Action-Planning and Execution from Multimodal Cues: An Integrated Cognitive Model for Artificial Autonomous Systems

Zenon Mathews, Sergi Bermúdez i Badia and Paul F.M.J. Verschure

Abstract Using multimodal sensors to perceive the environment and subsequently performing intelligent sensor/motor allocation is of crucial interest for building autonomous systems. Such a capability should allow autonomous entities to (re)allocate their resources for solving their most critical tasks depending on their current state, sensory input and knowledge about the world. Architectures of artificial real-world systems with internal representation of the world and such dynamic motor allocation capabilities are invaluable for systems with limited resources. Based upon recent advances in attention research and psychophysiology we propose a general purpose selective attention mechanism that supports the construction of a world model and subsequent intelligent motor control. We implement and test this architecture including its selective attention mechanism, to build a probabilistic world model. The constructed world-model is used to select actions by means of a Bayesian inference method. Our method is tested in a multi-robot task, both in simulation and in the real world, including a coordination mission involving aerial and ground vehicles.

Zenon Mathews
SPECS, Institut Universitari de l'Audiovisual, Universitat Pompeu Fabra, Barcelona, Spain
e-mail: zenon.mathews@upf.edu

Sergi Bermúdez i Badia
SPECS, Institut Universitari de l'Audiovisual, Universitat Pompeu Fabra, Barcelona, Spain
e-mail: sergi.bermudez@upf.edu

Paul F.M.J. Verschure
SPECS, Institut Universitari de l'Audiovisual, Universitat Pompeu Fabra, Barcelona, Spain
Institució Catalana de Recerca i Estudis Avançats (ICREA) Barcelona, Spain
e-mail: paul.verschure@upf.edu

1 Introduction

The rapid task-dependent processing of sensory information is among the many phenomenal capabilities of biological nervous systems. Biomimetic robotics aims at capturing these kinds of capabilities of biological systems to construct more advanced artificial systems. Indeed, in recent years, the control design of artificial autonomous systems have seen a shift from mere symbolic artificial intelligence (sense-plan-act) to newer concepts like embodiment, situatedness and contextual intelligence [27]. In this context, visual attention and selective attention for task-related performance have been relocated as a functional basis for behavior-based control [27, 6]. Selective attention systems usually rely on the notion of an attentional window or "spotlight of attention", which is defined as a subset of the sensory data perceived by the information processing system [21]. This "spotlight of attention" is forwarding a selected subset of sensory data to higher-order processes that plan and trigger the responses of the system. The constraints imposed by limited resources make such solutions to information bottlenecks of great interest for artificial autonomous systems. Common autonomous system tasks in robotics such as collission-avoidance, navigation and object manipulation all give a prime role of machine attention to find points of interest. At the same time, streams of multi-modal sensory data provide new challenges for systems of selective attention. E.g. currently available humanoid robots such as the iCub robot have more than five sensor modalities and more than fifty degrees of movement freedom [3]. At the same time the design of novel autonomous information processing systems has seen an increasing interest in mimicking the mechanisms of selective attention observed in biological systems [20, 31, 17]. Such systems are sometimes designed to learn by acting in the real-world, under utilization of attentional strategies [20, 27]. However, autonomous system selective attention systems are still in their early stages [14, 27]. Moreover, the interplay among the different components of such complex systems are yet to be formalized.

In this context, we propose a general-purpose architecture for autonomous systems with multimodal sensors, employing biologically inspired mechanisms enabling it to alternate between volitional, top-down and reflex driven bottom up actions to maintain coherence of action. Our model is based on an extension of the Distributed Adaptive Control (DAC) architecture proposed earlier for controlling behavioral systems both in simulations and in the real-world [29, 30]. In our experiments we explore how, for a given complex task, a world-model can guide top-down attention in order to perform actions that are either computed using Bayesian inference or a stochastic path planning method. Our model is based on a so called push-pull mechanism for selective attention. This model is based on psychophysical studies that have proposed such a role for the the extrastriate cortex in the processing of complex visual stimuli [22]. The push-pull mechanism is integrated in our model to allow for optimal data-flow between the different subsystems supporting load balancing. We use this model to develop attentional modulation of sensory data inside the DAC framework. In the next sections we define the mathematical foundation for data association, world model building and decision making and introduce

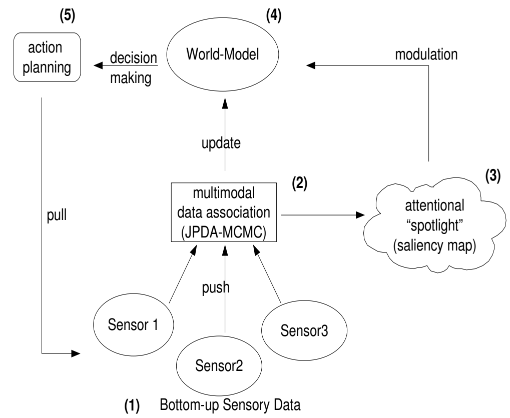

Fig. 1 Model Architecture and the Push-Pull Data Flow: 1) Multimodal sensory data are forwarded/*pushed* by the sensors in real-time. 2) Multimodal stimuli are associated to already existing targets or new targets are created. 3) The attentional mechanism modulates the relevance of target representations in the world model depending on the current task. 4) A probabilistic representation of the relevance of the targets is maintained in the world-model. World-model based decision making generates motor actions using a concrete action generation mechanism.

our neural implementation for selective attention. Finally we discuss a real-world multi-robot task and a robot swarm simulation task to validate the systems performance.

