



# Basic Agents for Visual/Motor Coordination of a Mobile Robot

Maria C.Garcia-Alegre, Felicidad Recio

Instituto de Automática Industrial  
Consejo Superior de Investigaciones Científicas  
28500 Arganda. Madrid. Spain  
maria@iai.csic.es

## Abstract

Present work addresses the guidelines that have been followed to construct Basic Behavioral Agents for visually guided navigation tasks within the framework of a Hierarchical Architecture . Visual and motor interactions are also described within this generic framework that allows for an incremental adaptation of knowledge, to be reused in tasks of ever increasing complexity. Basic Locomotion Agents as, Stop&Backward, Avoid, and Forward are implemented as Fuzzy knowledge - based systems to embody the uncertainty and imprecision inherent to real system-environments. Basic Visual Agents as, Saccadic, Contour, and Center have been developed under a space-variant representation in an anthropomorphic approach. Coordination of Basic Agents has been demonstrated with a caterpillar type mobile robot in tasks related to qualitative descriptions of visual stimuli.

## Introduction

For years A.I. research has devoted most of its efforts to complex problem solving models disregarding real worlds, facing simple environments without noise and uncertainty that could be precisely modeled. Emphasis has been put on explicit knowledge, rational selection, and problem solving strategies (Nillson 1971). In contrast to these traditional approaches , reactive and behavior-based systems offer a methodology to design structures, with low level of deliberative complexity but embedded in complex, uncertain and noisy environments.

As a consequence during last decades an ever increasing interest has appeared in the design and construction of models that account for relevant interactions with real environments and its significance in the proposed solutions (Maes 1990). Such architectures lead to a firm coupling between perception and action in distributed systems able to deal with scarce knowledge and resources, in dynamic environments .

*Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of ACM. To copy otherwise, or to republish, requires a fee and/or specific permission.*  
*Agents'97 Conference Proceedings, copyright 1997 ACM.*

One of the problems involved in such architectures is that of action selection, and the distinction between planning and reaction, that direct most research to hybrid system that fuse both components (Hayes-Roth 93), (Horswill & Stein 1993). The action selection paradigm is discussed in this work, as the branching mechanism in a Hierarchical Architecture of Behaviors. A framework is proposed to incrementally build up a hierarchy of Behaviors from a set of basic and well-fitted visual and locomotion Agents, that embrace deliberative and reactive capabilities.

## Initial concepts and definitions

The rapid growth of interest in Multi-Agents systems, during last decades, has brought an associated difficulty concerning basic terms and definitions. The wide range of areas, Artificial Intelligence, Robotics, Software Engineering, in which these terms are applied, give rise to a lack of consensus in the meaning of terms as, Agent, Agency, Behavior, Autonomy. They are frequently used to express quite different notions with different implications.

Here are some of the definitions of the concepts that will be used from now on:

*Agent:* Represents the formal unit of the Architecture and hold for both deliberative and reactive aspects without any constraint on its complexity, thus acting as an encapsulation tool for functional modularization (Garcia-Alegre, Bustos, & Guinea 1995). In this work *Agent*, *Agency* and *Behavior* will be indistinctly used in the sense of a control / computational process addressed to reach and/or maintain a goal (Mataric 1994).

*Basic Agents:* Are defined as the elementary building blocks that are needed to solve a specific class of problems, for example collision-free motion. In order to generate ever complex Agents a simple software structure is contemplated having three main properties: Modularity, Flexibility and Incremental development. Modular in the sense of being able to be expanded through heterogeneous hardware to avoid getting out of resources. Flexible to elude "a priori" constraints on the size and complexity of each software element. Incremental development to dynamically add and delete Agencies without disturbances in the system.

*Level:* Is conformed by a set of Agents that share a common language composed of a set of terms related to perception, knowledge representation, and actions. A level gravitates around a specific representation of the system-environment, and its Agents are modulated by higher level ones. Levels may have a short or long term memory to optimize the observable behavior and compensate for an excess of activity. This memory may support learning capabilities to improve the overall observable behavior competing the level.

