19930016882

N93-26071

A Vision Architecture for the Extravehicular Activity Retriever

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1992

Johnson Space Center

Prepared by:	Michael Magee, Ph.D.
Academic Rank:	Professor
University & Department:	University of Wyoming Computer Science Department Laramie, Wyoming 82071-3682

NASA/JSC

Directorate:	Engineering
Division:	Automation and Robotics
Branch:	Intelligent Systems
JSC Colleague:	Thomas W. Pendleton
Date Submitted:	August 7, 1992
Contract Number:	NGT-44-005-803

ABSTRACT

The Extravehicular Activity Retriever (EVAR) is a robotic device currently being developed by the Automation and Robotics Division at the NASA Johnson Space Center to support activities in the neighborhood of the Space Shuttle or Space Station Freedom. As the name implies, the Retriever's primary function will be to provide the capability to retrieve tools, equipment or other objects which have become detached from the spacecraft, but it will also be able to rescue a crew member who may have become inadvertently de-tethered. Later goals will include cooperative operations between a crew member and the Retriever such as fetching a tool that is required for servicing or maintenance operations.

This report documents a preliminary design for a Vision System Planner (VSP) for the EVAR that is capable of achieving visual objectives provided to it by a high level task planner. Typical commands which the task planner might issue to the VSP relate to object recognition, object location determination, and obstacle detection. Upon receiving a command from the task planner, the VSP then plans a sequence of actions to achieve the specified objective using a model-based reasoning approach. This sequence may involve choosing an appropriate sensor, selecting an algorithm to process the data, reorienting the sensor, adjusting the effective resolution of the image using lens zooming capability, and/or requesting the task planner to reposition the EVAR to obtain a different view of the object.

An initial version of the Vision System Planner which realizes the above capabilities using simulated images has been implemented and tested. The remaining sections describe the architecture and capabilities of the VSP and its relationship to the high level task planner. In addition, typical plans that are generated to achieve visual goals for various scenarios will be discussed. Specific topics to be addressed will include object search strategies, repositioning of the EVAR to improve the quality of information obtained from the sensors, complementary usage of the sensors and redundant capabilities.

INTRODUCTION

Vision systems that provide autonomous or semi-autonomous robots with information that describes their surrounding environment or objects in that environment should be able to plan and execute actions that solve visual problems efficiently and effectively. From a software architectural design standpoint, the highest level or supervisory planner is called the Task Planner (Figure 1). The Task Planner oversees the actions of several subplanners, one of which is the Vision System Planner. Each of these subplanners can be considered to be an expert with special knowledge regarding how to solve problems within its particular domain. When commanded to do so by the Task Planner, a subplanner will determine a method for achieving the specified goal given its knowledge of the current state of the world and it will then communicate the result of executing the planned action back to the Task Planner.

Although each subplanner is subservient to the Task Planner, it may nevertheless ask for assistance from the Task Planner if such assistance would help it achieve the specified goal. For example, if the Task Planner requests the Vision System Planner to recognize an object and the robot on which the vision hardware is mounted is poorly positioned to sense the object, the VSP may request the Task Planner to cause the robot to be moved. If the Task Planner honors the request from the VSP, it would then send commands to other subplanners (involving navigation and control) to move the robot so that the Vision System can accomplish the objective originally requested by the Task Planner.

The Vision System module itself (Figure 2) should be a self-contained entity capable of accomplishing many types of objectives such as object detection, recognition, tracking and pose estimation. A typical plan that would be formulated to achieve one of these goals would involve choosing an appropriate sensor, selecting an algorithm to process the data, and communicating the results or a request for assistance to the Task Planner. The remaining sections discuss a suggested architecture for such a Vision System within the context of the Extravehicular Activity Retriever.

VISION SYSTEM PLANNER DESIGN CONSIDERATIONS

The planning mechanisms developed are founded on the assumption that there should be at least two visual sensors which provide intensity (color) and range images. There are several reasons why such a multisensory approach is desirable, three of which are particularly significant. First, the availability of sensors with complementary capabilities permits the VSP to select a sensor/algorithm combination that is most appropriate for

