INTELLIGENT MANIPULATION TECHNIQUE FOR MULTI-BRANCH ROBOTIC SYSTEMS

Quarterly Progress Report

for Period September 1, 1990 - November 30, 1990

Alexander Y.K. Chen and Eugene Y.S. Chen

Scientific Research Associates, Inc.
50 Nye Road, P.O. Box 1058
Glastonbury, CT 06033

Prepared for

Jet Propulsion Laboratories
National Aeronautics and Space Administration
Pasadena, CA 91109

December, 1990

(NASA-CR-187651) INTELLIGENT MANIPULATION TECHNIQUE FOR MULTI-BRANCH ROBOTIC SYSTEMS Quarterly Progress Report, 1 Sep. - 30 Nov. 1990 (Scientific Research Associates) 26 p N91-14603
CSCL 13I G3/37 Unclas 0319016

Summary

To our knowledge, the following technical items cover all the attainable subjects which can be addressed in this program. The proposed robotic system has two major advancements in mechanical design. The modularization consideration attempts to incorporate the actuating device directly at each active link. The investigation of the modular approach is very important to fully utilize the flexibility of reconfiguration capability of general robotic systems. The second advancement is to organize multiple redundant degrees of freedom in a systematical fashion to extend the dexterity of the robotic system. With fourteen active actuators to form the head and two arms, the resulting upper body of the proposed robot presents an electromechanical system with unprecedented complexity, in which the proposed spatial planning technique (i.e. AISP) will be employed to drive the system.

A novel approach for developing real-time robot vision is one of the main results under the framework of SANE (Sensor-Actuator NEtwork). Dual CCD cameras and one frame grabber, together with two rotational degrees of freedom constitute the basic apparatus for prototyping the proposed vision system. Multiple ultrasonic sensors and proximity sensors would build up a sensor data fusion network which can directly and indirectly interface with actuators. The resulting robot behavior will be the first prototype to demonstrate the functionality of SANE.

Dynamic Knowledge Evolution (DKE) is the main thrust of integrating the artificial intelligence technique to generate the learning capability in robot intelligence development. The traditional robotics techniques such as spatial planning, kinematic planning, dynamic calculation and control, error correction and fault tolerance, will be utilized as the building blocks of passive robot knowledge base. Submission control architecture and neural network typed parallel distributed processing (PDP) are two possible ways of engagement in system integration. Detailed development is part of the on-going process of software engineering.

The final advancement of this program is the construction of a versatile MuMicS (Multiple Microcomputer System). For realizing the GIC-LICs robot intelligence system, multiple microcomputers need to be linked together. The related data/address bus structure presents a unique solution for performing various serial/parallel operations in DMA, cache memory, coprocessor access, RAM/ROM memory management, timing and clock, and I/O interface.

Software development for achieving the project's goals, as were declared in the last quarterly report, is the focal point of this period of R&D activities. In order to maintain sufficient flexibility for accommodating various applications, many related technology fronts were tackled, and different levels of progress at various areas have resulted. The efforts required for establishing the required development framework are beyond what was originally anticipated. However, despite limited resources, it is believed that the best effort has been made to catch up on the planned schedule.

During this period, the upper body of SRAARS (SRA Advanced Robotic System) has been completely assembled. The upper body includes two articulated arms and a robot head with two CCD cameras as the robot's eyes. There are fourteen degrees of freedom in the upper body part, which are actuated by ten DC motors, two

stepping motors and two solenoids for controlling the pneumatic power. A temporary mechanical stand is currently under construction to support the upper body mechanism. The motion control electronic development is behind schedule due to the overall consideration of MuMicS, which has been reviewed several times but which has not yet come to any final decision. The scarcity of available force/torque sensors and tactile sensors has also contributed to the delay of finalizing the hardware implementation of SANE framework. However, some of this information has just arrived, which will be very helpful in the pursuit of the finalization of SANE realization in the near future. As for the robot vision development, the major bottleneck of current work is again the software development and the integration with other parts of the robotic system. In all, there have been some results in the analytical development, software engineering, system design and hardware implementation.

Analytical Development

The analytical development of the intelligent manipulation technique during this period concentrated on two sections: the kinematics planning and the robot learning. As a bridge between spatial planning and dynamics control, the intelligent kinematics planner (InKiP) should function as a operational buffer to smoothing out the transition from the position control (low impedance) to the force control (high impedance). Moreover, there are essentially three modes in robot manipulation: the inertial mode, the resistive mode and the capacitive mode. The InKiP design should have the capacity of flexibly shifting from one mode to another without causing any disturbance on the system kinematics. When the manipulation is in the inertial mode, collision avoidance is not concerned. There will be sufficient free space for arbitrary kinematics planning. Even with potential obstruction of sight, the determination of collision-free traveling path will be dealing with the case of open convex set. For the resistive mode of robot manipulation, the end-effectors are either very close to or have physical contact with the interacted objects. For those cases of physically contacting the environment, the external interactive forces are limited to the order of frictional effects in the resistive mode. When the robot manipulation is entering the capacitive mode, the external interactive force or torque becomes an integral part of the active force control, which may play an important role for accomplishing the assigned tasks. The manipulation of human arms, hands and fingers is a typical example of abling to successfully transform from one mode to another without abrupt kinematic fluctuations. However, the abrupt fluctuation of the associated kinematic profile does happen in the behaviors of infants, elders, or the physically challenged. Hence, the successful development of InKiP not only would advance the performance of the robotic system, widen the applicable range of robotics, but also would improve the functions of manipulation aids for the elders or the physically challenged.

A. The Intelligent Kinematics Planner (InKiP)

The analytical content of InKiP construction has two major components: the kinematic spline theory and the adaptive logic annealing process. As commonly known in robotics research, there are two domains for planning and execution of robot manipulation. The three-dimensional global work space describes the actual robot maneuvers and the interface with external environment, and the N-dimensional joint variable space illustrates the states of the employed control mechanism.

The mapping between these two spaces is not a one-to-one mapping in general. However, by utilizing the uniquely determined relationship between the first-order time derivatives at both domains through the transformation Jacobian, the resolved motion rate control was then developed. However, the computational requirement for calculating the Jacobian matrix of each point along the motion trajectory is substantial, which makes the real-time trajectory planning unfeasible. This restriction becomes more detrimental when the number of degrees of freedom of the robotic system and the complexity of the system's configuration increase. A certain interpolation operation would have to be employed such that part of the trajectory planning workload can be carried over to the joint space. Since the joint space is the direct control domain, an on-line planning function can be readily implemented and its effect on the control stability can be easily assessed. As long as the resulting performance in the global work space is within the acceptable tolerance range, the planning of the traveling path is preferable in the joint space than that of the work space. If the required robot performance in the work space is specified by a sequence of three-dimensional locations without timing-dependency, then employing various existing trajectory planning techniques in the joint space would be sufficient. But, if the required motion trajectory of the robotic system in the work space is timing dependent or even timing critical, then there are two choices for performing the trajectory planning. One can plan the path with kinematic consideration in the global work space first, then transform the results into the corresponded joint space, which may encounter some nonsmooth or unfeasible joint variable profile; or one can plan the kinematic profile of motion trajectory in the work space with the consideration of the kinematic smoothness in the joint space. The kinematic spline theory is developed to realize the later proposition.