### To a Hierarchy of Levels

A hierarchy of levels is proposed to handle complexity, where each level is associated to a degree of abstraction, time-response,

and representation. In this hierarchy representation is understood as a means of internal coordination at each level (Van den Velde 1995).

On the other hand, the main properties postulated for a level reduces development times and allows several users to program different modules concurrently from different machines, to use the common existing ones. Levels in the hierarchy communicate through bi-directional channels, that have not an associated semantics, and consequently the protocols must be defined directly between communicating Agents.

The aim that underlies this structure is to ease the incremental process of new Agents construction, by reusing all the functionality offered by the existing ones, starting with the set of Basic Agents.

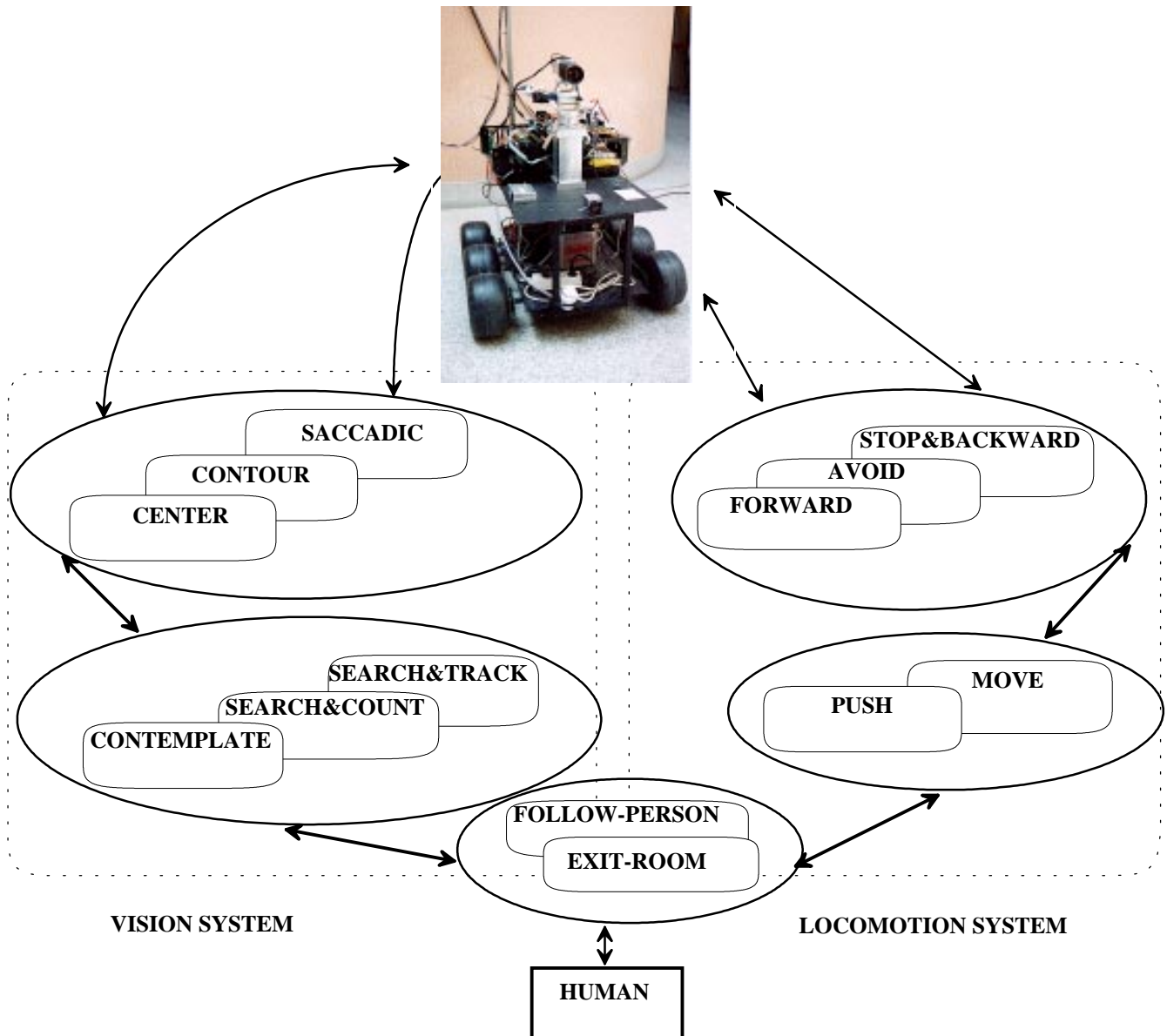


Figure 1. A Hierarchical Architecture of Behaviors

From a global Perception-Action perspective, there are two information flows: goal-driven top-down and event-driven bottom-up. The first one depends on the arbitration mechanism to use lower level abilities of the system. The second one rely on the upwards information propagation of the robot-environment state.

*Arbitration:* Is the mechanism by which a Basic Agent allows the instantiation of its own behavior. Arbitration is a run-time process which depends on the goals of the system and the situation it finds itself in. Agents of a level are arbitrated by Agents of a higher level that requests some solution of the class of the competed problems, in order to reach their goal. In practice there is no guarantee that the request will be achieved, hence time-out mechanisms and alternative strategies must be available to avoid the freezing of the whole system.

*Propagation:* Corresponds to the perception information flow from lower to higher levels. An Agent propagates information when it is explicitly requested to do so or when an arbitrated demand can not be accomplished within the required constraints. This framework has been kept as simple as possible under the assumption that complexity must grow in parallel for all three: robot, environment and tasks to be autonomously performed.

In Figure 1, it is schematically shown the developed architecture for a robot that has to achieve some visually guided tasks, in an office environment. Most Agents have already been implemented, but EXIT-ROOM and SEARCH&COUNT are still in progress. Behaviors are initiated under the requirements of a mission directed by a human being