2 Methods

2.1 Model Architecture

Our model is capable of filtering the currently relevant information for a given task from the multimodal sensory input and then select an optimal action in a Bayesian fashion, thereby updating its existing world model. The bottom-up multimodal sensory data is continuously *pushed* by the individual sensors to the data association mechanism, which associates the multimodal stimuli to already existing targets or creates new targets. The result of this is forwarded to the world-model but also to the saliency computation module. In parallel a goal-oriented attentional "'spotlight" generation modulates the relevance of target representations in the world model so that depending on the current task the representation of relevant targets is enhanced. In the world model the relevance of the individual targets is represented probabilistically by means of Gaussian distributions. The decision making module operates on this world model and selects motor actions and attention signals that bias further sensory processing, i.e. top-down attentional control, which are then sent to a planning process.

2.2 *Managing Bottom-Up Multimodal Sensory Data*

The question we address in this section is how an autonomous entity manages the amount of multimodal sensory information it receives continuously in particular the, so called, data association or data alignment problem. In the following, *stimulus* refers to a single modal observation (or data unit) and *target* means a well-defined physical object that also exists in the same space as the autonomous entity. Targets are perceived by the autonomous entity through the multimodal stimuli they evoke.

2.2.1 Joint Probabilistic Data Association

The Joint Probabilistic Data Association method (JPDA) has been successfully used for solving data association problems in a various fields such as computer vision, surveillance, mobile robots, etc. JPDA is a single-scan approximation to the optimal Bayesian filter, which associates observations to known targets sequentially. JPDA thereby enumerates all possible associations between observations and targets at each time step and computes the association probabilities β_{jk}, which is the probability that the j-th observation was caused by the k-th target. Once such association probabilities can be computed, the target state can be estimated by Kalman filtering [9]. Such a conditional expectation of the state is weighed by the association probability. In the following, let x_t^k indicate the state of target k at time step t, ω_{jk} the association event where the observation j is associated to target k and $Y_{1:t}$ stays for all the observations from time step 1 to time step t. Then the state of the target can be estimated as

$$E(x_t^k|Y_{1:t}) = \sum_{\omega} E(x_t^k|\omega, Y_{1:t})P(\omega|Y_{1:t}) \tag{1}$$

$$= \sum_{j} E(x_t^k|\omega_{jk}, Y_{1:t})P(\omega_{jk}|Y_{1:t}) \tag{2}$$

where ω_{jk} denotes the association event where observation j is associated to target k and ω_{0k} denotes the event that no observation is associated to target k. Therefore the event association probability is

$$\beta_{jk} = P(\omega_{jk}|Y_{1:t}) \tag{3}$$

JPDA uses the notion of a validation gate and only observations inside the validation gate for each target are considered. A validation gate is computed for each target using the kalman innovation of new observations. For further mathematical details of JPDA see [9].

β_{jk} can be computed by summing over the posterior probabilities and the exact calculation is NP-hard, which is the major drawback of JPDA [15]. This is due to the fact that the number of association events rise exponentially in relation to the number of observations. We therefore implemented a Markov Chain Monte Carlo method to compute β_{jk} in polynomial time [23] similar to the proposal by Oh and Sastry in [26].

The Markov Chain Monte Carlo (MCMC) method is used in our system to estimate the association event probabilities β_{jk} in polynomial time and with good stability. We only consider feasible mappings from data to target, i.e. the ones that respect the validation gate criteria for the JPDA. The algorithm starts with one such feasible mapping and a Markov chain is generated. MCMC is used for computing β_{jk} in real-time as its time complexity is polynomial with respect to the number of targets. For details of the MCMC approximation of β_{jk}, its convergence and stability see Oh and Sastry [26]. The stimuli have to be associated to existing targets, or if the stimuli is spatiotemporally distant from existing targets, i.e. outside all validation gates, new targets have to be created.

2.3 Goal Oriented Selective Attention

The sensor processing performed by many biological nervous systems is far superior to what can be achieved by human engineered systems. For instance, recent research has shown that the retina transmits between one and ten million bits of information per second, which is about the same rate as an Ethernet connection could support, to the brain [4]. Here we explore how attentional selection can add functional advantages to behavioral systems that deal with large amounts of sensory data. In particular, we consider here goal-oriented top-down selective attention as an information bottleneck that filters the most relevant sensory data, depending on the current task of the system [5, 22]. Such an information bottleneck, that changes dynamically with the system's task, is critical for the survival of biological organisms as the incoming sensory data clearly overwhelms the available limited computational resources. Psychophysiological research suggests that selective attention is load-dependent, i.e. how many unattended stimuli are processed depends on the degree to which attentional resources are engaged by an attended stimulus [**?**]. This delivers evidence for a load-dependent push-pull protocol of selective attention operating at intermediate processing stages of the sensory-data. Such a push-pull protocol has behavioral effects for an autonomous system: when the attentional load is low the system can allocate motor and computational resources for unattended targets. Our architecture makes use of this load-dependent push-pull mechanism allowing the acting system to switch between volitional, reflexive and explorative behaviors.

For the implementation of the selective attention mechanism we use the IQR simulation system [10]. IQR allows implementing large neural networks for real-time applications and interfacing them to real world devices [8, 28]. As suggested by Itti and Koch [21], we implemented a set of neuronal feature filters with excitatory, inhibitory and time-delayed connections between them for the computation of salient points. The feature filters are modulated by the current state of the system, e.g. if the system is running out of power, the feature filters for the charger have stronger excitatory influence on the salience computation. This computation delivers goal-dependend salient target locations which are a subset of the total number of targets the data association mechanism has computed before (see figure 2). In the experiment section examples of such feature filters are discussed in more detail.