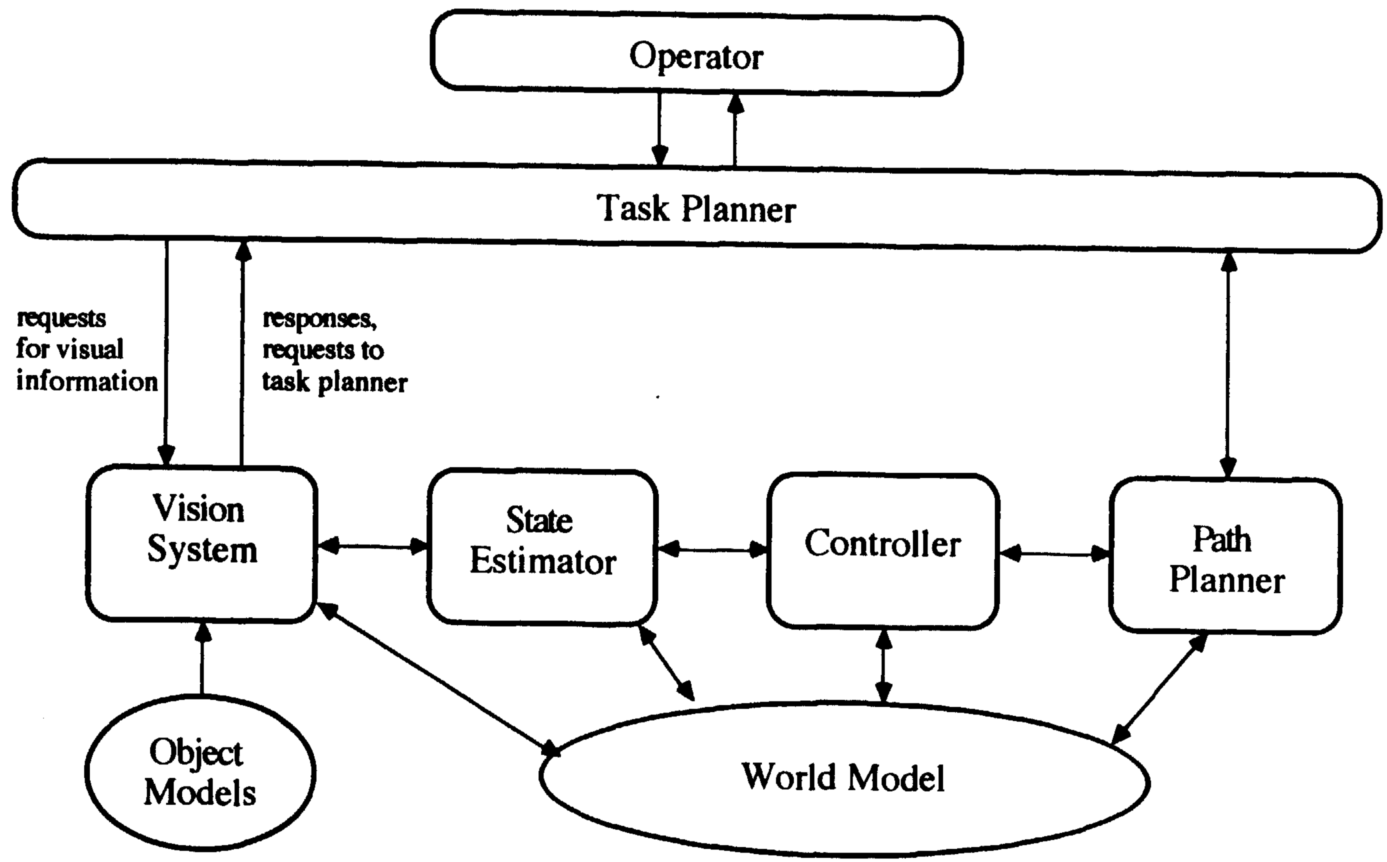

Figure 1: Planning System Architecture

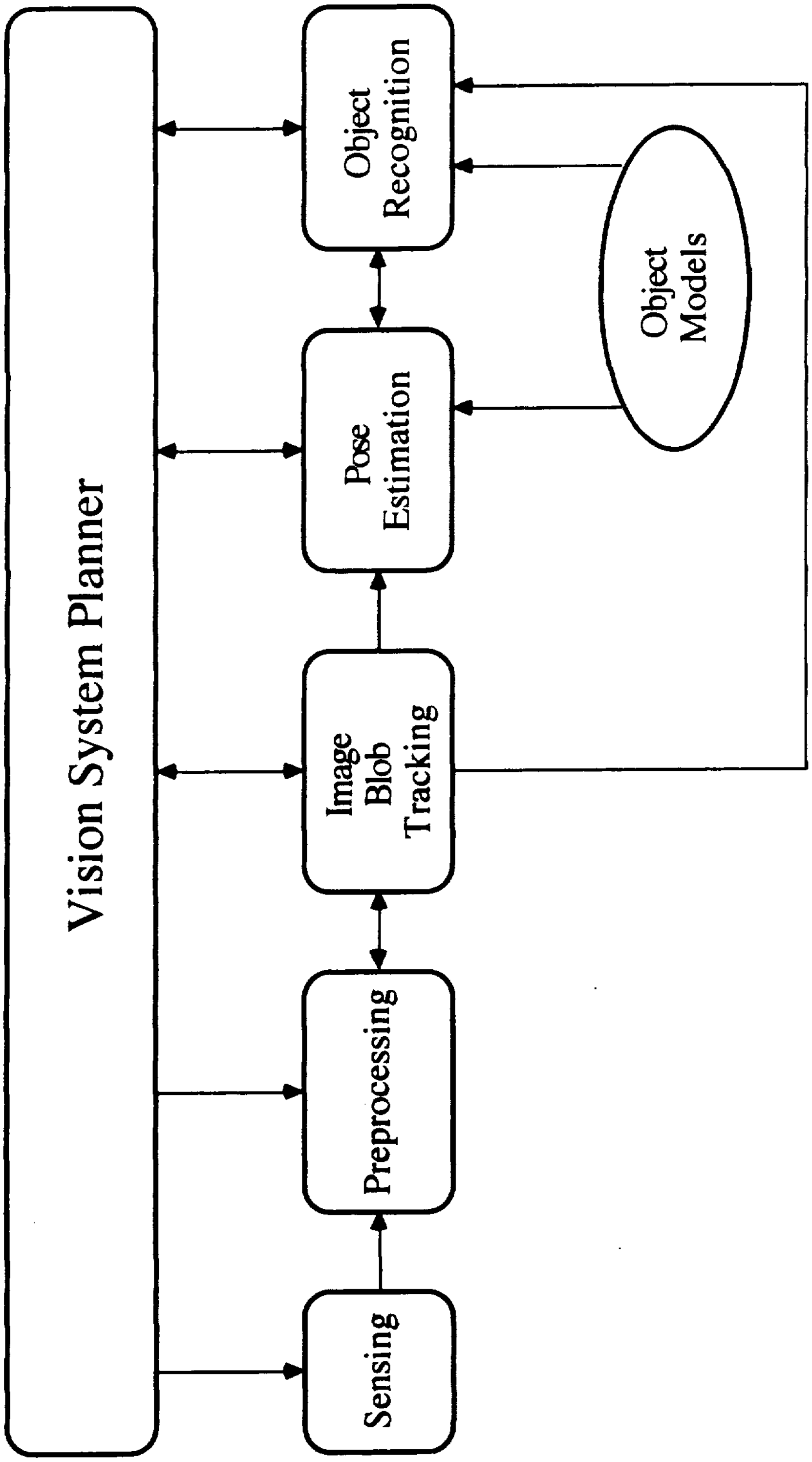

Figure 2: Vision System Architecture

achieving the current visual goal as specified by the task planner. Second, if the sensor that the VSP would normally select as its first choice to achieve the goal is either unavailable or inappropriate for usage because of some current constraint, it may be possible to perform the desired task using the other sensor to achieve the same goal, albeit perhaps by accepting a penalty in performance. Finally, instances may occur for which it is desirable to verify results from two different sensory sources.

The first of the above motivations addresses achieving the visual goal in the most effective manner by allowing the VSP to choose among sensors with complementary capabilities. For example, if it is desired to distinguish between two objects of similar structure with the color of the objects being the primary differentiating feature, then it is apparent that the color camera should be used as the primary sensor. On the other hand, if the size and/or geometry of the objects are most useful for determining identity, then it is important to be able to expeditiously extract and process three-dimensional coordinates. Clearly, this is a task that would be most appropriately assigned to the laser scanner.

The previous example involving the need for three-dimensional coordinates is illustrative of a case in which the primary sensor (the laser scanner) is engaged to extract the required information. However, there may be cases for which the laser scanner cannot be used to obtain range information because (a) the object to be processed is covered with a highly specularly reflective material thus preventing acquisition of good return signals, (b) the laser scanner is currently assigned to another task, or (c) the laser scanner is temporarily not functioning properly. For such instances, it is highly desirable to provide a redundant capability by using the other sensor if possible. The classical method for determining three-dimensional coordinates from intensity images involves a dual (stereo vision) camera setup in which feature correspondences are established and the stereo equations are solved for each pair of feature points. Although the current simulated configuration has only one intensity image camera, this alternative mechanism for computing range values is in fact possible for the VSP to achieve by requesting the task planner to reposition the EVAR such that the camera's initial and final positions are offset by a known baseline distance. Of course, there is a penalty in performance if the (pseudo) stereo vision method is chosen, since the EVAR must be moved and feature correspondences computed. However, it is nevertheless important to have such a redundant sensing capability for the reasons previously mentioned and to be able to independently verify the results obtained from one sensor or to increase the confidence of those results.

