tanulmányok 30/1974

THE BUDAPEST ROBOT
— Pragmatic intelligence —

by

Z. Vassy and T. Vámos

Computer and Automation Institute,
Hungarian Academy of Sciences
Budapest, Hungary

A kiadásért felel

Arató Mátyás

Abstract

A status—report outlines the basic ideas of a hand—eye project. A laser input uses the scanning microscope principle. The system is taught by an operator for special classes of industrial recognition and assembly tasks, after that minicomputer—controlled robot uses various methods (brute—force, grammatical—bayesian heuristic etc.) combined and separate depending on the task's speciality and complexity. Teaching, feature extraction, recognition, hand and input—device control are parts of a user—extensible software—system with a common dictionary—data base for semantic information. A motion picture is intended to be presented.

1/ Introduction and basic philosophy

The new breed of robots in the last decade brought this old utopistic idea in the real closeness by recent advances in technology and software science. Needs and approaches are very different, a simple fixprogrammed grasping manipulator or a highly intelligent self—contained man—like system working in the space under unknown conditions far enough to get a direct control from the earth.

The project to be reported, is a typical approach of a fairly developed small country:

a) to reach an economically feasible solution to replace some class of missing working power with a real industrial start in the next ten years — i.e. an extrapolation of requirements, prices and wages of this period;

b) to integrate every solution and idea of worldwide progress within the limits of the above conditions and in the frames of a technically systematic realization;

c) that is, to do this job without too much preconception but with an open system, which can be used in an experimental way as soon as possible, but extended with new ideas sometimes contradictory with the earlier ones.

To realize these, we applied the following graspe in the maze of the problem:

d) a system hierarchy should be used depending on the problem's difficulties, i.e. we try to solve the task on a medium configuration minicomputer with one minidisc, this is the

economically feasible basis, the mini has an access to a medium scale computer and a man–machine–dialogue should do the rest of the job. This yields an optimization of efforts and resources and avoids exaggerated overcomplications.

e) much effort has been concentrated on input hardware, because the kernel of the system is the contour–detection of the object to be handled. The basis of every pattern recognition and action–decision is a list of contour–primitives, this could be fed into the system as definitely as possible.

f) the feature extraction being the starting point of every further analysis must be a very reliable but simple, relatively fast action.

g) the user of the system should have for disposal a bunch of various software devices to solve his specific task, depending on the complication and other circumstances. These procedures should all be based on the contour–primitive list, use the same data base, can easily be changed during the teaching period of one scenery, can have a combined application possibility, the user should be able to extend the procedures library.

h) concerning the industrial reality, a sophisticated, man–machine interaction is provided during the teaching of one special task (one scenery of action), where the user has a wide selection of hardware configuration and software procedures. The consolidated working period should use as simple and ragged system as possible; if it has any unwaited difficulties, refers back to the user.

For our research area the input and the software–system are the challenging fields, the hand itself is – until now – a practical choice from the available means.


2/ The input

Our first experiments were based on a closed–loop TV. The contour–finding was facilitated by a variable level discriminator giving an output of two levels (black and white). Illumination could also be altered. The noise on input date was too high, only very specially defined environments could be processed.
Based on the results of Stanford AI group and the Tokyo Electrotechnical Labs. (1, 2), using laser beam illumination, a very accurate contour detection could be realized. Additions to the cited predecessor–solutions:

a) the laser beam is deflected by a programmable 1000x1000 point acoustic deflector addressed by the computer. The deflector was developed by the Institute's holography–group (3, 4).

b) the system uses the principle of scanning electron microscope, with simple photodetectors placed in various directions.

c) with a few number of parallel operating photodetectors (minimum 3) we get a very substantial information on the illuminated spot, an optical information word.

d) extending this idea, a simple device is given for the calculation of the object's surface point spatial coordinates.

e) using a screen before the photodetector, calibrated for the laser's wavelength, the operations can take place in full daylight.

