

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221908247>

Multi-Task Active-Vision in Robotics

Chapter · March 2010

DOI: 10.5772/9308 · Source: InTech

CITATIONS

0

READS

166

4 authors:



Jorge Cabrera

Universidad de Las Palmas de Gran Canaria

77 PUBLICATIONS 385 CITATIONS

[SEE PROFILE](#)



Daniel Hernández-Sosa

Universidad de Las Palmas de Gran Canaria

92 PUBLICATIONS 646 CITATIONS

[SEE PROFILE](#)



Antonio Carlos Domínguez Brito

Universidad de Las Palmas de Gran Canaria

63 PUBLICATIONS 340 CITATIONS

[SEE PROFILE](#)



Eduardo Fernandez

Universidad Miguel Hernández de Elche

416 PUBLICATIONS 5,862 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Estudio del procesamiento cerebral de imágenes con valencia emocional en población pediátrica mediante el análisis espacio-temporal con electroencefalografía.

[View project](#)



ACUSQUAT [View project](#)

Multi-Task Active-Vision in Robotics¹

J. Cabrera, D. Hernandez, A. Dominguez and E. Fernandez
SIANI, ULPGC
Spain

1. Introduction

Vision constitutes the most important sense in the vast majority of animals. Researchers in robotic systems, where biological inspiration has been always a reference, frequently try to make use of vision as a primary sensor. The technological advances have notably favoured this, permitting the incorporation of cameras to a wide variety of robots, including even low cost models. Initially configured as passive devices, soon the same biological emulation led to the introduction of active stereo heads with several mechanical and optical degrees of freedom. However, the use of active vision is far from trivial and has and still is proposing challenging problems.

On the other hand, robots in general, and more specifically mobile robots, are extremely complex systems on their own. The “scientific pull” aiming at obtaining higher and higher levels of autonomy has contributed to create a great number of research lines, many of them remaining still open: obstacle avoidance, localization, navigation, SLAM.



Fig. 1. Museum robot with active binocular head

¹ This work has been partially supported by project PI2007/039 from Canary Islands Government and FEDER funds, and by project TIN2008-060608 from Spanish MICINN.

¿What issues arise when we consider a system that integrates both elements, that is, when a mobile robot is equipped with an active head (see Fig. 1)? Obviously, the answer is too broad and depends on the perspective under which that system is analysed.

Here we will centre the discussion on attention mechanisms in multitasking environments, complemented with some software engineering considerations. Attention is relevant in the proposed scenario as it constitutes a common strategy used to optimize limited computational resources when high sensor data volumes are available. On the other hand, multitasking (see Fig. 2) is also present as it is frequently used in robotics to reduce the inherent complexity of the programming of these systems. Finally, this programming complexity makes highly desirable the use of good software engineering practices to promote reusability and facilitate upgrading and maintenance.

This chapter is organized in two main sections. The first one analyzes the general problems that can be defined for active vision robotics and reviews some works that have been developed in the area. The second one illustrates the most relevant problems stated in this context through the analysis of the solutions adopted in a case study, the MTVS system. The aspects considered include the control of the attention, task coordination and software engineering. Some experimental results obtained with MTVS in a real robot application will also be showed. The chapter ends with the analysis of the main conclusions obtained and the references.

2. Multi-Task Active-Vision

As commented previously, when configured as an active device (Clark & Ferrier, 1992; Bradshaw et al., 1994) cameras constitute a versatile and powerful source of information. However, in a dynamic context with real-time constraints, active vision has important limitations that must be taken into account: mechanical latencies and processing times.

An active binocular head is normally equipped with 6 motors for mechanical degrees of freedom (neck pan/tilt, and pan/tilt for each eye) and 6 more for optical degrees of freedom (zoom, iris and focus for each eye). Taken an image from a lateral position, for example, may result in a relatively high waiting time until the motors reach and stabilize at the final commanded angle, and the frame is captured. After that, the usually high processing time characteristic of visual sensors must be added. As a consequence, in order to preserve system performance, head movements should be economized and the processing should be applied to reduced regions of interest inside the image.

