Task-Oriented and Situation-Dependent Gaze Control for Vision Guided Humanoid Walking

J. F. Seara[†], K. H. Strobl[††], E. Martín[‡], and G. Schmidt[†]

[†]Institute of Automatic Control Engineering
Technische Universität München
D-80290 Munich, Germany

[††]Institute of Robotics and Mechatronics
DLR – German Aerospace Center
D-82234 Wessling, Germany

[‡]Collège des Ingénieurs
49, Rue de l'Université
75007 Paris, France

Javier.Fernandez.Seara@ei.tum.de
http://www.lsr.ei.tum.de

Abstract. –This article presents various aspects of a gaze control scheme for visually guided humanoid robot navigation. The strategy is based on the maximization of the predicted visual information. For the *information management* a *coupled hybrid Extended Kalman Filter* is employed. Specific view direction control strategies for two concurrent objectives of different nature, obstacle avoidance and self-localization, have to be weighted and pursued in parallel. A combination of both shows the task dependence of the gazing strategy. The main goal of this work is to formalize and implement a decision strategy in order to achieve an intelligent task-oriented active vision system for a biped walking robot. It intends to explain the active vision decision making problem: *Where to look next?*, of an agent facing multiple and concurrent goals. The general approach rests upon the definition of a set of *Utility Functions* over the outcomes of the set of possible view directions. Next the various utility functions, i.e. *Agents*, representing different kinds of preference rankings over the predicted outcomes, are organized to solve the *Action/Selection* problem as a *Society of Minds*. Simulation results based on experimental experience demonstrate the applicability of the approach.

1 Introduction and Outline of the General Approach

Navigation is the science of directing the course of a mobile robot as it traverses the environment. Inherent in any navigation scheme is the desire to reach a destination without getting lost or crashing into anything [1]. McKerrow includes in his definition the two main tasks related to the problem of navigation: self-localization and obstacle avoidance, as indicated in Fig. 1.

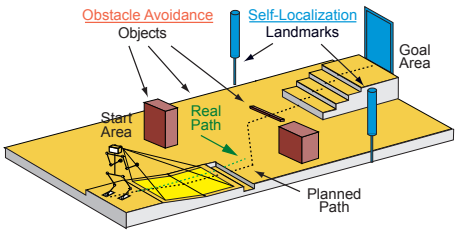

Fig. 1. Biped robot in scenario.

If a robot navigates by use of dead-reckoning information only, the uncertainty in position grows monotonically, and at a specific point, the locomotion goal can no longer be fulfilled. Thus, visual information or information from other external sensors is used to reduce the errors for safe and efficient navigation [2,3]. Perception guided robot navigation has been a relevant research topic in robotics for decades. The navigation strategies reported were primarily designed for wheeled robots due to early state of development of legged robots to work with. For humanoids, the problems of mechanical design and stabilization seem to be solved just recently [4]. The question of perception based biped locomotion has over the last years won interest worldwide with several groups working in parallel [5,6]. Obviously, the next key issue yielding abilities of autonomous locomotion behaviors is the creation of new navigation strategies, or the adaptation of the existing ones to the very peculiar characteristics of legged robots, where a high degree of coordination between vision and walking is required. The understanding of this *merging* of perception and locomotion is considered a primordial aspect when studying and expanding the *cognitive functions* of humanoid robots.

For a vision guided walking machine to execute a given locomotion task, the guidance system must provide appropriate step sequences to allow the robot to navigate and reach the specified goal position in spite of obstacles in the walking trail, see Fig. 1. Information about these obstacles is provided by an active vision system [7]. In order to increase the degree of autonomy of the robot, some intelligent means of controlling perception is required in order to guarantee that the humanoid is provided with the highest possible amount of relevant task-dependent information. This information acquisition and management can be accomplished by a gaze control scheme selecting the orientation of the sensor, so that the relevant objects of the scene are kept within the sensor's limited field of view. In this article a modular task-oriented and situation-dependent gaze control architecture is proposed, see Fig. 2. It comprises three major modules: (I) *Information Management*, (II) *Task-specific Gaze Strategy*, and (III) *Decision Scheme*.

In our research we consider an *information management* scheme, Fig. 2-(I), which registrates and administers the uncertainties associated with the different processes involved in perception based locomotion [8,9]. The *information/uncertainty management* employs a *coupled hybrid EKF*. Based on this scheme, a description of the accumulated information combined with the uncertainties in the measurements provide a basis for the task-oriented gathering of the relevant information, i.e. *task-specific gaze strategy*, Fig. 2-(II). Our approach to this information quantification is based on Shannon‘s Information Theory [10]. A well-known definition of information is: *Information is a measure of the decrement of uncertainty.* Since information must be redefined in the context of biped walking, we define the *Information Content* of a view situation as a measure of the degree to which perception under these terms is capable of reducing *Incertitude* ν. In order to make optimal use the perceptual resources, the observer must act in a task-oriented way and concentrate its abilities on the

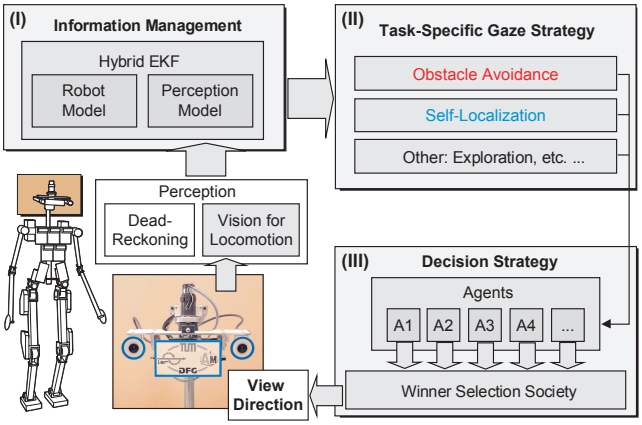

Fig. 2. Modular gaze control structure: (I) *Information Management*, (II) *Task-specific Gaze Strategy*, and (III) *Decision Scheme*.

current situation. This means that not at every instant high precision in all estimates involved in the locomotion process is required, i.e. not all uncertainties, although desirable, have to be minimized simultaneously. Note the difference between the problem considered here and some typical environment mapping problems, for details see [11]. Depending on the task some uncertainties become critical and other irrelevant and that is the reason to employ the term *incertitude*. This term indicates which uncertainties, i.e. estimates and measurements, are directly involved in the current task and must be therefore considered as relevant or critical. Thus, this term expresses which and how much the uncertainties are relevant for a given task. The specific definition of the *incertitude* requires a deep understanding of the nature of the problem considered.

With the defined concepts we present a biologically inspired approach to predictive gaze control for an active vision system based on the maximization of the predicted visual information, see [12]. The general approach takes into account the accumulated available information about the scene (for example a 3D-map), the current observer location, actual motion parameters, and the current task to predict the optimal pose of the visual sensor for a future view situation (*prediction* or *anticipation*, see [13]). An optimal pose is found by selecting the state of maximum *predicted information content* among a set of predicted states.

When navigating, one of the main goals is to follow a planned path and arrive at a desired location without getting lost. This means that the uncertainties related to self-localization must be kept small. These uncertainties are those describing the position of the robot in the world, which are necessary for navigating, e.g. path planning and correction. In order to achieve a precise knowledge about the pose of the robot, the perception system must focus on objects whose position in the world is more exactly known, i.e. landmarks. In this case, the goal for gaze control can be stated as gathering the largest amount of relevant

information so that the lowest level of the uncertainties in the position of the robot during its motion is guaranteed. The second main objective of safe locomotion is avoiding collisions with obstacles in the environment. This implies that the relevance of an obstacle for the gaze control depends in this case on the risk of collision that this object may represent for the robot locomotion. Given a local path through a scenario with different objects, there exist several points on this path in which the danger of colliding with some obstacle is bigger than in others. The goal of the gaze controller is to minimize the risk of collision in these *critical points*.

The next step in order to fulfill the requirements of autonomous locomotion must be to find a way of combining the two strategies by means of the definition of an optimal *decision strategy*, Fig. 2-(III), i.e. *Where to look next?* in order to optimally fulfil both navigation-tasks depending on the current situation: In other words, minimize self-localization *incertitude* as well as obstacle avoidance *incertitude*. The *decision maker*, i.e. an *Agent* or a group of *agents*, is in front of an *Action/Selection* problem (here, the selection of a view direction) with two different objectives, of different natures. These goals have to be weighted and pursued in parallel. The analysis of the action/selection problem in conventional terms showed clearly that one agent - definition of only one *Utility Function*, i.e. a simple *way of thinking* and fixed priorities - which could be optimal in a concrete situation, but could lead to risky states in different scenarios. In order to safeguard the flexibility of the decision maker, the principal objective of this work was then to find some kind of suboptimal behavior which could assure a secure and satisfactory view direction choice in a wide range of situations and scenarios. For the decision process, we propose a *Winner Selection Society*, which is a kind of decision making in a group of very different agents. Each agent proposes a view direction and how much it wants to choose it. Group decisions in large and homogenous groups also seem to be the best way to take decisions in current life, as very different ways of thinking are supposed to be listened to and respected.