## Basic Locomotion and Visual Agents

A great variety of motion and visual complex tasks can be described and accomplished in terms of dynamic sequencing of elementary Agents. These Agents are intended to be the minimum set needed to perform a diversity of tasks. Some specific desirable characteristics of an Agent are efficiency in its goal-directed behavior, reactivity to unexpected events, and adaptability to dynamic and unpredictable environments.

For each domain, Vision and Locomotion, a set of Agents have been found that are called Basic in the sense that they are the minimal set required for generating other Agents in order to reach de desired goal-repertoire.

### Locomotion Agents

This Agents have been extensively studied all over last decades and correspond to basic reflex behaviors of approaching, running away and not colliding motion. This selection is based on the belief that intelligent complex behavior in a conventional ambience, would result from the combination, sequencing or inhibition of former Agents through simple arbitration rules. These Basic Agents are proposed as elements of a methodology for structuring such rules through synthesis and evaluation processes.

Three Basic Agents are selected at the locomotive domain, namely: STOP&BACKWARD, AVOID and FORWARD. They are proposed in a decreasing degree of emergency. All three are

embedded in a common skin, the Piloting Level, sharing a local world representation centered in the mobile robot platform. This representation is a mapping of ultrasonic, infrared, and tactile readings to distance and orientation of the obstacles in its vicinity.

**STOP&BACKWARD:** It is driven by the tactile and infrared sensors and is needed to recover from unexpected events due to sensor failures or very acute obstacles shapes. It appears as a branching in the actions, for identical perceptual conditions. As an example: IF close-to-contact conditions THEN FORWARD or IF close-to-contact conditions THEN STOP&BACKWARD, where first rule conforms a safe navigation and second one a push behavior. The appearance of more than one Agent at a level causes the existance of a higher level that arbitrates the action selection for identical sensory information.

**AVOID:** Is triggered by ultrasonic sensors. Its objective is the detection of a kind of objects: Obstacles. Supervises medium distance areas all around, and generates a detour action when a safety distance is reached. The selection of this value modulates the mobile behavior, from a very precautious to a "Fitipaldi" one. It also accords with a branching of FORWARD when different actions are required for the same conditions. An example would be: IF obstacle-foreseeing THEN FORWARD or IF obstacle-foreseeing THEN AVOID.

