August 1989

Active Perception and Exploratory Robotics

Ruzena Bajcsy
University of Pennsylvania

Recommended Citation

Ruzena Bajcsy, "Active Perception and Exploratory Robotics", . August 1989.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-65.

Active Perception and Exploratory Robotics

Abstract

Most past and present work in machine perception has involved extensive static analysis of passively sampled data. However, it should be axiomatic that perception is not passive, but active. Furthermore, most past and current robotics research use rather rigid assumptions, models about the world, objects and their relationships. It is not so difficult to see that these assumptions, most of the time, in realistic situations do not hold, and hence, the robots do not perform to the designer's expectations.

Perceptual activity is exploratory, which implies probing and searching. We do not just see, we look. We do not only touch, we feel. And in the course, our pupils adjust to the level of illumination, our eyes bring the world into sharp focus, our eyes converge or diverge, we move our heads or change our position to get a better view of something, and sometimes we even put on spectacles.

Similarly, our hands adjust to the size of the object, to the surface coarseness and to the hardness or compliance of the material. This adaptiveness is crucial for survival in an uncertain, and generally, unfriendly world as millenia of experiments with different perceptual organizations have clearly demonstrated. Although no adequate account or theory of activity of perception has been presented by machine perception research, very recently, some researchers have recognized the value of actively probing the environment and emphasized the importance of data acquisition during the perception including head/eye movement.

Because of the realization of today's inadequacies of robotic performances, we in the GRASP laboratory at the University of Pennsylvania for the past five years have embarked on research in Active Perception and Exploratory Robotics. What follows is an expose of our theoretical foundation and some preliminary results. First, we shall describe what we mean by Active Perception, then we shall argue that Perception must also include manipulation, and finally, we will present Exploratory Robotics as a paradigm for extracting physical properties from an unknown environment.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-65.

Active Perception And Exploratory Robotics

MS-CIS-89-65
GRASP LAB 191

Ruzena Bajcsy

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389

October 1989

Acknowledgements:
This research was funded in part by the United States Postal Service, BOA Contract: 104230-87-H-0001/M-0195, DARPA/ONR grant N00014-85-K-0807, NSF grant DCR 841-771, Air Force grant AFOSR F49620-85-K-0018, Army grant DAAG-29-84-K-0061, by DEC Corp., IBM Corp., and LORD Corp.

Active Perception and Exploratory Robotics

Ruzena Bajcsy
Computer and Information Science Department
University of Pennsylvania
Philadelphia, PA 19104

August 29, 1989

1 Introduction

Most past and present work in machine perception has involved extensive static analysis of passively sampled data. However, it should be axiomatic that perception is not passive, but active. Furthermore, most past and current robotics research use rather rigid assumptions, models about the world, objects and their relationships. It is not so difficult to see that these assumptions, most of the time, in realistic situations do not hold, and hence, the robots do not perform to the designer's expectations.

Perceptual activity is exploratory, which implies probing and searching. We do not just see, we look. We do not only touch, we feel. And in the course, our pupils adjust to the level of illumination, our eyes bring the world into sharp focus, our eyes converge or diverge, we move our heads or change our position to get a better view of something, and sometimes we even put on spectacles.

Similarly, our hands adjust to the size of the object, to the surface coarseness and to the hardness or compliance of the material. This adaptiveness is crucial for survival in an uncertain, and generally, unfriendly world as millenia of experiments with different perceptual organizations have clearly demonstrated. Although no adequate account or theory of activity of perception has been presented by machine perception research, very recently, some researchers have recognized the value of actively probing the environment and emphasized the importance of data acquisition during the perception including head/eye movement [3][7].

Because of the realization of today's inadequacies of robotic performances, we in the GRASP laboratory at the University of Pennsylvania for the past five years have embarked on research in Active Perception and Exploratory Robotics. What follows is an expose of our theoretical foundation and some preliminary results. First, we shall describe what we mean by Active Perception, then we shall argue that Perception must also include manipulation, and finally, we will present Exploratory Robotics as a paradigm for extracting physical properties from an unknown environment.

2 What is Active Perception?

In the robotics and computer vision literature, the term "active sensor" generally refers to a sensor that transmits (generally electromagnetic radiation, e.g., radar, sonar, ultrasound, microwaves and collimated light) into the environment and receives and measures the reflected signals. We include under the term Active Perception, Active Sensing as well. We believe that the use of active sensors is not a necessary condition on active sensing, and that active sensing can be performed with passive sensors (that only receive, and do not emit, information), employed actively. Here we use the term active not to denote a time-of-flight type sensor, but to denote a passive sensor employed in an active fashion, purposefully changing the sensor's state parameters according to sensing strategies. Putting it more succinctly, we are introducing a new paradigm for research in Machine Perception [4,5] called Active Perception. The new ingredients of this paradigm are taking multiple measurements and their integration, and the inclusion of feedback. Hence the problem of Active Sensing can be stated as a problem to control strategies applied to a data acquisition process that will depend on the current state of the data interpretation including recognition. The question may be asked: "Is Active Sensing only an application of Control Theory?" Our answer is: "No, at least in its simple version." Here is why: The feedback is performed not only on sensory data but on complex processed sensory data, i.e. various extracted features, including relational features. We do not have a complete descriptions of the states of the system. Furthermore the models that are used here are a mixture of numeric/parametric and symbolic information.