<u>(1). The Kinematics Spline Theory (KST)</u>

The principle of the kinematics spline theory (KST) is to adequately patch the piecewise trajectory planning elements such that the continuity and the differentiability can be maintained in the spatio-temporal domain. Moreover, for the consideration of robotics applications, the continuity and the differentiability of the motion trajectory will be examined in both the global work space and the joint space. The minimization procedure of AISP does not have a specific interpolation strategy, which makes the planned IKP solution unpredictable. The developed KST will be used to fill the gap to carry out the dual-space trajectory planning task. In general, a segment of trajectory in which one of the end effector is traveling is a segment of space curve with temporal correspondence. The simplest parametric representation is

$$g = g(u), \quad u_0 \le u \le u_1,$$

In the three-dimensional global work space,

$$g = g(x, y, z) = g(s)$$

where

$$ds = \left[dx^2 + dy^2 + dz^2 \right]^{\frac{1}{2}}$$

If the initial position is $S_0[x_0,y_0,z_0]$ and the final position is $S_1[x_1,y_1,z_1]$, then the line integral

$$g(s) = \int_{a}^{s} \nabla(g)\ d\alpha$$

$$a = \left[x_0^2 + y_0^2 + z_0^2\right]^{\frac{1}{2}}$$

$$\nabla(g) = \text{gradient of function } g$$

$$d\alpha = \left[\ \frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}\ \right]$$

In the N-dimensional joint space,

$$g = g(j_1, j_2, \ldots., j_N)$$

and for the temporal consideration,

$$g = g(t), \qquad t_I \le t \le t_f$$

The uniqueness of this segment of trajectory makes the following mappings uniquely determined:

$$u <-> s <-> (x,y,z) <-> (j_1, j_2, \ldots, j_N) <-> t$$

From the system dynamics point of view, unless the robot is in a complete halt mode, the spatio-temporal correspondence of every two adjacent points is dependent upon the kinematics of the predecessor point. Some of the common factors in determining this dependency include the moment of inertia of both the robotic links and the actuation devices, the capacity of those actuators, frictions, the gravitational force, the Coliolis and the centripetal forces. Since it would be impossible to constantly evaluate and measure all these terms for planning a feasible trajectory, the KST is applied to determine the transition from one control point to the next control point, where the associated two point boundary valued problem with constraints on both the velocity and the acceleration levels is solved at joint space with the spline technique, and by determining the terminal kinematic profile of joint variables from the inverse kinematics mapping of the work space trajectory, which itself is generated by employing the spline technique in the work space.

For each joint variable, an interpolation of order 3 is applied to the acceleration profile with the estimated maximal and minimal accelerations one the knots. Due to the unique characteristics of the convex hull, the interpolated acceleration will be contained within the specified range. If the Berstein-Bezier approach is employed, then the acceleration function would by

$$A(u) = \sum_{k=0}^{3} A_k \phi_{k,3}(u)$$

where

$$\phi_{k,3}(u) = \binom{3}{k} u^k (1-u)^{3-k} \quad , \quad 0 \le u \le 1$$

Since u is the normalized time variable, and if the task is to reach D_f at time f_f with the initial D_I at t_I, then

$$u = \frac{t}{\Delta t} \quad , \quad \Delta t = t_f - t_I$$

Then, the kinematics profile would be

$$A(t) = \sum_{i=0}^{3} b_i t^i$$

$$V(t) = F_0 + \sum_{i=0}^{3} \frac{b_i}{\Delta t} \frac{t^{i+1}}{i+1}$$

$$D(t) = D_I + V_o t + \sum_{i=0}^{3} \frac{b_i}{\Delta t} \frac{t^{i+2}}{(i+1)(i+2)}$$

where

$$b_0 = A_0$$

$$b_1 = 3(A_1 - A_0)$$

$$b_2 = 3(A_2 - 2A_1 + A_0)$$

$$b_3 = A_3 - 3A_2 + 3A_1 - A_0$$

A_0 is the initial acceleration at t_I

A_3 is the final acceleration at t_f

A_1 and A_2 are the control knots which represent the effect of actuator limitations on force/torque generation.

In order to achieve the final goal of reaching D_f at t_f, the following equation results.

$$\frac{\Delta D}{\Delta t} - V_0 = \sum_{j=0}^{3} C_j A_j$$

where

$$\Delta D = D_f - D_I$$

$$C_3 = \frac{1}{20} (\Delta t)^3$$

$$C_2 = -\frac{3}{20} (\Delta t)^3 + \frac{1}{4} (\Delta t)^2$$

$$C_1 = \frac{3}{20} (\Delta t)^3 - \frac{1}{2} (\Delta t)^2 + \frac{1}{2} (\Delta t)$$

$$C_0 = -\frac{1}{20} (\Delta t)^3 + \frac{1}{4} (\Delta t)^2 - \frac{1}{2} (\Delta t) + \frac{1}{2}$$

A_0, V_0 are known initial conditions of joint velocity and acceleration. If the final given acceleration is A_3, and assuming the acceleration constraint has the relationship $A_2 = -A_1$, then A_1 can be uniquely determined as

$$A_1 = \frac{(\Delta D/\Delta t) - V_0 - C_0 A_0 - C_3 A_3}{C_1 - C_2}$$

Also, on the other hand, if the acceleration constraints (i.e., A_1 and A_2) are given, then the final acceleration (A_3) can be determined accordingly. For the special case of a complete stop at both end points, that is, $V_0 = A_0 = A_3 = 0$, the acceleration constraint can be computed as

$$A_1 = \frac{\Delta D}{\Delta T} \cdot \frac{1}{C_1 - C_2}$$

As for the trajectory planning in the three-dimensional work space, not only the cubic spline can be directly utilized to interpolate the path along each individual axis, but also the orientation of the trajectory can be determined by the Serret-Frenet formulas. As summarized in the following matrix format, the transition of the tangent, the normal, and the binormal vectors can be calculated by the kinematic profile of the trajectory.

$$\frac{dA}{ds} = A \cdot Q$$

where

A = $[\mathbf{n}, \mathbf{b}, \mathbf{t}]$, the path orientation matrix

$\mathbf{h}$ = the normal vector

$\mathbf{b}$ = the binormal vector

1t = the tangent vector

$$Q = \begin{bmatrix} 0 & -\tau & k \\ \tau & 0 & 0 \\ -k & 0 & 0 \end{bmatrix}$$

k = the curvature

τ = the torsion

k and τ can be determined as

$$k^2 = \frac{\left[\mathbf{x}^{(1)} \times \mathbf{x}^{(2)}\right] \cdot \left[\mathbf{x}^{(1)} \times \mathbf{x}^{(2)}\right]}{\left[\mathbf{x}^{(1)} \cdot \mathbf{x}^{(1)}\right]^3}$$

$$\tau = \frac{\left[\mathbf{x}^{(1)}\ \mathbf{x}^{(2)}\ \mathbf{x}^{(3)}\right]}{\left[\mathbf{x}^{(1)} \times \mathbf{x}^{(2)}\right] \cdot \left[\mathbf{x}^{(1)} \times \mathbf{x}^{(2)}\right]}$$

$$\mathbf{x} = [x(t),\ y(t),\ z(t)]^T$$

$$\mathbf{x}^{(i)} = \frac{d^i \mathbf{x}}{dt^i}, \qquad (\ \cdot\ ) = \text{the inner product}$$

$$(\ \times\ ) = \text{the cross product}$$

However, the corresponding trajectory in the work space with the interpolation in the joint space will not be identical to the interpolated trajectory in the work space. The interpolation in the work space is used as the guideline in searching for the "reasonable" trajectory for joint space interpolation, which becomes more important when the adaptive logic annealing process is employed.

Considering the work space is divided into several subspaces by the hypersurfaces which are mapped from the geometrically singular surfaces in the joint space, from the spatial relationship to the kinematics, the planning of the motion trajectory would have to move from one subspace to another. Guided by the piecewise polynomial approximation and constrained by the limitations on the kinematics profile, the resulted trajectory would have the minimum transactions between subspaces which can reduce the possibility of redundant oscillations.

Besides the dual-space trajectory planning, the developed KST can be used to derive the on-line trajectory modification scheme which is crucial for interactive mode switching. The resulted on-line trajectory modifier will generate stepwise interpolation to accommodate the latest control parameters and determine the next motion command.