**FORWARD:** It is also driven by the ultrasonic sensors. Its aims the detection of a kind of situations, namely free-space. Moves the robot to a location specified in polar coordinates ( $\vartheta$ ,  $\rho$ ), and it is provided with time-out mechanisms.

These Agents are modeled as Fuzzy knowledge bases for decision making on the robot heading and speed, taking also into account proprioceptive sensors readings at a "clock-tic" of one hundred milliseconds. Key system-environment states are selected and defined as the relevant sensory patterns that trigger a Basic Agent, any other pattern not accounted for generates a default action: STOP.

In present approach the Basic Agents are triggered by mutually exclusive system-environment states, so as to completely skirt the exigency for explicit arbitration within a level. External and proprioceptive sensory variables are represented as Fuzzy Sets, and are defined through trapezoidal membership functions.

The set of Fuzzy rules that conform the Knowledge Base of the AVOID Agent to control the output variable  $\Delta v$  is displayed in Figure 2. These rules are designed keeping in mind simplicity:

**R1: IF** dist\_obst is FAR. AND. speed is LOW  
**THEN**  $\Delta v$  is POSITIVE

**R2: IF** dist\_obst is NEAR. AND. speed is HIGH  
**THEN**  $\Delta v$  is NEGATIVE

**R3: IF** dist\_obst is FAR. AND. speed is HIGH  
**THEN**  $\Delta v$  is ZERO

R4: **IF** dist\_obst is NEAR. **AND**. speed is LOW  
**THEN**  $\Delta v$  is ZERO

Figure 2. Knowledge Base for  $\Delta v$  Action in AVOID Agent

### Visual Agents

They are designed to selectively react to visual stimuli in a dynamic environment under the requirements of a specific task (Tistarelli & Sandini 1992). Human vision forms multi-resolutional images at the visual cortex due to the non uniform photo-receptors distribution. Inspired in the advantages of this biological vision system, the logarithmic-polar transformation has proved to be one of the most suitable approaches (Pardo et al. 1996) to implement this non-homogeneous sensing to make artificial vision systems space selective.

By taking this characteristic into account a foveated sensor has been implemented, mapping the image cartesian plane to an image cortical one, thus providing a space variant vision. This logarithmic-polar representation of the image is characterized by a variable resolution on the visual field, with a high resolution at the inner/foveal region that gradually decreases toward the periphery.

To quickly find a specified target in a scene, a visual system must use all the information available about the target. Furthermore, when using a log-polar representation the visual field for accurate recognition correspond to a very short area. Thus Searching and Tracking tasks must be divided in consecutively subtasks: Segmentation and Recognition of the object currently at the field of view center, Centering, and finally rapid motions to periphery to look for and to center new targets (Bustos 95). Based upon the former premises three Agents are considered as Basic Visual Behaviors:

**SACADDIC:** This Agent shift the focus of attention towards different scene locations, by selecting targets from the periphery of the visual field. Possible goals are all head/camera motions within the field of view, arbitrated by the higher level Agent conditions. With such a policy attention is focused only in small image regions.

**CONTOUR:** Extracts a closed contour from the edge filtered image. This contour is adjusted considering the inner and outer gradient directions, and the relative position of the contour adjacent pixels. Finally, the contour is centered in the visual field at a rate of 25 fps. Once stabilized, inner region is classified in accordance with the target description. Descriptions of objects are made in terms of gray level, elongation, orientation and size.