The article is organized as follows: After presenting in this Section the basic problem of gaze control in vision guided walking is defined. Section 2 gives an outline of the information management mechanism developed. It represents the starting point for the general gaze control strategy and different task-specific *predictive* view direction control presented in Section 3. In Section 4 the proposed decision strategy for gaze control is described. Section 4.4 demonstrates with simulations the applicability of this approach. The article closes with some concluding remarks in Section 5.

2 Information Management Mechanism

In [8] we presented an *information management* mechanism for active gaze control in the context of vision guided humanoid walking, cf. Fig. 2-(I). The *information/uncertainty management*, i.e. relationship between system state estimation and the active measurements, rests upon a *coupled* (considering cross-

covariances) *hybrid* (reflecting the discontinuous character of biped walking) *Extended* (nonlinear system) *Kalman Filter*. In this earlier contribution, we proposed a novel definition of the *state vector*, which will prove its suitability when trying to solve key problems of navigation. The *state variables* selected are: the pose of the robot's right or left foot frame $S_{F_{R,L}}$ with respect to the world frame S_0 – ${}_0\boldsymbol{x}_0$ for self-localization; and the position of the various objects i with respect to the foot frame $S_{F_{R,L}}$ – ${}_F\boldsymbol{x}_i$ for obstacle avoidance, see Fig. 4.

$$\boldsymbol{x} = \left[\, {}_0\boldsymbol{x}_0^T \;\; {}_F\boldsymbol{x}_1^T \;\; {}_F\boldsymbol{x}_2^T \cdots \right]^T, \text{ with } {}_0\boldsymbol{x}_0 = \left[\, x_0 \;\; y_0 \;\; \theta \,\right]^T \text{ and } {}_F\boldsymbol{x}_i = \left[\, x_i \;\; y_i \;\; z_i \,\right]^T. \quad (1)$$

Based on the *state vector* $\boldsymbol{x}$, a complete model of the *state functions* of both, the system and the measurement process have to be considered in order to obtain an appropriate formulation of the EKF. *Hybrid* system equations – both, for a step execution period and for standing foot change instants – can be easily defined for this *state vector* together with the robot model errors.

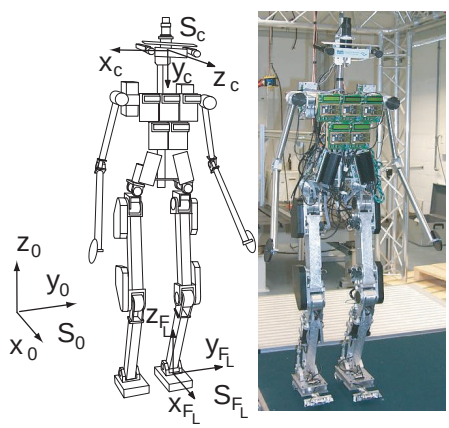

Fig. 3. Biped robot with camera reference frame S_C, foot reference frame S_{F_L} and world frame S_0.

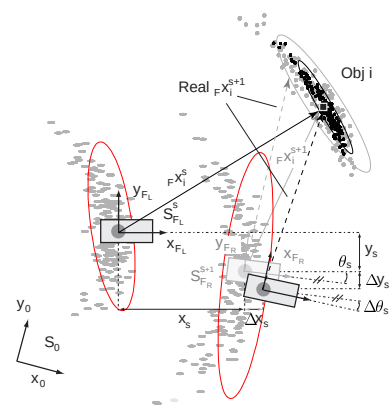

Fig. 4. Stepping process and dead-reckoning errors in biped walking and linearized estimates.

The biped robot model considered is a general model based on experience with real walking machines [5]. The robot model comprises kinematic transformations between different reference coordinate frames. ${}^C\boldsymbol{T}_{F_{R,L}}$ describes the transformation between the camera reference frame S_C and the foot frame $S_{F_{R,L}}$. This frame changes in each step, see $S_{F_L}^S$ to $S_{F_R}^{S+1}$ in Fig. 4, and remains attached to the currently standing foot for the step duration. This yields a modelling of the walking process taking into account the dead-reckoning step errors, since both, the relative position of the next foot and its orientation are considered as erroneous (errors $\Delta x_S, \Delta y_S$, and $\Delta\theta_S$). The effects of these errors on the estimates and the linearization approximation are shown in Fig. 4.

With the already explained process model for walking, the *state transition function* $\boldsymbol{f}$ for the novel hybrid filter formulation can be represented as follows:

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}\left(\boldsymbol{x}_k, \boldsymbol{u}_{k+1}, \boldsymbol{w}_k\right) = \boldsymbol{x}_k\left(1 - \gamma_{k+1}\right) + \boldsymbol{f}_s\left(\boldsymbol{x}_k, \boldsymbol{w}_k\right)\gamma_{k+1}\ , \quad (2)$$

where γ_{k+1} is a binary variable representing the current control vector: $\gamma_{k+1} = 0$ means no change in the coordinate frame, whereas $\gamma_{k+1} = 1$ means an instantaneous change in the coordinate frame. Consequently, the function $\boldsymbol{f}_s$ is defined as the transformation of the *state vector* $\boldsymbol{x}$ when the standing foot changes.

In addition, in this work, a visual perception model for a stereo-camera pair was considered and an appropriate measurement equation was chosen based both on the stereo-vision system properties and the object features detected by it. Its capability of measuring distances is used, and both, the reconstruction formulation and the error representation have been explained in detail in [9].

3 General and Task-Specific Predictive Gaze Control Strategy

As discussed in Section 1, an active vision system needs a view direction control mechanism to determine where to direct the sensor so that the interesting objects in the scene are within its field of view. A simple reactive controller neglecting the delay due to non-zero image acquisition and processing time seems therefore inappropriate for active vision application. Prediction proves to be an appropriate method to improve performance in such systems. For biological and technical considerations on prediction or *anticipation* see [14] and [13] respectively.

With the earlier defined concepts of *information content* and *incertitude*, we present a predictive general gaze control strategy for an active vision system based on the maximization of the predicted visual information. The approach is as follows: by use of the accumulated available information about the scene (for example a 3D-map, i.e. ${}_F\boldsymbol{x}_i$ and covariance matrix $\boldsymbol{C}_{ij}$), the current observer location ${}_0\boldsymbol{x}_0$, actual motion parameters, and the current task (which determines the relationship $\nu_i = f(\boldsymbol{x}, \boldsymbol{C}_{ij}, task)$), the gaze controller predicts the quasi-optimal pose of the visual sensor for a future view situation $\widehat{\boldsymbol{\Omega}}_\star$ (predicted variables are denoted by $\widehat{\ }$ and optimal ones by index $\star$). The optimal pose, i.e. position and orientation, but often and without loss of generality only orientation (view direction), is found by selecting the state of maximum predicted *information content*, $\widehat{\mathcal{IC}}_{max}$, among a set of predicted states. This can be expressed by:

$$\widehat{\boldsymbol{\Omega}}_\star = \arg\max_{\widehat{\boldsymbol{\Omega}}} \sum_{i=1}^{N} \widehat{\mathcal{IC}}_i(\widehat{\boldsymbol{\Omega}}, {}_0\boldsymbol{x}_0, {}_F\boldsymbol{x}_i, \boldsymbol{C}_{ij}, \nu_i), \tag{3}$$

subject to

$$\boldsymbol{\Omega}_{min} \leq \widehat{\boldsymbol{\Omega}} \leq \boldsymbol{\Omega}_{max} \quad \text{and} \tag{4}$$

$$\boldsymbol{r}(\dot{\boldsymbol{\Omega}}) = \boldsymbol{0}\,, \tag{5}$$

where Eq. (4) are upper and lower bounds for the active vision system parameters and Eq. (5) guarantees a continuous gaze behavior without abrupt view direction changes. The latter would disturb the perception process. Eq. (5) takes into consideration the observer's system dynamics and limits the performance (e.g. speed) of saccades.

In the following, as explained in Section 1, two task-specific *incertitude* formulations for self-localization and obstacle avoidance are presented, see Fig. 2-(II). Based on these formulations, two strategies appropriate for optimally solving these problems are stated.