(2). The Adaptive Logic Annealing Process (ALAP)

Planning the kinematics profile of the robot manipulation trajectory is a task with multiple possible approaches. Previous approaches of determining the profile with a specific optimization procedure or a heuristic search do not provide the sufficient spectrum for the robot to handle heterogenerous tasks. It requires an adaptive scheme to transform from one planning algorithm to another with smooth transition and consistent principles. The developing ALAP technique will meet this requirement. The considered methodology is the Self-Organized Neural Networks (SONN), which has three major components: the generating rule, the evaluation method and the search strategy. The entire logic of InKiP is modeled as a neural network structure. The generating rule would select the new function for a new node from a given set of prototype functions; the evaluation method will determine how well the current logic is performing in terms of trajectory planning; and the search strategy will provide other available alternatives if there is room to improve the system performance.

By utilizing the concept of maintaining the simplest functionality at each individual node, the ALAP will be suitable for developing decision machines with parallel distributed processing. For this particular application, the main integrity would be the development of geometric reasoning for the robotic systems with the consideration of system dynamics. The details of ALAP will be presented in the next quarterly report.

B. The Robot Learning

In order to manage the entire range of eight major operations defined as the robot brain functions, the robot learning mechanism has to be flexible, reasonable and sufficiently compact. More than often, the user's commands may not be precise, the encountered environment may not be identical to what has previously being measured, and some unexpected interruption may alter the defined operational sequence. An intelligent robotic system should be able to overcome these types of fuzziness up to a certain extent where the original assignment can still be partially recovered. A new technique is currently under development called FULOSONN, an acronym for FUzzy LOgic Self-Organized Neural Networks. The fuzzy concept includes formal analysis of the fuzziness in commands, control, searching and reasoning. The logic approach stands for the embedded expert system implementation with the rule-generation, the inferencing and the knowledge representation. The self-organized capability is formulated by the on-line modification of either the network topology or the connection functions, or both. The incorporation of neural networks would emphasize the flexible internal representation of the learning mechanism with or without supervision.

The fuzzy commands concept is a very friendly user interface. For advancing the coordination between the human operator and the robot, the development of recognizing the fuzzy commands is critical. A specific fuzzifier will be constructed to read these types of commands and then translate them into ranges of quantitative descriptions with proper probability consideration. By maximizing the associated likelihood function, the desired actions can be selected. The fuzzy control would provide the approximate control strategy to convert the processed fuzzy input into the defuzzified enforcing function. A combination of fuzzy control law and fuzzy logic process would normally be considered as the fuzzy controller design. Fuzzy searching is employed when a one-to-one mapping between the input domain and the output domain does not exist. Instead of the conventional heuristic search, the fuzzy searching considers the complete spectrum with fuzzily smoothing input information, which enlarges the probability of finding the feasible solution with high efficiency. Since this searching technique can translate fuzzy inputs into adequate executable outputs, it is also referred to as the "fuzzy inferencing". As for fuzzy reasoning, it is an augmented version of the hypothesis-and-test paradigm. Not only the postulated rules have certain probabilities of being true, but also the derived outcomes contain conditional tolerable grey areas. One major advantage of fuzzy reasoning is to provide a remedy to relax the rigidity of the traditional expert system logic. For many application cases, the capability of "justifiably bending the rules" may be highly desirable.

On the other hand, "common sense" is always the foundation of any behavior activation. The provision of logic stands for the operational rules for determining "what, when, where, who and how" of each action. As the organizer of executing multiple tasks, running multiple procedures and coordinating the usage of available resources, it has to be deterministic, predictable and yet adjustable to manage variable conditions. Disregarding it is defined by a single, thousands, or millions of rules, self-conflicting points and redundant loops should be eliminated. An efficient expert system would provide specific execution recommendation by processing the input information with induction/deduction and forward/backward chaining. Some of the most important logic processes include the linguistic and semantic logic processing, symbolic manipulations, the numerical computations and the stochastic processing.

The integrated framework of FULOSONN provides most of the required techniques to explicitly establish the robot's learning capability. Merging these four fronts, the on-board robot intelligence can incorporate the basic knowledge of robot manipulations via man-machine interface inputs and can establish the expert system for standard maneuvering, manage the fuzzy commands with fuzzy controller, activate the fuzzy searching for unknown situations, handle the conflicting cases with fuzzy reasoning, organize the knowledge representation in a parallel distributed fashion such that some constantly failing memory will be rectified in a superimposed structure of layers of neural networks, which has the self-organized capability to modify the network topology as well as the connection functions. A complete development of the FULOSONN framework is beyond the scope of this program. A simplified version of FULOSONN implementation will be developed to support the learning mechanism of the SRAARS prototype.

Overall System Design Update

After months of investigation and evaluations, the first MuMicS prototype is going to utilize the well-developed LAN (Local Area Networks) technology to connect the GIC and those LICs. The basic port communication (parallel or serial) is too limited to be useful. Other low end equipment interface specifications, such as RS232 or IEEE488, do not have sufficient control capability over the communication line, which would make the vital connections among GIC and LICs very vulnerable. Various bus structures, such as PC bus, PC-AT bus, STD bus, VME bus, MULTIBUS I or II, or FUTUREBUS, tend to be an internal connection arrangements, which are severely limited by the available hardware connectors and the insufficient system software support. In order to consider the expandability, the security of GIC-LICs communication and the integration of scalable local intelligence distribution, it was finally decided to employ the well-established PC LAN technology as the first implementation platform for demonstrating the MuMicS architecture. The considered PC LAN product is the Novell Netware. The selected connection would be a 16-bit Ethernet. Each subsystem will be equipped with a specific motion controller board, the driver board for operating various sensors and/or detection devices, some basic memory management circuitry and a network communication board. The procurement of the required hardware and software is expected to be accomplished within the next period.

Software Development

Conversion from C language programs to the object oriented C++ programs and the development of the embedded expert system, fuzzy logic and neural networks are two major targets in current software development. Both of them have extensive involvement in state-of-the-art technology digestion, searching for available tools and application design. Along these two lines of work, there is no specific result which can be presented at this moment. As an on-going process, the goal is that there would be some solid software modules being constructed in the next period.

The development of the robot vision system has been consolidated into one conference paper which had been presented at the SPIE's International Symposium on Advances in Intelligent Systems, November 5, 1990 in Boston MA. A copy of the paper is included in the Appendix.

Hardware Development

The completion of the upper body of the SRAARS is the major hardware accomplishment during this period. Figure 1 shows four different views of the upper body construction, Figure 2 shows a close-up of the robot head, and Figure 3 illustrates the evolutionary path of the SRAARS development.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