**CENTER:** Maintains the recognized contour in fovea by tracking the contour centroid at the inner sensor area. Periphery is dynamically defined as the outer region of the current segmented contour

The real-time dynamic interaction of these Agents is arbitrated by higher level Agents as: **SEARCH&TRACK** or

**SEARCH&COUNT**, and allows for the solution of a multiplicity of cases that concern visual stimuli guided navigation.

### On Visual and Motor Interaction: Approaching Visual Stimulus

Interaction of Basic Locomotion and Visual Agents in tasks related to navigation towards specified visual stimuli has been demonstrated (Recio 95) with the caterpillar type Robot: BOSS-IAI which has been built at the IAI/CSIC. The robot is endowed with a rotating ultrasonic sensor (1 dof) as an EAR, and a visual system composed of a CCD camera (512x512 EYE) on a Pan/Tilt unit (2 dof) acting as a NECK/HEAD. All sensors are located on the BODY, a six wheels platform (2 dof) with an on board Pentium/90 (BRAIN), Linux O.S., that communicates with host Unix machines via Ethernet, Figure 1. Unix processes are connected by sockets, and the communication software allows each level to behave as both a sockets server and client accepting and requesting the needed connections each time it is run.

The missions for this robot are all addressed to solve a class of problems, involved with a "safe navigation guided by visual stimuli". The achievement of these goals implies the accomplishment of a sequence of simpler subgoals. An instance of the class of problems already mentioned will be analyzed in terms of the proposed architecture in a dynamic environment, that takes profit of some system-environment knowledge.

Having the goal of "FOLLOW-PERSON", that actually corresponds to one of the possible Agents embodied in the highest level, the system will proceed from a top-down approach to arbitrate the activation of Agents of the adjacent lower level. It is assumed motion de-coupling between head and body.

Activation sequencing will interactively depend on the dynamic system-environment information flow that propagates bottom-up and top-down (goal modulation). External representation, associated to a higher level, allows its Agents the coordination of lower level ones, thus acting as a mechanism of internalization of external representations. Here is a description of the conditions that focus the attention of the FOLLOW-PERSON Agent and the tentative actions generated to accomplish the goal:

```
<->FOLLOW-PERSON [Attributes]
  [WHAT?(gray level, size, free-space, safety-distance),
  WHERE?(pan and tilt area, proximity-distance)
  HOW?(slow, fast, accurate, fuzzy for HEAD and BODY)]
```

```
-->FOLLOW-PERSON [Actions]<=>
```

```
<->SEARCH&TRACK [Attributes]
  [WHAT?(gray level, size),
  WHERE?(pan and tilt area),
  HOW? (slow, fast, accurate, fuzzy)]
```

+

```
<->MOVE [Attributes]
  [WHAT? (free-space),
  WHERE? (J=Head-Body°, r= proximity-distance),
  HOW? (slow, fast, accurate, fuzzy)]
```

43AND

```
-->SEARCH&TRACK [Actions] <=>  
While active goal & not time-out  
  if CONTOUR not recognized then SACCADIC  
to next target direction  
  else CENTER target
```

+

```
-->MOVE [Actions] <=>  
While free-space distance > proximity distance FORWARD  
  if free-space distance = proximity distance STOP  
  else AVOID or STOP&BACKWARD
```

```
-->REPORTS (to HUMAN)
```

The SEARCH&TRACK Agent will take into account target attributes before determining the actions sequences to be

tentatively triggered at a first glance. On success or failure it notifies the FOLLOW-PERSON Agent. This Agent arbitrates three Agents at the adjacent lower level, Figure 1. When a target is lost, a search in periphery is activated to the direction of the target disappearance. Here again the interaction of the three Agents along with its eventual arbitration generates an emergent Behavior able to cope with a new, more complex class of problems. FOLLOW-PERSON Agent relies on the SEARCH&TRACK Agent ability to recover from tracking failures by continuously alternating between searching in the scene area and smoothly pursuing the centroid of the recognized targets. Same happens with the MOVE Agent able to solve most unexpected events, thus dealing with a safe navigation.

In very extreme situations both will report on detected failures to FOLLOW-PERSON that will look for a new strategy or will pass some message to the human being. MOVE Agent ensures mobile robot safe motion avoiding obstacles under the direction and distance requirements. To obtain a local representation of the free-space in front, it modulates the ultrasonic sensors.