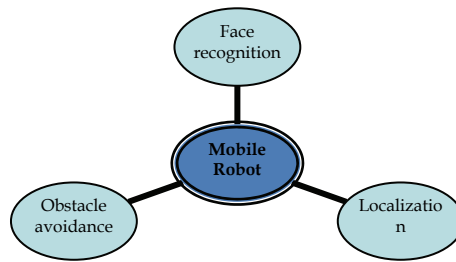


Fig. 2. Mobile robot multitask diagram

Autonomous robots have evolved in the last decade to systems able to develop more and more complex missions. As an example, humanoid robots demos are constantly improving, adding each time higher competence levels to their shows (always below expectations, inevitably).

As anticipated in the introduction, multitasking has been present in robotic systems from the beginning as a “divide and conquer” strategy, in terms of behaviours, actions, goals, etc (Arkin, 1998). In the proposed scenario, this fact automatically transforms an active vision head into a multipurpose shared sensor. Regarding again system performance, the question of which task gets the control of the head arises naturally and must be paid special attention by system developers.

2.1 Problem definitions

Several problems can be defined in this context. On one hand, the control of the gaze in an active vision system is usually formulated as a problem of detection of significant points in the image. The objective is to locate regions of interest on the image where processing can be focused. Under this perspective several aspects such as saliency, bottom-up vs. top-down control, computational modelling, etc, must be analyzed. No motors or, at most, only fast eye motors participate in this process, so the effects of mechanical delays are not so critical and can be neglected, if not ignored.

On the other hand, considering the shared resource nature of the sensor, the problem of management/coordination of the gaze in a multi-task environment should also be taken into account. The visual attention must be distributed among the tasks executing concurrently inside the robotic system. In this case, several motors, including the slower neck pan/tilt, can be operated, so the blind intervals induced by mechanical latencies play an important role and must be carefully handled.

Additionally, and from a more engineering point of view, is the fact that researchers and engineers involved in the development of vision systems are primarily concerned with the visual capabilities of the system in terms of performance, reliability, knowledge integration, etc. However, as important as that, is the problem of modular composition of vision capabilities in perception-action systems. Posing a simple analogy to clarify these issues,

when we execute programs that read from or write to files that are kept in the hard disk, we aren't normally aware of any contention problem and need not to care if any other process is accessing the same disk at the same time. This is managed by the underlying services, simplifying the writing of programs that can be more easily codified as if they have exclusive access to the device.

According to the previous comments, the main objectives that must be pursued by a vision system, especially in the robotic area, can be summarized as follows:

- The system should offer a reduced set of services or visual primitives in pre-categorical terms to the clients.
- The control of the gaze must be assigned on the basis of a simple model of scheduler, allowing the results to be easily interpreted externally.
- The client's tasks should be integrated in the system individually with no coordination involved.
- It must be possible to change dynamically the set of tasks (activity) that the system manages.
- Some level of response time guarantee should be desirable or, at least, the system should offer high priority attention modes.
- Avoid monolithic developments with arbitration code embedded.
- Facilitate the integration of new vision capabilities and promote reusability.

2.2 Related work

As a selection of related work, several studies, systems and applications can be commented. Here we will centre on integrated systems, although many of them rely on basic attention mechanisms (Kundur & Raviv 2000; Itti, 2005).

In a first group, the problem of active vision applied to autonomous or assisted car driving has been analyzed by several authors. At a primary level, there is a long tradition of works that focus on human reactions while performing driving activities (Shinar, 1978; Land & Horwood, 1995; Underwood et al., 2003; Rogers et al., 2005).

Ernst Dickmanns and colleagues (Pellkoer et al., 2001; Dickmanns, 2003) have studied the problem of gaze control in the context of their MarVEye Project, where an active multi-camera head (see Fig. 3) is used to drive a car in a highway. In that project, several areas of interest are promoted and ranked by different modules of the control architecture using a measure of information gain.