Fig. 1a/ shows two objects of very different reflecting surfaces and
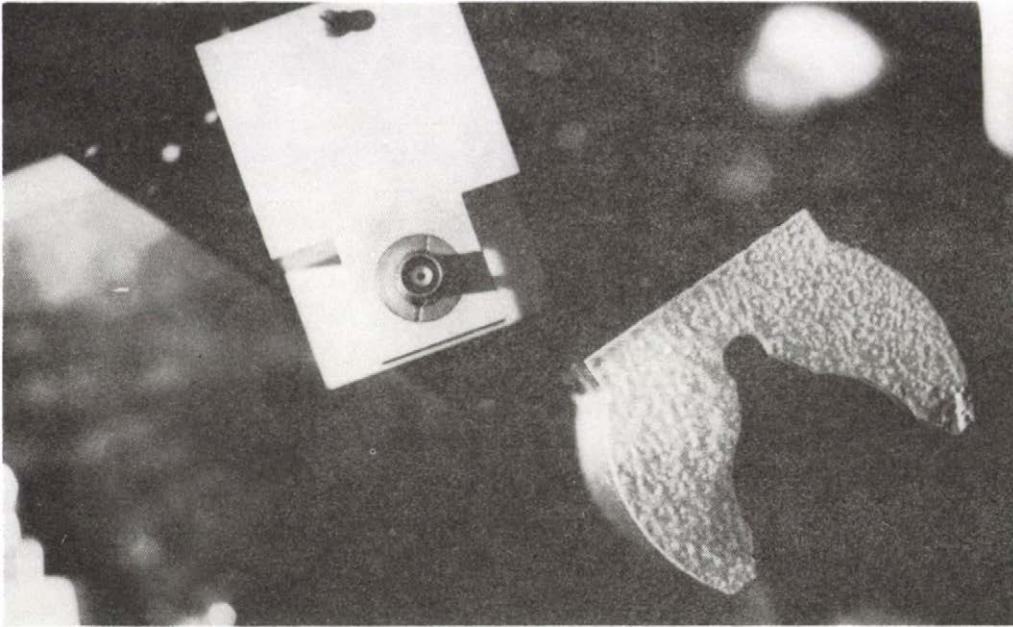
Side–view
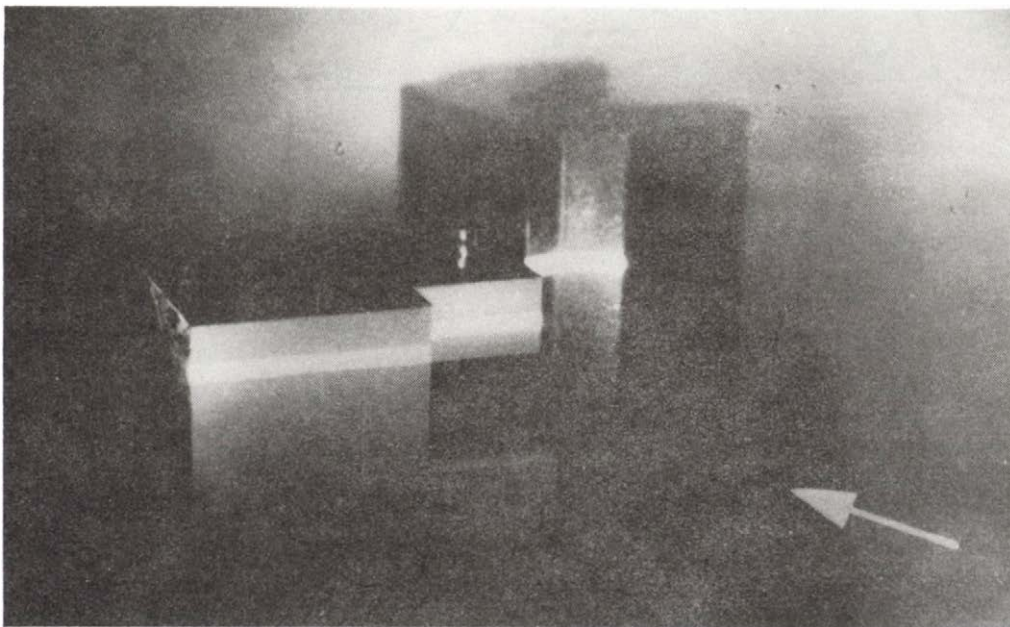
Top–view

