

Bold Hearts 3D Soccer Simulation 2011 Team Description

Sander van Dijk, Valerio Lattarulo, Alex Metaxas, and Daniel Polani

School of Computer Science
University of Hertfordshire, UK

Abstract. We aim at systematically developing a battery of principled methods to generate behaviours useful to achieve a viable RoboCup 3D gameplay. behaviours. This is particularly interesting and challenging in humanoids since any possible tactical, strategical or cooperative aspects can only be successfully tackled once the basic skills are in place. The construction of basic skills in humanoids is usually an intricate business that requires a large amount of hand-tuning. We aim to develop a systematic path towards reducing this amount of handtuning both on the perception as well as the actuation side. We approach this by combining principled methods, many grounded in information-theory, some in well-known Kalman- and particle-filtering, as well as hand-coded components. The long-term goal is to ultimately replace the hand-crafted structuring of the code by learnt frameworks. This Team Description Paper discusses the aspects the Bold Hearts team currently concentrates upon.

1 Base and Locomotion

Team Bold Hearts has competed in the RoboCup Soccer Simulation league since 2003. The first two years the team participated in the 2D competitions, in 2005 it joined the 3D community. At the beginning of 2009 a full restart of the team was initiated, after attracting Sander van Dijk to the team, former member of the successful team Little Green BATS¹. To get the Bold Hearts up to steam quickly, the new code is based on the libbats library released by the Little Green BATS².

A simple open loop oscillator model, based on [4], and similar to that released by the Little Green BATS, is momentarily used as a gait generator. This already results in a fast walking behaviour, however the lack of using any sensory feedback leaves open several possibilities for enhancing stability.

Firstly, we attempt to handle sideways stability. In [4] a method is presented to couple the phase of the oscillator to the natural phase of the robots dynamics. The underlying idea is that this natural phase ϕ_r can be inferred from the current location of the Centre Of Pressure (COP), x_{COP} , and its velocity, \dot{x}_{COP} :

$$\phi_r = \arctan\left(\frac{\dot{x}_{COP}}{x_{COP}}\right). \quad (1)$$

¹ See <http://www.littlegreenbats.nl/>

² See <http://www.launchpad.net/littlegreenbats/>

This is then combined with the oscillators static rate of change ω_c to determine the rate of change of the phase of the walking gait:

$$\dot{\phi}_c = \omega_c + K_c \sin(\phi_r - \phi_c), \quad (2)$$

where K_c is a coupling constant, which determines the strength of the coupling. When set to 0, the model reduces to a fully open loop controller.

Initial experiments have shown that the Force Resistance Preceptor used in the simulation has poor resolution, affecting the usability of the COP measure and thus the effectiveness of the phase coupling. To overcome this, we study the effect of different stability measures to replace the COP in (1), most notably the Zero-Rate of Angular Momentum (ZRAM) point, introduced by [2]. This measure takes into account additional postural information, resulting in a more detailed stability indication. Experiments are currently being done to test the usefulness of the ZRAM-point for phase coupling quantitatively.

Besides lateral stability, another important issue to handle is anterior/posterior stability. Currently, a basic method is used to lean the torso in the direction of movement to counteract angular momentum caused by the walking gait:

$$f = \min\left(\frac{v_y(t)}{v_{max}} + \alpha a_y(t), 1\right) \quad (3)$$

$$\theta(t) = f \cdot \theta_{max}, \quad (4)$$

where $v_y(t)$ is the agents desired forward velocity at time t , v_{max} its maximum velocity, $a_y(t)$ its desired forward acceleration, $\theta(t)$ the added torso pitch and θ_{max} the maximum torso pitch. Efforts are under way to make this method more adaptive and effective, using reinforcement learning methods and stability measures like the ZRAM-point.

2 Gait Optimization

The parameters of the walking gait have been optimized with a new optimization algorithm called the Alliance Algorithm (AA) [1]. The AA is based on the metaphoric idea that a certain number of tribes struggle to conquer an environment, which offers resources to survive. Two features characterize each tribe: the strength and the resources necessary for its survival. In order to increase their strength, the tribes may enter into alliances. Also an alliance is characterized by a strength and an amount of necessary resources, these two values depend on the features that characterize the tribes inside the alliance. The execution of the algorithm is based on the formation of alliances. The algorithm will end when the strongest alliance will be created: it will be able to conquer the environment, to meet its resource demand and force other alliances out of the environment. It is possible to perform several iterations of the algorithm starting with tribes that are influenced by the previous strongest alliance.

In this scenario there are different entities that play a fundamental role:

1. A tribe is a tuple composed of: a point in the solution space, a strength, a resource demand and an alliance
2. The set of alliances is a mutually disjoint partition of tribes. Each alliance represents a point in the solution space, defined by the tribes that compose the alliance
3. The strength of tribe/alliance is the value obtained with the objective function with the respective solution
4. The amount of resources needed by a tribe/alliance is the value obtained with the constraint function on the respective solution

The tribes that compose the strongest alliance represent the solution of the problem.