But one can say that Active Perception is an application of intelligent control theory which includes estimation, reasoning, decision making and control [6]. This approach has been eloquently defended for Computer Vision by Tenenbaum [21]:"Because of the inherent limitation of a single image, the acquisition of information should be treated as an integral part of the perceptual process...Accommodation attacks the fundamental limitation of image inadequacy rather than the secondary problems caused by it. "Although he uses the term "Accommodation" rather than "active sensing", the message is the same. Before we can outline the problem of active sensing more formally, we need to spell out the assumptions under which we are making the design.

The assumptions are that we have *a priori* available or we can extract:

1. Models of sensors and all subsequent processing modules, that i,: physics and geometry of the modules, including noise and uncertainty considerations.
2. The models of integration process of different modules, that is, combination rules and feedback.
3. Explicit specification of the initial and final state/goal and of the task.

If Active Perception is a theory, what is its predictive power? There are three components to our theory each with certain predictions:

1. Models at each processing level are characterized by parameters. These parameters are estimated using estimation theory and determine the lower bounds of performance.
2. The Combination rules again predict the lower bounds of the final outcome from the system.
3. The task model and the final state/goal specification guarantees the termination of the process and predicts the cost of accomplishing the task.

2.1 The Models

When we speak about models of sensors we are not restricted to the hardware only but also include various software modules that play a role in the processing chain. The following highlights of this work are worth mentioning.

Sensory models:

1. Physics models. These models represent the mathematical equations of principles that the sensors operates. The analysis of these models provides range for expected performance of the sensors if no other influences than physics are at work. Examples of these models are optics, illumination, radiance, and forces.
2. Geometric models. Here we get predictions from various aspects of geometry on the best possible values. Examples are: the geometry of a pair of stereo cameras predicts how resolution decreases as a function of distance [19].
3. Ideal Measurement or Signal Models. These models will help us analyze and predict the feasibility of detection of certain features. Examples of this case are: edge (step, linear or non-linear) and region (piece-wise constant, or linear or nonlinear, but monotonic) models [9, 17].
4. Noise or disturbance Models. Here we have considered not only the abnormal distribution (as everyone else has) but also abnormal distributions, symmetric or non-symmetric distributions of the random variables.

All these models provide upper and lower bounds for expected errors, resolution, and robustness, which is necessary for making certain decisions, in particular: "Do we need more data in order to get more accuracy? Can we afford to take more data based on some economy? Given the errors how do we combine different pieces of information in order to improve the overall performance?" (For details, see Hager [11])

The Models and Estimation theory have been very successfully applied by Zucker in 1985 [42]. In this basic work titled: Theory of Early Orientation Selection, Zucker used the model of a contour that comes from differential geometry. He divides the orientation selection process into three steps:

1. The measurement step-series of convolutions
2. The interpretation step of these convolution values. (This is a functional minimization problem.)
3. Finding the integral curve through the vector field

This decomposition into steps, having the parameters of each step explicit, allows Zucker to make clear predictions about where the contours will or will not be found. We very much agree with Zucker's criticism of the field for the lack of this kind of methodology! The very same flavor is in the paper of Leclerc and Zucker [14] where they study the edge detection of image discontinuities. The work of Binford and Nalwa [16] is again similar in flavor but applied to the modeling of edges or more general discontinuities.

2.2 A Concrete Example

A systematic and thorough approach to modeling, as it applies to Active Vision, is shown in the recent Ph.D. thesis of E. Krotkov [13] at the University of Pennsylvania. He has defined the task of determination of spatial layout using an agile camera system and two cues: range from focus and range from vergence. He has decomposed the problem into three subproblems:

1. Identifying an appropriate model M to represent the spatial layout of the environment;
2. finding effective methods for constructing M from vision data; and,
3. determining strategies for actively, dynamically, and adaptively setting sensor parameters for acquiring the vision data.

In this section, we shall review only the first subproblem. Krotkov modeled two characteristics of objects – extent and position – in the environment. This means encoding a map of location of objects with respect to the viewer. In order to accomplish the above, he had to model the details of the sensor (the camera) as well as the details of the computational process of obtaining range from focus and range from vergence.

It is not possible to go into all the details of the analysis but we can summarize the model as follows:

1. determine the optics of the lenses, the depth of the field, the accuracy of object distance, (in this setup the distance of the object is independent of the depth of field for distances 1-3 m.)
2. circle of confusion; its diameter depends upon the distance of the object plane from focusing distance. For a given distance between the image and detector planes the confusion circle is directly proportional to the diameter of the aperture, in this case diameter is 58mm.
3. the spatial resolution of the detector array is another limiting factor; (for the CCD chip used in this work the width of one photoreceptor is 0.03 mm and the focal length f=105mm determines the evaluation window size, typically 20x20 pixels).