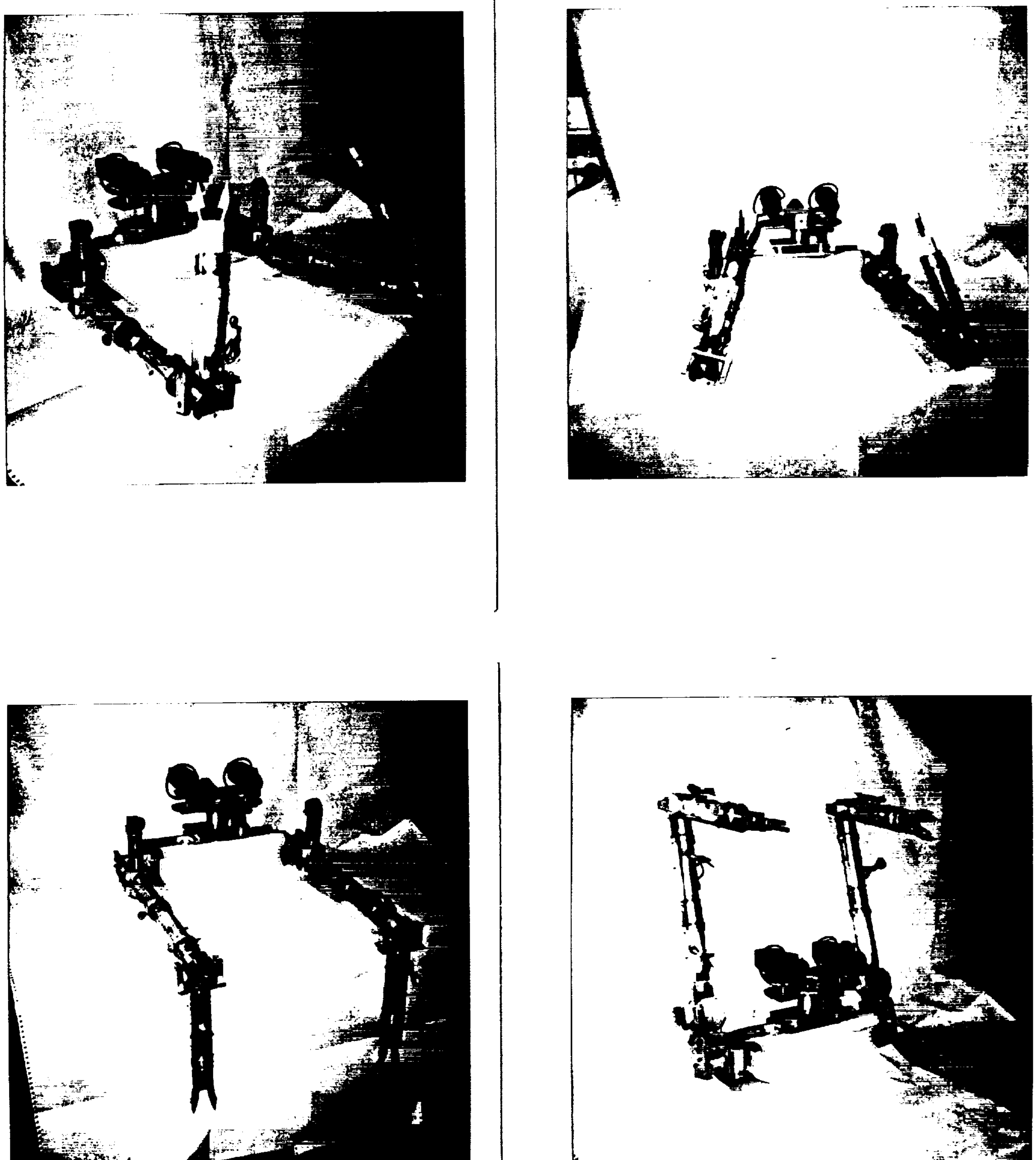


Fig. 1. Four views of the Upper Body of the SRAARS.

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

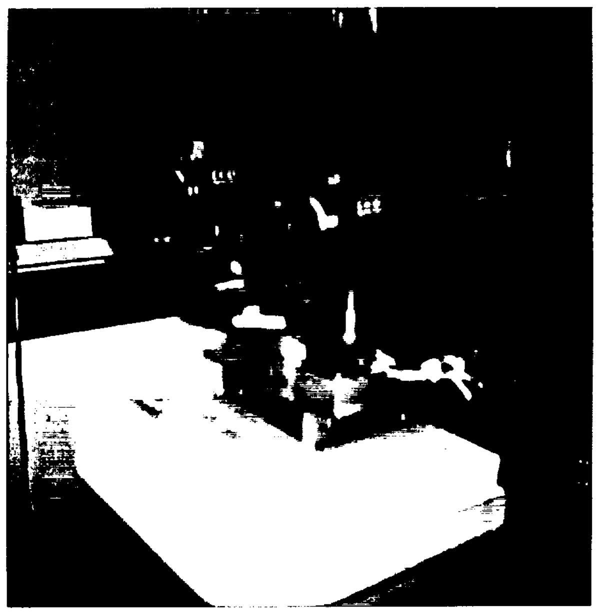
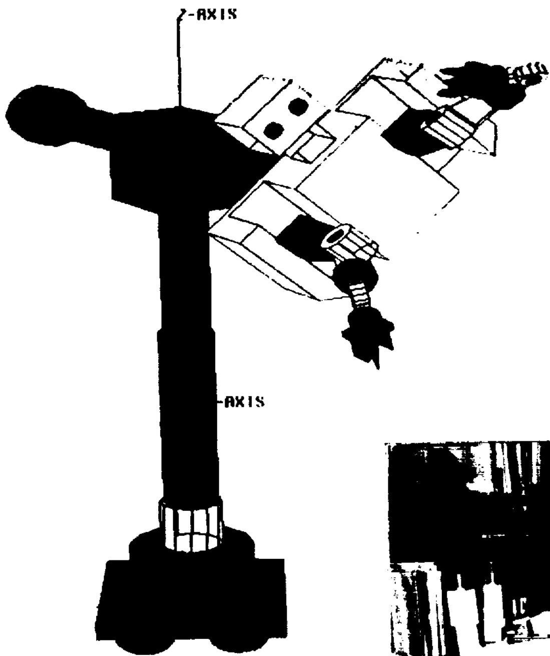


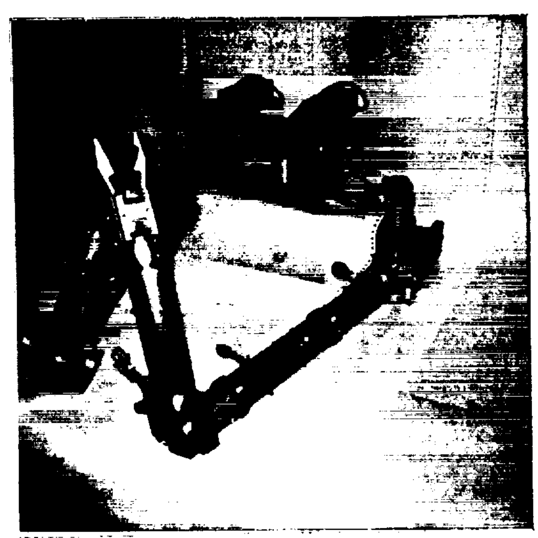
Fig. 2. The Head of the SRAARS.



1988, We had the computer graphics image of SRAARS

1989, We had the mockup model of SRAARS

1990, Now, we have the prototype of SRAARS upperbody



Advanced Robotics Group of SRA, Inc. @ Sept. 1990

Fig. 3. The Evolutionary Path of SRAARS development.

Intelligent Vision Process for Robot Manipulation

Alexander Y.K. Chen and Eugene Y.S. Chen

Scientific Research Associates, Inc.
50 Nye Road, P.O. Box 1058 Glastonbury, CT 06033

ABSTRACT

A new approach is introduced in this paper to deal with the problems of real-time machine vision and pattern recognition for robotic manipulations. This approach emphasizes three directions: (1) the developed algorithm has to be compact enough for embedded intelligent control implementation; (2) the computational scheme should be highly efficient for on-line robot reasoning and manipulations; and (3) the resulting system has to be sufficiently flexible to accommodate various working environments and to cope with some system shortcomings. The vertical integration of related vision hardware, image analysis software and analytical techniques (e.g. fuzzy logic and neural networks) together with the novel algorithms for robot eye-brain-hand coordination constitutes a unique robot vision system. The potential of more extensive hardware implementation is discussed and a wider spectrum of applications of the proposed robot vision system is envisioned.

1. INTRODUCTION

Both machine vision and robotics are multi-disciplinary, computer-related and application-oriented. To merge these two fields, the following challenges were experienced: (i) the computational requirements for vision information gathering and image analysis is substantial; (ii) a certain format of implicit representation of robotic manipulations in vision database is required; (iii) the geometric reasoning for robot eye-brain-hand coordination is essential to vision-based robot control.

Employing CCD or CID techniques, the resulting vision signals of one camera would represent a two-dimensional map of a three-dimensional projection. In order to extract information from the vision signals, a certain procedure of sorting, filtering, or transformation is required. From low level processing[1] to high level analysis[2], there are quite a few algorithms being developed in the field of machine vision. However, most of them would need either sophisticated computational schemes or a substantial amount of memory space, or both--which prevents a direct implementation of those techniques in robotic system applications. Furthermore, some of the foundations of utilizing machine vision in robot manipulations are not yet available, e.g. vision-based motion analysis, virtual construction of three-dimensional objects, and linkage representation with incomplete vision information, etc. Also, it is essential to establish a reasoning mechanism which can digest vision information as well as other sensor inputs for the purpose of robot manipulations. In this paper, a novel approach is presented to introduce the integration of vision, intelligence and manipulation of a robotic system.