3.1 Self-Localization *Incertitude*

In order to follow a path and avoid getting lost the uncertainties related to self-localization must be kept small. These uncertainties are those describing the position of the robot in the world, i.e. ${}_0\boldsymbol{x}_0$ in world frame S_0, cf. Fig. 4, necessary for navigating, e.g. path planning and correction. In order to achieve a precise knowledge about the pose of the robot, the perception system must focus on those objects whose position in the world S_0 is more exactly known, e.g. landmarks with absolute position ${}_0\boldsymbol{L}$. According to the first order stochastic position error representation in this work, the position uncertainty of the objects in the foot frame $S_{F_{R,L}}$ and the robot position in world frame S_0 are represented by means of 3D/2D-ellipsoids, see Fig. 4. The size of the largest independent axes of these ellipsoids are proportional to the squares of the *eigenvalues* of the covariance matrix.

Self-localization *incertitude* definition is stated by means of *task-dependent* translations of the uncertainties in the robot position. The process is then to map the robot position covariance matrix $\boldsymbol{C}_{\boldsymbol{x}_0\boldsymbol{x}_0}$ into the self-localization *incertitude* ν_0. *Incertitude* at the step s, ν_0^s is then defined as the mean of the position uncertainties of the present foot position S_F^s in the world frame S_0. This can be expressed by:

$$\nu_0^s = \frac{1}{2}\sum_{j=1}^{2}\sqrt{e_j^s} \tag{6}$$

where e_j^s are the *eigenvalues* of $\boldsymbol{C}^s_{\boldsymbol{x}_0\boldsymbol{x}_0}$. Although $\boldsymbol{C}_{\boldsymbol{x}_0\boldsymbol{x}_0}$ comprises the position and orientation covariances, the latter have not been included in this definition. Note that this does not imply any change in the strategy; it is a way of defining the uncertainties to be reduced. Simulations based on this defintion have demonstrated that, due to characteristics of the stereo measurement system, the robot orientation is precisely estimated whenever the position uncertainties are low.

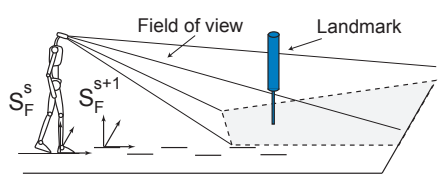

Fig. 5. Definition of *incertitute* for Self-localization.

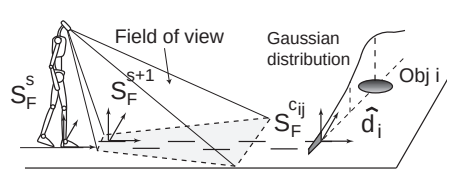

Fig. 6. Definition of the *probability of collision*.

Gaze Control Strategy for Self-localization

In this case, the goal for the gaze control can be stated as gathering the largest amount of relevant information so that the lowest level of the uncertainties in the position of the robot during its motion is guaranteed. As explained earlier, the objective is then to define the optimal camera orientation $\widehat{\boldsymbol{\Omega}}_\star$ that maximizes

the predicted acquired $\mathcal{IC}_0$ in the measurement process in the next step $s+1$, cf. Fig. 5. This can be expressed as a maximization of the *incertitude* reduction:

$$\begin{aligned}\widehat{\boldsymbol{\Omega}}_{\star}^{s+1|s} &= \arg\max_{\widehat{\boldsymbol{\Omega}}} \widehat{\mathcal{IC}}_0^{s+1}\left(\widehat{\boldsymbol{\Omega}},\, \boldsymbol{x}^{s},\, \boldsymbol{C}^{s},\, {}_0\boldsymbol{L},\, \nu_0^{s+1|s}\right) = \\ &= \arg\max_{\widehat{\boldsymbol{\Omega}}} \left(\nu_0^{s+1|s} - \hat{\nu}_0^{s+1|s+1}\right) \;=\; \arg\min_{\widehat{\boldsymbol{\Omega}}} \hat{\nu}_0^{s+1|s+1}\end{aligned} \tag{7}$$

3.2 Obstacle Avoidance *Incertitude*

The concurrent main objective for safe locomotion is avoiding collisions with obstacles in the environment. This implies that the relevance of an obstacle for gaze control depends on the risk of collision that this object may represent for the robot locomotion. Thus, an obstacle avoidance *incertitude* definition must include the planned local path in order to be able to optimize the resulting gaze behavior [15]. Given a local path through a scenario scattered with various objects, there exist several points on this path in which the danger of colliding with some obstacle is higher than in others. These points are denoted as *critical points* for collision and are defined as the points on the path where the distance d_i between the standing foot S_F and the object i has a local minimum, taking into account the discrete character of the walking process. These points and their critical distances – when the robot passes by for the j-th time in the instant c_{ij}, cf. Fig. 6 – can be characterized both by the position ${}_F\boldsymbol{x}_i^{c_{ij}}$ of the object i relative to the predicted foot frame $S_F^{c_{ij}}$ and its covariance matrix $\boldsymbol{C}_{ii}^{c_{ij}}$. The estimation of these parameters in the *critical points* rests upon the propagation of the system state and can be achieved by means of a linearization of the system equations, see Section 2. Information inputs for these estimates are previous measurements (in step s and before). The presented aspects lead to a definition of obstacle avoidance *incertitude* is proposed corresponding with the concept of *probability of collision*. This *incertitude* is $\nu_{ij}^{s} = \widehat{P}_i^{c_{ij}}$, i.e. the predicted *probability of collision* with the object i in c_{ij}, where c_{ij} is a descriptor for a critical point. This gives a measure of the *security* when crossing the object i the j-th time. $\nu_{ij}^{s|s}$ should reflect the uncertainty in the position of the object i in the instant c_{ij} from ${}_F\widehat{\boldsymbol{x}}_i^{c_{ij}|s}$ and $\widehat{\boldsymbol{C}}_{ii}^{c_{ij}|s}$. In c_{ij}, the critical parameters are the predicted object distance $\widehat{d}_i$ and its corresponding uncertainty represented by $\widehat{C}_{\widehat{d}_i}$ – which can be easily obtained from $\widehat{\boldsymbol{C}}_{ii}^{c_{ij}|s}$. With both the *predicted* distance $\widehat{d}_i$ and the statistical deviation $\sigma_{\widehat{d}_i}$, the probability $\widehat{P}_i^{c_{ij}}$ of collision with the object can be estimated by:

$$P_i = \int_{-\infty}^{0} \frac{1}{\sigma_{d_i}\sqrt{2\pi}}\, e^{\frac{-(x-d_i)^2}{2\sigma_{d_i}^2}}\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\widehat{d}_{ij}}^{\infty} e^{\frac{-x^2}{2\sigma^2}}\, dx \tag{8}$$

P_i can be calculated either numerically or from look-up tables. In the present case, where there is only an interest in the reduction of these values, and not in

their evaluation, it may be replaced by $P_i \sim \sigma_{d_i}/d_i$, which means a substantial reduction of the computational efforts.

Gaze Control Strategy for Obstacle Avoidance

The task of the gaze control can be defined as follows: given the current situation described by a planned local path, several *critical points* for collision, the present knowledge about the environment, and the characteristics of the walking machine. The goal consists in minimizing the risk of collision, i.e. the obstacle avoidance *incertitude* ν_{ij}^{s} in order to gather sufficient information either for passing the object safely or planning an alternative path around it. This definition also holds in the case of complex paths, since multiple critical points may appear for each object. The objective is then to define the optimal camera orientation $\widehat{\boldsymbol{\Omega}}_\star$ that maximizes the predicted acquired $\mathcal{IC}$ in the measurement process in $s+1$ for the obstacle avoidance task. This can be expressed as a maximization of the *predicted incertitudes* reduction:

$$\begin{aligned}\widehat{\boldsymbol{\Omega}}_\star^{s+1|s} &= \arg\max_{\widehat{\boldsymbol{\Omega}}} \sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \widehat{\mathcal{IC}}_{ij}^{s+1} \left(\widehat{\boldsymbol{\Omega}},\, \boldsymbol{x}^{s},\, \boldsymbol{C}^{s},\, \nu_{ij}^{s+1|s}\right) = \\ &= \arg\max_{\widehat{\boldsymbol{\Omega}}} \sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \left(\nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}\right) = \arg\min_{\widehat{\boldsymbol{\Omega}}} \sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \hat{\nu}_{ij}^{s+1|\,s+1} \end{aligned} \tag{9}$$

where N_{obj} is the number of objects present and N_{c_i} the estimated number of *critical points* per object. $\widehat{\mathcal{IC}}_{ij}^{s+1}$ is the predicted acquired $\mathcal{IC}$ about the objects' relevance in the *critical points* c_{ij}, and equals the difference $\nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}$. The first term represents the predicted *probability of collision* with the present knowledge, and the second represents the probability if a measurement is taken under the conditions $\widehat{\boldsymbol{\Omega}}$ in the next step $s{+}1$. Since all the *incertitudes* ν_{ij} have the same meaning, i.e. predicted *probabilities of collision*, these are directly comparable. All the reductions have the same significance, although some will be reduced more than others.