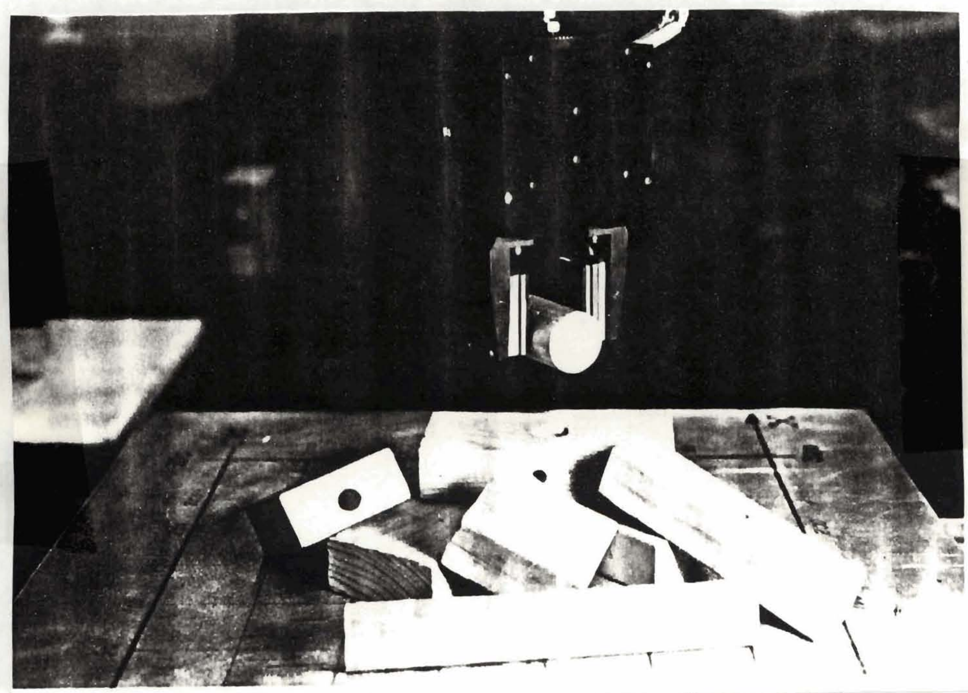
4. determine how to measure the sharpness of focus with a criterion function after analyzing defocus as an attenuation of high spatial-frequencies and experimentally comparing a number of possible criterion functions the method based on maximizing the magnitude of the intensity gradient was chosen. It proves superior to others in monotonicity about the mode and in robustness in the presence of noise. Then the Fibonacci search technique is employed to optimally locate the mode of the criterion function.

5. Finally the distance to an object point, given the focus motor position of sharpest focus is modeled by the thick lens law.

All the above predictions were experimentally verified on more than 3,000 points.
A very similar exercise can be presented, although, will not be for lack of space, is the modeling of the physical relationships for the vergence controller and the modeling of the line finder that is being used for matching the two stereo pairs of lines.

3 Perception using Manipulation

The motivation for this approach is the observation that it is impossible to discern movable and removable object/parts without manipulating them. This problem is rather broad though fundamental in Perception. In order to make some progress, we have limited ourselves to a subproblem which is how to decide that two objects are detachable [22]. We postulate that this cannot be decided only by vision, or in general, by any noncontact sensing. An exception to this is the case when the objects/parts are physically separated so that the noncontact sensor can measure this separation or one knows a great deal of *a priori* knowledge about the objects (their geometry, material, etc.). We assume no such knowledge is available. Instead we assume that the scene is reachable with a manipulator. Hence, the problem represents a class of problems of segmentation that occur on an assembly line, bin picking, organizing a desk top, etc. Figure 1 shows a scene of a pile of objects to be segmented.



What are the typical properties of this class of problems?

1. The objects are rigid. Their size and weight is such that they are manipulable with a suitable end effector. Their numbers on the scene are such that in a reasonable time each piece can be examined and manipulated, i.e., the complexity of the scene is bounded.
2. The scene is accessible to the sensors, i.e., the whole scene is visible, although some parts may be occluded, and reachable by the manipulator.
3. There is a well defined goal which is detectable by the available sensors. Specifically, the goal may be: an empty scene, or an organized/ordered scene.

The segmentation problem as is specified above is a subclass of a more general problem of a disassembly task that we wish to address in the future. As for any perceptual theory, the theory of segmentation using manipulation must have the following components: models of sensors, world/scene models, task/utility models, and models of actions. The segmentation process is formulated in terms of graph-theoretic operations that are mapped into corresponding manipulatory actions.

1. **Models of sensors:**

   these include the characterization of the non-contact sensor such as the spatial resolution, signal to noise ratio, the physical parameters of the different end effectors, such as the vacuum succession cup, the size of the spatula for pushing objects, the span of the gripper, and the maximum allowable weight and or force.

2. **Models of objects:**

   specified in terms of their geometry, size and substance.

3. **The Model of our world:**
   this work is limited to arrangement of objects thrown at random on a plane, called a heap. Then a scene is a (partial) view of a heap. The objects in the scene are represented as nodes in the digraph and the arcs denote : on-top-of relation. It is important to emphasize that this digraph represents relations of only the visible surface segments, i.e., as they appear through the visual sensor which is not always the same as the physical objects and their surface segments. The true physical arrangements of objects on the scene as well as the part-whole relations of objects are not known.