In an attempt to make the entire robot vision package as compact, efficient and flexible as possible, the system is divided into three phases. In Phase I, the images are acquired by two CCD cameras synchronously and stored in the computer memory for processing. In Phase II, several essential image analysis procedures take place, e.g. edge detection, feature extraction and object status identification. In Phase III, the results of pattern recognition are fed into the virtual representation of the working environment to determine the updated relationship between the robot manipulator and the obstacles, where the virtual representation includes the solid models of the robot body and the related nearby objects. The procedure of verification and validation between the current status and the new status will be activated. The associated motion analysis would provide critical information in supporting the robotic manipulations. For implementing this new approach to the parallel distributed processing (PDP) platform, the Phase I and the Phase II will be localized in one Local Intelligence Center (LIC), and the Phase III will be performed at the Global Intelligence Center (GIC). The details of the intelligence allocation can be found in Ref. 3.

This paper is organized in the following sequence: a new approach for preprocessing low level vision signals is introduced in the next section which presents an effective and simple method to store useful vision raw materials. The employed stereoscopic viewing system is explained in the third section, which is used mainly for determining the object location and its size. The associated pattern recognition procedure is presented in the fourth section. The interactive coordination between the manipulator controller and the vision system is explained in the fifth section. Some of the preliminary results are shown in the sixth section. Conclusions and future researches are included in the last section.

2. AN EFFECTIVE IMAGE REPRESENTATION FOR ROBOT VISION

Low level vision processing schemes have been studied and discussed extensively in the past decade[4-7]. Almost all the methods require substantial computations and the results still remain case-specific. These conditions may be appropriate for a dedicated image analysis task, but will not be suitable for robot vision applications. Since the focusing point of robot vision is somewhat different than that of the ordinary image analysis, a new approach of image representation is developed which is designed for real-time mobile vision applications.

Before the procedure of edge detection, a step of preconditioning the input signals is essential to any vision system. In our laboratory, the employed image frame grabber is capable of performing the first stage of offsetting and restoring. A simple histogram algorithm is then employed to smoothe out the gray levels and enhance the image contrast. An adaptive single threshold is considered to convert the pixel map of 256 gray levels into several pixel maps of binary values. By comparing the modified gray level of each pixel with those of its right and its lower adjacent neighbors, a new pixel map is generated by assigning the corresponding value of each pixel to be "1" if the difference is greater than the given threshold, or the value "0" if otherwise. A global search of objects can be carried out with interactive threshold adjustment and binary maps comparison.

The ideal edge representation in a pixel map is a precise curve description of each visible edge projection. However, the actual converted binary pixel map would normally have a certain width of pixel band to represent each edge. The thinning process[8] is one possible way for improving this situation. Another difficult task is the reconstruction of some edges from the broken pieces. Noise filtering and edge localization are also very important. Some of those techniques can be found in Refs. 9-11. The resulting two-dimensional array of pixel locations of value "1" represents the image of interested objects, which can be explored in searching for unknown objects or can be examined to determine the status of some specified targets. This type of two-dimensional arrays can be considered as signatures of moving objects as well. Figure 1 shows an example of a graphic display of an image representation, which contains one two-dimensional object and two three-dimensional objects. The detailed procedure of feature extraction and object recognition will be explained in later sections.

3. OBJECT STATUS IDENTIFICATION WITH STEREOSCOPIC VIEWING SYSTEM

Edge level image representation is regarded as the final result of the low-level vision process. In order to perform feature extraction and object recognition, many previous approaches can be found in the survey of Besl and Jain[2]. For robot vision development, it is proposed that the complexity of high-level vision processing should depend upon the intelligence level of the robot. Since the utmost goal of robotic system development is to establish its autonomy such that it can execute simple tasks with effectively simple operations and perform delicate jobs with more sophisticated skills, it is therefore crucial to maintain the high-level vision processing flexible enough to respond to different levels of robot tasks. The proposed object status identification method will examine the tasks of determining simple two-dimensional objects first. The three-dimensional object recognition and the associated occlusion problem will then be discussed.

Assuming that the geometric descriptions of the designated targets are already available, the purpose of object status identification is to determine the existence of the specified targets, the locations of the targets once they have been identified, and the relationship between the targets and the vision system. The relevant parameters for describing the locations of the objects are their size, their positions and the orientations of the objects as shown in the image. The capability of zooming is also very important to minimize the computational requirement of status identification. One possible approach for determining the optimal window size and location is to utilize the cluster analysis.

After the zooming procedure, there are four parameters should be determined for calculating the object location: the positions of the object's reference points in the left and the right views, the distance between the pivoting points of the two cameras, and the vergence angle of the cameras. The center of the line segment specified by the pivoting points of two cameras is selected to be the origin of the coordinate system of the robot vision. A preliminary study has been done to determine the correspondence between the pixel maps and the three-dimensional object location with different vertical and horizontal view angles of cameras. The calculated object location is a function of the positions of the referenced points, the actual object size, the vergence angle and the view angles of cameras. The configuration of the developed robot vision system is shown in Figure 2 and the physical realization is displayed in Figure 3. There are two actuators dedicated to the viewing mechanism of the robot vision by providing the asynchronous and synchronous rotations of both cameras. Other necessary degrees of freedom would be supported by the motion of the robot's shoulder. The associated stereoscopic motion analysis has been considered in Ref. 12. However, utilizing the moving capability of both the vision system itself as well as the connected robotic system has not yet been explored. The proposed technique would tackle the motion stereo as part of the sensor-actuator network (i.e. the SANE framework[3]). Some of the basic relationships among variables and parameters are listed as follows:

The object position in the vision coordinates is P(x,y,z), which is determined in the following manner:

$$y = L/(\tan\alpha + \tan\beta)$$

$$x = -\frac{L}{2} + y\tan\beta = \frac{L}{2} - y\tan\alpha$$

$$z = yz_r/(w\cos\alpha) = yz_\ell/(w\cos\beta)$$

where

$$\tan\alpha = \frac{\tan V_a - (x_r/w)}{1 + (x_r/w)\tan V_a}$$

$$\tan\beta = \frac{\tan V_a + (x_\ell/w)}{1 - (x_\ell/w)\tan V_a}$$

$$\tan\alpha + \tan\beta = 2\left[1 + \frac{x_r x_\ell}{w^2}\right]\tan V_a + \frac{x_\ell - x_r}{w}\left[1 - \tan^2 V_a\right]$$

$$\cos\alpha = \frac{w\cos V_a + x_r\sin V_a}{\sqrt{w^2 + x_r^{\,2}}}$$

$$\cos\beta = \frac{w\cos V_a - x_\ell\sin V_a}{\sqrt{w^2 + x_\ell^{\,2}}}$$

$$x_r = d_{hp} \cdot P_{xr}$$

$$x_\ell = d_{hp} \cdot P_{x\ell}$$

$$z_r = d_{vp} \cdot P_{zr}$$

$$z_\ell = d_{vp} \cdot P_{z\ell}$$

Notations: V_a = Vergence angle

L = Distance between the pivoting points of two cameras

d_{hp} = Horizontal distance between two adjacent pixels in the referenced window (in here, d_{hp} = 2.63 mm)

d_{vp} = Vertical distance between two adjacent pixels in the referenced window (in here, d_{vp} = 2.09 mm)

(P_{xr}, P_{zr}) = The corresponded pixel location of the object in the right referenced window

$(P_{x\ell}, P_{z\ell})$ = The corresponded pixel location of the object in the left referenced window

w = The distance between the pivoting point of the camera and the corresponded referenced window (in here, w is set to be 1 m).

The signs of V_a, α and β are all set to be positive if the angles are inward, and negative if they are outward.