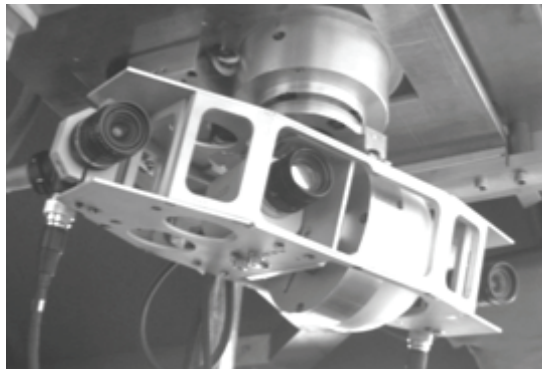


Fig. 3. MarVEye trinocular prototype (from Dickmanns et al.)

Other groups have explored the problem of gaze arbitration in the humanoid robots scenario, both in simulation and with real robots. Javier Seara and colleagues (Seara et al. 2001; Seara et al. 2003) have experimented with a biped robot that used a combination of two tasks to visually avoid obstacles and localize itself (see Fig. 4). The decision of where to look next was solved in two stages. Firstly, each task selects its next preferred focus of attention as that providing the largest reduction of uncertainty in the robot localization, or in the location of obstacles. In a second stage, a multiagent decision schema, along with a winner-selection society model, was used to finally decide which task was granted the control of gaze.

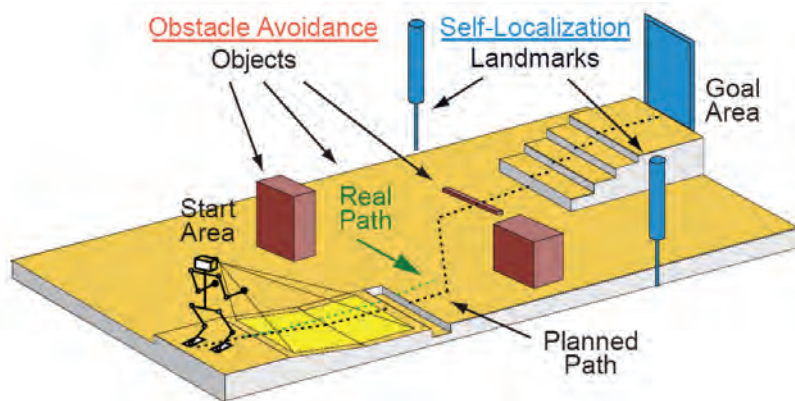


Fig. 4. Biped robot in scenario (from Seara et al.)

Nathan Sprague and colleagues (Sprague et al., 2005; Sprague et al. 2007) have designed a simulation environment where a biped robot must walk a lane while it picks up litter and avoids obstacles (see Fig. 5), using vision as the only sensor. These capabilities are implemented as visual behaviours using a reinforcement learning method for discovering the optimal gaze control policy for each task.



Fig. 5. Walter simulation (from Sprague et al.)

3. MTVS – A case study

Motivated by the previously stated description of the problem we have designed and implemented MTVS (Multi-Tasking Vision System), a proposal of architecture for active-vision systems in multi-tasking environments. MTVS has been designed to deal with the scheduling of concurrent visual tasks in such a way that resource arbitration is hidden to the user.

Similar in spirit to the systems described in the previous section, the aim of MTVS is two-fold: contribute to build a vision system more consistent from an engineering point of view, and to take a first step towards systems where the vision becomes integrated in an action context with higher semantic and cognitive level (an “intelligent” way of looking).

More in detail, MTVS pursues the following objectives:

- The assignment of the gaze control to a task is based on a simple scheduler model, so that the behaviour can be easily interpreted by an external observer.
- The client tasks are integrated in the system individually with no coordination requirements.
- The set of tasks managed (the activity) can change dynamically.
- The clients should not assume any a priori response time guarantee, though the system can offer high-priority attention modes.
- The system offers services based on a reduced set of visual primitives, in pre-categorical terms.

3.1 System architecture

The figure 6 shows an example with two clients and the basic elements making up the vision system architecture: a system server, a task scheduler and the data acquisition subsystem. Basically, the clients connect to the system server to ask for visual services with a given configuration (client A active). In response, the system server launches both a task-thread, to deal with internal scheduling issues, and a devoted server-thread that will be in charge of the interaction with the external client. The scheduler analyzes the tasks demands under a given scheduling policy and selects one to receive the gaze control. In combination, a second covert scheduler checks for compatibility between tasks to share images among them (FOA's overlapping). The data acquisition subsystem processes the different sensor data streams (images, head pose and robot pose) to generate as accurately as possible time stamps and pose labels for the served images.