Figure 1/a

1/b the well recognizable salient points of the edges on the photodetector's oscilloscope.
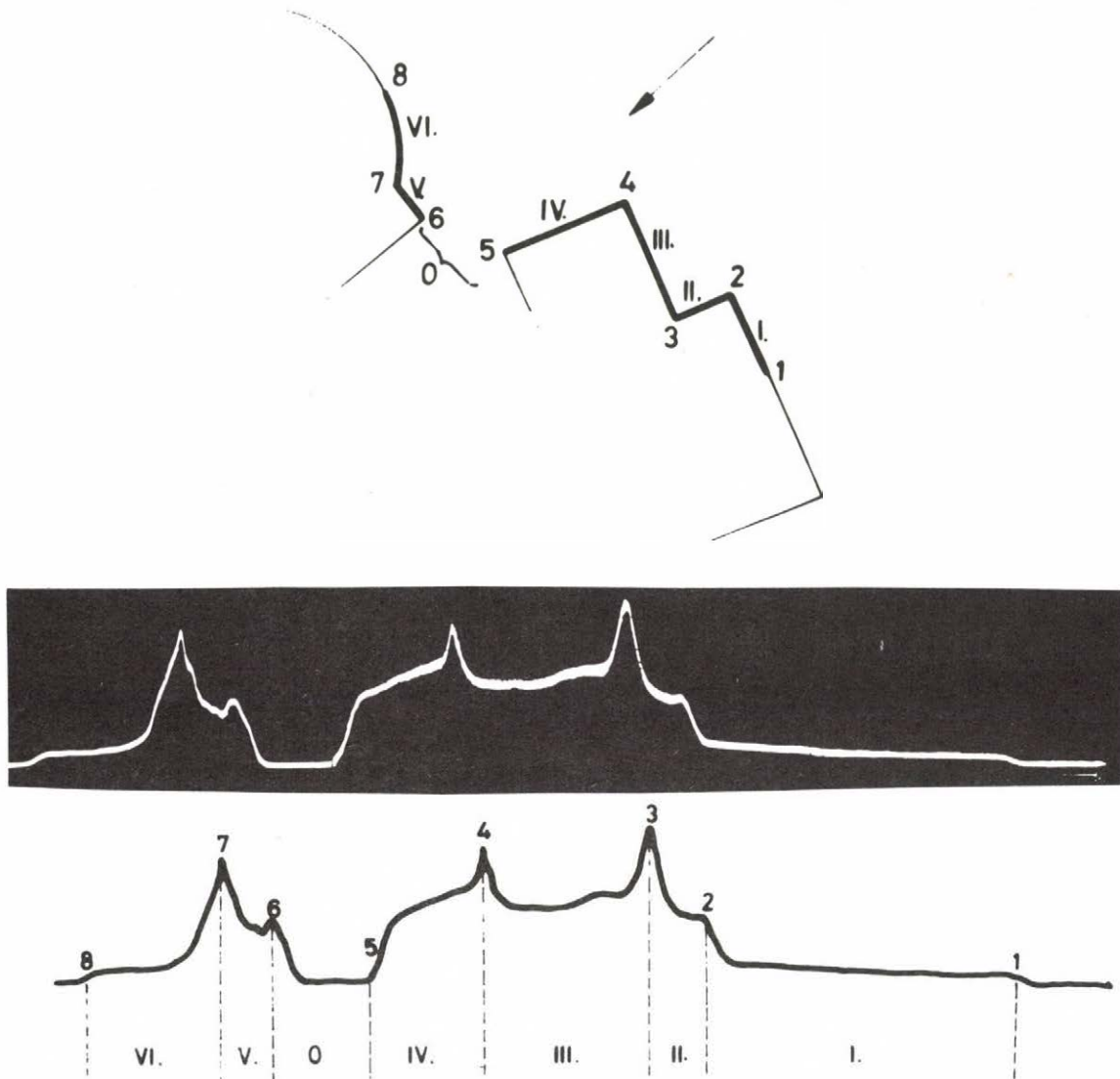

Figure 1/b

3/ Feature extraction

This has been reported earlier (5). Short recapitulation: only four kinds of primitives are used: straight line, arc, node, undefined. The result of this procedure contains a base list of node coordinates, node– pointers to joining nodes, primitive names, centre–coordinates attached to arcs.