Once the object has been located in the stereoscopic vision, it can be extracted from the background and rescaled into the size of a standardized template for further analysis. The polar coordinates transformation of the object map is known to have the advantage of determining the shape of the object if it is a closed-form representation. By scanning the extreme points of the transformed curve, both the shape and the orientation of the object can be obtained. Due to the effect of discretization, the selected template size would affect the precision of object recognition. The resulted output map, namely the polar table, contains the information of the distances from the edge points to the origin of the polar coordinates. Hence, the polar table virtually converts the rotational relationship in the original coordinates into a translational relationship in the polar system. Therefore, any rotation of a closed-form object representation can be analyzed with this technique.

In order to demonstrate the concept of the polar table approach, a simple two-dimensional case is employed with a template size of 21x21. Three types of basic geometric shapes are used: Triangle, Circle and Square. These shapes have their unique corner numbers: 3, 0 and 4. Once the polar table is formed, the corner number can be readily obtained. An algorithm of counting the local maximal points and their locations is derived, and the results would directly indicate the shape as well as the orientation of the given object. Figure 4 shows the standard templates and their corresponding polar tables. This simple technique can be used to perform the initial calibration of the eye-hand coordination. Also, it can be extended to consider any object with a unique two-dimensional shape or projection.

Another similar approach for dealing with this problem is complex-log conformal mapping[13]. Although complex-log mapping is capable of converting both the rotation and the scaling into a translational representation, the stretching problem caused by the logarithm operations presents a more sensitive situation in determining the distance between the observed objects and the vision system. The polar table approach shown above appears to be a more efficient method in determining the orientation and the shape of the observed objects.

4. THREE-DIMENSIONAL OBJECT RECOGNITION

The focus of three-dimensional object recognition in robot vision is to reason the geometric relationship among concerned objects, such as the designated targets, the affected obstacles and the robot body. In order to verify the geometric relationship between two edges, two surfaces, or two solid volumes, the comparisons of distances, positions, orientations, shapes and sizes constitutes a crucial scheme in determining the consistency among geometric representations, where the reasoning rules are merely general guidelines for engaging sequential comparisons. The resulting object recognition technique is a progressive procedure in which simple objects can be readily identified and complex relationships among intertwined objects will be determined by examining the correlation among images from different view angles. Moreover, the input from other types of sensors and the interactions with the manipulation of the robotic system can also be utilized to carry out those delicate three-dimensional object recognition tasks.

5. MACHINE VISION AND ROBOTIC MANIPULATION

Normally, machine vision and robot manipulation are treated as two independent subjects. Only a few papers (for example, Ref.14) incorporate the vision information into the robot's motion control loop. The objects which the vision system is applied to are basically two-dimensional, well-defined configurations. On the other hand, there are some developments[15-16] which utilize the vision information to guide the navigation of the mobile robot. For most of those existing systems additional sensor input is required, the working environment is assumed to be uniform, and the robot manipulation is considered separately if there is any. What we propose here will be the first attempt to incorporate the vision process into the robot intelligence for geometric reasoning. Since the real-time reasoning mechanism would directly relate to the robot manipulation control, only a minimum set of vision information is constantly updated to ensure the performance of robot operations. Also, within the reasoning process, the symbolic representations of objects are employed to keep the processing kernel size as compact as possible. Hence, after the object status has been identified, the information will be converted into the status of symbolic objects which will then be used in the procedure of motion decisions. The entire functional flow chart is shown in Fig. 5.

6. SOME EXPERIMENTAL RESULTS

For demonstrating the complete process of the developed robot vision system, a series of pictures are taken to represent the result after each procedure. Figure 6a shows the original image. Figure 6b displays the discretized picture after image capturing. Figure 6c illustrates the result after the histogram processing. Figure 6d depicts the result after threshold scanning. Figure 6e shows the result after the thinning process. Figure 6f presents the final edge representation as the output of the low-level vision process. Figure 7a demonstrates the process of cluster analysis. Figures 7b and 7c show the two views of the corresponded stereoscopic vision. Figure 7d illustrates the object within the pre-determined template as the result of the zooming process. Figure 7e shows the identifiable template, and Fig. 7f is the resulted polar table. The object is identified as the given object "A" with the size 24.2 cm^2, the distance 1.055 m, the position (–8.6,105.5,–9.3) cm and the orientation (175.1°,177.7°,117.5°) [in Eulerian Angles]. By continuously taking several snapshots, the mobility status of the object "A" can be determined. If the observed object is unknown, then the mobility of the vision system can be commanded to vary either the vergence angle or the angle of the central line or both. By comparing the results of several images from different angles, the object recognition process will be activated to categorize the observed object. With two SONY XC-57 CCD cameras, Metrabyte MV1 frame grabber, INTEL 25MHZ 80386DX microprocessor and 80387-25 math coprocessor, the overall vision processing time is less than 8 seconds. It can be significantly improved with more advanced hardware devices--which are already commercially available. Hence, this new vision system presents a very attractive module for advanced robotic system development.

7. CONCLUSIONS AND FUTURE RESEARCHES

A new approach for developing the real-time robot vision system is presented in this paper. One of the major advantages of the new approach is to establish a flexible vision processing sequence such that the robot intelligence can discriminate simple tasks from those complicated jobs and would be able to activate only the necessary procedures to accomplish the assignment. New techniques in edge detection, thinning process, stereoscopic viewing, object status identification and potential unknown object recognition are developed. The vertical integration of all the new developments and the available related techniques has resulted in a powerful robot vision system, which will be incorporated into the development of SRAARS[3]. With a total of eighteen active joints, constituting two articulated mechanical arms, one robot head with vision, a lower body with cylindrical workspace and a potentially mobile base, the utilization of the vision capability would be an important part of the multi-tasking, real-time learning and parallel scalable robot intelligence.

It is believed that the most appropriate principle for robot vision development should be based upon the same concept found in those biological systems. In the past two decades, the results of scientific research concerning biological neural systems has indicated some fascinating properties in providing the generic capability of distributed memory and learning. The techniques of generating the artificial neural networks (ANN) have been vigorously investigated throughout the years. Some of those ANN models have shown a certain extent of image pattern recognition capability[17-19]. Methods such as the back-propagation algorithm[20] or the Boltzmann machine[21] are capable of identifying known objects under supervision. Others such as the Adaptive Resonance Theory (ART)[22] or the self-associative memory model[23] address the unsupervised learning mechanism. With currently available solid-state technology, the hardware implementation of ANN circuitry has become an important step towards the actual realization of ANN theories. The new robot vision system presented herein can incorporate the ANN techniques and the fuzzy logic process to further strengthen its capability in fault tolerance. It will be pursued in the near future.

REFERENCES:

1. Nitzan, D., "Three-Dimensional Vision Structure for Robot Applications," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 3, May 1988, pp. 291-309.

2. Beal, P.J. and Jain, R.C., "Three-Dimensional Object Recognition", *Computing Surveys*, Vol. 17, No. 1, Mar. 1985, pp. 75-145.

3. Chen, A.Y.K., "The Real-time Learning Mechanism of the Scientific Research Associates Advanced Robotic System (SRAARS)," presented at The Fifth Conference on Artificial Intelligence for Space Applications, Huntsville, AL, May 22-23, 1990.

4. Lunscher, W.H.H.J. and Beddoes, M.P., "Optimal Edge Detector Design I: Parameter Selection and Noise Effects" and "Optimal Edge Detector Design II: Coefficient Quantization", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, Mar. 1986, pp. 164-187.

5. Lacroix, V., "A Three-Module Strategy for Edge Detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, Nov. 1988.

6. Nalwa, V.S. and Binford, T.O., "On Detecting Edges", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6. Nov. 1986, pp. 699-714.

7. Torre, V. and Poggio, T.A., "On Edge Detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, March 1986, pp. 147-163.

