Complex behaviour generation on autonomous robots: A case study

M.C.Garcia-Alegre, P.Bustos, D.Guinea
Instituto de Automática Industrial
Consejo Superior de Investigaciones Científicas
Arganda del Rey. 28500. Madrid
Spain.e-mail:maria@iai.es

ABSTRACT

This paper proposes a framework for developing distributed architectures aimed at complex behavior generation in autonomous robots. Without engaging to radical "mentalistic" or behaviorist ideas, the system developed is a useful tool to explore scalability, modularity and complexity issues in building autonomous robots. To show its validity, a case study is presented in which a mobile robot endowed with a vision system performs the "goto target" task. Several topics, including the identification of classes of problems that might constitute building blocks of more general forms of intelligence in autonomous robots, the use of space-variant vision or the problem of "getting stuck" for a robot, are addressed.

1. INTRODUCTION

Technological Development is fed from all efficient solutions found when solving specific problems addressed by human activity. Robotics and Artificial Intelligence have proven as quite young areas but strongly spread out in applications as: Manufacturing, Mining, Agriculture, Health, Transport, Energetic resources, Services. With one foot on each of the two former areas, an attracting research field is actively emerging during last years, with a main concern: The the organization and emergence of autonomous behavior in artificial systems.

During the last ten years, robotics researchers have maintained a strong debate around the action selection problem [1]. In one side, the "mentalistic" school has defended the need of global explicit representations of the robot, the environment and the relation between them, in order to plan ways to solve problems. This planning capability reduces to a search process on a state-space [2]. By searching, the robot predicts the outcome of its own actions as far in the future as it is needed. The process ends when the robot and the environment are finally "imagined" in some desired situation. From this point some Executive process takes on the task of moving the robot checking a list intermediate perceptual conditions that must be met. Although many variations have been introduced to this scheme since its origin back in the 60's, the main idea remains: given a perfectly designed and updated forward model of the environment and of the robot, a search algorithm and a set of heuristics expressing our knowledge about the inverse model, should find the way to any goal state. Maybe, the problem is not the idea itself, but trying to apply it everywhere and as the solely mechanism of intelligence. Furthermore, less radical approaches to symbolic intelligence propose a shift to this central dogma, suggesting that the key role of high level problem solving is to obtain the most adequate representations for a specific problem that avoid extensive searches [3], [4].

On the other hand, behavior-based A.I. tries to find the inverse models of certain classes of problems considered as basic elements of intelligence such as: moving without colliding, searching and tracking stimuli, reaching and grasping objects and, more recently, some combinations of these ones. The results of this research have remarked the need of more detailed study of the dynamic interactions between the agent and its environment [5],[6],[7]. When taken into account , it turns out that this dynamics simplify dramatically the machinery needed to perform complex tasks. The change in the way of thinking that this approach pursues is to focus on solutions to classes of problems rather than to build complex searching machinery that relies on ideal forward models of uncertain and unpredictable robots and environments.

As it becomes obvious, it is necessary some sort of predicting capabilities in order to avoid doing always what you already know how to do, even if it is perfectly correct under certain situations. The reason is

0-7803-2559-1/95 $4.00 © 1995 IEEE

adaptation to environments and situations not contemplated in the designed solutions (i.e. behaviours) [8], [9], [10]. A pure behavior-based built robot is trapped in the class of problems which guided its design [11]. The way out of this will probably need representation, but in being so, we might be forced to accept the behavior as our primitive building element. This fact entails new problems such dynamically grouping behaviors to perform more complex tasks, representing concepts as behaviors that encapsulate the sequence of actions and sensations needed to perceive some object or class of objects, new metrics to manipulate this concepts, and so on.

In the remain of this paper we propose a framework for building hierarchical architectures in which it building elements are agencies. The aim of this system is to explore some of the questions stated above. To show the validity of this approach we have built a mobile robot endowed with a pan-tilt camera which is to achieve a class a problems known as the "goto target task".

2. A FRAMEWORK FOR BUILDING AGENCIES

In order to quickly develop complex sets of interconnected behaviors, we propose a simple software structure that meets three important requirements: modularity, flexibility and dynamic developing. By modularity we mean the capacity of being expanded through heterogeneous hardware, avoiding the problem of getting out of resources. Flexibility is used here in the sense of avoiding "a priori" constraints on the size and complexity of each software element. Dynamic developing means the possibility of adding and deleting agencies from the system without having to restart everything each time. This feature has reduced developing times dramatically and, furthermore, allows several users to program different modules concurrently from different machines, even when these new agencies use common existing ones.