Thus, one can assert that the essence of the strategy is to minimize the uncertainty in the objects position at the key instants of time – *critical points*, thus doing the best that the accumulated information allows to do at each instant. Future measurements are obviously not taken into account due to the enormous increase in alternatives this would imply. For more details and practical considerations refer to recently published work on this topic [15].

3.3 Implementation and Simulation

To test the validity of the approach described so far, simulation experiments were carried out in which a local path with stochastic deviations was generated for the robot by driving its motion model with Gaussian noises: This simulates a local path composed by straight lines connected by various corners, and the robot's noisy motion over it. Different goals for gaze control are simulated and both, the

effects of this goal-oriented perception and the functioning of the *information management* mechanism are demonstrated.

Simulation Layout

The robot is commanded to perform position-based navigation along a path (from *Start* to *Goal*, see Fig. 7 and Fig. 8), specified by a sequence of step parameters. The scenario comprises an irregular distribution of feature points – representing obstacles ($Obj\,1$ and 2, marked with squares) and landmarks (marked with asterisks) – of which the robot is able to make simulated noisy stereo visual measurements.

For the simulation we assume a walking motion resulting in a constant head movement in finite 3D-space starting with straight ahead motion in x-direction, and performing steps translating its foot frame $S_{F_{R,L}}$ according to $[x_S \ \pm y_S \ 0]^T$. "+" is used to change from the right to the left foot, and "-" otherwise. Assumed values are x_S=0.3 m and y_S=0.2 m. The parameter θ_S describing the changes in orientation is varied after some steps in order to follow the planned local path. In the simulation three sources of white Gaussian noise for the walking process have been considered: $\sigma_{\Delta x_S} = 0.005\,m$, $\sigma_{\Delta y_S} = 0.005\,m$, and $\sigma_{\Delta\theta_S} = 3°$. Errors in the kinematic chain modelling have not been taken into account. The active vision system takes an image of the scenario every 0.02 m, as the robot moves. After a sequence of 14 measurements (marked with crosses), a step is performed, i.e. the foot frame $S_{F_{R,L}}$ changes to the other foot. Assuming the measuring process to take place in camera frame-rate, this corresponds to a walking speed of $0.6\,m/s$, wich is typical for current bipeds [16,17]. For the measurement process a Gaussian non-homogeneous (*foveated*) pixel noise with a standard deviation depending on the projecting point has been considered [18]. The active perception system has only two rotational degrees of freedom: *pan* and *tilt*, i.e. $\boldsymbol{\Omega} = \{pan\,tilt\}$. The gaze control system makes a prediction once per step and selects the new view direction $\boldsymbol{\Omega}_\star$.

Simulation Results

In order to demonstrate the task-dependent gaze behaviors resulting from the different specific strategies, results from two different simulations are presented, see Fig. 7 and Fig. 8. Real and estimated robot head motion and feet positions in the world frame are depicted. The operation of the *information management* mechanism is illustrated by means of the estimated head motion and feet position estimates – by the 90% confidence ellipses. The resulting gaze direction in each step is indicated by arrows.

The simulation in Fig. 7, shows the resulting behavior when considering the self-localization task only, cf. Section 3.1. Here the goal is to gather as much information as possible in order to reduce the uncertainties in the robot position relative to world, i.e. reduce the predicted *incertitude* ν_0^{s+1}. It is evident, that in this case the perception of landmarks becomes rather important and it is absolutely dominating the behavior. It is noteworthy that the fixation during a longer period of some specific landmarks does not seem to be adequate, but

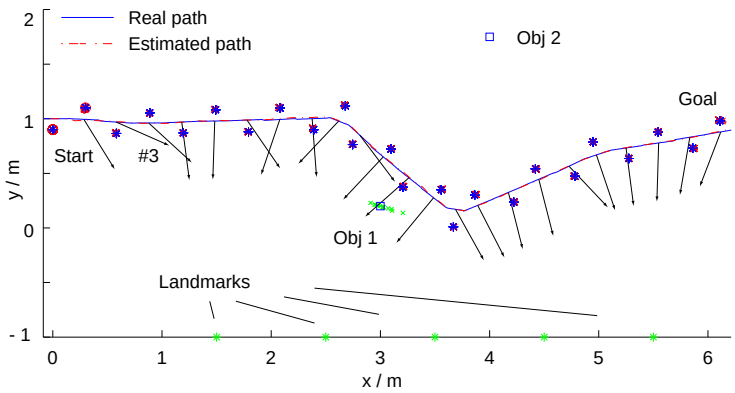

Fig. 7. Simulated scenario and view direction behavior for self-localization task.

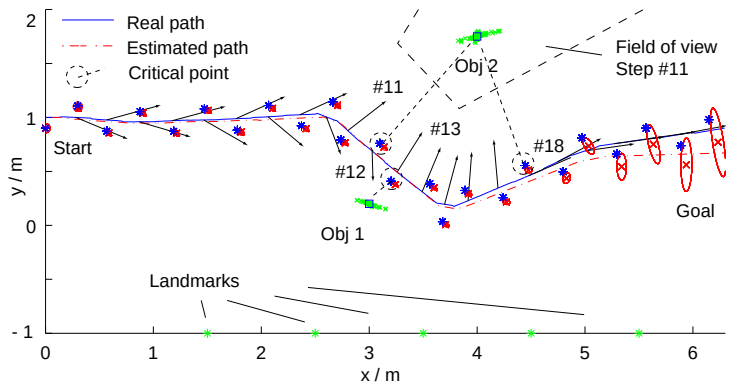

Fig. 8. Simulated scenario and view direction behavior for obstacle avoidance task.

maintaining a level of knowledge about different ones, contributing in this way to a greater extent to the localization under the non homogeneous precision characteristics of a stereo vision measurement system. Nevertheless, the self-localization goal does not lead to a gaze behavior which neglects measuring the objects, since these may be used to improve the localization of the robot relating to the information about well situated objects, which act in some way as landmarks, see measurements of $Obj\,1$ in $Step\,\#3$.

The simulation in Fig. 8, shows the behavior when considering obstacle avoidance task only. The goal is here to gather as much information as possible to reduce the *probability of collision*, as defined in Section 3.2. The objects are focused relative to the presence of *critical points* for collision. For $Obj\,1$ $Step\,\#13$ is critical, and steps #12 and #18 for $Obj\,2$. It is obvious that in the vicinity of these *critical points*, these objects become more relevant. At the last critical point of the scenario, $Step\,\#18$, no more obstacles are relevant and the view

direction remains in its rest position. From this point on, the functioning of the *information management* mechanism can be observed. No extra information is gathered by the perception system and therefore the quality of estimation of the robot position becomes worse. In *Step* #11, in addition to the arrow, the field of view on the ground is depicted in order to give an indication of the measurement range. Analogous to previous consideration in the simulation for self-localization, here the fact that the task reduces to obstacle avoidance does not imply that the gaze control neglects to considered the landmarks. The reduction of the uncertainties in the position of the robot may improve the localization of different objects with respect to the foot frame S_F in a better way than, for example, measuring just a single object.

4 Gaze Control Decision Strategy

After the definition of the two task-specific strategies for obstacle avoidance and self-localization, an adequate way of combining both strategies is proposed. For this purpose, a *decision strategy*, Fig. 2-(III), in order to optimally fulfil both navigation-tasks depending on the current situation is presented. The *decision maker*, i.e. an *agent* or a group of *agents*, is in front of an action/selection problem, i.e. the selection of a view direction, with two different objectives of different natures: minimizing self-localization *incertitude* as well as obstacle avoidance *incertitude*. Taking a decision in an action/selection problem consists on balancing judgements about uncertainties with preferences for possible outcomes in front of multiple conflicting objectives. In order to safeguard the flexibility of the decision maker, the principal objective of this work is to find some kind of suboptimal behavior which could assure a secure and satisfying view direction choice in a wide range of situations and scenarios. For the decision process, we propose a *winner selection society*, which is a kind of decision making in a group of rather different agents, as indicated in Fig. 2-(III). In order to represent the preferences of the different agents, appropriate utility functions have been defined. According to these functions, each agent proposes the view direction it wants to choose and how much it wants to choose it. For an introduction on decision strategies and related work on decision making see [19,20,21,22]; for concepts related to Utility Theory refer to [23,24].