8. Fu, K.S., Gonzalez, R.C. and Lee, C.S.G., Robotics Control, Sensing, Vision, and Intelligence, McGraw-Hill, 1987.

9. DeMicheli, E., Caprile, B., Ottonello, P. and Torre, V., "Localization and Noise in Edge Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 10, October 1989, pp. 1106-1117.

10. Chen, J.S., Huertas, A. and Medioni, G., "Fast Convolution with Laplacian-of-Gaussian Masks," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 4, July 1987, pp. 584-590.

11. Deriche, R., "Fast Algorithms for Low-Level Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, January 1990, pp. 78-87.

12. Mutch, K. M., "Determining Object Translation Information Using Stereoscopic Motion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, Nov. 1986, pp. 750-755.

13. Wechsler, H. and Zimmerman, G. L., "2-D Invariant Object Recognition Using Distributed Associative Memory," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, Nov. 1988.

14. Feddema, J.T. and Mitchell, O.R., "Vision-Guided Servoing With Feature-Based Trajectory Generation," *IEEE Trans. Robotics and Automation*, Vol. 5, No. 5, October 1989, pp. 691-700.

15. Yachida, M., Ichinose, T. and Tsuji, S., "Model-Guided Monitoring of a Building Environment," in Proceedings of the 8th Int'l. Joint Conf. on Artificial Intelligence, IJCAI, 1983, pp. 1125-1127.

16. Kuan, D., Phipps, G. and Hsueh, A.-C., "Autonomous Robotic Vehicle Road Following," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, Sept. 1988, pp. 648-658.

17. Sivilotti, M.A., Mahowald, M.A. and Mead, C.A., "Real-time Visual Computations Using Analog CMOS Processing Arrays", Advanced Research in VLSI: Proceedings of the 1987 Stanford Conference, Cambridge, MA, MIT Press.

18. Lu, T., Yu, F.T.S. and Gregory, D.A., "Self-organizing Optical Neural Network for Unsupervised Learning", *Optical Engineering*, Vol. 29, No. 9, Sept. 1990.

19. Fukushima, K., Miyake, S. and Ito, T., "Neocognitron: A Neural Network Model for a Mechanism for Visual Pattern Recognition", *IEEE Trans. System, Man and Cybernetics*, Vol. SMC-13, 1983, pp. 826-834.

20. Rumelhart, D.E., Hinton, G.E. and Williams, R.J., "Learning Internal Representations by Error Propagation", *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol. 1., Rumelhart and McClelland, Cambridge, MA, MIT Press.

21. Ackley, D.H., Hinton, G.E. and Sejnowski, T.J., "A Learning Algorithm for Boltzmann Machines", *Cognitive Science*, Vol. 9, 1985, pp. 147-169.

22. Grossberg, S. ed., Neural Networks and Natural Intelligence, A Bradford Book, The MIT Press, Cambridge, MA, 1988.

23. Kohonen, T., "Self-organized Formation of Topologically Correct Feature Maps", *Biological Cybernetics*, Vol. 43, 1983, pp. 59-69.

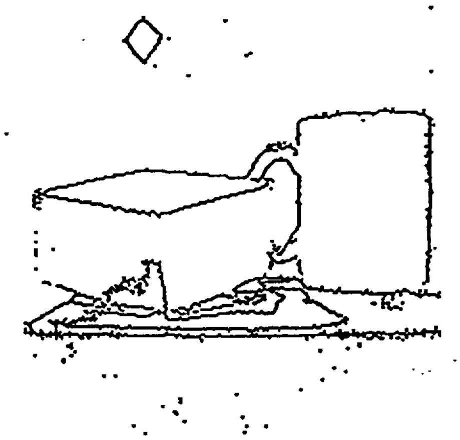


Figure 1. Image After Thinning in Low-level Vision

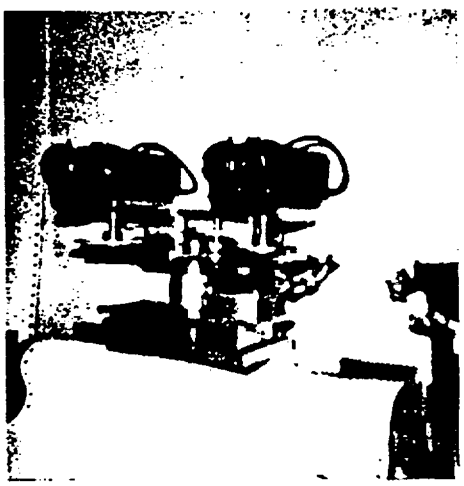


Figure 3. The Developed Robot Vision System

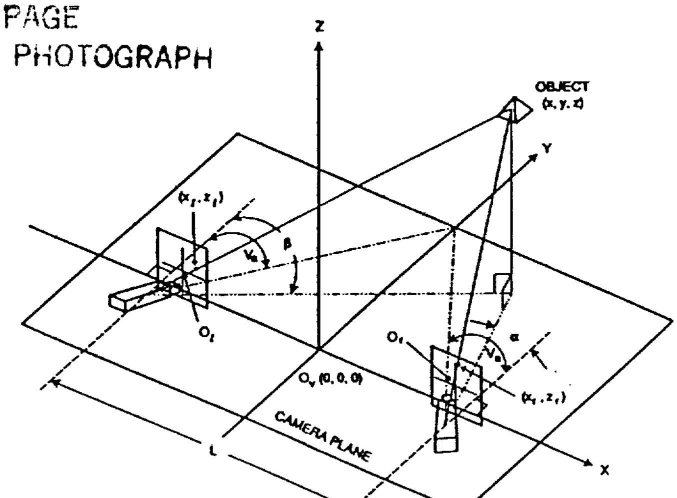


Figure 2. The Three-Dimensional Positioning Diagram in the Developed Stereoscopic Viewing System.

Fig. 4a. Template I

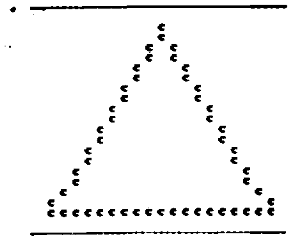


Fig. 4b. Template II

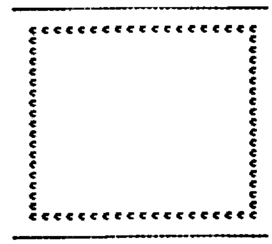


Fig. 4c. Template III

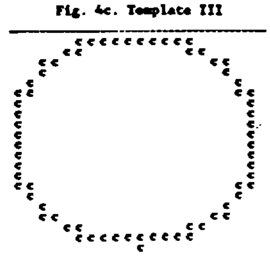


Fig. 4d. Polar Table of Template I

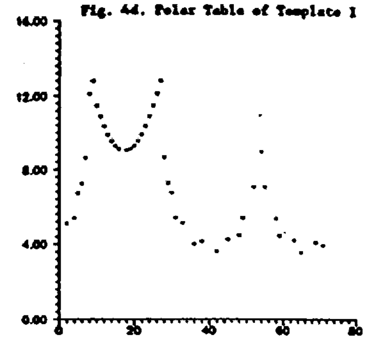


Fig. 4e. Polar Table of Template II

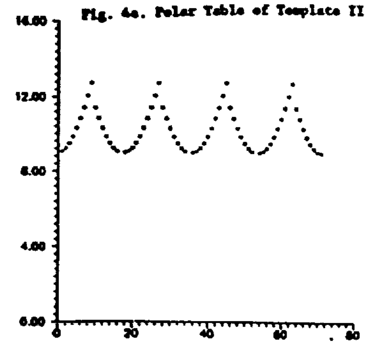
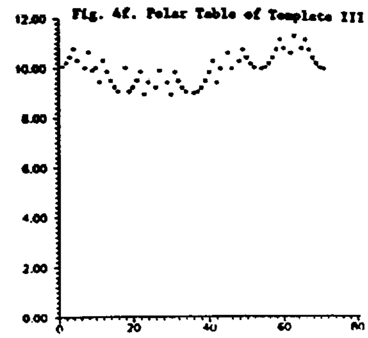