4/ The system concept

The software system was based earlier on a CII 10010 mini. A Data General NOVA–1200 with a 16 x 16 w core memory and a 2 MB disc is used now. Fig. 2 indicates the scheme.
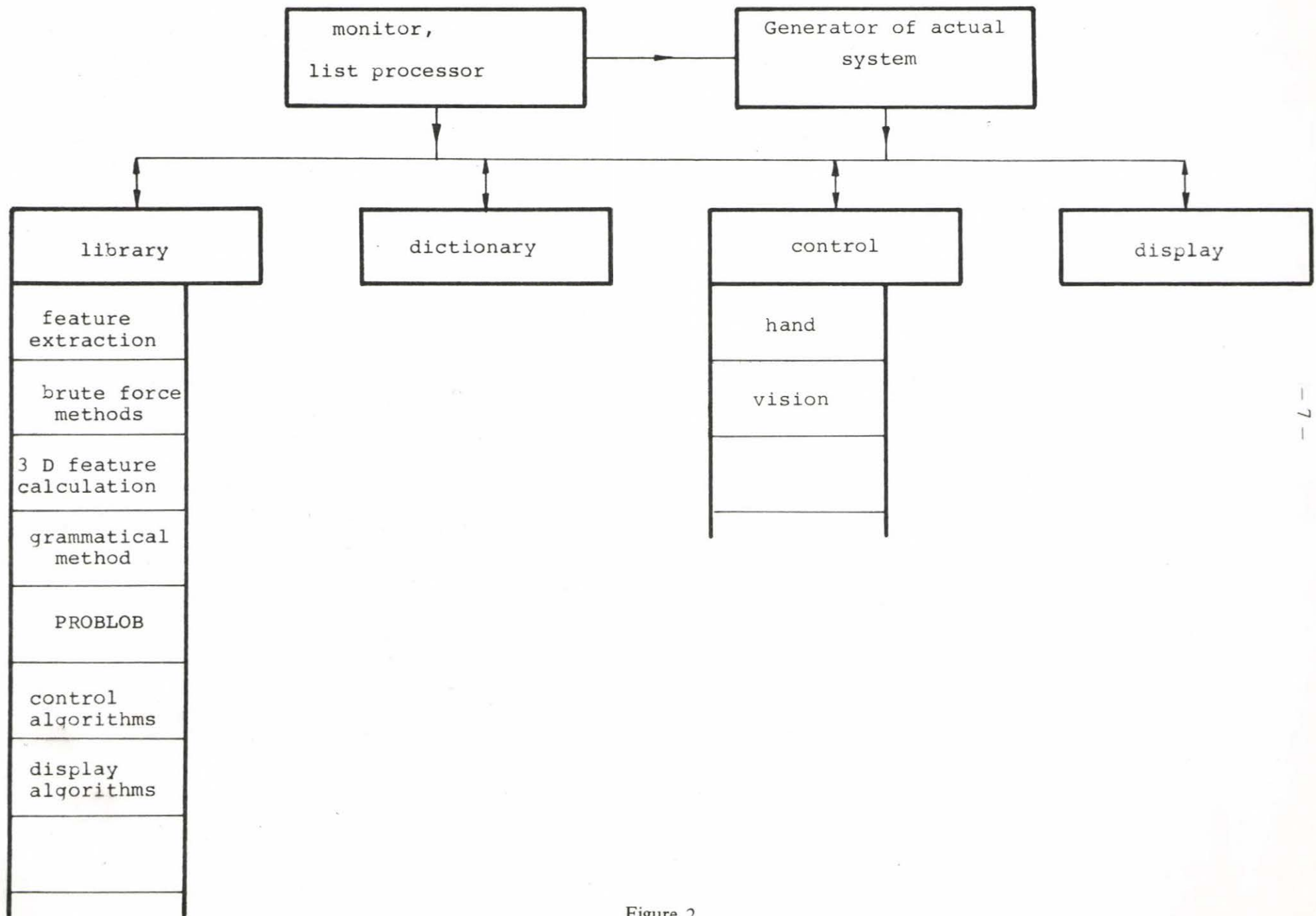
Figure 2

It containes a monitor used by the teacher–operator. Man–machine interaction is done by a display, which can also visualize the recognition phases (e.g. contours drawn by the feature–extraction base–lists etc.). The monitor handles the library residing on the disc. The library contains all the procedures to be used, e.g. the feature extraction routines, the brute–force method (see later), grammars, other means of recognition, hand–control routines, display routines. A dictionary with semantic information is also added. The system should have a Generator of Actual System, which compiles all the procedures, information, lists and open files to be used after the learning of one specific task for the actual job. This should be organized by the principles of § 1, entry h).

The operator–system dialogue is especially provided for the teaching period, but is partly maintained for the working phase too, when the system comes up to its intelligence–limitations, or any major change is foreseen by the supervisory control.