In the present study the set of actions $\boldsymbol{A} = \widehat{\boldsymbol{\Omega}}$, designates the set of all possible view directions the agent can choose for the next step, $a \in A$ being a feasible alternative. The first attribute of the multi-attributive consequence vector O, o_O is the reduction of self-localization *incertitude*, the second attribute o_{ij}, represents the reduction of obstacle avoidance *incertitude*. In general the proposed utility functions can be defined as:

$$u = u(\nu_0^{s+1|s} - \hat{\nu}_0^{s+1|s+1}; \sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}) = u(o_O, o_{ij}); \qquad (10)$$

4.1 Proposed Agents

Utility functions characterizing the agents represent different kinds of preference rankings over the predicted outcomes, i.e. different *ways of thinking*. For each agent some characteristics which represent the attitude of the agent when deciding are explained. These characteristics describe the form and range of the utility and sub-utility functions. The three main characteristics are: (a) monotonicity, (b) independence, and (c) attitude towards risk. (a) Monotonicity describes the character of the preferences of the agent; for example monotonous decrease means that the agent always prefers a smaller amount of *incertitude* to a larger one. (b) Utility or additive independence expresses the grade of independence between both attributes: self-localization and obstacle avoidance. Utility independence essentially asserts that the preferences over outcomes with different values of a given attribute do not depend on the values to which the remaining attributes are set. Additive independence implies that preferences over values of both attributes are independent of each other, and the strength of these preferences remains unchanged for all possible values of both attributes. (c) An agent being risk averse – opposite to risk prone – means that it has a conservative attitude when taking a certain decision, which can be expressed by the concavity or convexity of the utility function. For more details about these aspects and normalization of sub-utility functions refer to [25].

Additive Agents: $AA_{\boldsymbol{r}}$

Three characteristics describe this first group of decision makers: Monotonous decrease, additive independence assumption, and risk neutrality. The additive independence assumption allows to represent the multi-attribute utility function for navigation for these agents as a sum of two sub-utility functions:

$$\begin{aligned} u(O) &= (1-r)u(o_O) + ru(o_{ij}) = (1-r)u(\widehat{\mathcal{IC}}_0^{s+1}) + ru\Big(\sum_{i=1}^{N_{obj}}\sum_{j=1}^{N_{c_i}} \widehat{\mathcal{IC}}_{ij}^{s+1}\Big) \\ &= (1-r)u(\nu_0^{s+1|s} - \hat{\nu}_0^{s+1|s+1}) + ru\Big(\sum_{i=1}^{N_{obj}}\sum_{j=1}^{N_{c_i}} \nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}\Big). \end{aligned} \tag{11}$$

u is linear due to risk neutrality and monotonicity. It is known that $u^\star = a + bu$ is equivalent to u, with a and b constants, thus the utility function can be written as:

$$\begin{aligned} u^\star(O) &= (1-r)u^\star(\hat{\nu}_0^{s+1|s+1}) + ru^\star\Big(\sum_{i=1}^{N_{obj}}\sum_{j=1}^{N_{c_i}} \hat{\nu}_{ij}^{s+1|s+1}\Big) \\ &= (1-r)u^\star(o_O^\star) + ru^\star(o_{ij}^\star) \end{aligned} \tag{12}$$

Note that between this group r $= \{0, 0.25, 0.5, 0.75, 1\}$ of decision makers, there are two extreme cases which only consider $u(o_O)$ (r = 0) and $u(o_{ij})$ (r = 1). For both cases, no mixture of strategies is present: the first one follows a

pure self-localization strategy, the second one a pure obstacle avoidance strategy, see Section 3.

Multiplicative Agent: MA

The MA was adapted from the proposal of D. Rimey's utility function for navigation: $U(a) = V(a) \,/\, C(a)$ (value of an action per cost of an action) [21,22]. This approach could be characterized with the key terms: monotonous decrease, risk neutrality, and utility independence. Some necessary considerations for normalization must be made: $u(o_{ij}^{worst}; o_0^{best}) = 0$ and $u(o_{ij}^{best}; o_0^{worst}) = 0$, which means that choosing the worst option for self-localization *incertitude* or choosing the worst option for obstacle avoidance *incertitude* implies directly an outcome O with $u(O) = 0$. According to [25] the utility function can then be written as:

$$\begin{aligned} u(O) = \; & K \; u(o_O) \; u(o_{ij}) = K \; u(\widehat{\mathcal{IC}}_0^{\,s+1}) \; u(\textstyle\sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \widehat{\mathcal{IC}}_{ij}^{\,s+1}) \\ & = K \; u(\nu_0^{s+1|s} - \hat{\nu}_0^{s+1|s+1}) \; u(\textstyle\sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}), \end{aligned} \tag{13}$$

where K expresses the weight of one of the two strategies over the other. u is strategically equivalent to $u^\star$ *iff* $u^\star = a + bu$. For taking a decision $u^\star$ as well as u, can be used:

$$u^\star(O) = u(\nu_0^{s+1|s} - \hat{\nu}_0^{s+1|s+1}) * u\Big(\sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \nu_{ij}^{s+1|s} - \hat{\nu}_{ij}^{s+1|s+1}\Big). \tag{14}$$

Note, that the assumptions for the worst option in order to become the multiplicative utility function are not appropriate in the navigation problem. In a situation where the agent is well orientated and has a low self-localization *incertitude*, but has a high obstacle avoidance *incertitude* and needs to reduce it, it does not seem very reasonable to reject the outcome $O = (o_0^{worst}; o_{ij}^{best})$, which would be the case if $u(O) = 0$. The reason why this kind of agents was studied is because they are commonly used method in multi-attribute decision making [22].

Risk Averse Agent by Self-Localization: RAA

The RAA agent can be defined by the three main characteristics: monotonous decrease, additive independence, and risk aversion. Risk aversion expresses a conservative attitude towards a decision. In this approach it is also considered that the decision of the agent is independent of its actual *incertitude* state, and so the consequence space is reduced to $O^\star = (o_0^\star; o_{ij}^\star) = (\hat{\nu}_0^{s+1|s+1}; \sum_{i=1}^{N_{obj}} \sum_{j=1}^{N_{c_i}} \hat{\nu}_{ij}^{s+1|\,s+1})$. By doing this, a globally acting agent acting is proposed. This assumption permitted to define a risk aversion coherent function over the outcomes, which would not be evident when thinking about *incertitude* reductions. Following the additive independence assumption, the utility function can be decomposed as in Eq. (11). Hence, if comparing this agent with AA_{r} agents, the weights of the sub-utility functions depending on the scenario will not be emphasized, whereas the global attitudes these sub-utility functions should present. Thus, both goals

actually have the same importance. Looking closer in the definition of both sub-utility functions, it can be stated:

$$u^\star(O) = u^\star(\ \hat{\nu}_0^{s+1|s+1})\ +\ u^\star(\sum_{i=1}^{N_{obj}}\sum_{j=1}^{N_{c_i}}\hat{\nu}_{ij}^{s+1|\,s+1}\ ) = u^\star(o_O^\star)\ +\ u^\star(o_{ij}^\star) \qquad (15)$$

For all possible view directions, the anchors consequences were not chosen locally for every step as by the previous agents, but following global assumptions that were considered for self-localization purposes in navigation:

- Based on numerous simulations, the worst admissible case for the global self-localization sub-utility function was considered, when the *incertitude* in self-localization reached a value of ν_0 =*worst*. This self-localization value corresponds to a utility of $u(o^{worst}) = 0$.
- The best consequence for the self-localization task, is evidently a self-localization *incertitude* of 0. This value corresponds to a utility of $u(o^{best}) = 1$. The agent needs to be able to give the optimal importance to the possibility of getting lost, giving the right scaled sub-utility for every state of the global ν_0 space of possible consequences.

To achieve this objective [25] was used in order to define a sub-utility function, and a decreasing risk aversion approach as the best way to represent preferences was chosen. Simulations were performed to distinguish some characteristic values of self-localization *incertitude*. Those values were ranked and then approximated by a characteristic form of a risk averse utility function presented in [25], a cubic polynomial.

Linear Approximation Agent: *LAA*

It is obvious that it is not a good idea allowing all agents to make decisions without considering the present *incertitude*, i.e. ignoring what happened until the current step. Therefore, when studying the risk aversion utility function independent of the actual *incertitude* states, a further approximation of the risk averse approach is defined. An evaluation of the *incertitude* at the present step is used to predict the weight of the self-localization linear utility function. Risk aversion is only present in the weighting factors, not in the sub-utility assessment. This agent is described by the assumptions: monotonous decrease, utility independence, and risk aversion. The additive independence assumption allows for this kind of agents to represent the multi-attribute utility functions for navigation as a sum of two sub-utility functions as for the first group of agents, AA_{r}.

4.2 Winner Selection Societies

For the decision process, we propose a *winner selection society* (*an abstract decentralized model of mind*), which is a kind of decision making in a group of very different agents, see Fig. 9. In order to represent the preferences of the different agents, appropriate utility functions have been defined in Section 4.1. According to these functions, each agent proposes its preferred view direction (*what*) and the grade of its preference (*how much*). This model is inspired by the work of Mark Humphrys [20]. Humphrys considers how agents might organize themselves in the absence of a global reward.