Fig. 4f. Polar Table of Template III



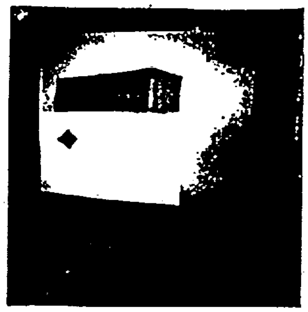


Fig. 6a. The Original Image

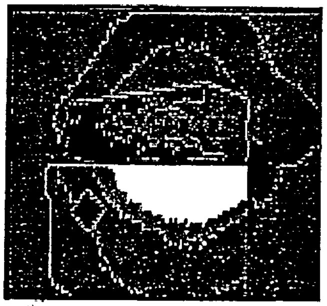


Fig. 6b. Image after Discretized Capture

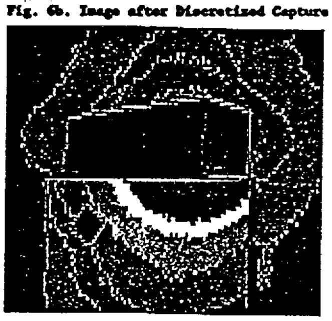


Fig. 6c. Image after Histogram Process

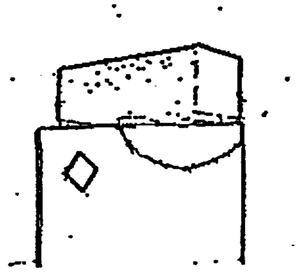


Fig. 6d. Image after Threshold Scanning

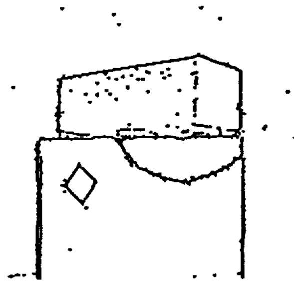


Fig. 6e. Image after Thinning Process

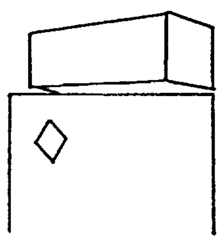


Fig. 6f. The Resulted Edge Representation

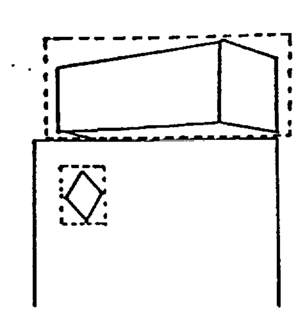


Fig. 7a. The Approach of Cluster Analysis

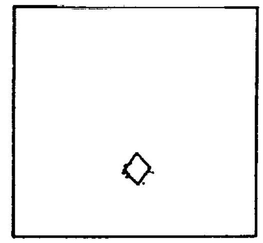


Fig. 7b. The Left Eye View of The Object

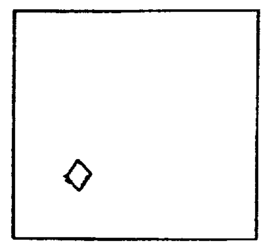


Fig. 7c. The Right Eye View of The Object

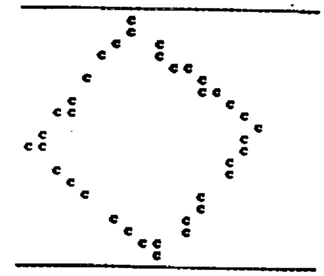


Fig. 7d. The Right View in Template

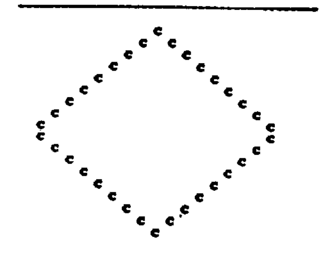


Fig. 7e. The Standard Template

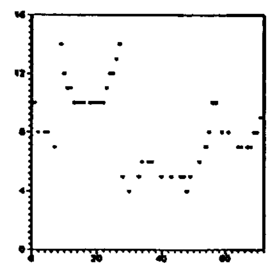


Fig. 7f. The Polar Table of The Right View

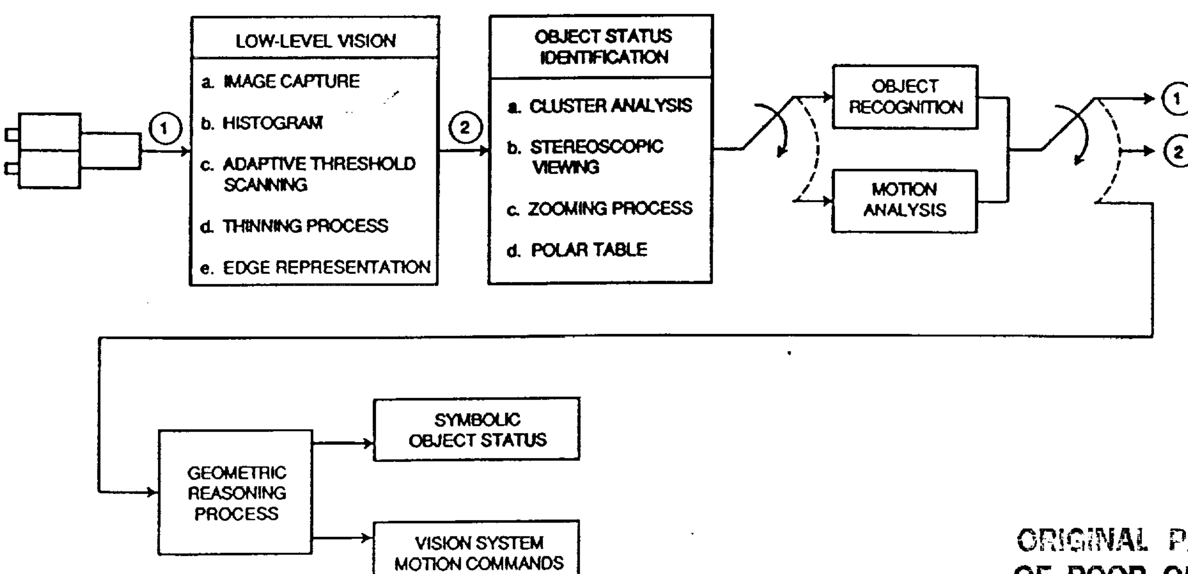


Figure 5. The Functional Diagram of the New Robot Vision System.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle
Intelligent Manipulation Technique for Multi-Branch Robotic Systems - Quarterly Progress Report

5. Report Date
December 1990

6. Performing Organization Code

7. Author(s)
Alexander Y.K. Chen, and Eugene Y.S. Chen

8. Performing Organization Report No.
R90-900077-4

9. Performing Organization Name and Address
Scientific Research Associates, Inc.
50 Nye Road, P.O. Box 1058
Glastonbury, CT 06033

10. Work Unit No.

11. Contract or Grant No.
NAS7-1072

12. Sponsoring Agency Name and Address
National Aeronautics and Space Administration
Washington, DC 20546-0001
NASA Resident Office - JPL

13. Type of Report and Period Covered
Quarterly Report
9/1/90 - 11/30/90

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract
New analytical development in kinematics planning is reported. The Intelligent Kinematics Planner (INKIP) consists of the kinematics spline theory and the adaptive logic annealing process. Also, a novel framework of robot learning mechanism is introduced. The FULOSONN integrates fuzzy logic in commands, control, searching and reasoning, the embedded expert system for nominal robotics knowledge implementation, and the self-organized neural networks for the dynamic knowledge evolutionary process. Progress on the mechanical construction of SRAARS and the real-time robot vision system is also reported. A decision has been made to incorporate the LAN (Local Area Network) technology in the overall communication system.

(SRA Advanced Robotic System)

17. Key Words (Suggested by Author(s))
Kinematics Planning, Spline Theory, Machine Learning and Local Area Network

18. Distribution Statement
Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	25	