   The scene can be classified based on the analysis of the digraph into the following categories: Empty, if there are no nodes in the graph; Dispersed, if there no arcs in the graph; Ambiguous, if there is a cycle in the graph; Overlapped, if there are at least two nodes connected with one arc in the graph; and, Unstable, this category is not tested by the analysis of the graph but through analysis of the contact point/line of the object with the support plane. If this contact is point or line it is classified as unstable.

4. **Task models:**

   The final goal of the process. An example of a final goal can be the empty scene and the intermediate goals then can be those scenes that are more simply measured by a cost/benefit function. This cost/benefit function entails the cost of performing the particular manipulation, and the benefit is measured via the estimate of the outcome of the manipulation with respect to the final goal, i.e., emptying the scene.

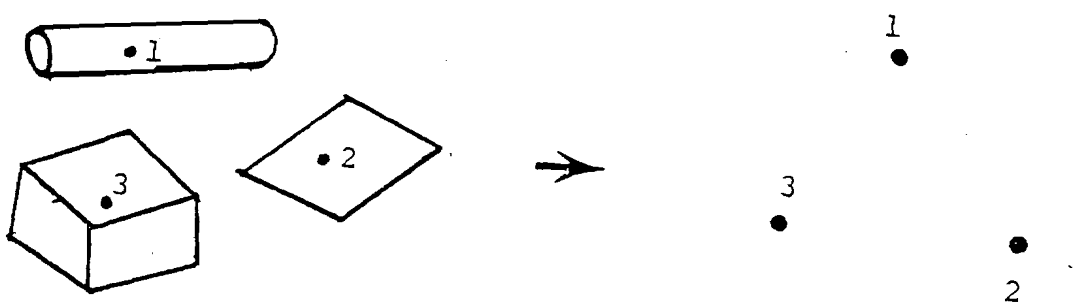
5. **Models of Action:**

   Parametrises the scene/ object /manipulation interaction.

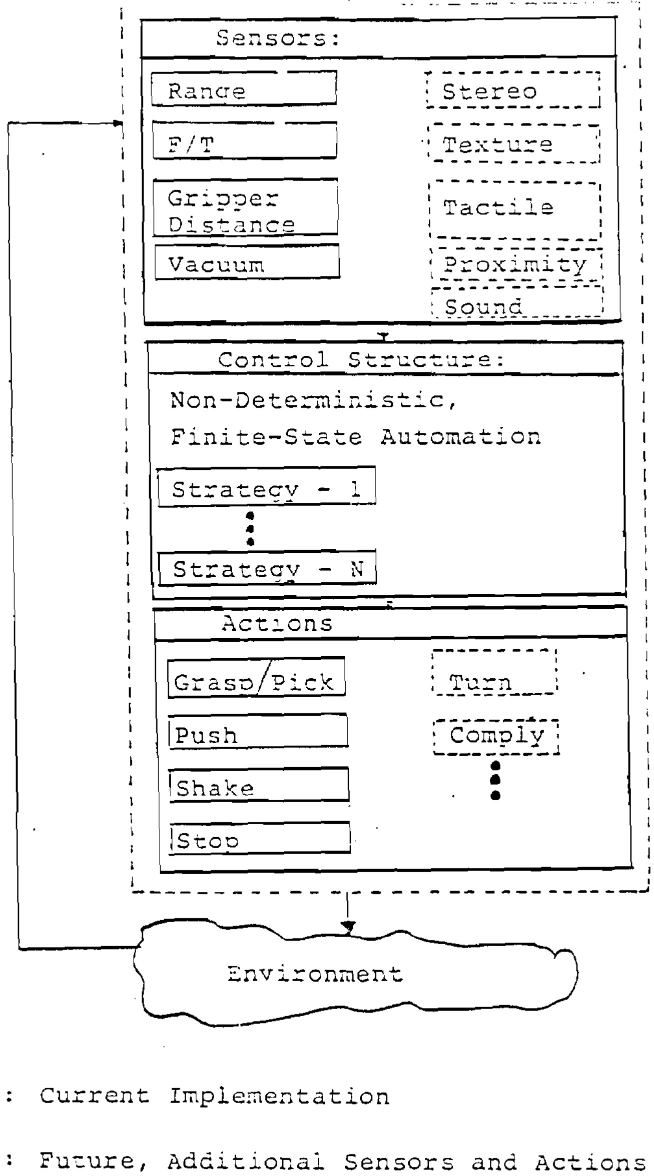
In principle there are two kinds of Actions:

1. Sensing Action, i.e., data acquisition action (look and/or feel), and,
2. Manipulatory action

[illegible] example of a graph of a dispersed scene.