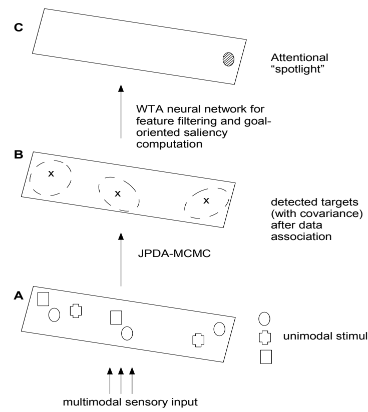

Fig. 2 From multimodal sensory input to attentional saliency: The multimodal sensory input **(A)** to the JPDA-MCMC algorithm is associated to targets in the world-model**(B)**. The goal-dependend saliency computation filters output the most salient targets depending on the current state of the system and the task at hand **(C)** .

2.4 World Model and Decision Making

The world model of an attention-guided behaving system should ideally contain representations of the targets it attends to, but also the unattended targets. Such a world model, or dynamic memory, allows the system to plan its actions depending on the top-down attentional load and the bottom-up sensory input. In this section we discuss the building, maintenance and the use of such a world model for decision making.

The data association and attentional mechanisms deliver a constant input to the world model. Our world model contains the spatial and temporal information of a total set of targets with the attended ones being represented more saliently than the unattended ones. We define Θ_s^t as the relevance of a certain target s at time t, and we are interested in the following conditional probability:

$$P(\Theta_s^t | F_s^t(\Theta_s^{t-1}) A_t(s)) \tag{4}$$

where $F_s^t(\Theta_s^{t-1})$ and $A_t(s)$ are two time-dependend functions which weigh the target s. For example $F_s^t(\Theta_s^{t-1})$ evaluates the spatial proximity of the target if there is at least one onset stimulus associated to this target and decays the current weight of the target otherwise. Whereas $A_t(s)$ evaluates the goal-dependend attentional saliency of this target. By computing the joint distribution of these relevance probabilities for all targets s the system can perform the motor action appropriate for the most relevant targets. The following subsection elaborates the update of these relevance probabilities.

Let us assume that we can compute relevance probabilities of individual targets as shown above in eq. 4. Given these individual target relevances we are interested in the fused relevance distribution:

$$P(\Theta^t|F^t(\Theta^{t-1})A_t) \tag{5}$$

We express this probability as the normalized sum of probabilities of individual relevances:

$$P(\Theta^t|F^t(\Theta^{t-1})A_t) = \sum_s P(s)P(\Theta_s^t|F_s^t(\Theta_s^{t-1})A_t(s)S) \tag{6}$$

where random variable $S \in 1...n$, n being the number of targets and $P(s)$ indicates the probability of this target. As $P(s)$ is uniformly distributed over all targets this allows for normalization.

Given the relevance distribution a decision that is optimal in the Bayesian sense is computed. We are therefore interested in the following probability distribution of action:

$$P(Action|F_s^t(\Theta_s^{t-1})A^t(s)) \tag{7}$$

where F implements a time-dependend decay function for utility. This probability distribution can be computed using the Bayesian rule, given apriori information about the environment the autonomous system is acting in.

2.5 Action Planning and Execution

2.5.1 Bayesian Optimal Action

The world model of an attention-guided behaving system should ideally consist of the *items* it perceives at the moment, but also possibly *items* perceived in the past [18, 13, 12]. We formulate an optimal Bayesian decision making method for generating actions based on a transient/dynamic memory that allows the system to plan its actions depending on current and past stimuli.

Multimodal stimuli from different sensors of the autonomous system is associated using the JPDA method discussed above. This method creates *items* in a memory of which the utility probability is computed. In the following we derive the equations from general equations 4, 5 and 6.

Let us assume that the motor action consists simply of choosing a direction of motion $\gamma \in 0..360$ and a travel distance $\psi \in 1..10$. The best action is then chosen in the direction γ of the most relevant *item* at distance ψ in the world-model. As in the general equation 7, here are interested in computing the most relevant direction of motion γ and distance ψ. Therefore we are interested in the probability:

$$P(\gamma\psi|F_s^t(\Theta_s^{t-1})A^t(s)) \tag{8}$$

As F is a function of distance d and time t we can express F as the known distance d to the *item*, the time t since the time of a previous stimulus associated to this *item*, the relative orientation γ_i and the attentional weights a_i for each *item*. For n *items* we formalize the above probability as:

$$P(\gamma\psi|d_1,...d_n t_1,...t_n \gamma_1,...\gamma_n a_1,..,a_n) \tag{9}$$

We consider the conditional probability 9 as if there were only one *item* i and without attentional inputs a_i. Supposing conditional independence for angle and distance domains we do the following decomposition:

$$P(\gamma\psi|d_i t_i \gamma_i) = P(\gamma|d_i t_i \gamma_i)P(\psi|d_i t_i \gamma_i)P(t_i d_i \gamma_i)$$

We formulate the probability distributions $P(\gamma|d_i t_i \gamma_i)$ and $P(\psi|d_i t_i \gamma_i)$ as Gaussian distributions:

$$P(\gamma|d_i t_i \gamma_i) = \mathcal{N}(\gamma_i, \frac{d_i t_i}{c_1}) \tag{10}$$

where the Gaussian is centered on the angle γ_i at which the *item i* is located. The standard deviation is a function of time t_i at which this *item* was last perceived and the distance d_i at which this *item* is. This allows gradualforgetting (time decay) what has been perceived in the past, as past information is always prone to changes in a dynamic world.