With respect to increasing the confidence of computed results, a typical scenario might involve a case in which the EVAR is close enough to a target object to hypothesize its class based on color, but too far away to definitively recognize its geometric structure using laser scanner data. In this case, the VSP would tentatively identify the object (using color) and would advise the task planner to move closer to the object so that a laser scanner image with higher resolution can be obtained. The confidence of the initial hypothesis would then be strengthened (or perhaps weakened) depending on the conclusion reached by processing the range data at close proximity.

The fundamental architecture for the Vision System includes modules which are designed to detect, recognize, track, and estimate the pose of objects. Upon receiving a request from the main task planner to achieve one of these objectives, the Vision System Planner determines an appropriate sequence of goals and subgoals that, when executed, will accomplish the objective. The plan generated by the VSP will generally involve (a) choosing an appropriate sensor, (b) selecting an efficient and effective algorithm to process the image data, (c) communicating the nominal (expected) results to the task planner or informing the task planner of anomalous (unexpected) conditions or results, and (d) advising the task planner of actions that would assist the vision system in achieving its objectives. The specific plan generated by the VSP will primarily depend on knowledge relating to the sensor models (e.g. effective range of operation, image acquisition rate), the object models (e.g. size, reflectivity, color), and the world model (e.g. expected distance to and attitude of objects). The next section presents the resulting plans generated by the VSP for several different scenarios.

RESULTS

The operation of the VSP that was designed and implemented can best be understood by examining the plans that it generates for various scenarios.

Scenario 1:

State of the world:
Three objects are somewhere in front of the EVAR. One of them is an Orbital Replacement Unit (ORU) with a known color.

Command received by the VSP:
Search in front of the EVAR for an ORU.

Plan generated by the VSP:

1. Search the hemisphere in front of the EVAR by activating the color camera, fixing the effective focal length and spiralling outward from the center until the object is found.

2. If the ORU is found, terminate the (spiralling) search and iteratively refine the estimate of where the object is located by adjusting the sensor gimbals toward the object and reduce the field of view (telephoto zoom) until the object is centered and large in the image.

   If the ORU was not found, the VSP reports failure, in which there are several actions that could be taken. First, the forward hemisphere could be rescanned at higher magnification (a slower process since more scans will be required). Second, the forward hemisphere could be rescanned with increased illumination (requiring a decision to be made regarding the desirability in terms of overall objectives and power consumption by the illumination source). Finally, the VSP could request the Task Planner to rotate the EVAR by 180 degrees and scan the rear hemisphere.

Scenario 2:

State of the world:
Same as Scenario 1

Command received by the VSP:
Determine the distance to the ORU, no sensor specified.

Plan generated by the VSP:

1. Locate the ORU as in Scenario 1 using the color camera.

2. Examine the object model for an ORU and determine which sensor is the most appropriate to be used. In this case, since an ORU is not specularly reflective, the laser scanner is chosen.

3. Examine that part of the laser scanner image that corresponds to the region belonging to the ORU in the color image and compute the distance to those range image elements.

Scenario 3:

State of the world:
Same as Scenario 1

Command received by the VSP:
Determine the distance to the ORU, but force the estimation of distance using single camera lateral stereo vision.

Plan generated by the VSP:

1. Locate the ORU as in Scenario 1 using the color camera.

2. Move the EVAR left a known distance, take an image, and record the location of the ORU in that image. Then move the EVAR right a known distance, take an image, and record the location of the ORU in that image.

4. Using triangulation (stereo vision with two cameras separated by a known baseline distance) compute the distance to the ORU.

Scenario 4:

State of the world:
Same as Scenario 1

Command received by the VSP:
Determine the distance to the ORU and move toward the ORU along the optical axis of the color camera until the EVAR is a specified distance (D) away from it.

Plan generated by the VSP:

1. Locate the ORU as in Scenario 1 using the color camera.

2. Estimate the distance to the ORU (D_{oru}) using the laser scanner.

3. Compute a vector along the optical axis of the color camera whose length is (D_{oru} - D). Transform that vector into EVAR coordinates and move to that position, maintaining the same attitude.

Scenario 5:

State of the world:
Same as Scenario 1

Command received by the VSP:
As in Scenario 4, determine the distance to the ORU and check to determine whether any other objects in the field of view are closer to the EVAR than the ORU prior to moving toward it.

Plan generated by the VSP:

1. Locate the ORU as in Scenario 1 using the color camera.

2. Estimate the distance to the ORU using the laser scanner.

3. Search the range image for values that lie outside of the region containing the ORU and report a potential obstacle if any of the values indicate distances between the EVAR and the ORU.