The purpose of the manipulatory actions for this paper is to exhaust a physical disturbance, being either global (as shaking is) or local (a pushing/pulling). In view of our formulation of the segmentation problem as a graph generation/decomposition problem we classify the manipulatory action in relationship to the operation that apply on the digraph. There are two such operations: the node removal, which means in terms of manipulation the removal of an object from the scene, the arc removal which in turn translates into object displacement in the scene so that the relationship of on-top-of does not hold anymore between the two objects. Putting it another way: an isomorphism exists between the manipulation actions and graph decomposition operations [39]. Our approach is to close the loop between sensing and manipulation. The manipulator is used to simplify the scene by decomposing the scene into visually simpler scenes. The manipulator carries the contact sensors to the region of interest and performs the necessary exploratory movements that will determine the nature of the mechanical binding between objects in the region. Perception-Action interaction is modeled by a Non-deterministic finite state Turing Machine. The model of sensing, manipulation and control is a Non-deterministic Turing Machine (NDTM) as we show below in Figure 3.



The physical world (scene) is the "tape" of the machine, the "read_from_tape" actions are the sensing actions and the "write to_tape" actions are the manipulation actions. The model is a Turing machine because the manipulation actions constantly change the physical environment (tape) and hence its own input. The above model is non-deterministic because of the non predictable state of the scene after each manipulatory step. From this, of course, follows also the non-deterministic control of actions. In addition to the non-determinism of the control strategies, the automaton has finite states, which are determined by the finite numbers of recognizable scenes and the finite number of available actions.

We believe that this model is quite general providing that one can quantize the scene descriptions and/or the sensory outputs into unique and mutually exclusive states, and, of course, one has only a finite number of manipulatory actions.

There are several advantages to the formalisms of the non-deterministic finite state Turing machine. The advantages are:

1. the sense-compute-act formalism allows the control problem to be partitioned in time and complexity. At any given time, the system deals only with present state and present input, produces an output which is a function of current state and current input and moves to a new state. Current state encodes information about past history of states and actions of the machine and its environment. Current sensory input is not deterministic (noise in sensory data). The next state of the NDTM is not deterministic because the machine modifies its tape via actions whose outcome cannot be known *a priori* (push and shake actions).

2. the theoretical tools needed to prove correctness of the machine's behavior have long been established and tested.

   Path sensitization and graph de-cyclization algorithms exist, [10, 12, 8] to prove:

   (a) the goal state is reachable

   (b) the state transition diagram does not contain deadlock states, or cycles.

   (c) it facilitates error handling. If additional states need to be defined to deal with non-anticipated error conditions, then these states can be simply inserted. The fourth advantage is that it is modularand allows insertion of new sensors, actions and feedback conditions.

3. it makes debugging easy. The sixth advantage is that it allows a system to be developed incrementally.

One disadvantage is that the number of states and transitions needed to represent the machine and its environment increases as more sensors are added. Addition of more sensors implies increased complexity.

4 Exploratory Robotics.

Much of the work in Robotics until now has been by and large conducted in the so-called "knowledge driven" framework. The justification for this approach was the fact that in the industrial environment the geometry, material, environmental conditions and the task are:

1. quit constrained

2. known *a priori*

3. well controllable.

However this is, not the case in many other situations and applications of robots in underwater, mine and outer space explorations. The common denominator to all of these cases is that the robot must be able to explore and adapt to unconstrained and unknown environments. This is the motivation for the investigation of Exploratory Robotics.

4.1 Definition of the Problem.

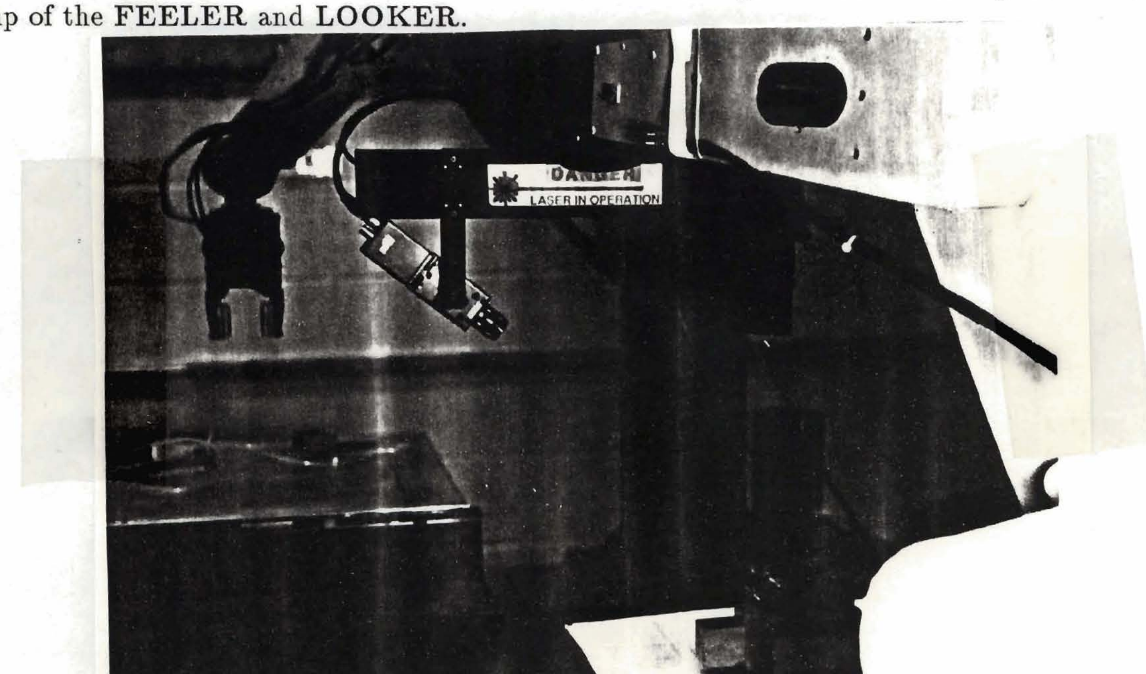
We wish to investigate to discover the necessary components/modules that must be embedded into a Robot with Exploratory Capabilities. These ideas came from our collaborative efforts between R. Klatzky and S. Lederman, see [15] In other words, what sensors, exploratory procedures, data processing, data reduction and interpretation capabilities for a given **TASK** must such a robot have. In full generality, this task is formidable. Hence, we shall limit ourselves to two more specific tasks:

1. Exploration of surface properties of ground for mobility purposes.
2. Exploration of an object for manipulatory purposes.

In the first task, we shall consider surfaces made from materials such as dirt/soil/sand, rocks/concrete, pebbles/gravel, metals, wood, glass/ceramics, rubber/polymers, and viscous mixtures (like mud). We shall not consider vegetables, textile, liquid, and like materials.

In the second task we will limit ourselves to objects by size and weight. This limitation will be determined by the size and flexibility of the end-effector, i.e., we shall consider objects that are graspable. This will exclude liquid, for example, but not deformable objects like a cable or a rubber ball. We shall also investigate objects that have two rigid parts joined by a hinge.

For both of the tasks, the robot will be equipped with one six-degree freedom manipulator and a range finder and/or a pair of CCD cameras, called the **LOOKER**, and ONE six-degree freedom manipulator and a hand, called the **FEELER**. The **LOOKER**, depending on the need, can also have a color camera system or any non-contact electromagnetic wave measuring detector (infrared as one possibility). The **FEELER** has a force/torque sensor in its wrist and a hand with three fingers and a rigid palm. Each finger has one and one-half degrees of freedom. Figure 4 shows the setup of the **FEELER** and **LOOKER**.



The sensors on the hand include:

- a position encoder and force sensor at each joint of the finger
- a tactile array at each of the finger tips and on the palm
- a thermo-sensor on the palm,
- an ultrasound sensor on the outer side of the hand.

In addition the Hand has available various tools that it can pick up under its control.
Both the **FEELER** and **LOOKER** are under software control of strategies for data acquisition and manipulation. For Task One, we consider a model of a foot with a planar sole as one tool that will act as the probe for testing the surfaces for mobility.

4.2 Exploration of Surface Properties.

The Problem
Given a surface, we wish to establish procedures to determine physical and geometric properties with minimal *a priori* information so that an object like a robot or vehicle can move on this surface. The basic assumption is that the surface is much larger than the robot and, at least, locally flat so that there is space to move around. The flatness assumption is relative to the size of the robot:

- the surface variation from a planar surface must be no more than 10

We do not consider the problem of obstacles.

Scientific fields older than robotics have investigated how to measure the attributes of the materials listed in the table. They are: mineralogy, geology, soil science, civil engineering (for testing soil in preparation for building), and material sciences in general. Tests from these fields share the following procedures: take samples into the laboratory and perform a multitude of tests, if necessary, perform destructive tests, such as for brittleness or penetrability, or even for deformability, and perform excavations of layered surface (examples of geologists).

The question for this research is which of these procedures are applicable for our domain. The procedure in:

1. can be applied in the robotic context. One can design a robot in such a way that it can carry with it a small testing kit.
2. is harder to envision though as part of the calibration process can be executed.
3. is totally inapplicable since the robot will not have time to perform excavation before it moves.

We examine those Exploratory Procedures (EPs) which will allow the robot to:

- stand firmly on the surface (static stability) and
- move on the surface in a stable manner (dynamic stability).

Further Assumptions

In order to further constrain the interpretation of the measurements we eliminate the effect of the geometry, that is, we assume that both the **LOOKER** and the **FEELER** are perpendicular to the examined surface.

Exploration for static stability: Exploratory Procedure for surface firmness versus penetrability. The penetrable surface can be deformable, compressible, either, or both. As an example, whereas penetrable objects such as soil, sand, pebbles, viscous mud, and rubber/polymers are deformable; only soil, sand, and pebbles are compressible (see the table).

This EP will utilize a cooperative effort between vision and force guided penetration. The **FEELER** exerts controlled and recorded force on the surface while the **LOOKER** observes the surface. If the surface has not changed under the given force, then it is firm; if it deforms then it is penetrable. It can be either deformable or compressible. The test for discriminating the latter two, is to use the **LOOKER** observation of the resulting surface after the **FEELER** has withdrawn the penetrating force. If the surface has not changed from the previous image then we have a deformable surface (just like mud would stay); otherwise we have a compressible surface. Naturally, this is not a sufficient test, especially when the measurement indicates no firm surface. Other tests like measure of pressure, surface roughness and viscosity must be carried out. Which ones are necessary and sufficient will be one of the topics of this research.