The term agent does not refer to the mind but to a part of it, i.e. it can be considered a piece of semi-autonomous software. It has its origin in Minsky's *Society of Minds* [26], in which multiple agents inhabit the same body. Agents, which have more autonomy than traditional modules or procedures, behave as autonomous actors. They are not necessarily structured hierarchically, and do not need to be cooperative.

In Humphrys' model an agent will control the entire body on its own if allowed. It is generally frustrated by the presence of other agents. Each agent must send its actions to a switch which will either obey or refuse it. Humphrys proposes to make the agents organize themselves off a dumb switch, rather than making this sort of switch simple. His model is organized from the point of view of the agent: they do not know about the global system or the mission of the creature as a whole. Each must live in a local world: they do not have explicit knowledge about the existence of any other agents. Next, some concepts relating to the way of represent *What* does the agent want to do, and *How much* does it want to be listened to, will be presented. For background information, refer to [20].

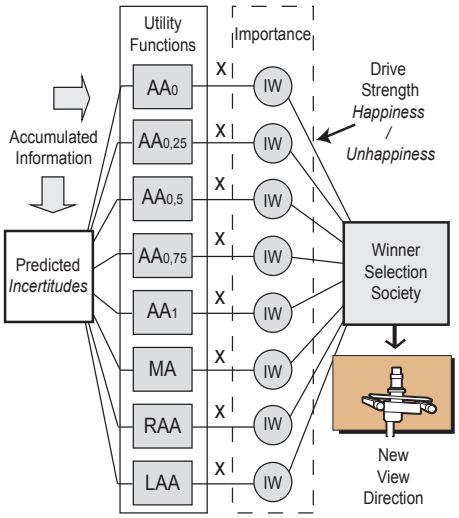

Fig. 9. Structure of the decision process: Agents and Societies.

– ***What does the agent want to do ?***: Each agent presents a set of view directions, which are classified with scores between 0 and 1, depending on the performance gain of the agent. The *predicted reward* $\boldsymbol{D}$ of each agent A_i at step s is $D_i(s) = p_i(o(a_i), s)\ U_i(o(a_i, s))$: the expected value of the utility distribution, if assuming it will propose the action a_i which maximizes its expected utility. In this case, $p_i(o(a_i)) = 1$ is supposed and thus $D_i(s) = U_i(o(a_i), s)$. All agents share the same set of actions, i.e. all agents

have to choose a concrete view direction among the same set of possible pan and tilt angles. If another agent is taking action a_k, an estimate of the *actual reward* $F_i(s) = U(o(a_k), s)$ is already available. Then the *Drive Strength* or *Weight* (dynamic $\boldsymbol{W}$-Values) can directly be assigned: $\boldsymbol{W} = \boldsymbol{D} - \boldsymbol{F}$, with $D_i = U_i(o(a_i), s)$ and $F_i = U_i(o(a_k), s)$. It can be described also as: $W_i(s) = d_{ki}(s)$, where d_{ki} is the *expected loss* that agent k, (A_k), causes for agent i, (A_i), if it is obeyed in state s. The losses that agents will cause to other agents will be defined in a $n * n$ matrix called *Value matrix*, $\boldsymbol{V}$. The elements of this matrix are all between 0 and 1 as the expression of the $\boldsymbol{D}$ and $\boldsymbol{F}$ for every agent were scaled before calculating $\boldsymbol{W}$. Then, the expression of $\boldsymbol{W}$ can be specified by the agents answering the question: *How much* does it want to be listened to?

– ***How much does the agent want to be listened to ?***: For this purpose, a new *importance value* $\boldsymbol{IW}$ of every separated task in each step is calculated and the values obtained are multiplied with the previously explained $\boldsymbol{W}$-Values to calculate the real $\boldsymbol{W}$ for every agent. $\boldsymbol{IW}_{self}$, the importance of the self-localization task, is obtained regarding the utility of the current value of self-localization *incertitude* obtained in the previous step. The *RAA* agent, conservative in front of the self-localization task, uses the risk averse utility function defined for the self-localization *incertitude* to express its $\boldsymbol{IW}_{self}$.
For the obstacle avoidance task, no utility function regarding the obstacle avoidance *incertitude* for every step can be correctly defined in absolute terms because of the dependence of *incertitude* values range on the number of objects present in the scenario, and of the location of the objects relative to the walking machine. Therefore, another kind of weighting $\boldsymbol{IW}_{objs}$ is proposed. In this case, the grade of improvement is evaluated, i.e. improvement between choosing for next step the predicted worst possible view direction for the obstacle avoidance task and the predicted best one. Depending on the magnitude of this difference, it is possible to assure if the current situation is important or not to pursuit the obstacle avoidance task. Therefore $\boldsymbol{IW}_{objs}$ is defined as proportional to this difference: $\boldsymbol{IW}_{objs} = f(o_{ij}^{best} - o_{ij}^{worst})$. The higher the possible improvement, the more important is the step for the obstacle avoidance task. Here were defined the $\boldsymbol{IW}_{self}$ for the self-localization task, and the $\boldsymbol{IW}_{objs}$ for the obstacle avoidance task. The $\boldsymbol{IW}_i$ for every agent, or *how much* does every agent want to be listened to, is then calculated as a function of both definitions.

$$\begin{cases} AA_p & IW(AA_p) = (1-p)IW_{self} + pIW_{objs} \\ MA & IW(MA) = IW_{objs} \text{ , similar to } \quad AA_1 \\ RA & IW(RA) = \quad 0.5IW_{self} + 0.5IW_{objs} \\ LA & IW(LA) = \quad 0.5IW_{self} + 0.5IW_{objs} \end{cases} . \tag{16}$$

Note that the $\boldsymbol{IW}$-Value of an agent grows, as long as it or a similarly thinking agent, did not become the winner for a longer period, representing a way of considering *history* when taking the decision. On the other hand, after being listened, its $\boldsymbol{IW}$ quickly decreases.

4.3 Proposed Societies

The different agents must be organized in a society determining when and how much to listen the different proposals. This organization is based on the definitions of *Happiness* and *Unhappiness*. There are five basic approaches to organizing action/selection without reference to a global reward. A deep analysis of other action/selection methods can be found in [20]. Some political and economical interpretation will be made here in order to better understand how the high level problem can be solved with the *society of minds* method.

Anarchy or Maximize Best Happiness

Here actions are promoted using a *static* form of $\boldsymbol{W}$, i.e. other agents are totally ignored. The agent A_i promotes its preferences over actions with the same strength no matter, if any competition. Then each one suggests its best proposal.

$$\text{(happiness)} \quad \boldsymbol{W}_i \;=\; \boldsymbol{IW}_i * \boldsymbol{D}_i, \qquad \text{(proposal)} \quad \max_{a \in A} W_i(s,a) \tag{17}$$

The switch becomes then a simple gate letting through the agent, with maximum of highest $\boldsymbol{W}$-values. This strategy is only interested in the best possible *individual happiness*.

$$\max_{a \in A} \max_{i \in 1,\ldots,n} W_i(s,a) \tag{18}$$

Liberal Democracy or Minimize Worst Unhappiness

Here, actions are promoted using a *dynamic* form of $\boldsymbol{W}$. How the agent suggests its actions depends on the current competition. Using such a measure of $\boldsymbol{W}$, does not mean that the agent will need explicit knowledge about who it is competing with. It needs only to have local knowledge: what state it is in, what action a it suggested.

$$\text{(unhappiness)} \quad \boldsymbol{W}_i \;=\; \boldsymbol{IW}_i * (\boldsymbol{D}_i - \boldsymbol{F}_i), \tag{19}$$

where $D_i = U_i(o(a_i), s)$ and $F_i = U_i(o(a_k), s)$. That is $W_i(s) = d_{ki}(s)$. d_{ki} is the expected loss that agent k, (A_k), causes for agent i, (A_i), if it is obeyed in state i. Here the *Value Matrix* $\boldsymbol{V}$ is identical with $(\boldsymbol{D} - \boldsymbol{F})$. The winner selection strategy is then:

$$\min_{a \in A} \max_{i \in 1,\ldots,n} W_i(s,a). \tag{20}$$

Firstly, the winner selection strategy looks for every agent, where and how it suffers when it suffers the most. Then, this strategy looks from the different suffering actions, which proposal is the one that makes the suffering agent suffer the less. This proposal is the one that will be executed. That is, the winner selection strategy is interested in avoiding a very suffering agent. Every action selection will suppose that there will be some agents which will suffer because of the decision taken. The action is chosen, for which the suffering agent, suffers the less. In economic theory, this would be the equivalent of a Rawlsian social welfare function [27], where the value of an allocation depends of the worst off agent. Here, the agents are listened, not because the winner selection strategy

cares about the different parts of the mind, but because individual agents are wanted to have the ability to raise their voices and to be heard above the others.