Similarly for the distance domain, the Gaussian is centered on the distance d_i of the *item* and the standard deviation is again a function of time t_i, allowing a time decay.

$$P(\psi|d_i t_i \gamma_i) = \mathcal{N}(c_2 d_i, c_3 t_i d_i) \tag{11}$$

And we assume the uniform distribution for the joint probability $P(t_i d_i \gamma_i)$ as we do not have any prior information about possible correlations between those random variables.

$$P(t_i d_i \gamma_i) = \mathcal{U} \tag{12}$$

where c_1, c_2 and c_3 are constants.

We now take the utilities of all the *items* into account for the computation of the total utility as shown in equation 6. We include the attentional components a_i and consider the following conditional probability distribution:

$$P(\gamma\psi|d_1,...d_n t_1,...t_n \gamma_1,...\gamma_n a_1,..,a_n) = \tag{13}$$

$$\sum_i \frac{a_i}{a_{tot}} P(\gamma|d_i t_i \gamma_i) P(\psi|d_i t_i \gamma_i) P(t_i d_i \gamma_i) \tag{14}$$

where a_{tot} is the sum of all attentional components a_i, which are the attentional saliencies for individual *items* depending on their detected remaining charge.

According to this formulation the attentional components a_i weigh the shares of the individual *item*s to the joint conditional probability distribution, i.e. attention modulates the world model, which is expressed as a probability distribution that changes in each step with the sensory input.

2.5.2 Stochastic Action Execution

In order to navigate through an arbitrary environment, an autonomous robot should make use of the knowledge about the environment. We evaluate our model in a benchmark task that contains dangerous objects such as mines and obstacles, where the autonomous robot should navigate from a given point A to point B avoiding the obstacles. In our approach, we decided to employ a metric or grid-based approach [19, 25]. In this framework, each cell in a grid represents a part of the field. Cells in an occupancy grid contain information about the presence of an obstacle. The cells are then updated relying on the information received by sensors. The value associated to a cell represents the degree of belief of the presence of an obstacle. However, in our case we had to extend this environment representation and adapt it to our needs. In fact, as previously explained, we have to deal with some issues. In particular, our environment is characterized by:

1. A high degree of uncertainty regarding the positions of both the robot and of the obstacles due to the position information resolution
2. Obstacles can be added on the fly provided updated information
3. We need to decide the cruise speed of the robots within each cell according to the probability of finding obstacles

Hence, instead of associating to each cell the likelihood of a cell being an obstacle, we associate to it the probability of colliding with an obstacle/mine in the part of the field represented by the cell. We subsequently employ such a probability to control the speed of the autonomous agent.

We are interested in reaching a known goal position while minimizing the probability of incurring in obstacles and maximizing exploration to improve the knowledge about our environment. Nevertheless, we have to consider two important factors inherent to our problem. The former is that that the available time to complete a task is limited, and the latter is that it may not exist a path not containing dangerous zones. Therefore, the two objectives of shortening the path to the solution, and minimizing the probability of incurring in mines could be in contrast. Hence, the output of our planning algorithm should be a sufficiently short path that reduces as much as possible the probability of entering into dangerous zones.

In order to implement the path-finder we employed a variation of a stochastic hill-climbing (or stochastic gradient decent) [11] algorithm boosted via a taboo search technique [16]. These algorithms are guided by a heuristic that combines the danger probability with the distance to the goal. In particular, the value of the heuristic function at each point of the grid is a weighted sum of the probability of encountering mines with the distance to the goal. This allows combining the fact that we aim at avoiding dangerous zones while not increasing too much the length of the path.

The Stochastic Hill-Climbing (SHC) algorithm works as follows. Departing from an initial state (a cell in our grid) a set of possible successors states are generated. Each of such successors has associated a value of the heuristic function estimating how good such a state is (how close to the goal and how dangerous it is).

However, the HC algorithm suffers from the limitation that local minima are very likely to produce an infinite loop. In order to overcome such a limitation SHC has been introduced. In SHC the successor that is chosen is not always the one minimizing a given heuristic function, but rather there is a given probability that the second-best, third-best, or one of the others is selected. This algorithm continues until the goal is reached. In this way, most of the local minima can be escaped and a random exploratory behavior is introduced in the path planning. Nevertheless, broader local minima are still a big problem. This may happen when we move on states that are cyclically connected. The mechanism underlying taboo search keeps track of some of the states that have already been visited and marks them as taboo, that is forbidden.

2.6 Test Scenarios

2.6.1 A Combined Micro-Aerial (MAV), Unmanned Ground Vehicle (UGV) and Human Rescue Mission

In this testbed we test

> multirobot coordination, the construction of a world-model from local perception and its online update using a distributed human-robot interactive system to complete a real-world task.

In this benchmark we use a heterogeneous group of UGVs and MAVs in which multiple robots are used to create a shared knowledge base about the world they are interacting in. The mission plan will consist of a sampling phase in which robots equipped with camera systems (UGV and MAV) will be driven to sample the environment and gather information. Our model has previously been employed for a real-world outdoor mission using MAV and UGVs [7]. The cognitive system will generate the plan instructions to be executed autonomously by the UGVs that can also be monitored and manually updated using the world model interface and 3D representation. The system allows for an eventual manual control of all robots.The MAV is a commercial quadcopter made by Ascending technologies, Germany (Fig

1). It consists of a cross like structure with four independent motor controllers that run 4 propellers that provide the necessary lift force during approx. 30 minutes continuous operation. The total weight of the MAV is about 600 grams with the battery pack. The quadcopter has been equipped with an additional wireless camera system for remote pilotage and inspection. With the additional 150 grams of the camera system, the autonomy is reduced to about 10-15 minutes. The range of the wireless video link is approx. 800 meters. However, a UGV has been designed as a mobile repeater station for all video signals providing additional 800 meters. On the base, a pilot, using a Head Mounted Display system, controls the robot from the image provided from its camera. The MAV can be remotely turned off during flight operation, turning it ballistic on demand.