5/ Brute–force method

The degree of sophistication should be matched with the depth of the problem. A simple attack can be successful in a short time, if the objects to be recognized are sufficiently different; if not, the class of objects can be preselected as a shortcut for the more powerful methods which are at our disposal in the library (§ 1, entry g). This pragmatism is very similar to the human recognition strategy. The method takes the basic–list of § 3 as a start, computes the magnitude of the line–length and the angles of the contouring figures. Taking the shortest primitive as a basis, the contour is reconfigured by lines of equal length rounded off integer multiples. The straight lines and the arcs have different weights, a node also depending on the edges starting from the node. The angles have also an integer code, changing by a nonlinear scale, which stresses more the right angles, sharp vertices, apices, as characteristic changes of directions. As angles between arc sector, we take the deviation of chords. All these very short computations yield two new lists of integers: weighted lines and weighted angles. Let us call them secondary base–lists. A comparison of this list, ratios of sums and other far simple examinations, informs about the articulation, compactness, oblongation, rotundness, apices and other qualitative differences (Fig. 3).
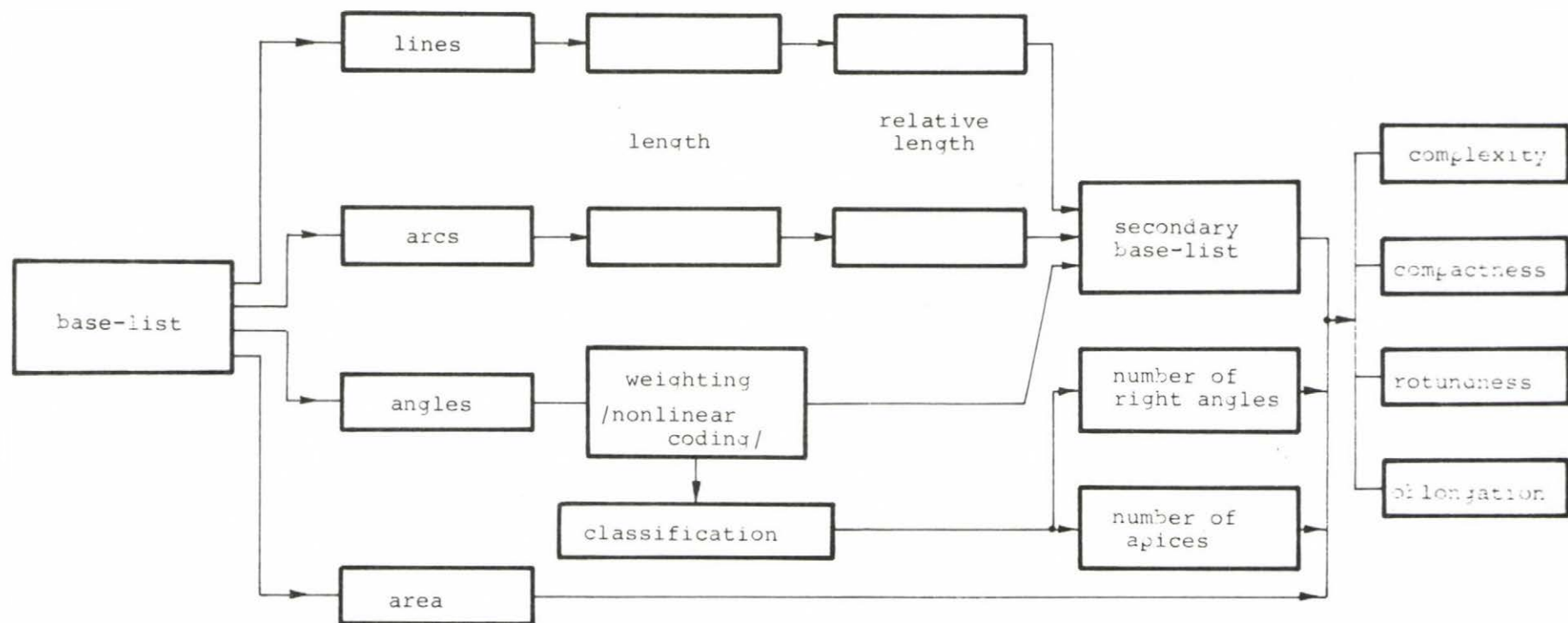
Figure 3

Such rough comparison routines can be created by the user during the learning period, looking at the visualized secondary base—lists. Computer—time of 6 different characteristics for about 100 primitives containing contours, lasts only a few seconds. The strength of the brute—force method is very much combined with its trivial weakness. Just the above characteristics (angles, oblongation etc.) are the most sensitive for the direction of view. This means that only the combined methods can be reliable.