Table 1 below summarizes the attributes and their relationship to different classes of materials.

MATERIALS AND THEIR SALIENT ATTRIBUTES

ATTRIBUTES	CLASSES OF MATERIALS							
	Metals	Rocks Concrete	Glass Ceramics	Rubber Polymers	Wood	Soil Sand	Pebbles Gravel	Viscous Mixtures
Penetrability	No	No	No	No	No	Yes	Yes	Yes
Deformability	No	No	No	Yes	No	Yes	Yes	Yes
Hardness	Yes	Yes	Yes	Yes	Yes	No	No	No
Brittleness	No	Yes	Yes	No	No	No	No	No
Compressibility	No	No	No	No	No	Yes	Yes	No
Compressive Strength	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Surface Roughness	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Thermal Conductivity	Yes	No	Yes	No	No	No	No	No
Electrical Conductivity	Yes	No	Yes	No	No	No	No	No
Magnetic Permeability	Yes	No	Yes	No	No	No	No	No
Optical Properties	Yes	No	Yes	No	No	No	No	No
Viscosity	No	No	No	No	No	Yes	Yes	Yes

Yes - Attribute exists, is measurable and is a distinguishing property.

No - Attribute does not exist or is not measurable or is not a distinguishing property.

4.3 Exploration of Graspable Objects

The Problem

We wish to find the following properties of the graspable object: material (its hardness and surface texture), its density, temperature, weight and size, rigidity versus flexibility, and finally its gross shape for identifying the graspable points.

In order to accomplish this task one needs two modes of exploration: A Static Mode and A Dynamic Mode. In the Static mode the object is stationary and the **LOOKER** and the **FEELER** can **Look** and **Feel** around the object. During the Dynamic mode the object is being grasped and manipulated, for example lifted or shaken. In the Static mode we can establish the following attributes: size, shape, temperature and hardness/surface texture. In the dynamic mode the remaining attributes are established: the weight, density and the rigidity vs. flexibility.

The Static Exploratory Procedures applied on objects

Following the work of Allen [2] and Stansfield [20] we accept their findings that blind touch is unproductive and the tactile exploration should be guided by vision. Hence we begin with the **LOOKER** which will give us the position, gross shape and size of the object. Using the superquadric fitting to the visual three-dimensional data developed by Solina [34], we get further parameterization of the data, that is: the orientation, extent in three orthogonal planes (the size), and estimate of the surfaces (whether they are planar or second order surfaces) of the object. Then following Stansfield's EPs for hardness and surface texture and using the **FEELER** we can estimate the material of the object. In addition, by measuring the conductivity of the material (by another similar low level EP), we can further distinguish the material as metal or non-metal. All these properties are passed to the next stage – the Dynamic mode.

The Dynamic Exploratory Procedure applied on objects

As mentioned before, the dynamic EPs will measure weight, density and rigidity. EPs for weight and density: Grasp the object and lift it to a height H. The exerted force divided by approximately .9 (gravitational force) will give the weight of the object. The weight divided by the volume (calculated from the shape parameters) is the density of the material.

The more sophisticated Exploratory Procedure is the test for rigidity. Another assumption: consider objects either rigid, bent, or as two parts connected with a hinge. This again involves a cooperation between the **LOOKER** (the vision) and the **FEELER** (with force-guided probe). There are several strategies that must be followed in a few specified orders:

1. Consider an object which is being translated or rotated on the table by pushing (we know the magnitude and direction of the exerted force). If the new image can be accounted for by rigid transformations for this manipulation, then the object is rigid; otherwise the change must be examined.

2. examination of the change: parts are rigid but their spatial relationship has changed. or the object is bent, i.e., a deformation has occurred.

3. The case of rigid parts indicates that there is either one fixed point of rotation, or one fixed line of rotation. In either case we have identified a hinged object.

4. in the case of a bent object, compute the amount of bend.

5 Conclusion

We have defined Active Perception as a problem of an intelligent data acquisition process. For this, one needs to define and measure parameters and errors from the scene which in turn can be fed back to control the data acquisition process. This is a difficult though important problem. Why? The difficulty is that many of the feedback parameters are context- and scene-dependent. The precise definition of these parameters depends on a thorough understanding of the data acquisition devices (camera parameters, illumination and reflectance parameters), algorithms (edge detectors, region growers, 3D recovery procedures) as well as the goal of the visual processing. The importance, however, of this understanding is that one does not spend time on processing and artificially improving imperfect data but rather on accepting imperfect, noisy data as a matter of fact and incorporating it into the overall processing strategy. The second point we made in this paper is that manipulation is an essential part of perceptual process. The hand is as the eye: a sensory device. Subsequently, one needs to consider not only signal processing modules but also basic manipulatory action called exploratory procedures as an essential ingredient of perceptual theory [24]. The third and last point we are making is a case for Exploratory Robotics. Today, it is assumed that the size and shape of the object is sufficient for grasping purposes. It should be very apparent that unless one knows what materials are being used the system may be easily fooled. And even if we know the material of the outer surface, we do not know the inside, which may very dramatically change the weight, and hence, the grasping strategy. Our research aims to fill this gap. The question of rigidity is also very crucial when a grasping strategy is considered. Furthermore, the tests for hinges and bending are the first tests towards testing the functionality of an object. In the test for rigidity, we need to further explore what changes will occur when other controlled manipulatory actions will be applied on such objects, for example, lifting or rotating the object in space. All these steps are part of a general examination of the object, finding stable positions, etc. All these tests lead to understanding of what the necessary components are for a general purpose Perceptual Theory.