The solution adopted is the use of UNIX processes connected by sockets. Communication software allows each process to behave as a sockets server and client, accepting and requesting the needed connections each time it is run. Over this framework we now define a terminology that will be used in the description of the case study:

Agency

Any computational process addressed to reach and/or to maintain a goal. It represents the formal unit in the Architecture and covers either deliberative or reactive algorithms in any representation. It does not impose any restriction on the complexity of the algorithms that hosts, acting as a encapsulation tool for functional modularization. Each Agency is composed of both computation and communication code. The term Behavior , will be used indistinctly as a synonym of Agency.

Communication channels

Any two agencies can communicate through a bi-directional channel. Agencies that already accept modulation can serve new connections to any other agency that requests it. The idea is, thus, to ease the process of incrementally build new agencies using all the functionality offered by the existing ones. Channels have not semantics and the protocols must be defined directly between communicating agencies

Basic Agencies

Agency or set of agencies that solve a specific class of problems, for example collision-free navigation, searching, etc. The functionality of a basic agency is a general resource to be used by other elements of the architecture. The mechanism by which a basic agency allows the instantiation of its own behaviour is called "modulation.".

Information flows

From a global perspective, there are two flows of information: a goal-driven top-down one and a event-driven bottom-up one. The first one relies on the modulation mechanism to use lower level abilities of the system. The second one propagates upwards information about the actual behavior of the robot in interaction with its environment.

Modulation

An agency modulates another when it requests some solution of the class of problems that the requested one can achieve. In practice there is no guarantee that the request will be achieved. Time out mechanisms and alternative strategies must be available to avoid the freezing of the whole system

Propagation

An agency propagates information when it is requested explicitly to do so or when a modulated request could not be accomplished within some specified conditions.

This framework is intentionally kept simple because of the idea that complexity must grow in parallel for the robot, its machinery, the environment and the tasks it performs.

3. A CASE STUDY: APPROACHING VISUAL TARGETS

The proposed methodology has been implemented and demonstrated on a mobile robot, fully designed at the IAI-CSIC: BOSS. The "Body", is a mobile platform endowed with a rotating ultrasonic "Ear" and a vision subsystem composed of a CCD camera (512x512) mounted on a pan/tilt head (2 dof) acting as an "Eye". The vehicle "Brain" is an on board Pentium/90 under LINUX operating system, that communicates with hosts UNIX Machines via an Ethernet link, Fig.1.

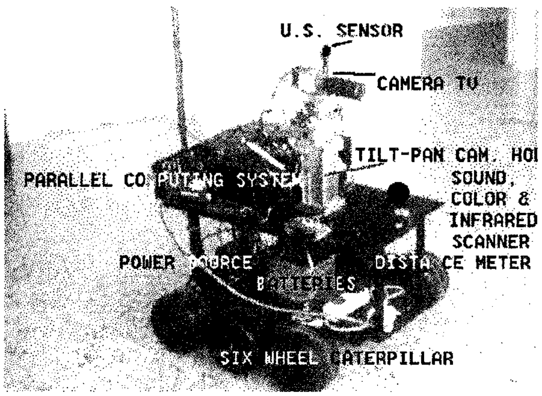

Fig. 1 BOSS prototype robot

The goal task for this robot is to "go towards specified visual stimuli". The accomplishment of this high level task implies the achievement and keeping of many other simpler goals. To show the machinery built to achieve this task we will proceed with a top-down analysis of the subtasks involved, describing the agency or agencies designed to accomplish it. A global view of the architecture is shown in Fig.2.

Reaching targets

To make the robot approach and specified target the following tasks must be accomplished:

search the target
track the target
pilot the robot towards the target

The three of them get activated with the following policy:

WHILE target not reached and not time-out
IF target is located
THEN TRACK target and PILOT direction
ELSE SEARCH target

GOTO target Agency: The goal of this agency is to reach targets specified by the human. In this case the human acts as a higher level agency that modulates the GOTO agency. The communication between them is implemented through a simple vocabulary to describe characteristics of the desired targets and of the piloting. Examples of these terms are: fast, regular, slow, rectangle, still, moving, short, medium, tall, black, left, right. The agency accepts goals as sentences describing a target and, maybe, a way of approaching it such as "GO fast towards tall moving rectangle".

