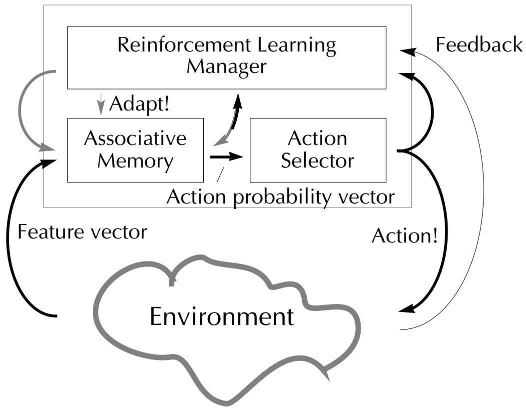

figure 2 : Reinforcement learning

the decision are exclusively extracted from the actual flow of laser-measurements:

- The length of the actually scanned edge.
- The sum of the absolute angles between the actually scanned edge-fragment.
- The mean distance between the edge and the background.
- The number of retries during the edge-scanning (the edge may be "lost" while the trace-procedures).

In order to keep the input space small, we use only these four statistical features, despite the fact that one may found a dozen other ones.

According to the reinforcement learning scheme shown in figure 2, we define an action probability vector consisting of the different possible actions attributed with the probability this action is to be chosen. The actions used here are: "Stop scanning immediately" and a couple of "Try another x elementary measurements, before asking me again". The probability attributes are generated by an associative memory module. We use a dynamic supervised feature map system (Growing cell structures [8]) instead of multilayer backpropagation or cascade correlation, because of high stability and adaption speed.

To allow trials with the several strategies, the final action selector determines the action to test by the action probabilities and a random number generator. This prevents the action with the highest probability to be chosen always.

The feature vector, the action probability vector together with the selected action, and the efficiency feedback is used to adapt the mapping in the associative memory. The probability of the selected action is increased resp. decreased by a certain factor according to the efficiency feedback and finally the other probabilities are scaled in order to keep the sum of all probabilities at one.

The number of found and useful edges is increased in any of our tests applying this module, whereas the amount of improvement depends widely on the actual environment. Therefore we are limited to qualitative results here.

5. Conclusion

Our visual search system is concipated as a general framework for efficient adaptive control of highly focused visual sensors mounted on mobile systems. Encouraged by the significant improvements in comparison to recursive edge-following, ongoing work will show the feasibility for exactly specified benchmark problems.

References

[1] Subutai Ahmad, Stephen Omohundro - 1.7.1991
Efficient Visual Search: A Connectionist Solution
Paper at the 13th Conference of the Cognitive Science Society, Chicago

[2] Subutai Ahmad, Stephen Omohundro - 25.7.1990
A Network for Extracting the Locations of Point Clusters Using Selective Attention
Paper at the 12th Conference of the Cognitive Science Society at MIT

[3] Jonathan A. Marshall - 01.07.91
Challenges Of Vision Theory: Self-organization Of Neural Mechanisms For Stable Steering Of Object-grouping Data In Visual Motion Perception
Invited paper, in Chen S-S (Ed.), Stochastic and Neural Methods in Signal Processing, Image Processing, and Computer Vision, Proceedings of the SPIE 1569, San Diego, July 1991, 200-215.

[4] Jörg Bruske, Ewald von Puttkamer, Uwe R. Zimmer - 01.12.93
SPIN-NFDS Learning and Preset Knowledge for Surface Fusion - A Neural Fuzzy Decision System -
Proc. of the ANZIIS ´93, Perth, Western Australia, December 1-3, 1993

[5] Herman Keuchel, Ewald von Puttkamer, Uwe R. Zimmer - 01.08.93
SPIN - Learning and Forgetting Surface Classifications with Dynamic Neural Networks
Proc. of the ICANN ´93, Amsterdam, The Netherlands, August 1993

[6] Martin Stecher
Strategien zur Effizienzsteigerung bei der Erstellung eines geometrischen Weltmodells für AMR
Diploma-thesis – University of Kaiserslautern - Computer Science Department - 3/94

[7] Bernd Fritzke - 1.5.1993
Growing cell structures - a self-organizing network for unsupervised and supervised learning
Technical report 93-26, International Computer Science Institute, Berkeley