In experiments performed with this algorithm, the objective function is the distance traversed by an agent in a fixed amount of time. With the Alliance Algorithm a solution was found that achieved a walking speed of 0.93 m/s, an increase of 36% over the gait used by our team in 2010. Also, the algorithm found such solutions significantly faster than the genetic algorithm that was used and optimized by our team over the last few years. Full results have been submitted to the RoboCup 2011 Symposium. Currently work is done to develop new gait generators that can be optimized further with the AA.

3 Localization and Active Vision Through Infotaxis

One of the challenges introduced in 2009 for the Robocup 3D Simulation teams were the restrictions placed on the vision sensor. The previous two years the simulated robots were equipped with perfect 360-degree omni vision cameras, making the environment fully accessible. From this year, however, a restricted vision sensor is introduced, similar to that used in the spheres version of the simulator until 2006. This sensor has a range of 120 degrees on both the horizontal as the vertical axis and supplies noisy data about the objects within its field of sight. The next sections describe methods and directions of research we use to handle this new challenge.

3.1 Localization

We supply the agents with a localization mechanism that maintains their global location in world coordinates. Many tasks can be achieved with only relative position information, for instance to kick the ball into a goal the relative position of the agent to the ball and to the goal is sufficient. Global coordinates however make it easier for the agent to deduce more about the world, like the trajectory of the ball and whether the team is in an attack or defence situation. In this section we will describe the Kalman filter localization method, a traditional method used to solve prediction problems, as described in [3] and [7].

With this method, the agents estimated location is described by a multivariate normal distribution $N(\mathbf{x}, \Sigma)$ with means \mathbf{x} , here a 6-dimensional vector

depicting the agents 3D coordinates and velocity, and covariance matrix Σ . After each time step this estimate is refined in two steps: first a prediction is made based on the dynamics of the environment and the agents actions, secondly this prediction is updated by integrating observations.

Predict In the prediction step at timestep k the mean $\mathbf{x}_{k|k1}$ and covariance matrix $\Sigma_{k|k1}$, where $(\cdot)_{k|l}$ means at timestep k , given all observations up to and including timestep l , are determined as follows:

$$\mathbf{x}_{k|k1} = \mathbf{A}\mathbf{x}_{k1|k1} + \mathbf{B}\mathbf{u}_{k1} \quad (5)$$

$$\Sigma_{k|k1} = \mathbf{A}\Sigma_{k1|k1}\mathbf{A}^T + \mathbf{Q} \quad (6)$$

where \mathbf{A} is the state transition model relating the state of the previous timestep to that of the current timestep, \mathbf{u}_k is the control vector at timestep k reflecting the agents actions, and \mathbf{Q} is the process noise.

Update The update step uses observations of landmarks at the current timestep, \mathbf{z}_k , to refine the estimate:

$$\mathbf{K}_k = \Sigma_{k|k1}\mathbf{H}^T(\mathbf{H}\Sigma_{k|k1}\mathbf{H}^T + \mathbf{R}_k)^{-1} \quad (7)$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\mathbf{x}_{k|k1}) \quad (8)$$

$$\Sigma_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\Sigma_{k|k1} \quad (9)$$

where \mathbf{H} is the observation model relating an observation to a location world coordinates and \mathbf{R}_k is the observation noise covariance matrix. The matrix \mathbf{K}_k is the gain or blending factor that minimizes the a posteriori error covariance. Note that the observation noise model depends on the current timestep, since the noise when observing far away landmarks is larger than with nearer objects.

4 Single-Landmark Observations

An observation consists of agent coordinates in the global coordinate system. These can be obtained through triangulation or trilateration of the observed locations of several landmarks. However, the gyroscopic sensor of the agent gives sufficient information to achieve an observation by using only a single landmark. To achieve this, we maintain the current rotation matrix \mathbf{T}_k of the agent relative to the field. After beaming we set $\mathbf{T}_k = \mathbf{I}$. In all subsequent timesteps we update the matrix using the angular velocity measurement $\dot{\theta}_k$, given by the gyroscopic sensor. Firstly, the previous rotation estimate is used to transform this measurement from the local to the global coordinate frame:

$$\dot{\theta}'_k = \mathbf{T}_{k-1}\dot{\theta}_k \quad (10)$$

Based on this a rotation matrix Θ_k is constructed, describing the rotation since the last time step in the global coordinate frame. Finally, this matrix is used to obtain the new global rotation matrix estimate:

$$\mathbf{T}_k = \Theta_k\mathbf{T}_{k-1} \quad (11)$$

With this new estimate a local observation can be transformed into a global location observation, enabling accurate localization based on a single landmark.