Two custom made tracked robots (50 x 30 x 20) and a standard RC wheeled vehicle (50 x 40 x 30), equipped with wireless camera systems constitute the Unmanned Ground Vehicle (UGV) support team. The tracked robots are equipped with standard GPS, compass, Metal-thin oxide chemo-sensors (supplied by Alpha MOS

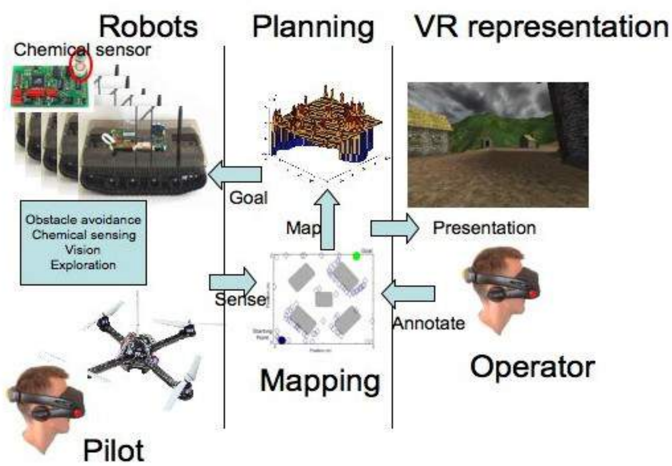

Fig. 3 Grand scheme of the integrated autonomous dynamic mapping and planning approach. In this task, robots are used as autonomous dynamic sensing platforms that contribute information to a central mapping and planning stage (cognitive system). The planning stage defines goal positions the robots should attempt to reach using their local proximal sensory-motor capabilities, e.g. collision avoidance, mine detection, etc. The aerial vehicle is guided by a human pilot, and the information gathered by this method is added to the world model. The state of the world model is transformed into a 3D representation of the task area. To this representation objects and terrain features are added in real-time. The human operator inspects the 3D model and makes decisions on future actions while making his own annotation in the language of the virtual world. See text for further explanation.

SA, France) and ultrasonic sensors to provide way-point based navigation, and to generate a world model and planning (cognitive layer) (fig.2).

The communication among all robots goes through the automatically generated world map. This map lives in a base station, which is used to communicate via a radio link and instruct the different robots. Additionally, the world model has a user interface that allows the operators to contribute to by adding supplementary information such as about obstacles and mines that will be taken into account by the cognitive system while generating the path planning.

The whole mission and world model status is represented online on a 3D model of the mission that allows the operators to freely navigate through the whole field and have a remote visualization of the scenario from the base station.

2.6.2 Simulation of a Robotic Swarm for Rescue Missions

The objective of this benchmark is to test

> world-model construction and the exploitation of the world-model for robot coordination and action generation that is optimal in a Bayesian sense.

With benchmark 1 we tested the capability of our model for multirobot coordination, creating and maintaining a world model, whereas here we specifically test how from local perception a world-model can be constructed, which is subsequently used to generate actions, which are optimal when solving a multiple goal task under limited resources constraints. For this purpose, we consider the following robot swarm scenario. A swarm of robots are on a common mission in a given environment. A rescue robot is equipped with our model, and is involved in the specific

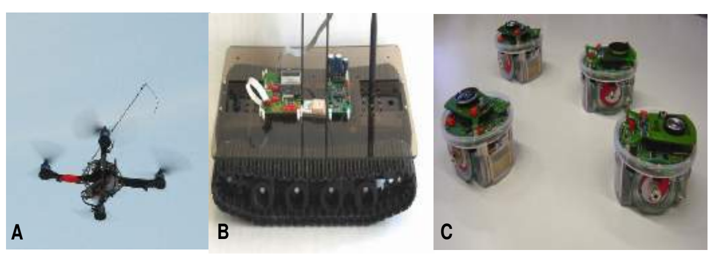

Fig. 4 The MAV and UGV platforms. **(A)** Quadcopter (Ascending Technologies) [1] and **(B)** custom build tracked ground vehicle. A wireless link connects the robotic platforms to the computing unit and/or human operators. We incorporated a camera, a GPS, ultrasonic sensors, a LI-PO battery, a compass and chemo-sensors. See text for further explanation. **(C)** A group of mini UGVs used for indoor testing of the multi-robot mission. We use the EPuck robot which features multimodal sensors and wireless communication [2].

task of aiding expired (i.e. broken-down or out of charge) agents. This means that the rescue robot first has to localize the expired agents using its sensors and approach them for repair or recharge. The rescue robot is equipped with a limited number of distance-measurement sensors like sonar and laser range scanners, with which it has to scan the environment and localize the agents to accomplish the given task. From time to time, also the rescue robot has to go back to the base station to recharge itself. Solving this multiple goal task involves multimodal data association, goal-driven selective attention generation for attending to the most vital subtask at the moment and maintaining a dynamic world model, which is used to compute the optimal action in the Bayesian sense. We implemented this testbed in a simulation with N robots and M sensors for the rescue robot.

items in this testbed are robots involved in the common mission. *Attentional saliency* here is proportional to the detected remaining power of this *item*, giving a high utility for reapproaching nearly expired *items*. The rescue robot computes the overall utility distribution of going in a certain angle and distance from the individual utilities of *items*. We simulate the data from different range sensors and use this as multimodal stimuli to the rescue robot. This creates *items* of which the utility probability has to be computed.