6/ Grammatical analysis

The recognition of a pattern can sometimes be directly reached by a brute—force method, or can be facilitated by restricting the further search to some plausible choices. Nevertheless a deeper structural analysis either for recognition or for the directing of the robot's activity is in most cases inevitable. Our method is based on the general form of rewriting rules suggested by Evans (6). The essence of his idea is, that the set of the possible geometric relations is not limited by the formalism of the grammar; these can be built in any kind of relations defined as subroutines or logical functions. This offers a very useful flexibility for the system. In our experimental work so far, we defined about 10 types geometric relations, e.g. coincidence of points, parallelity or perpendicularity etc. of lines or chords of arcs, relations between the length of two or more lines, relations between the extension of some substructures etc. A special list processing language, LIDI—72 (7) developed in our institute originally for automatic circuit design, is used for programming the grammar and the grammar handling algorithm. The grammatical analysis does not work autonomously, i.e. the system generally does not run as a syntax—directed one, instead we use it as a special tool among others that is mobilized when it seems adequate to the given task.

7/ Structural levels of the description of objects

The first level is a list of *picture primitives* (see § 3), i.e. lines, arcs, modes and undefined elements all attached with their quantitative attributes. Extraction of these is common and unified in all cases. Second level built on the list of primitives, is that of generalized picture primitives (GPP-s). GPP-s are simple but characteristic substructures containing 2—5 primitives. They are fixed for a given pattern class — e.g. for machine parts, see Fig. 4 —, recognition and description of them is done by grammatical analysis. The user can write their grammar in the LIDI—72 input language. GPP-s also have quantitative attributes that can be used in synthesizing higher—level structures. Third level is that of so—called *lobes*.
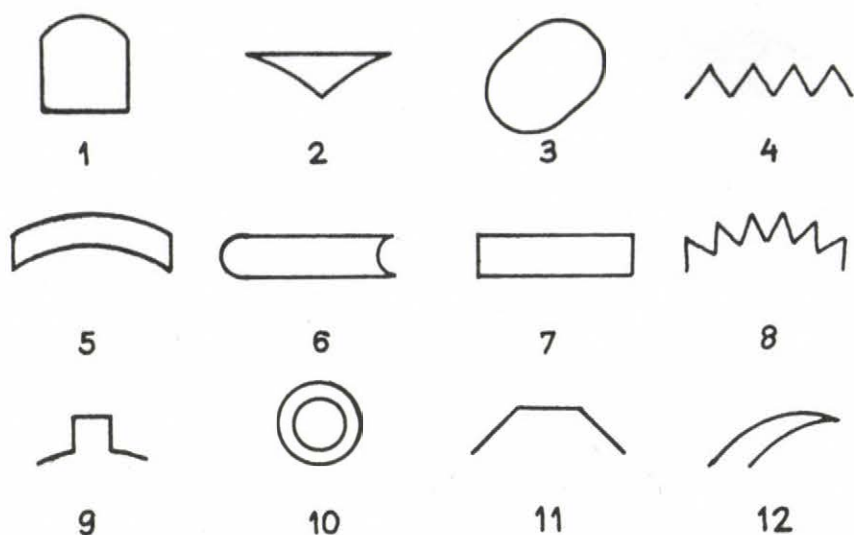
Figure 4