Democracy or Collective Methods

Collective methods will all generate a similar sort of behavior: they keep the majority of the agents happy, but at the expense of a small minority. Even if one agent is facing a very risky situation, it may still normally not be able to overcome the opinion of the majority. This kind of societies seem normally to be the more appropriate ones when talking about a society of humans and therefore majoritarism is often chosen to take a group decision. Majoritarism, is however not always the optimal solution if we are interested in avoiding minorities to be completely excluded from the decision making.

- **Maximize Collective Happiness:**
  This third proposed *society of minds* is a collective method which uses a simple form of $\boldsymbol{W}$. The agent simply suggests his preferences over the actions. Over all the different actions, it also chooses the best one. This action will be included into the set of possible actions analyzed.

$$\text{(happiness)} \quad \boldsymbol{W}_i \;=\; \boldsymbol{IW}_i * \boldsymbol{D}, \tag{21}$$

  Here the *Value Matrix* $\boldsymbol{V}$ is the same as $\boldsymbol{D}$.
  The winner selection law is then:

$$\max_{a \in A} \sum_{i=1\ldots n} \boldsymbol{W}_i(s,a) \tag{22}$$

  where A is the set of proposed actions.
  For every proposed action, it is calculated how much the total reward obtained will be; the expected reward of every agent is summed and multiplied by the importance the agent has in this situation. Then the proposal is executed which maximizes the global reward.

- **Minimize Collective Unhappiness:**
  The counterpart of the previous method, uses a *dynamic* form of $\boldsymbol{W}$. Every agent proposes its optimal action. Then the $\boldsymbol{W}$ is calculated by:

$$\text{(unhappiness)} \quad \boldsymbol{W}_i \;=\; \boldsymbol{IW}_i * (\boldsymbol{D}_i - \;\boldsymbol{F}_i), \tag{23}$$

  where $D = U_i(o(a_i), s)$ and $F = U_i(o(a_k), s)$
  That is, $W_i(s) = d_{ki}(s)$; where d_{ki} is the expected loss that agent k, (A_k), causes for agent i, (A_i), if it is obeyed in state S. The action selection process is then:

$$\min_{a \in A} \sum_{i=1\ldots n} \boldsymbol{W}_i(s,a) \tag{24}$$

  The losses that an actual proposed action causes to the global set of agents are summed. Then the proposal is chosen which yields the least global suffering.

Global methods can be compared with the Benthamite social welfare function, when talking about economical theories, which can be summarized in *the greatest happiness for the greatest number*.

4.4 Implementation and Simulation

To evaluate the capabilities of each of the presented agents, different simulations for different scenarios and different walking machines were carried out. With the objective of finding a *global and right* decision maker, i.e. one *super-brain* was searched. The simulations proved that for each specific situation, the optimal decision maker varies: depending on the number of objects and landmarks, the errors taken into account, or the environment distribution..., it seems more appropriate to prefer one special agent. From this finding, it became clear that it is needed to choose the agent depending on the scenario and on the situation. For the special group of agents in Section 4.1 simulations were made separately in order to study when a optimal *mixture* of goals was achieved and when the attitude is the optimal one. Depending on the position of the objects or landmarks with respect to the field of view of the humanoid, their number, and the considered visual error sources, the simulations results vary, and so the *best*, the safest solution ranges from r = 0 to r = 1 – pure self-localization to pure obstacle avoidance.

Simulation Layout In the presented simulations, the scenario and the characteristic of the robot and walking model are analogous to the layout of the simulations presented in Section 3.3. In general, the robot is commanded to perform position-based navigation along a path (from *Start* to *Goal*, see Fig. 7), specified in form of a sequence of step parameters. An irregular distribution of feature points – representing obstacles (Obj 1,Obj 2 and Obj 3) and landmarks – of which the robot is able to make simulated noisy stereo visual measurements, are present in this scenario. In order to show the situation-dependent gaze behaviors resulting from the different specific strategies, sample simulation results are presented in Fig. 10 and Fig. 11. Here, only the estimated robot head motion and not the feet positions in the world frame are depicted. The resulting gaze behavior is indicated in each step by arrows.

Simulation Results

In this section the simulation results for one selected scenario are presented. The presented simulations Fig. 10 show the resulting view directions for every agent separately in the same scenario. Note the differences between the agent view directions selection: for the two extreme cases, the additive agents with r = 0 and r = 1, the differences become, as expected, clearer. It is evident, that in the case of the self-localization task the perception of landmarks becomes very important, dominating absolutely the behavior. Nevertheless, the self-localization goal does not lead to a gaze behavior which neglects measuring the objects, since these may be used to improve the localization of the robot relating to the information about well-situated objects, which act in some way as landmarks [8,15]. In the case where the goal is to avoid the obstacles, the

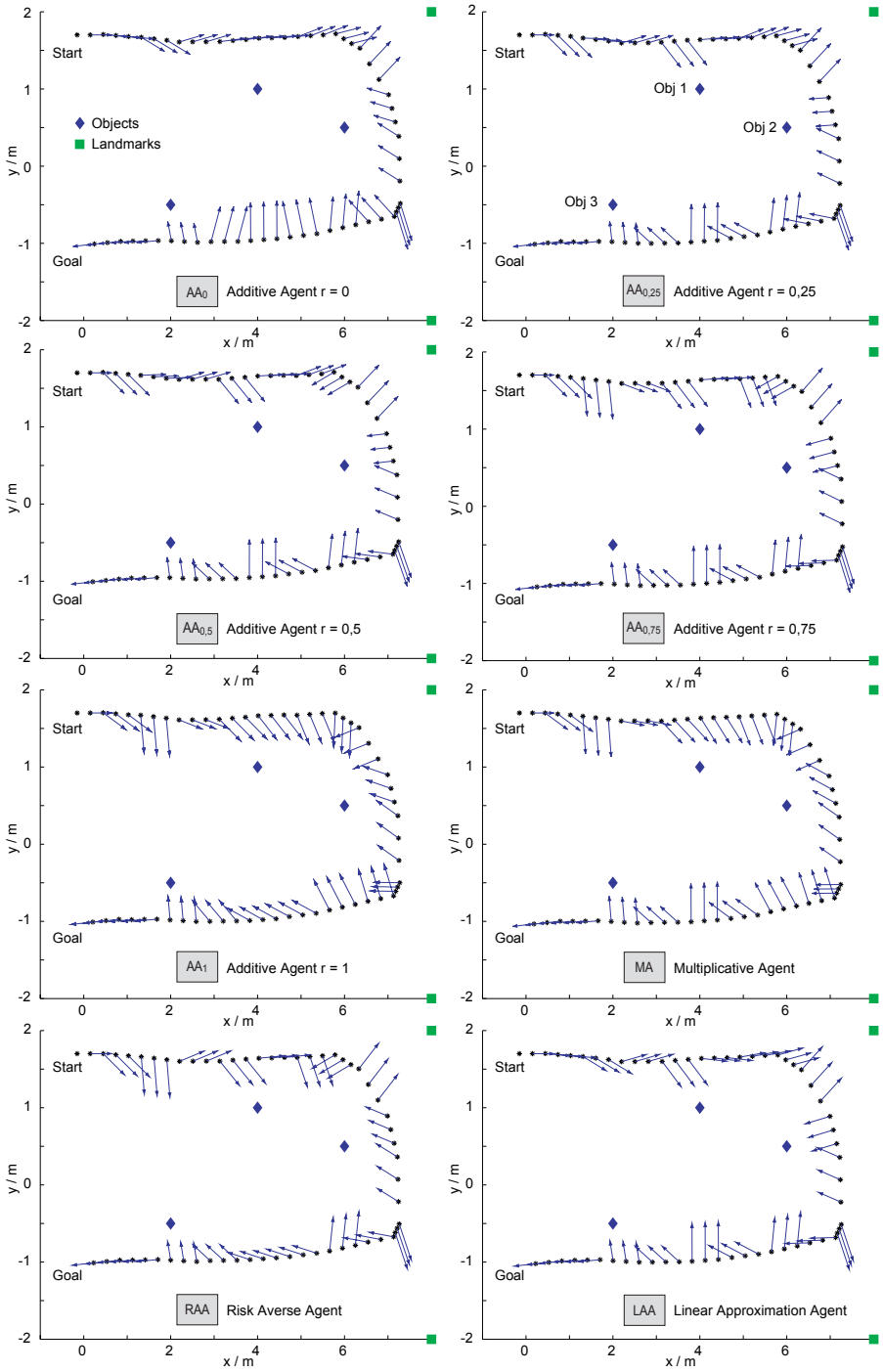

Fig. 10. Simulated scenario with walking motion and view direction behavior for different agents.

objects are focused related to the presence of *critical points* for collision. In this particular scenario the multiplication agent MA has a similar behavior as the r = 1 agent (obstacle avoidance agent). The linear approximation of the risk averse function, the LA agent, behaves on the other hand like the self-localization agent, and also similar to the risk averse one.