3 Results

3.1 Static World, Multirobot Coordination

In the first benchmark we investigated a static real-world environment that had to be explored using multiple robots. As discussed earlier we have outdoor ground and aerial vehicles for an exploratory mission involving avoiding dangerous mines and obstacles while developing a map. We showed that multi-robot collaborative exploration is achieved using our model. First we test the multirobot mission using the e-Puck robots in an indoor environment which allows scaling and testing of individual components of our model. We setup a 2 x 2 meter arena where a single robot was said to reach a feeder located at the center position from randomly starting points. Thus we measured the mean positioning error resulting from the PID controller after 10 runs, being it about 4 cm. Consequently, some objects where placed in the arena to obstruct the direct path from the starting point to the goal position, figure 5. In this case, both the PID and the obstacle avoidance (the reactive layer of the Distributed Adaptive Control) were necessary to accomplish this task. We employed multiple robots to perform this task. The results show that the multirobot autonomous control permitted the robot to reach its goal in all cases. Nevertheless, the resulting paths were not the optimal ones as evidenced by the run durations (median 170 seconds). Additionally, we observe that the gain of using a number of robots to explore the environment does not report very good results if there is no strategy on how to use the acquired information and how can this be share among robots.

Fig. 5 Multi-robot trials: Traces of multiple robots and runs under the control of the robot autonomous layer. The robots are released at the starting point and their autonomous (PID + reactive) layer allow them to reach their goals in the presence of obstacles.

In order to improve the performance of both exploration and goal oriented behavior, we implemented the previously described autonomous control architecture. Preliminary tests were done by letting a single robot explore the environment. While the robot was performing the task, the world model was created with its sensory information (proximity sensors) and was also improved online with new information. The then generated world model contains in this case the detected contours of elements within the test environment, figure 6. This information is therefore very valuable in order to plan strategies at collective of individual level, which are then executed by the group of robots controlled by the multirobot autonomous control system.

We tested the reliability of the world model with the path planning system in order to generate routes through the test arena. The experiment consisted in this case of a goal oriented exploration of the environment, in which a robot had to reach a goal. Every run the robot was performing the task, the multirobot autonomous model was improving its world model at the same time the planning could generate more reliable paths through the arena, figure 7A. The results show how over runs, the length of the robot trajectory gets shorter and closer to an optimum, figure 7B. In the case of multiple robots, the generation of the world model and therefore the planning strategy would improve even faster since all the robots would collaborate by contributing with their local sensory information to the multirobot autonomous control architecture.

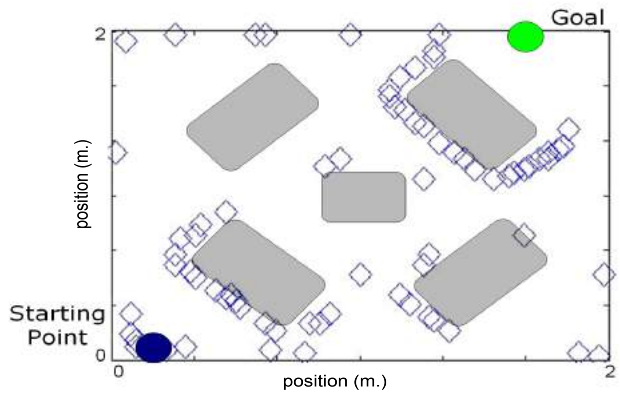

Fig. 6 World model generation from multirobot coordination: Generated world model resulting of the goal oriented and exploratory behavior of a robot after a number of test runs. Goal, robot position and objects are represented.

3.2 Dynamic World, Partial Perception

Subsequently we investigated how the system behaves in a dynamic environment using a single agent. The partial knowledge acquired by the agent is then used to construct the world-model and robot coordination and world-model exploitation in a robot swarm simulation is evaluated. We used 10 simulated agents who move around the experimental field of a 600 by 600 m. window, who start with a maximum speed and maximum energy but slow down as the energy drops. The energy-drop is proportional to the covered distance. Their direction of motion is arbitrary but always inside the arena. The rescue robot, controlled by our model, always starts from the base station and alternates between exploration and exploitational time slots. During exploration it moves about randomly in the field to detect the agents. Thereby the multimodal sensor fusion and attentional saliency computation delivers input to update its world model. During the exploitation time slot, the rescue robot performs the intelligent motor actions as described earlier. We compare the performance of the rescue robot using the world model in the exploitation phase with a system when not using it. When the world model is not used to compute an intelligent action, the rescue robot is in constant exploration. For each category 5 trials each with 5000 time-steps were carried out. A probabilistic world model computed as is shown in figure 8.