Every lobe is an unordered list of GPP-s contained by a given object. Relations between the GPP-s are not taken here into account. A lobe is highly characteristic for an object, at least if the GPP-s are properly chosen, but as an effect of noise, shadows, alterations etc. it may generally occur, that an object has more lobes with different probabilities. For this reason in a learning phase the system constructs the field of conditional probabilities of that a given object can be identified by a given lobe (Fig. 5). In the recognition phase the PROBLOB heuristic program – Preliminary Recognition of OBjects by LOBes – issues a list based on the learned conditional probabilities and the lobe actually found on the picture, the so–called "measured lobe". This last tells to the organizer program which objects or group of objects are most probably seen on the picture. By the aid of this heuristics the system can avoid the vast and unsuccessful effort to synthesize structurally one by one a number of objects that are known in its world, instead it can with high probability start probing immediately in the best direction. These heuristics have been motivated on the ideas of Fu (8) on stochastic languages and Zadeh (9) on fuzzy sets, although much modified for the special purpose. In details we refer (5).
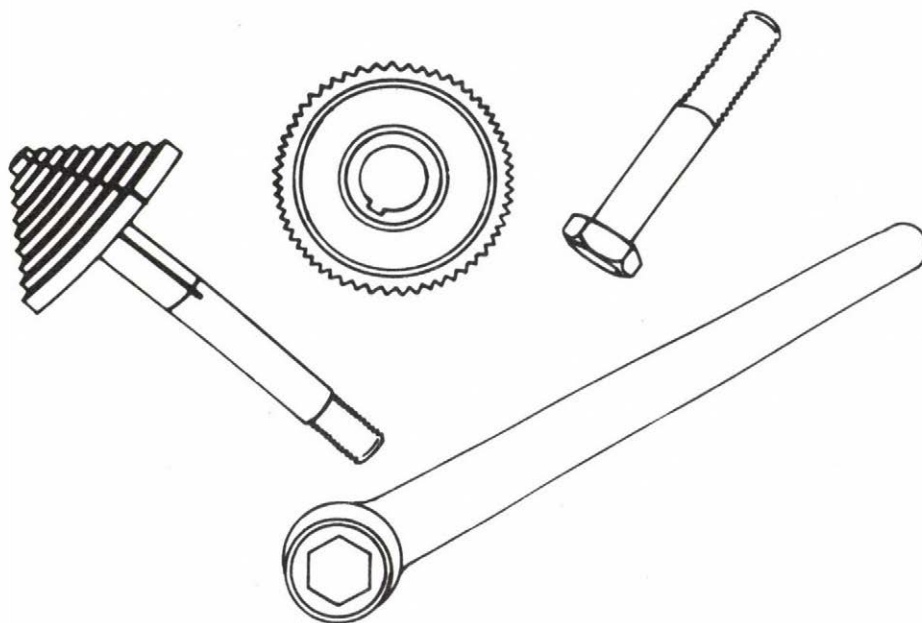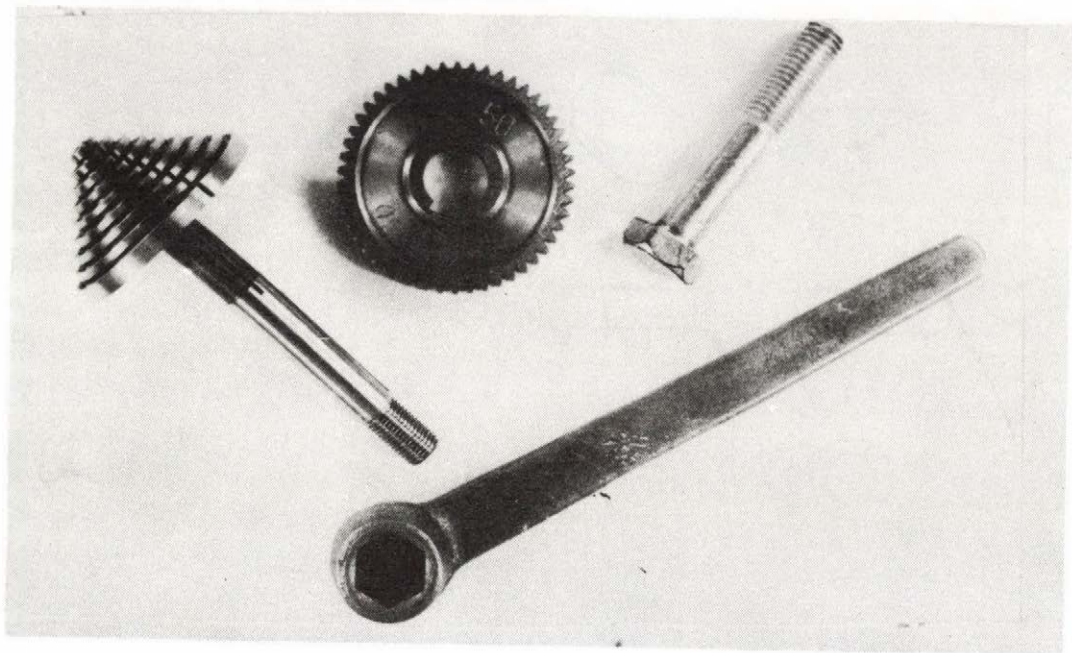
L1 = /1, 2, 3, 4/
L2 = /1, 3, 4/
L3 = /4, 5, 6, 7/
L4 = /4, 5, 7/
L5 = /8, 9, 10/
L6 = /6, 11, 12/
L7 = /6, 11/

Conditional probabilities after
learning:

p/H, L1/ = 0,81
p/H, L2/ = 0,18
p/H, Others/ = 0,01
p/R, L3/ = 0,65
p/R, L4/ = 0,22
p/R, Others/ = 0,13
p/G, L5/ = 0,96
p/G, Others/ = 0,04
p/K, L6/ = 0,43
p/K, L7/ = 0,49
p/K, Others/ = 0,08

/H/ Hexagonal head screw;   /R/ Ring fastener;
/G/ Gear;   /K/ Key-wrench.