ACKNOWLEDGMENTS

This research was funded in part by the United States Postal Service, BOA Contract: 104230-87-H-0001/M-0195, DARPA/ONR grant N0014-85-K-0807, NSF grant DCR 8410771, Air Force grant AFOSR F49620-85-K-0018, Army grant DAAG-29-84-K- 0061, by DEC Corporation, IBM Corporation, and LORD Corporation.

References

[1] Albus, J., Barbera, A. & Fitzgerald, M. (1982) *Programming a Hierarchical Robot Control System*, Proceedings of the 12th International Conference on Industrial Robots, Paris, France.

[2] Allen, P.K. (1987) *Robotic Object Recognition Using Vision and Touch.* Norwell, MA.: Kluwer Academic Publishers.

[3] Aloimonos, J. & Badyopadhyay, A. (1987) *Active Vision*, Proceedings of the 1st IEEE International Conference on Computer Vision, London, U.K.: Computer Society Press, pp. 35-54.

[4] Bajcsy, R. (1985) *Active Perception vs. Passive Perception*, Proceedings of the 3rd Workshop on Computer Vision: Representation and Control Bellaire, MI., Computer Society Press, pp. 55-59.

[5] Bajcsy, R. (1988) *Active Perception*, Proceedings of the IEEE, 76, pp. 996-1005.

[6] Bajcsy, R., Gupta, A. & Mintz, M. (1989) *Research on Symbolic Inference in Computation Vision*, Technical Report, University of Pennsylvania.

[7] Ballard, D.H (1987) *Eye Movements and Spatial Cognition*, Computer Science Dept. Univ. of Rochester, TR 21, November 1987.

[8] Deo, N. (1974) *Graph Theory with Applications to Engineering and Computer Science*, Prentice-Hall Inc., Englewood Cliffs, NJ 07632.

[9] Haralick,R.M. (1981) *The Digital Edge*, Proceedings of IEEE Computer Conference Pattern Recognition and Image Processing, Computer Society Press, pp. 285-291.

[10] Hartmanis, J. & Stearns, R. (1966) *Algebraic Structure Theory of Sequential Machines*, Prentice-Hall Inc., Englewood Cliffs, NJ 07632.

[11] Hager, G.D. (1988) *Active Reduction of Uncertainity in Multi-Sensor Systems*, Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA 19104.

[12] Kohavi, Z. (1970) *Switching and Finite Automata Theory*, McGraw-Hill Book Company, New York.

[13] Krotkov, E. (1987) *Exploratory Visual Sensing for Determining Spatial Layout with and Agile Camera System*, Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA 19104.

[14] Leclerc ,Y.G. and S.W. Zucker (1989): *The Local Structure of Image Discontinuities in One Dimension*, IEEE Transactions on PAMI, Vol-9, No. 3, pp. 341-355

[15] Lederman, S.J. & Klatzky, R.L. (1987) *Hand Movements: A Window into Haptic Object Recognition*, Cognitive Psychology, 19, pp. 342-368.

[16] Nalwa, V.S. & Binford, T.O. (1986). On Detecting Edges. IEEE Transactions on PAMI, Vol. 8. pp. 699-714.

[17] Pavlidis, R and Liou, Y.T. (1988) *Integrating Region Growing and Edge Detection*, submitted to *ICVPR.*

[18] Solina, F.(1987), *Shape Recovery and Segmentation with Deformable Part Model* PhD Dissertation, University of Pennsylvania, Philadelphia, PA 19104.

[19] Solina, Franc (1985) *Errors in stereo due to quantization*, GRASP lab. memo, Univ. of Pennsylvania, Philadelphia, December 1985.

[20] Stansfield, S.A. (1988) *A Robotics Perceptual System Utilizing Passive Vision and Active Touch*, International Journal of Robotics Research, 7, pp. 138-161.

[21] Tenenbaum, J. M. (1970) *Accommodation in Computer Vision* Stanford University Ph.D. Thesis.

[22] Tsikos, C. (1987) *Segmentation of 3-D Scenes Using Multi-Modal Interaction Between Machine Vision and Programmable, Mechanical Scene Manipulation*, Ph.D. Dissertation, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104.

[23] Zucker, S.W. (1985) *Early Orientation Selection: Tangent Fields and the dimensionality of their support*, TR-85-13-R, Computer Vision & Robotics Laboratory, Dept. of EE, McGill Univ. Montreal, Quebec, Canada.