To assert the performance of the system, we evaluate the number of recharged agents during each trial and also at the total expiry time of all agents together in each run and observe a significant improvement when using the probabilistic world model and motor action selection. WM indicates the use of world model and non-

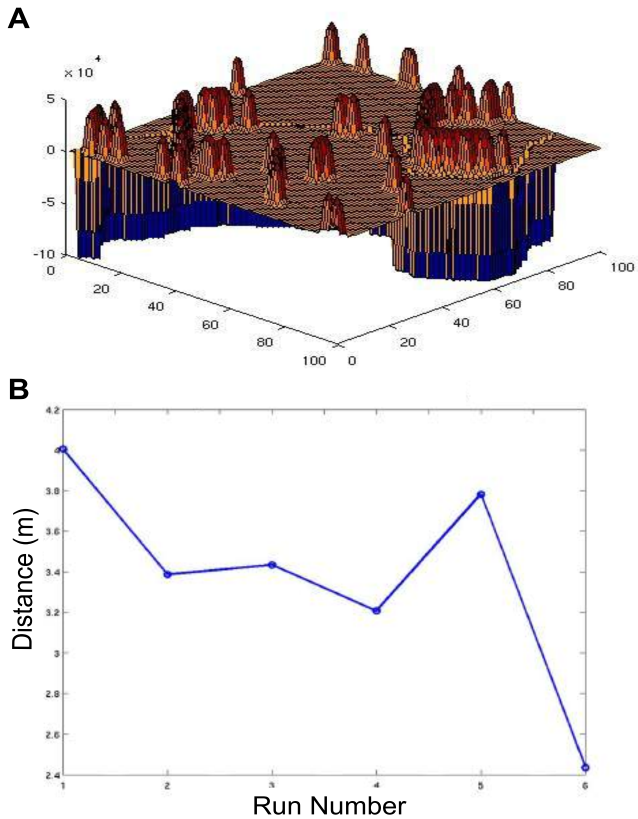

Fig. 7 Exploration and Path Planning. A) Online generated robot trajectory: The positive values indicate the cost related to go through a specific position of the arena, the computed trajectory is represented by negative values. **B) Task performance vs number of runs:** evolution of the performance of the planning with the number of test runs. The decrease in the traveled distances shows that the world model is more complete and accurate, and therefore it results in a more optimal robot path.

WM indicates the use of a reactive system that explores the robot arena without a world-model or attentional mechanisms; see figures 9 and 10.

4 Conclusions

We have proposed an integrated model of multimodal data association, attentional saliency computation, world-model construction and maintenance and action selec-

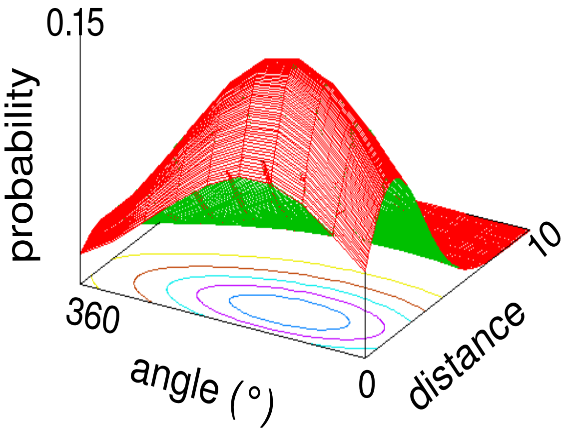

Fig. 8 World model probability manifold example: Angles range from 0..360 and distances range from 1 to 10. Higher salient experiences are represented with higher probabilities. This world model suggests the most probable action as the one that leads to the expired agents, which were perceived to be running slow in the past. This probability distribution is computed at each time step before an intelligent motor action decision is made.

tion for artificial autonomous systems. Our model is based on biological principles of perception, information processing and action selection, and is incorporated in the Distributed Adaptive Control framework for automated control of sensory signals using both top-down and bottom-up attention. Our model suggests how the different subsystems of an artificial autonomous system can interplay seamlessly in an integrated framework. We demonstrated the use of our model in a multirobot coordination task, where a common world-model for the multiple robots is created, maintained and used to compute optimal actions for the individual robots. We have shown how to generate trajectories for individual robots using a multirobot exploration of the environment and how the performance of the system augments with increasing exploration. The first testbed was for a static environments for multirobot collaboration. In the second testbed we evaluated the possibility of computing a global world-model from local perceptions in a robot swarm experiment where we used a dynamic, partially visible environment. Selective attention mechanisms are employed to focus the information processing capacities on the currently most rele-

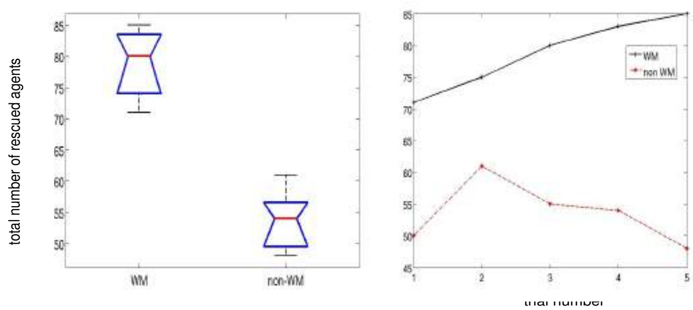

Fig. 9 Number of recharged agents: The use of our autonomous system control with the world-model achieves higher number of recharged agents in all trials when compared to a system that does not possess a world model.

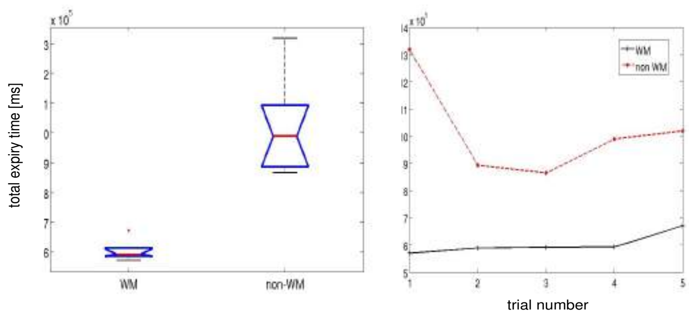

Fig. 10 Total expiry time of agents: The use of our autonomous system control with the world-model achieves much less expiry time of agents in all trials when compared to a system that does not possess a world model.

vant task. We have shown that our model performs significantly better than a system without a world-model in the given rescue mission. The modularity of our architecture allows for customizing the individual components of the model for the given task. In further work we will evaluate the capability of the model for the control of various autonomous systems such as our insect inspired robotic foraging model [24], and for the binocular visual processing in the humanoid robot iCub [3].