Figure 5

Synthesizing the objects' geometric structure is the fourth level of the description. The system is not interested in the whole structure with all details, because for its practical aims only a few well—defined details are essential. The role of the description of the objects' structure as a whole is merely to give a possibility of finding these important details. For this reason the geometric structure of the objects is defined only in terms of GPP-s, i.e. by relations between the GPP-s' quantitative attributes. This structure thus can be imagined as a special map on which only the GPP-s are placed. All other details can then be located relative to the coordinate system defined by means of the GPP-s' attributes. The whole operation is very similar to the human strategy; if we find a very big creature, the conclusion can be drawn with a high probability that it is an elephant or a rhinoceros; the next glance without going into any other detail, concentrates on the place, where the proboscis has to be.

8/ Dictionary

This is the semantic data base relying very much on the results of Winograd (10) but not so much conversation oriented, but for robot activity. It should contain the names of objects (and similar tools), various descriptions (brute—force characteristics, grammatical), hints on the method to be used, subroutine names and data for handling (calculation of the centre of gravity, grasping, locating), class relations (basic features of scenery, related objects, e.g. the search for a nut and a key—wrench for a screw), essential pre- and post—history of the object (location, working etc.). (Fig. 6)

DICTIONARY		
IDENTIFIERS		
CLASSIFICATION	OBJECT CLASS	
	SCENERY	
	RECOMMENDED METHODS	
	OTHERS	
DESCRIPTIONS	BASE LIST	
	GRAMMAR	
	CHARACTERISTICS BRUTE FORCE	
	OTHERS	
ASSOCIATED IDENTIFIERS	RELATED OBJECTS	
	TOOLS	
	OTHERS	
CONTROL PROCEDURES	HAND	
	CENTRE OF GRAVITY	
	GRASP	
	VISION	
HISTORY	LAST POSITION	
	LAST ORIENTATION	
	USUAL POSITION	
	OTHERS	

Figure 6

The dictionary has also two editions as to § 1, a more comprehensive one for one working area with a number of similar successive tasks and a minimalized abbreviation for one scenery after the teaching period for one specific task.

9/ Hand

As indicated before, this is the last step of the project. A broad class of practical hands are developed in the Institute both pneumatic and electrical, to cover the various needs of industrial automation. One or two of these is to be applied in this project.

10/ Conclusions and comments

A laboratory experiment for a hand—eye system shows the feasibility of intelligent minicomputer—controlled robot as a practical and economical tool in the next few years. A laser—photodetector input appears to be superior to the TV camera offering also the possibility of obtaining more precise contour detection and information on 3 D features. A practical man—machine interaction helps in the restriction of the task to obtain a limited—intelligence device. A pragmatic combination of various pattern recognition methods results in a new approach for a user and application oriented, flexible and extendible software.
A film presenting both the ideas and the experimental tools is intended to be shown at the Congress.

11/ Acknowledgements

The authors express their gratitude to Mrs. V. Galló for the construction of the grammar and the grammatical analysis of algorithms.

Reference

[1] Ishii, Y., Nagata, T.: Feature Abstraction of Three Dimensional Objects by Laser Tracker.
Preprint of 12th SICO Annual Conference, 1973, p. 269 (in Japanese)

[2] Agin, G. J., Binford, Th. O.: Computer Description of Curved Objects.
3rd International Joint Conference on Artificial Intelligence, Stanford,
Aug. 1973, pp. 629—635.

[3] Tőkés, Sz. et. al: An acousto—optical deflector. Internal report of Computer and Automation
Research Institute, Hungarian Academy of Sciences.

[4] Tőkés, Sz. et. al: Optical 3 D Sensor for Robots. Internal Report of Computer and
Automation Research Institute, Hungarian Academy of Sciences.

[5] Vámos, T., Vassy, Z.: Industrial Pattern Recognition Experiment — A Syntax Aided Approach. Proceedings of the 1st International Joint Conference on Pattern Recognition, Washington, 1973.

[6] Evans, T. G.: A Grammar—Controlled Pattern Analyzer. IFIP Congress '68, Appl. 3., North—Holland Publ. Comp., Amsterdam, 1968.

[7] Uzsoky, M., Fidrich, I.: LIDI—72. List Processing System. Computer and Automation Institute, Hungarian Academy of Sciences, Report No. 16/73.

[8] Huang, T., Fu, K. S.: On Stochastic Context—Free Languages. Information Sciences, Vol. 3 (1971), pp. 201—224.

[9] Zadeh, L. A.: Fuzzy Sets. Information and Control, Vol. 8, 1965, pp. 338—353.

[10] Winograd, T.: Understanding Natural Language. Acad. Press, New York, 1972.