4.1 Information Gathering

During a match an agent that focusses solely on the ball will receive sufficient observations of landmarks to be able to localize effectively using the method described in the previous section. However, there are more interesting objects in the field. Especially with the current increase in team size, coordination and keeping track of the opponents players becomes an important issue, hindered by the fact that an agent can only pay attention to a small part of the field at the same time. Therefore we search for active vision strategies which optimize the amount of useful information gathered by the agent. To do this we will use the infotaxis strategy which locally maximizes the expected rate of information gain[6]. The information gain resulting from an observation can be measured by the decrease of the entropy $H(f)$ of the distribution $f(\mathbf{x})$. In our case of multivariate normal distribution we have:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \quad (12)$$

$$H(f) = \int_{-\infty}^{\infty} f(\mathbf{x}) \log f(\mathbf{x}) d\mathbf{x} \quad (13)$$

$$= \log\left(\sqrt{(2\pi e)^N |\boldsymbol{\Sigma}|}\right), \quad (14)$$

where N is the number of dimensions, in our case $N = 6$. The problem we need to solve is which action $a \in \mathcal{A}$ of the possible actions \mathcal{A} to take to maximize the decrease in entropy:

$$a_k = \arg \max_a E[\Delta H(f)|a] \quad (15)$$

$$= \arg \min_a E[H(f)_{k+1} H(f)_k |a] \quad (16)$$

$$= \arg \min_a E[H(f)_{k+1} |a] \quad (17)$$

$$= \arg \min_a E[\log\left(\sqrt{(2\pi e)^N |\boldsymbol{\Sigma}_{k+1|k+1}|}\right) |a] \quad (18)$$

$$= \arg \min_a E[|\boldsymbol{\Sigma}_{k+1|k+1}| |a] \quad (19)$$

$$= \arg \min_a E[\mathbf{R}_{\mathbf{k}+1|\mathbf{a}} |a], \quad (20)$$

where $E[X|y]$ denotes the expected value of X given y . This basically means that the agent should take the action that results in an observation with the highest accuracy, which can be estimated with the believe of current state state.

4.2 Ongoing Research

There are several ways to continue from here and multiple problems we are or will be researching. Firstly, the choice of set of actions A is important to get the best results. If for instance it consists of turn head n degrees left/right the agent may focus on a single set of landmarks, unwilling to sweep over empty areas, even though that may lead to observing better landmarks. Secondly, the current model of infotaxis assumes that all information is equally valuable. However, in football this is not the case; information about the location of the ball may be worth more than where your keeper is. Moreover, the relative value of different types of information could change during a game. A tradeoff has to be made to decide on which target to focus, e.g. by limiting A to actions that will not lose sight of the ball or by alternating between the targets based on the current value of the information about each to the agent. To optimize the latter case we will use another information theoretical principle, relevant information, which measures the amount of information present in a random variable that is relevant for an agent's optimal strategy [5]. This amount gives an indication which variable should get more attention.

5 Team Coordination

With growing team sizes, the importance of the coordination of the agents' positions increases. Due to its simplicity we have opted for a centralised system with the keeper responsible for coordinating the remainder of the team.

Initially, each agent is assigned a position according to his uniform number. As the game evolves, agents may drift from their assigned position when executing certain tasks, such as attacking the opponents goal. Our system allows the keeper to reassign each agent so that it assumes the position closest to it. This allows agent to swap positions dynamically in response to changing game developments, closing gaps in the formation that would have otherwise been left open until the departed agent returns.

Further work is planned to explore other models of coordination, such as distributed systems, and to develop responsiveness to game situations (e.g. winning/losing, corner-kicks).

References

1. V. Calderaro, V. Galdi, V. Lattarulo, and P. Siano. A new algorithm for steady state load-shedding strategy. In *12th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, pages 48–53, 2010.
2. Ambarish Goswami and Vinutha Kallem. Rate of change of angular momentum and balance maintenance of biped robots. *International Conference on Robotics and Automation*, 2004.
3. Kalman, Rudolph, and Emil. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.

4. J. Morimoto, G. Endo, J. Nakanishi, S. Hyon, G. Cheng, and D. Bentevegna. Modulation of simple sinusoidal patterns by a coupled oscillator model for biped walking. In *Proceedings of the 2006 IEEE international conference on robotics and automation*, pages 1579–1584, 2006.
5. Daniel Polani, Thomas Martinetz, and Jan T. Kim. An information-theoretic approach for the quantification of relevance. In *ECAL '01: Proceedings of the 6th European Conference on Advances in Artificial Life*, pages 704–713, London, UK, 2001. Springer-Verlag.
6. Massimo Vergassola, Emmanuel Villermaux, and Boris I. Shraiman. 'infotaxis' as a strategy for searching without gradients. *Nature*, 445:406–409, 2007.
7. Greg Welch and Gary Bishop. An introduction to the kalman filter. Technical report, Chapel Hill, NC, USA, 1995.