Acknowledgements The authors wish to thank Fabio Manzi, Ramón Loureiro, Andrea Giovanucci, Ivan Herreros, Armin Duff, Joan Reixach Sadurní, Riccardo Zucca, and Wendy Mansilla for the joint work on the multirobot coordination project.

References

1. http://www.asctec.de
2. http://www.e-puck.org/
3. http://www.robotcub.org
4. How much the eye tells the brain. Current Biololgy **16**, 1428–1434 (2006)
5. Search goal tunes visual features optimally. Neuron **53**, 605–617 (2007)
6. Arkin, R.: Behavior-based robotics (1998)
7. i Badia, S.B., Manzi, F., Mathews, Z., Mansilla, W., Duff, A., Giovannucci, A., Herreros, I., Loureiro, R., Reixac, J., Zucca, R., Verschure, P.F.: Collective machine cognition: Autonomous dynamic mapping and planning using a hybrid team of aerial and ground based robots. 1st US-Asian Demonstartion and Assessment of Micro-Aerial and Unmanned Ground Vehicle Technology (2008)
8. i Badia, S.B., Pyk, P., Verschure, P.: A biologically based flight control system for a blimp-based UAV. Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on pp. 3053–3059 (2005)
9. Bar-Shalom, Y., Fortmann, T.E.: Tracking and data association. Boston Academic Press (1988)
10. Bernardet, U., Blanchard, M., Verschure, P.: Iqr: A distributed system for real-time real-world neuronal simulation. Neurocomputing pp. 1043–1048 (2002)
11. Blum, C., Roli, A.: Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Comput. Surv. **35**(3), 268–308 (2003)
12. Botvinick, M.M., Plaut, D.C.: Short-term memory for serial order: a recurrent neural network model. Psychological review **113**(2), 201–233 (2006)
13. Byrne, P., Becker, S., Burgess, N.: Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychological review **114**(2), 340–375 (2007)
14. Coelho, J., Piater, J., Grupen, R.: Developing haptic and visual perceptual categories for reaching and grasping with a humanoid robot. Robotics and Autonomous Systems **37**, 195..218 (2001)
15. Collins, J., Uhlmann, J.: Efficient gating in data association with multivariate gaussian distributed states. Aerospace and Electronic Systems, IEEE Transactions on **28**(3), 909–916 (1992)
16. Cvijovicacute, D., Klinowski, J.: Taboo search: An approach to the multiple minima problem. Science **267**(5198), 664–666 (1995)
17. Dickinson, S.J., Christensen, H.I., Tsotsos, J.K., Olofsson, G.: Active object recognition integrating attention and viewpoint control. Computer Vision and Image Understanding **67**, 239..260 (1997)
18. Dominey, P.F., Arbib, M.A.: A cortico-subcortical model for generation of spatially accurate sequential saccades. Cerebral cortex (New York, N.Y.: 1991) **2**(2), 153–175 (1992)
19. Elfes, A.: Using occupancy grids for mobile robot perception and navigation. Computer **22**(6), 46–57 (1989)
20. G. Billock C. Koch, D.P.: Selective attention as an optimal computational strategy. Neurobiology of Attention p. 18..23 (2005)
21. Itti, L., Koch, C.: Feature combination strategies for saliency-based visual attention systems. Journal of Electronic Imaging **10**(1), 161 (2001)
22. M. A. Pinsk, G.D., Kastner, S.: Push-pull mechanism of selective attention in human extrastriate cortex. Journal of Neurophysiology **92** (2004)
23. Mathews, Z., Bermúdez i Badia, S., F. M. J. Verschure, P.: Intelligent motor decision: From selective attention to a bayesian world model. 4th International IEEE Conference on Intelligent Systems **1** (2008)
24. Mathews, Z., Lechón, M., Calvo, J.B., Dhir, A., Duff, A., i Badia, S.B., Verschure, P.F.: Insect-like mapless navigation using contextual learning and chemo-visual sensors. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS09) (2009)
25. Moravec, H.: Sensor fusion in certainty grids for mobile robots. AI Mag. **9**(2), 61–74 (1988)

26. Oh, S., Sastry, S.: A polynomial-time approximation algorithm for joint probabilistic data association. pp. 1283–1288 vol. 2 (2005)
27. Paletta, L., Rome, E., Buxton, H.: Attenion architectures for machnine vision and mobile robots. Neurobiology of Attention p. 642..648 (2005)
28. Pyk, P., i Badia, S.B., Bernardet, U., Knüsel, P., Carlsson, M., Gu, J., Chanie, E., Hansson, B.S., Pearce, T.C., Verschure, P.F.M.J.: An artificial moth: Chemical source localization using a robot based neuronal model of moth optomotor anemotactic search. Autonomous Robots (2006)
29. Verschure, P.F., Voegtlin, T., Douglas, R.J.: Environmentally mediated synergy between perception and behaviour in mobile robots. Nature **425**(6958), 620–624 (2003)
30. Verschure, P.F.M.J., Althaus, P.: A real-world rational agent: unifying old and new ai. Cognitive Science A Multidisciplinary Journal **27**(4), 561–590 (2003)
31. Y. Jiang N. Xiao, L.Z.: Towards an efficient contextual perception for humanoid robot: A selective attention-based approach. 6th World Congress on Intelligent Control and Automation (2006)