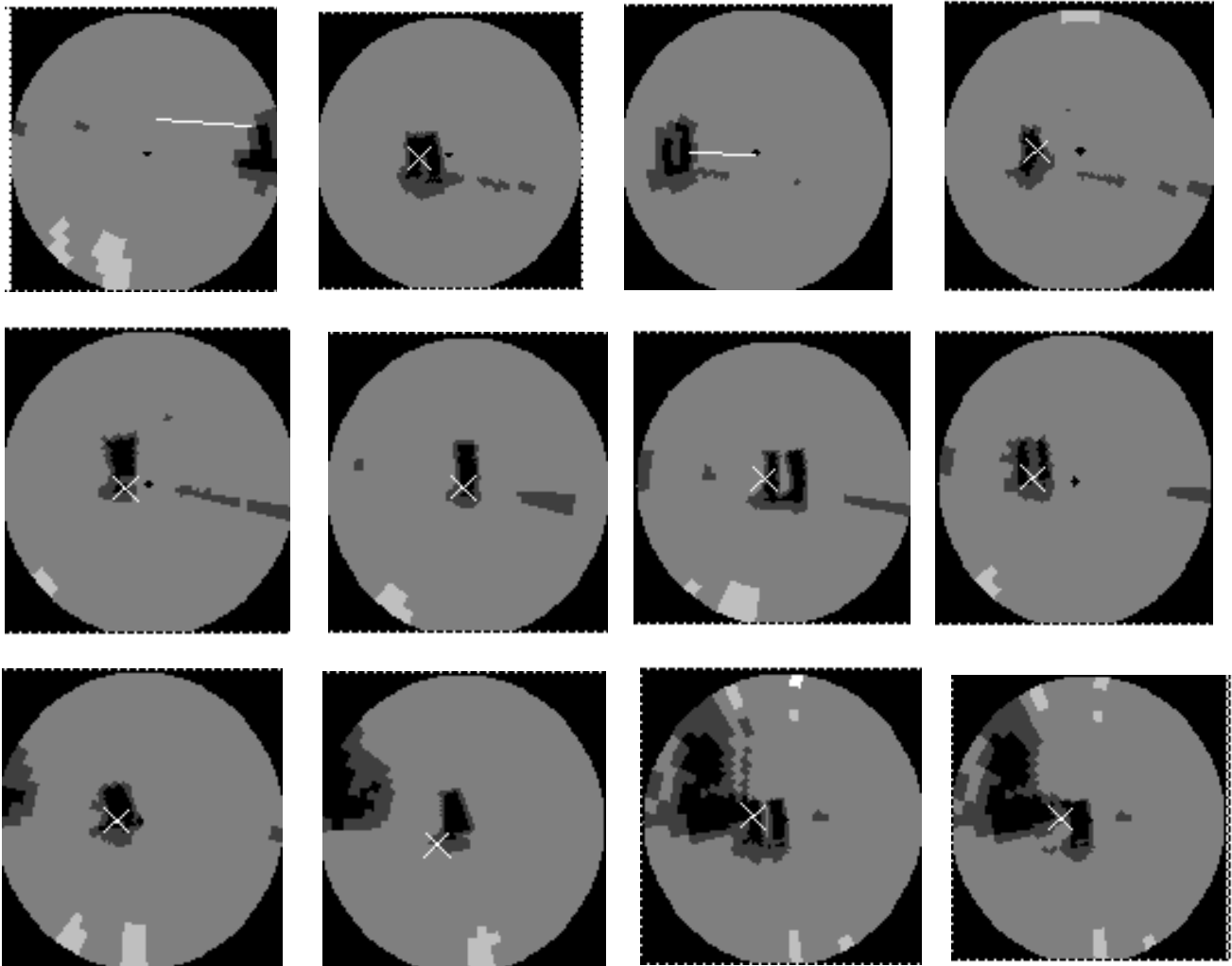


Figure 3. A tracking sequence of FOLLOW-PERSON Agent

An alternative to the MOVE Agent is offered by the PUSH Agent, which is defined as:

```
PUSH [Attributes]
[WHAT? (obstacle)
WHERE?(J=obstacle-direction, r= target_d, push_d),
```

```
HOW? (slow, fast, accurate, fuzzy)]
```

```
PUSH [Actions] <=>
if free-space distance<= safet_d .AND. curren_d < (target_d -
safety_d)
THEN AVOID or STOP&BACKWARD
else FORWARD
```

Here the second premise is concerned with the direction fixed at the PUSH attributes, being first premise implicated in the rest of the directions. A tracking sequence of the FOLLOW-PERSON Behavior is shown in Figure 3.