3.2 Visual services

Clients can connect to the vision system and use it through a number of pre-categorical low-level services. The MTVS services are built around basic visual capabilities or primitives that have also been explored by other authors (Christensen & Granum, 1995):

- WATCH: Capture N images of a 3D point with a given camera configuration.
- SCAN: Take N images while the head is moving along a trajectory.
- SEARCH: Detect a model pre-categorically in a given image area.
- TRACK: Track a model pre-categorically.
- NOTIFY: Inform the client about movement, colour or other changes.

Except for WATCH, the rest of primitives can be executed discontinuously, allowing for the implementation of interruptible visual tasks.

The clients also regulate their activity in the system by means of the messages they interchange with their devoted server. Currently, the following messages have been defined for a task: creation, suspension, reconfiguration (modify parameters, change priority, commute primitive on success) and annihilation.

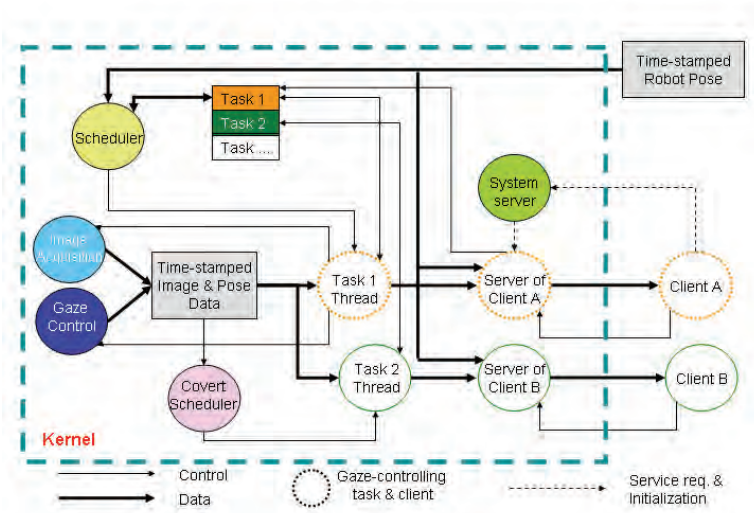


Fig. 6. Control Architecture: example with two clients

3.3 Scheduling policies

Several scheduling policies have been implemented and studied inside MTVS. This analysis has considered two main groups of schedulers: time-based and urgency based schedulers.

Time-based schedulers

Three types of time-based schedulers have been studied: Round-Robin (RR), Earliest Deadline First (EDF) and EDF with priorities (EDFP). The prioritized RR algorithm revealed rapidly as useless in a dynamic and contextual action schema. First, it makes no sense to assign similar time slices to different tasks, and second, the time assigned used for saccadic movements, especially when a slow neck is involved, becomes wasted.

The EDF algorithm yielded a slightly better performance than RR, but was difficult to generalize as visual tasks are not suitable for being modelled as periodic tasks.

The best results of this group were obtained by the EDFP algorithm combining critical tasks (strict deadline) with non-critical tasks. Each time a task is considered for execution and not selected its priority is incremented by a certain quantity (Kushleyeva et al., 2005).

Urgency-based schedulers

The concept of urgency is well correlated with a criterion of loss minimization, as a consequence of the task not receiving the control of the gaze within a time window. This measure can also be put into relation with uncertainty in many visual tasks.

Two schedulers have been studied in this group: lottery (Sprague & Ballard, 2003) and max-urgency. The lottery scheduler is based in a randomized scheme where the probability of a task being selected to obtain the gaze control is directly proportional to its urgency. Every task has the possibility of gaining the control of the gaze, but the random unpredictability can sometimes produce undesirable effects.

The max-urgency scheduler substitutes the weighted voting by a direct selection of the task with higher urgency value. This scheme has produced acceptable results provided that the urgency of a task is reduced significantly after gaining the control of the gaze (similar to an inhibition of return mechanism).

