

A Knowledge-based Interactive Robot-vision System

T. Vámos
Computer and Automation Institute,
Hungarian Academy of Scs.
1111 Budapest, XI., Kende u. 13/17
Hungary

M. Báthor and L. Mérő
Computer and Automation Institute,
Hungarian Academy of Scs.
1111 Budapest, XI., Kende u. 13/17
Hungary

A robot-vision project is reported which incorporates several existing AI methods and some new results. The ambition of the project is a system which can economically complete various intelligent tasks within the scope of mini- and microcomputers. The tuned composition of the applied methods provides a new and powerful approach to RandD, engineering and workshop-operation.

1. INTRODUCTION

This paper reports on a robot-vision lab for RandD and for experimentation of specialized industrial applications. The system combines many well-known methods of recognition with some own ones /a fast template method, a quasi-parallel edge-finder etc./ but this combination provides a sophisticated goal-oriented approach for real-life problem solving.

2. PICTURE PROCESSING, RECOGNITION

The system is based on man-machine communication; the goal is the recognition of 3D industrial objects whose number is limited to 10-20 for each task. The objects are taught either by building a geometric model or by a computer-aided transformation of the real situations viewed during the learning period to create a standard model description similar to the artificial one. The camera input /Fig. 1/ is processed in overlapping pixels by a simplified version of the Hueckel-operator using only two linear templates. An estimator of the fit coordinates a weight-attribute to the stroke. These weights and some fuzzy-like parameters which are attributed to the higher picture-hierarchy are relevant elements of the heuristic search. The hairy stroke-picture /Fig. 2/ is used both by the brute-force shape esti-

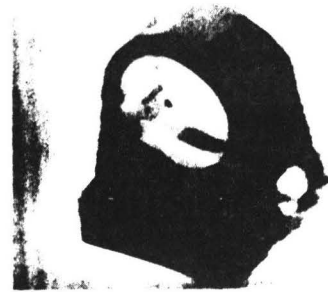


Fig.1. Digitized input /negative picture/

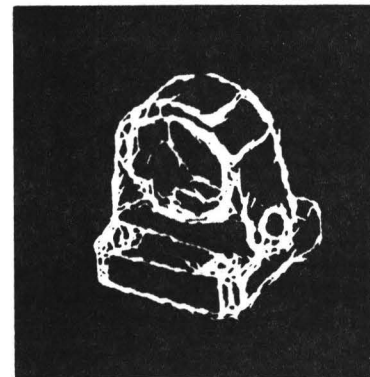


Fig.2. The strokes found in the picture

mators which provide a first guess and, by the further picture processing. Nodes

of the contour are supposed where the slope-variances of the strokes attain local maxima /Fig. 3/. Streaks along the contour lines are detected by a quasi-

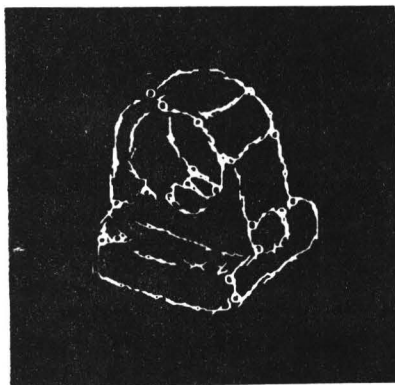


Fig.3. The streaks of strokes with the nodes found

parallel minimum-cost algorithm operating in parallel between every supposed node-pair connected by paths.

This algorithm is based on a mathematical model of the expected streaks. The requirements with the paths are expressed by four graph-theoretic criteria which mean that every contour-line should be optimally approximated by a streak and each of them by only one streak. A potential-field growing algorithm has been developed in order to find the streaks and it can be proved that the streaks provided by this algorithm satisfy the criteria of the mathematical model of the streaks. The streaks found using this algorithm are displayed on Fig. 3.

The next preprocessing algorithm merges the streaks either into line segments by uniting the approximately identically directed consecutive strokes or into arcs by comparing the slope differences. This interpretation is generally weak, due to the uncertainty of the whole process; thus a probability value is assigned to each possible interpretation, considering that each streak may have various interpretations /Fig. 4/. This value serves for the ordering of hypotheses in the linguistic recognition.

Matching the processed picture with the model and the whole knowledge base, decision on the ambiguities due to multi-

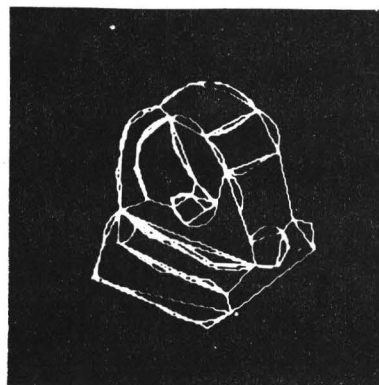


Fig. 4. All possible interpretations of the streaks

ple or failing edges and areas are solved by linguistic methods. Three accelerating heuristic tools are applied in directing the search on the problem-graph: a priori information, the weights with the calculated fuzzy-like estimation attributes of the picture hierarchy and the distance measures between pictures and models. The estimation attributes are calculated from the weights of the strokes, considering in each stage /node, streak, edge, picture part/ the reliability of the fit, decomposition and concatenation. Several distance measures are distinguished: brute-force characteristics with their special task-oriented combinations, distances similar to those used by Fu /3/, distances obtained from statistics of learning, special measures preprogrammed by the operator with the help of his personal experience.