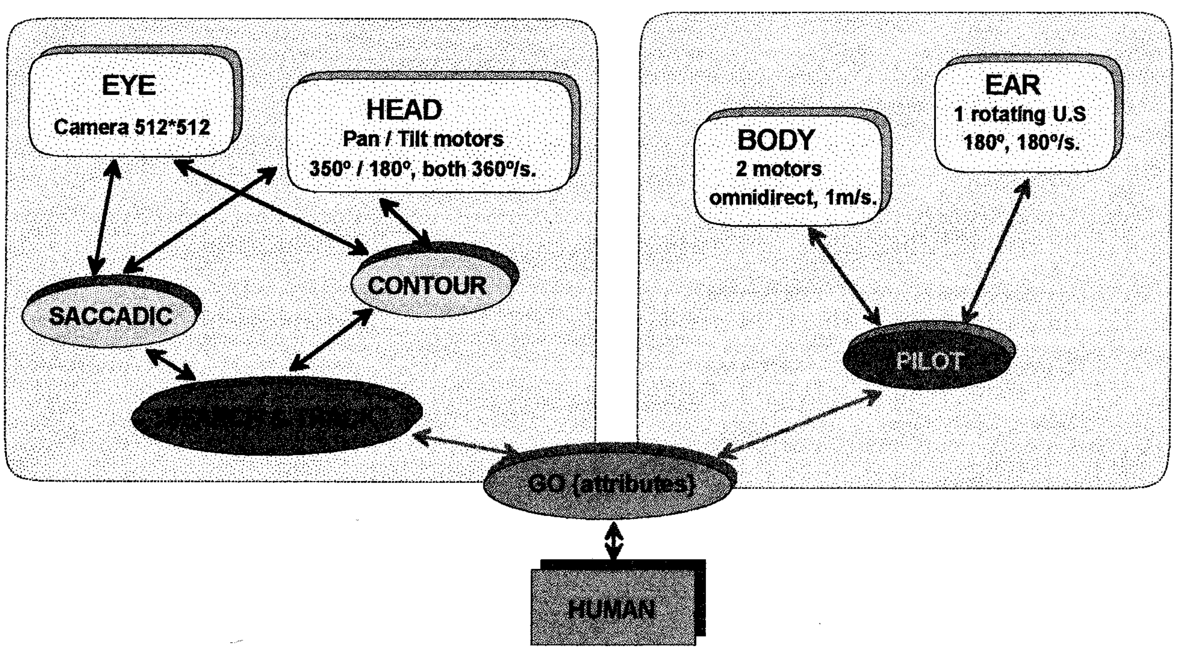

Fig. 2 Proposed distributed architecture

To achieve this goals, it modulates the lower agency SEARCH&TRACK sending it the description of the desired target and waits for completion or for a time-out event. If the target is located, the agency activates PILOT telling it to turn the robot until the camera axis is aligned with the robot longitudinal axis. If the action was possible, the next command will be a forward move, which will be maintained until the distance to the target reaches a pre-defined threshold. PILOT might not be able to even start the action and will notify upwards, or might need to avoid an obstacle. In this case the camera and robot axis will go out of line again, given that SEARCH&TRACK are still active.

The GOTO agency will take care of this situation by "reminding" the PILOT to turn in the opposite direction. As soon as the obstacle is surpassed, PILOT will accept this modulation correcting the robot orientation. This example shows how it is, in principle, possible to obtain a long term coordinated behavior from the simple interaction of two simpler ones, through the environment and a supervisor.

Collision-free navigation

Moving the robot without colliding with objects is a well known task demonstrated in many existing robots. Basically, it is needed some estimation of the distance to surrounding objects and a way to control the movement of the vehicle. Behavior-based approaches to this problem combine both low times of response and simplicity.

PILOT Agency: This Agency uses the information from lower level agencies EAR and BODY to ensure mobile robot safe motion avoiding obstacles. While not avoiding an obstacle, it accepts modulation from the GOTO agency specifying changes of direction. To obtain a local representation of the free space in front of the robot it modulates de EAR agency, which in turn performs a five step sweep of the ultrasonic support. When an obstacle becomes to close it triggers an avoiding strategy, turning left or right depending on previous modulations, or even backwards if the obstacle is too close.

EAR Agency: Controls the positioning of the ultrasonic support that is placed in the front part of the robot. It connects to a specific microcontroller-based board through a high speed (20 Mbits) channel, that holds the circuitry necessary to control a Futaba servo motor and to trigger the ultrasonic sensor.