3.4 Experiments

A set of experiments were carried out to analyze the behaviour of MTVS on a real robotic application. The basic experimental setup consists of two ActivMedia Pioneer robots, one with the basic configuration and the other mounting an active monocular vision system (see Fig. 7). The active head is formed by a Directed Perception PTU (as a neck with two degrees of freedom) and a motorized Sony EVI-G21 camera (eye with two mechanical and two optical degrees of freedom).



Fig. 7. Robot and active-vision system

Two main tasks were combined along the different experiments: target following and obstacle avoidance. The target following task commands the active vision robot (pursuer) to detect and follow a special square target mounted on other robot (leader), trying to keep a predefined constant distance between them. The obstacle avoidance task looks for coloured

cylinders on the floor, estimating, as exactly as possible their 2D position. Kalman filtering is used to model both target and obstacles positions.

One-task experiments

As a reference for the maximum expected performance for each task some experiments were designed involving only one task.

Experiment 1: Follow Robot only

In this experiment, the leader robot is commanded to move forward at a constant speed of 200 mm/sec, while the pursuer must try to keep a constant separation of 2 meters. Several tests have been conducted along the main corridor of our lab following a 15 meters straight line path. The pursuer was able to stabilize the reference distance with a maximum error around 150 mm as shown in figure 8.

This experiment determines the base performance level for the follow task.

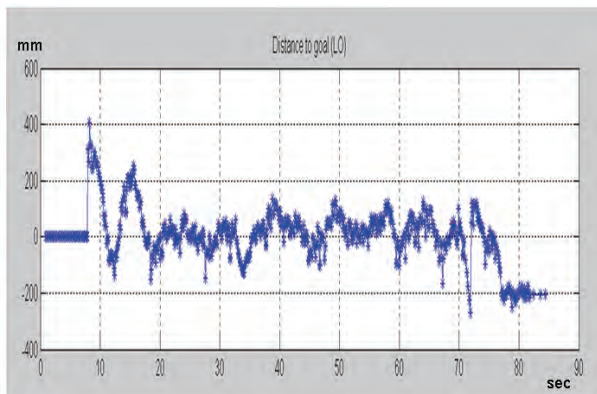


Fig. 8. Robot and active-vision system

Experiment 2: Obstacle avoidance only

The active vision robot is commanded to explore the environment looking for objects (yellow cylinders), trying to reduce their position uncertainty below a predefined threshold. The robot moves straight-line inside a corridor formed by 8 cylinders equally distributed in a zigzag pattern along the path.

The figure 9 illustrates the robot path and the different detections for each localized object, including their first (larger) and minimum uncertainty ellipses. The results show how the robot was able to localize all the objects with minimum uncertainty ellipses ranging from 100 to 200 mm in diameter.

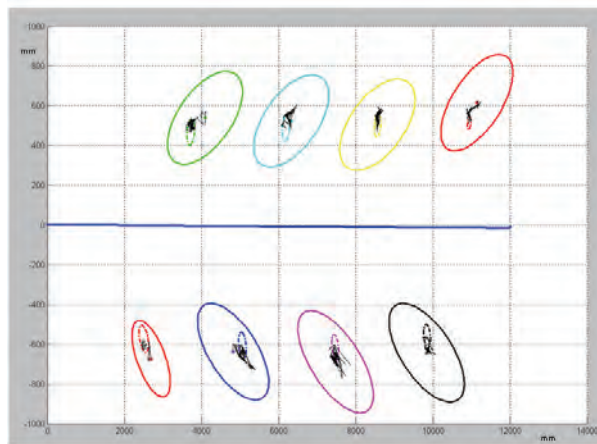


Fig. 9. Obstacle avoidance-Only experiment

This experiment determines the base performance level for the obstacle avoidance task.

Multiple-task experiments

The multiple-task experiments consider a scenario in which each task computes its desired camera configuration and urgency and asks the MTVS scheduler to obtain the gaze control. The scheduler uses this information to select where to look next and how to distribute images. The obstacle avoidance task is extended to classify special configurations of objects as “doors” (two objects aligned perpendicularly to robot initial orientation with a predefined separation).

The urgency of the following task is computed as a function of the distance error, the robot velocity and the time. This urgency increases as the distance between the robots differs from the reference, the velocity is high and the elapsed time since the last image was received becomes larger.