The process of recognition includes the determination of orientation, namely the calculation of the transformation matrix which transforms the position of the model to the real position of the object. Modelling is used not only in the ways initiated by Roberts and later Baumgart and others but as a general tool for man-machine communication during i/ the system design, ii/ the task learning and iii/ the robot operation. Each level is different in intelligence; none of the levels can disturb its superior one but each has a systematic downward compatible interactive apparatus based on a graphic display. The graphics is 3D-2D oriented: the internal representation and manipulations refer mostly to the 3D model /Fig. 5/, the different 2D pic-

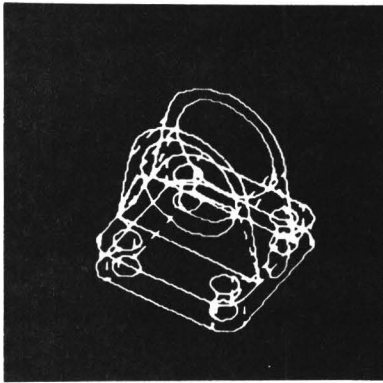


Fig. 5. 3D wire-frame model of the object

tures to be matched with the scene are derived from that using a fast hidden line elimination algorithm. This means that all usual 3D operations are implemented /rotation, distortion, unification etc./, adapted specially to our requirements, i. e. real time, low storage capacity.

3. EXPERIMENTAL RESULTS

Robot control and hardware are usual not to be detailed here. The system uses a 16 bit 40 kB mini with a 2 μ s cycle-time /VIDEOTON R-10 licence after the French CII-Mitra-15/, minidisc, magtape, graphic display.

The whole recognition procedure needs about 30-50 seconds to recognize an object shown e. g. in Fig. 1. The running times of the particular algorithms are the following. The hidden line algorithm is extremely fast, it provided e. g. Fig. 6 within 2 seconds. The edge-finding operator needs about 4 seconds for processing a TV picture input. The algorithm which assembles the strokes into streaks requires 10-15 seconds. The grammatical searching procedure of finding the possible faces and the possible object drawings in the streaks' interpretations takes 5-8 seconds. The other on-line algorithms take less than 1 or 2 seconds altogether.

Experiments have proved that the indicated heuristic search results in a time requirement depending statistically linearly on the object's complexity, number of objects, number of ambiguities /edges,

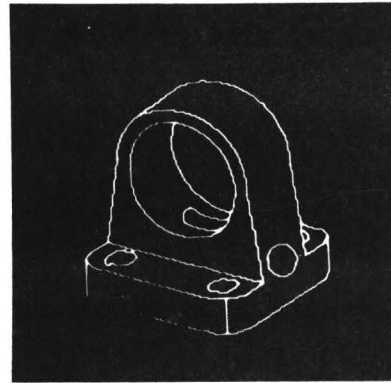


Fig. 6. A hidden line projection of the perfect model

interpretations, shadows, partial observations/. A successful guess /e.g. a brute-force estimator/ can cut short occasionally all further trials. In unfortunate cases the interaction of the operator is also possible.

4. APPLICATIONS

The system is primarily assembly-oriented. Other applications of our results that already work are being realized recently: a recognition system in a bus-body factory identifying the metal sheets for controlling a painter robot and a man-machine system in neuro-biological research which should detect neuronal nets in microscope-sections.

REFERENCES

- 1/M.Báthor:Hidden-Line Algorithm for a Robot Experiment.Ph.D.Thesis,Budapest, 1977
- 2/B.G.Baumgart:Geometric Modelling for Computer Vision.Stanford AIM-249, 1974
- 3/K.S.Fu:Stochastic Tree Languages and Their Applications to Picture Processing.Int.Symp. on Multivariable Anal., Pittsburgh, 1978
- 4/V.Galló:A Program for Grammatical PR.4th IJCAI /Tbilisi/, 1975, pp. 628-634
- 5/L.Mérő:A Quasi-Parallel Contour Following Algorithm.AISB/GI Conf.on AI,Hamburg,1978, pp. 189-194
- 6/T.Vámos:Industrial Objects and Machine Parts Recognition,in Applications of Syntactic PR. /Ed.K.S.Fu/,Heidelberg,1977,pp. 243-267
- 7/T.Vámos:Automatic Control and AI/invited survey/.Prepr.of 7th IFAC World Congress /Helsinki/, 1978, 4, pp. 2355-2369