BODY Agency: Accepts commands to modify the speed of the robot. It communicates with a specific microcontroller-based board through a high speed (20 Mbits) channel. This board holds dedicated PID chips that perform the speed control of each wheel.

Searching and tracking targets

In order to quickly find a specified target in the scene, a visual system must use all the information available about the target. Furthermore, when using a foveated image the portion of the visual field where accurate recognition can take place is usually small. Taking into account these facts, the searching task can be divided into two subtasks: segmenting and recognizing the object currently centered in the field of view, and finding new targets locations in the periphery of the visual field to redirect the focus of attention. Tracking is performed by centering the contour of the recognized object in the field of view. When the target is lost, the same searching procedure in the periphery can be activated, constraining the search to the direction by which the target disappeared. Again, the interaction of two agencies along with its eventual coordination, allows for a solution to a new, more complex class of problems.

SEARCH&TRACK Agency : This agency's goal is the location and tracking of specified visual stimuli. On success or failure it notifies to the GOTO agency. To accomplish it task it modulates two lower level agencies: SACCADIC and CONTOUR with the following policy:

```
WHILE active goal and not time-out
        if CONTOUR not recognized
        SACCADIC next target direction
```

GOTO agency relies on SEARCH&TRACK ability to recover from tracking failures by continuously alternating between searching the scene and smooth pursuing the centroid of a recognized target. In case of definitive failure, GOTO would be notified and whether new information would be supplied to SEARCH & TRACK to restart the search, or a message would be passed up to the human.

SACCADIC Agency: This Agency shift the focus of attention towards different scene locations. To achieve this, it selects targets from the periphery of the visual field using the constraints suggested by higher level Agencies, and commands target positions to the HEAD agency.

Periphery is dynamically defined as the outer region of the current segmented contour. It is connected to EYE Agency to get visual images, to HEAD Agency to command changes in the position, and to SEARCH&TRACK Agency to receive information about current goals. Possible goals are all camera motions within the field of view, but can be modulated by constraints that exclude certain regions of the periphery (gray level, size, shape, direction)

CONTOUR Agency: Continuously extracts a closed contour from the edge filtered image. Initially, it explores each radial direction on the log-polar (see low-level vision section) image searching for a pre-defined edge threshold. This contour is further on adjusted accounting for the contour inner and outer gradient direction, and the relative position of neighbor contour pixels, Fig.3. At the same time, the contour is centered in the visual field by means of small movements of the camera. This control loop actually tracks the segmented object up to a speed limit established by the 12 Hz. frame processing rate. Once the contour becomes stable, the inner region is classified according to the description of the target. Classification is done by checking conditions in increasing order of computational cost. This mechanism allows for discarding non valid targets with maximum efficiency. CONTOUR, currently accepts descriptions of objects in terms of gray level, elongation, orientation and size. This description space is enough for locating a relevant number of locations at the laboratory. CONTOUR is connected to the EYE, HEAD, and SEARCH&TRACK agencies.

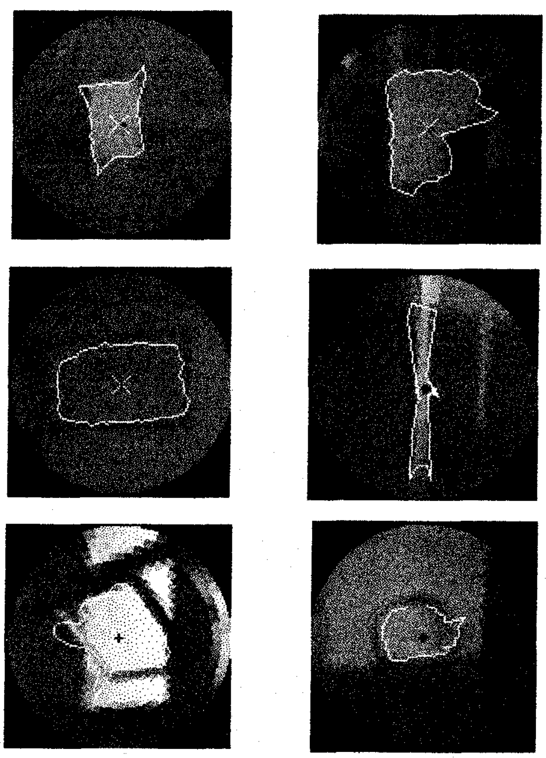

Fig. 3 Extracted contours of some objetcs:

box, occluded wall, blackboard, table leg & ring

Low-level Vision

In this subtask we include the initial transformations on the image as well as the position control of the camera. Also, other low level control loops could be include here, such as focus and iris control.