Unexpected behaviors sometimes emerge as a result of the interaction among, the system-environment, the proposed goal, and the explicit coordination algorithms. Experiments on observable behavior sweeping a wide range of initial conditions and scenarios, point to coordination as the main cue for well suited behavior generation.

## Discussion

The selection of a set of Basic Behaviors for Visual and Locomotion tasks is proposed as the elementary building blocks within a framework that aims the organization of mobile robot behavior in complex, dynamics and unpredictable environments. These Basic Agents are reused, arbitrated and coordinated by other Agents that perform more complex tasks, in an office-like environment.

The Basic Locomotion Agents for the Robot are formulated by means of fuzzy rule-based systems approaching human solving strategies. They generate smooth trajectories to attain a location within the allowed tolerances. Environmental and structural knowledge is injected as a set of attributes to help cutting the perceptual vision and motion space to a reduced area of expectations, in order to reach real time behaviors.

The proposed architecture is an appropriate framework for agent-environment interactions descriptions, that allows for an ease incremental growing up of the class of problems that can be tackled with the available perception system.

## Acknowledgments

Present work is funded in completion by Spanish CICYT Grants: TAP94-0171 and TIC96- 1392-C02- 01. Authors want to thank G.S.Mezquita for valuable help in sensors integration , and D.Guinea and P. Bustos for fruitful discussions.

## References

- Bustos, P., Recio, F., Guinea, D., Garcia-Alegre, M.C. 1995. Cortical Representations in Active Vision on a Network of Transputers, In Proceedings of First ECPD Intern. Conference on Advanced Robotics and Intelligent Automation, 259-264.
- Garcia-Alegre, M.C., Bustos, P., Guinea, D. 1995. Complex Behaviour Generation on Autonomous Robots: A Case Study. In Proceedings of IEEE on System, Man, and Cybernetics, 1729-1734. Vancouver, B.C.: IEEE Press.
- Hayes-Roth, B. 1993. Opportunistic Control of Action in Intelligent Agents. *IEEE Trans. on System, Man and Cybernetics*, (23),6, 1575-1587.
- Horswill, Ian, Stein L A. 1994. Life after planning and reaction. AAAI 1994 Fall Symp. on Control of the Physical World by Intelligent Systems.
- Luck, M., d'Inverno M. 1995. A Formal Framework for Agency and Autonomy. In Proceedings First International Conference on Multi-Agent Systems, 254-260, San Francisco, CA.: AAAI Press.
- Maes, P. ed. 1990. *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. Cambridge, Mass.: MIT Press.
- Mataric, M. 1994. Interaction and Intelligent Behavior. Ph.D. diss., Dept. of Elect.Eng. & Computer Science, MIT.
- Nillson, N.J. 1971. *Problem Solving Methods in Artificial Intelligence*. McGraw Hill.
- Panerai, F., Capurro, C., Sandini, G. 1995. Space variant vision for an active camera mount. LIRA-TR 1/95. LIRA-DIST, University of Genoa.
- Pardo, F., Boluda, J.A., Pérez, J.J., Felici S., Dierickx, B., Scheffer, D. 1996. Response properties of a foveated space-variant CMOS image sensor. ISCAS96.
- Tistarelli, M., Sandini G. 1992. Dynamic Aspects in Active Vision. *CVGPI: Image understanding*, (56), 1, 108-129.
- Recio, F. 1995. Tracking Monocular. Informe Técnico. 09/95, Dept. of Systems, IAI/Consejo Superior de Investigaciones Científicas.
- Van den Velde, W. 1995. Cognitive Architectures-From Knowledge Level to Structural Coupling. In: *The Biology and Technology of Intelligent Autonomous Agents*. (Luc Steel ed.). Springer Verlag.