The urgency of the obstacle avoidance task is computed separately for three possible focus of attention: front (the urgency increases when the robot moves towards visually unexplored areas), worst estimated object (the urgency increases as the position of a previously detected object is not confirmed with new images), and closest door (the urgency increases with narrow doors).

The first simple multiple-task experiments try to illustrate the sharing images capability of MTVS. In experiment 5 a more complex scenario including doors is analyzed.

Experiment 3: Obstacle avoidance and robot following competing for the gaze (following priority)

In this experiment, the control of the gaze is only granted to the avoidance task when both the leader speed and the distance error are low. Typically, the following task performance is

not affected significantly, but the avoidance task degrades yielding few objects localization with poor precision. As an example, the upper plot of the figure 10 presents the results of a non sharing run where only half the potential objects (all right sided due to the position of the closest obstacle) have been detected with large uncertainty ellipses. As the lower plot of the figure shows, the sharing of images permits a much better behaviour of the obstacle avoidance task.

Experiment 4: Obstacle avoidance and robot following competing for the gaze (obstacle avoidance priority)

In this experiment, the localization task has the priority, and the control of the gaze is only released in the best possible scenario, that is, all the objects have been precisely detected or left behind the robot position.

In the no-sharing context, the target robot goes away with no reaction from the pursuer, as the camera sensor is captured exclusively by the localization task. In the image-sharing mode, some initial frames can also be used by the following task, and the pursuer moves to reduce the gap. As the robot approaches to the first objects' position, the pan angle becomes larger and the images are not valid for the following task. Finally, the target robot also escapes from the pursuer.

Experiment 5: Localize doors and robot following competing for the gaze (narrow and wide doors)

The configuration of objects used for this experiment consists of a set of four "doors": two narrow type (600 mm width) and two wide type (1500 mm width).

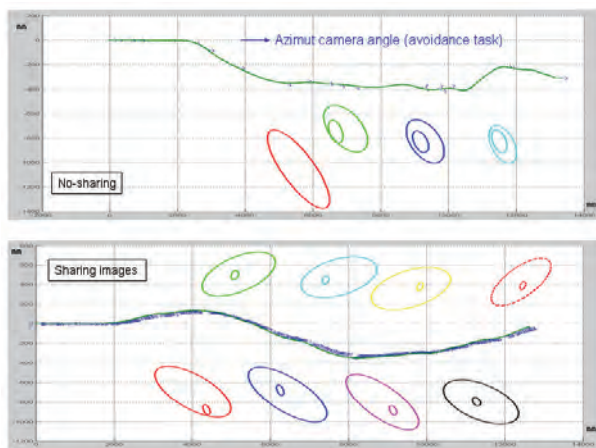


Fig. 10. Follow (priority) and avoidance experiment

All doors are located straight line in front of the robot, the first one (wide) three meters ahead and the rest every 1.5 meters, alternating narrow and wide types. The leader robot is commanded to move at constant speed crossing the doors centred.

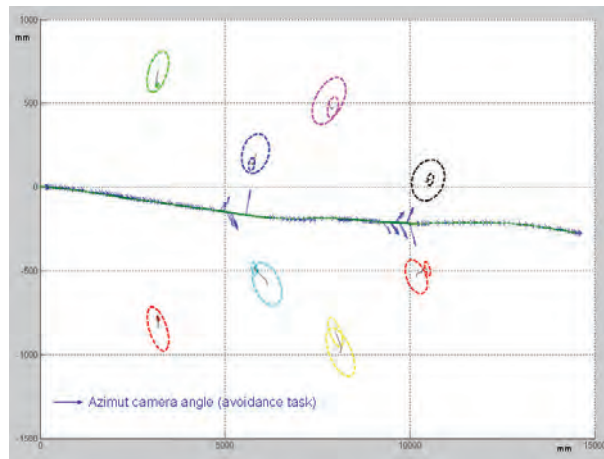


Fig. 11. Narrow and wide doors experiment

The figure 11 illustrates how the camera is pointed to both sides when crossing narrow doors. As a consequence of this behaviour, the pursuer robot slows down when approaching a narrow door until the doorframe position has been estimated with the required precision (compare final error ellipses for narrow and wide doors). After traversing the door, the robot accelerates to recover the desired following distance from the leader.