EYE Agency:

This Agency carries out all image transformations required to extract relevant information related to the 2D structure of the scene. It uses a Transputer pipeline of three stages summing up nine processors. The first stage computes a log-polar transform that simulates a space-variant retina-like sensor [12], [13]. The output of these sensor is a 64x32 (angular and eccentricity resolution respectively) image that keeps a good resolution near the center of the visual field (fovea) while being of a very reduced size. The second stage computes a Laplacian of the log-polar image.

HEAD Agency:

This agency is similar to the BODY agency, in the sense that it uses and analogous control board. It differs in that HEAD admits position control commands and also maintains the pixel to encoder calibration map in order to accept position requests directly from visual coordinates.

4. CONCLUDING REMARKS

This case study shows how a rather complex behavior generation system can be developed from a simple computational framework. This can only be accomplished if we study the interaction dynamics of our robot, its machinery and the environment, obtaining cues to develop new abilities from the already existing ones. So far, complexity in behavior generation can only be faced by a double direction analysis: top-down by identifying classes of problems suitable of becoming building blocks of larger aspects of intelligence, and bottom-up by combining existing abilities and studying its interactions through the environment. Much more work remains to be done with real , more complex robots. However, the development life-cycle of real robots is extremely tedious. An effort has been made to ease this problem by combining simplicity and modularity in a computational framework. Also, a general class of problems for visually guided robots has been used as case study , finding simple solutions that take into account the reusability of existing ones and the interaction dynamics of the robot and its environment.

5. ACKNOWLEDGMENTS

Present work is funded in completion by Spanish CICYT Grants: TAP92-0574 and TAP94-0171. Authors want to thanks Felicidad Recio and Jose Vicente as responsible of the vision and mobility subsystems, and Miguel Schneider and Lourdes Agapito for valuable discussions and suggestions.

6. REFERENCES

[1] P.Maes, "The Dynamics of Action Selection", in *IJCAI89*, Detroit, MI, 1989, pp.991-997.

[2] J.S.Albus, H.G.McCain, R.Lumia, "NASA/NBS Standard Reference Model for Telerobot Control System Architecture (NASREM)", *NIST Techn.Note 1235*, April 1989.

[3] N.J.Nilsson, "Problem Solving Methods in Artificial Intelligence", *McGraw Hill*, 1971.

[4] A.Newell, H.A.Simon, "Computer Science as Empirical Inquiry: Symbols and Search", *Comm. of the ACM*, 19-3, 1976, pp.113-126.

[5] P.E. Agre, "The Dynamic Structure of Everyday Life", *Ph.D. Thesis*, MIT, 1989.

[6] D.Chapman, "Planning for Conjunctive Goals", *Artificial Intelligence*, 32, 1987, pp.333-377.

[7] D. Chapman, "Vision, Instruction and Action", *MIT Press*, 1992.

[8] M.J. Mataric, "A Distributed Model for Mobile Robot Environment-Learning and Navigation", MIT EECS Master's Thesis, *MIT AI Lab.Tech.Report AITR-1228*, May 1990.

[9] M.C.Garcia-Alegre, A.Ribeiro, J.Gasos, J.Salido, "Optimization of Fuzzy Behavior-based Robots Navigation in Partially Known Industrial Environments", *Proc. IEEE Intern.Conf.on Industrial Fuzzy Control & Intell.Systems*, Houston, TX, 1993, pp.50-54.

[10] M.C. García-Alegre, P.Bustos, D.Guinea, " The Behavioural Agents Approach in a Hierarchical Architecture for Robotics", *Proc. Intern. Workshop on Advanced Automation ECLA94*, Madrid, Dec.1994.

[11] R. Brooks, "A Robust Layered System for a Mobile Robot", *IEEE J. of Robotics and Automation*, Vol.2, No.1, March 1986, pp. 14-23.

[12] G.Sandini, F.Gandolfo, E.Grosso, M.Tistarelli, "Vision during Action", in: *Active Perception*, (Y.Aloimonos Ed.), LEA Pub., 1993, pp.151-190.

[13] P.Bustos, F.Recio, D.Guinea, M.C.Garcia-Alegre, "Cortical Representations in Active Vision on a Transputer Network", *First ECPD Conf.on Advanced Robotics and Intell.Automat.*, Sept.1995, Athens.