The different *Winner Selection Societies* presented in this article have been simulated, evaluated and compared in very different scenarios and situations. For evaluation, four different criteria were studied: local scaled *incertitude* rates – for self-localization and obstacle avoidance, global utility, mixture of the task and evaluation of the constraints (risks). Fig. 11 shows the results of one of this simulations. In this concrete scenario the Collective method *Minimize Global Unhappiness* assures a real pursuit of both tasks in parallel. The results obtained for both the self-localization task and the obstacle avoidance task are nearly the same as those obtained separately for the self-localization agent and the obstacle avoidance agent, respectively.

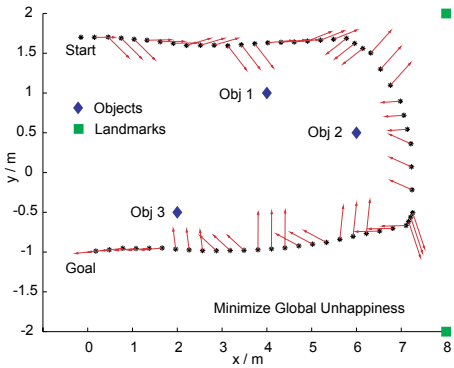

Fig. 11. View direction behavior for *winner selection society* Minimize Global Unhappiness.

Conclusions on Winner Selection Societies

After a detailed study of the different organizations, we found that the best society scheme would be the democratic method which assures global minimal unhappiness, i.e. *Minimize Global Unhappiness* society. The best society scheme depends clearly on the scenario and the objects position. This method presents the highest flexibility and is nearly always the best option that can be chosen.

5 Conclusions

In this article we presented a general approach to gaze control for navigation in the context of vision guided humanoid walking. Specific strategies for controlling the view direction for obstacle avoidance and self-localization tasks built the base for the developed general gaze control scheme, and show the generalization possibilities of this approach.

A redefinition of the concept of information and uncertainty in the context of vision guided humanoid navigation allowed the inclusion of the planned local path-dependence and the self-localization requirements. The general strategy is based on the maximization of the predicted visual information and employs for *information management* a *coupled hybrid* EKF. The resulting view direction behavior shows the dependence of the gazing strategy on the present task.

The next step for fulfilling the requirements of autonomous locomotion must be to find a way of compiling the two strategies by means of the definition of an optimal decision strategy, i.e. *Where to look next?* in order to optimally fulfill both navigation-tasks: minimize self-localization *incertitude* as well as obstacle avoidance *incertitude*. The agent is in front of an action/selection problem, the selection of a view direction, it has to take a decision with two different objectives of different natures. These goals have to be weighted and pursued in parallel. The general approach to this decision problem is to define different utility functions over the outcomes of the set of possible view directions. These utility functions i.e. *agents*, which represent different kinds of preferences rankings over the predicted outcomes, will then be organized to solve the action/selection problem as a *society of minds*.

Acknowledgments: This work was supported in part by the German Research Foundation (DFG) within the "Autonomous Walking" Priority Research Program.

References

1. P. J. McKerrow, *Introduction to Robotics*, Electronic systems engineering series. Australia Addison-Wesley Publishing, 1998. 1
2. I. Moon, J. Miura, Y. Yanagi, and Y. Shira, "Planning of vision-based navigation for a mobile robot under uncertainty," in *Procs. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Grenoble, France, 1997, pp. 1202–1207. 2
3. A. J. Davison, *Mobile Robot Navigation Using Active Vision*, Ph.D. thesis, Robotics Research Group, Department of Engineering Science, University of Oxford, October 1999. 2
4. H. Inoue, "Explorations of Human Robot Applications," in *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Tokio, Japan, November 2001, pp. 497–499. 2
5. O. Lorch, A. Albert, J. Denk, M. Gerecke, R. Cupec, J. F. Seara, W. Gerth, and G. Schmidt, "Experiments in vision-guided biped walking," in *Procs. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Lausanne, Switzerland, 2002, pp. 2484–2490. 2, 5
6. K. Nishiwaki, S. Kagami, J. J. Kuffner, K. Okada, Y. Kuniyoshi, M. Inaba, and H. Inoue, "Online humanoid locomotion control by using 3d vision information," in *Procs. of the I8th Int. Symp. on Exp. Robotics*, Sant'Angelo d'Ischia, Italy, July 2002. 2
7. Y. Aloimonos, "Introduction: Active Vision Revisited," Tech. Rep., University of Maryland, 1999. 2
8. J. F. Seara, K. H. Strobl, and G. Schmidt, "Information management for gaze control in vision guided biped walking," in *Procs. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Lausanne, Switzerland, 2002, pp. 31–36. 2, 4, 21
9. O. Lorch, J. F. Seara, K. H. Strobl, U. D. Hanebeck, and G. Schmidt, "Perception Errors in Vision Guided Walking: Analysis, Modeling, and Filtering," in *Procs. of*

the IEEE Int. Conf. on Robotics and Automation ICRA, Washington DC, USA, 2002, pp. 2048–2053. 2, 6

10. C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 397–423, 623–659, 1948. 2
11. S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millenium*, G. Lakemeyer and B. Nebel, Eds. Morgan Kaufmann, 2002, to appear. 3
12. J. F. Seara, O. Lorch, and G. Schmidt, "Gaze Control for Goal-Oriented Humanoid Walking," in *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Tokio, Japan, November 2001, pp. 187–195. 3
13. R. Möller, "Perception through anticipation - an approach to behaviour-based perception," in *New Trends in Cognitive Science*, ASoCs Technical Report 97-01, Ed., Vienna, 1997, pp. 184–190. 3, 6
14. A. E. Patla, E. Niechwiej, V. Racco, and M. A. Goodale, "Understanding the contribution of binocular vision to the control of adaptive locomotion," *J Exp Brain Research*, vol. 142, no. 4, pp. 551–561, Sep. 2002. 6
15. J. F. Seara, K. H. Strobl, and G. Schmidt, "Path-Dependent Gaze Control for Obstacle Avoidance in Vision Guided Humanoid Walking," in *Proceedings of the IEEE International Conference on Robotics and Automation ICRA*, Taipei, Taiwan, 2003, to appear. 8, 10, 21
16. M. Gienger, K. Loeffler, and F. Pfeiffer, "Towards the Design of a Biped Jogging Robot," Seoul, Korea, 2001, pp. 4140–4145. 10
17. S. Kagami, K. Nishiwaki, J. J. Kuffner, Y. Kuniyoshi, M. Inaba, and H. Inoue, "Online 3d vision, motion planning and bipedal locomotion control coupling system of humanoid robot: H7," in *Procs. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Lausanne, Switzerland, 2002, pp. 2557–2562. 10
18. L. Matthies and S. A. Shafer, "Error modeling in stereo navigation," *IEEE Journal of Robotics and Automation*, vol. 3, no. 3, June 1987. 10
19. K. P. Murphy, *Dynamic Bayesian Networks: Representation, Inference and Learning*, Ph.D. thesis, UC Berkeley, Computer Science Division, July 2002. 13
20. M. Humphrys, *Action Selection Methods using Reinforcement Learning*, Ph.D. thesis, Trinity Hall, Cambridge, 1997. 13, 17, 19
21. R. D. Rimey, "Where to Look Next Using a Bayes Net: An Overview," in *DARPA Image Understanding Workshop*, 1992, pp. 260–267. 13, 14
22. R. D. Rimey and C. M. Brown, "Control of Selective Perception Using Bayes Nets and Decision Theory," *International Journal of Computer Vision*, 1994. 13, 14, 15
23. M. Luce and M. Raiffa, *Games and Decisions*, John Willey and Sons, Inc. New York, 1957. 13
24. M. von Neumann and M. Morgenstern, *Theory of Games and Economic Behaviour*, Princetown University Press, New York, 1947. 13
25. M. Raiffa and M. Keeney, *Decisions with Multiple Objectives: Preferences and Tradeoffs*, Wiley series in probability and mathematical statistics. John Wiley and Sons, 1979. 13, 14, 15
26. M. Minsky, *Society of Mind*, A Touchstone Book, Simon and Schuster, 1988. 17
27. M. J. Rawls, *A Theory of Justice*, The Belknap Press of Harvard University, 1971. 19