4. Conclusions

In this chapter we describe the main problems associated with the integration of active vision and multitasking. This configuration, though attractive, must be properly handled by means of simplification strategies that cope with its inherent complexity. Besides, the programming of this kind of complex system is prone to conclude often in monolithic ad-hoc solutions. These problems are illustrated through the analysis of MTVS, a prototype system that proposes an open architecture for the integration of concurrent visual tasks.

In MTVS the client's requests are articulated on the basis of a reduced set of services or visual primitives. All the low level control/coordination aspects are hidden to the clients simplifying the programming and allowing for an open and dynamic composition of visual activity from much simpler visual capabilities.

Regarding the gaze control assignation problem, several schedulers have been implemented. The best results are obtained by a contextual scheme governed by urgencies, taking the interaction of the agent with its environment as organization principle instead of temporal frequencies. Usually, a correspondence between urgency and uncertainty about a relevant task element can be established.

The internal structure of MTVS, its organization in terms of visual primitives and its separated scheduling mechanisms contribute to obtain modular software applications that facilitate maintenance and promote software reuse.

5. References

- Arkin, R. (1998). *Behavior-Based Robotics*, MIT Press
- Bradshaw, K.; McLauchlan, P.; Reid, I. & Murray, D. (1994). Saccade and Pursuit on an Active Head/Eye Platform in *Image and Vision Computing*
- Christensen, H. & Granum, E. (1995). Control of perception in *Vision as process*, Springer-Verlag
- Clark, J. & Ferrier, N. (1992). Attentive visual servoing in *Active Vision* MIT Press
- Dickmanns, E. (2003). An Advanced Vision System for Ground Vehicles, *Proceedings of 1st Workshop on In-Vehicle Cognitive Computer Vision Systems (IVC2VS)*, Graz, Austria
- Itti, L. (2005). Models of bottom-up attention and saliency in *Neurobiology of Attention*, Elsevier Academic Press
- Kundur, S. & Raviv D. (2000). Active vision-based control schemes for autonomous navigation task in *Pattern Recognition*, Elsevier Academic Press
- Kushleyeva, Y.; Salvucci, D.D. & Lee, F.J. (2005). Deciding when to switch tasks in time-critical multitasking in *Cognitive Systems Research*
- Land, M. & Horwood, J. (1995). Which parts of the road guide steering? in *Nature*
- Pellkoer, M.; Ltzeler, M. & Dickmanns, E. (2001). Interaction of perception and gaze control in autonomous vehicles, *Proceedings of SPIE: Intelligent Robots and Computer Vision XX: Algorithms, Techniques and Active Vision*, Newton, USA
- Rogers, S.; Kadar, E. & Costall, A. (2005). Drivers' Gaze Patterns in Braking From Three Different Approaches to a Crash Barrier in *Ecological Psychology*, Lawrence Erlbaum Associates
- Seara, J.; Lorch, O. & Schmidt, G. (2001). Gaze Control for Goal-Oriented Humanoid Walking, *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 187-195, Tokio, Japan
- Seara, J.; Strobl, & Schmidt, G. (2002). Information Management for Gaze Control in Vision Guided Biped Walking, *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 187-195, Tokio, Japan
- Seara, J., Strobl, K.H., Martin, E. & Schmidt, G. (2003). Task-oriented and Situation-dependent Gaze Control for Vision Guided Autonomous Walking, *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Munich and Karlsruhe, Germany
- Shinar, D. (1978). *Psychology on the road*, New York: Wiley
- Sprague, N. & Ballard, D. (2003). Eye movements for reward maximization in *Advances in Neural Information Processing Systems*, Vol. 16, MIT-Press
- Sprague, N.; Ballard, D. & Robinson, A. (2005). Modeling attention with embodied visual Behaviours in *ACM Transactions on Applied Perception*
- Sprague, N.; Ballard, D. & Robinson, A. (2007). Modeling Embodied Visual Behaviors in *ACM Transactions on Applied Perception*
- Underwood, G.; Chapman, P.; Brocklehurst, N.; Underwood, J. & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers in *Ergonomics*