Proceedings of the 10th International Workshop on
Robotics in Alpe-Adria-Danube Region (RAAD'01)
Vienna, Austria, May 2001 • CD-ROM RD075

Investigations into Goal-Oriented Vision-Based Walking of a Biped Humanoid Robot

G. Schmidt, O. Lorch, J. F. Seara, J. Denk, and F. Freyberger

Institute of Automatic Control Engineering,
Technische Universität München,
D-80290 Munich, Germany
ViGWaM@lsr.ei.tum.de

Abstract: Goal-oriented vision-based legged locomotion requires a high degree of coordination between perception and walking. The way how this coordination can be established remains a fundamental and rarely studied problem in legged robotics. Our investigations into this field are outlined in this article by presenting recent results in three major research directions. The approaches developed are experimentally validated in a rapid prototyping environment comprising real vision components and a dynamically simulated walking machine. It is shown how perception results are employed for step sequence adaptation in the closed-loop controlled walking machine visualized in an augmented reality display.

1 Introduction and Research Directions

Numerous research groups worldwide concentrate on the design and stabilization of humanoid walking machines. Others deal with the development of vision algorithms for environmental perception. A combination of both research fields seems to be rare; some ideas in this direction have been reported by [6, 2]. Most walking machines still lack the ability of intentional, perception-based goal-oriented locomotion behavior when compared to biological or human walking systems.

Our research is directed to the interaction between vision and biped walking. To support studies on this topic, we have developed an upgraded emulation environment presented on RAAD'99 [3]. This set-up allows the validation and extrapolation of our vision-based control approaches for biped walking. Data from real visual perception and image processing are coupled with a **vi**sion-**g**uided virtual **wa**lking **m**achine — ViGWaM. Its operation is visualized in an augmented reality display, where an image of an external camera observing the real scenario is overlayed by a virtual 3D graphic visualization of ViGWaM.

Investigations carried out in our group can be subdivided in three major directions: (i) development of strategies for *walking pattern synthesis and concatenation* to a stable step sequence to allow goal-oriented walking; (ii) algorithms for robust *image processing* and obstacle reconstruction considering the head motion in 6 DOFs during walking; (iii) biologically motivated studies on intention based *gaze control* with the objective to increase visual information in a task and goal dependent way.

The article is organized as follows: Section 2 summarizes our approach to the synthesis and concatenation of various walking patterns. The cooperation between image processing and gaze control is outlined in Section 3. Results obtained in the emulation environment as explained in Section 4 will be reported in Section 5.

2 Walking Pattern Synthesis and Concatenation

Autonomous motion of a walking machine requires the ability to gait adaptation with respect to obstacles. In order to avoid unstable abrupt gait pattern modifications and to walk smoothly, predictive, real-time path and step planning algorithms are necessary. In addition, the disturbances of the environment perception sensor as caused by walking motion must be kept small.

Usually the problem of executing a defined step with a biped while simultaneously ensuring its stability is handled by preplanning of suitable trajectories for all joints. To the best of our knowledge it is not possible to accomplish this task in real-time for a 3D-biped. We solve this synthesis problem offline by optimization of individual walking primitives for level steps with different step lengths, for stair climbing and for striding over an obstacle. These primitives are stored in a database and can be exported on demand. The walking primitives are computed in such a way that ViGWaM remains statically stable at each point of the path. Dynamic stability [5] is assured by executing the statically stable walking primitives with low speed. During online execution a step sequence planner appropriately selects and concatenates the stored walking primitives to a walking pattern, step sequence resp., using the sensor information about the type of the next obstacle and its relative distance. More details on the optimization and concatenation problem can be found in [4].

As a result of the preplanned primitives, the discrete path of the head in 6 DOFs during walking, here with $N = 21$ sampling points, can be obtained, see Figure 1. Head position is given by $[x_H\ y_H\ z_H]^T$ and head orientation by roll-pitch-yaw angles $[\alpha_H\ \beta_H\ \gamma_H]^T$ with respect to the coordinate system in Figure 3. With a spline interpolated continuous path and a fixed execution speed, a trajectory for the head motion can be generated and emulated in our experimental set-up.

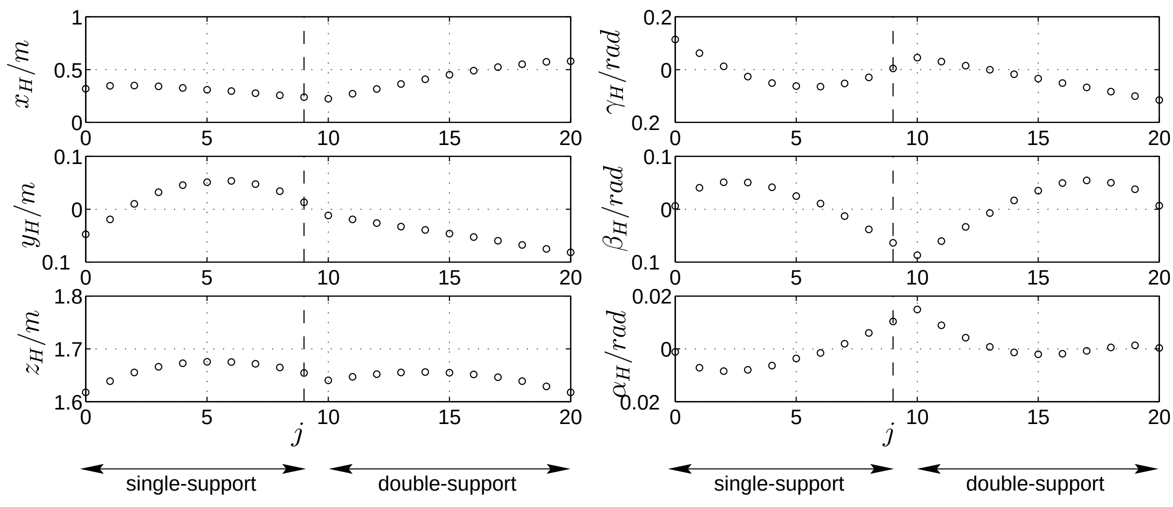

Figure 1: Path of the head of the walking machine for a single walking primitive.

3 Image Processing and Gaze Control

Walking in an environment with obstacles requires a reliable, predictive classification and pose estimation of all obstacles in the walking trail. On the other hand, the resulting head and camera motions during walking, see Figure 1, may cause severe problems for image processing algorithms. Furthermore, online perception capabilities are essential for performing a smooth step sequence adaptation where image processing time is a critical factor. On one hand, image analysis for scene interpretation is computationally expensive and cannot be performed in real-time. Tracking algorithms, on the other hand, are fast, but have less interpretative potential for object recognition. To merge the advantages of both approaches, a cascaded image processing architecture for "vision for action" has been developed. Figure 2 shows the proposed architecture with the image processing blocks and the gaze control blocks highlighted. In the outer loop for static scene analysis, online capability is not required. Here, static images, or just regions in them, are analyzed slower than camera frame rate. The results are integrated into a 3D-map of the walking environment. All relevant data of the objects are stored in this map: type and size of the obstacle, pose in the world, visibility, and estimated accuracy, for details see [3].

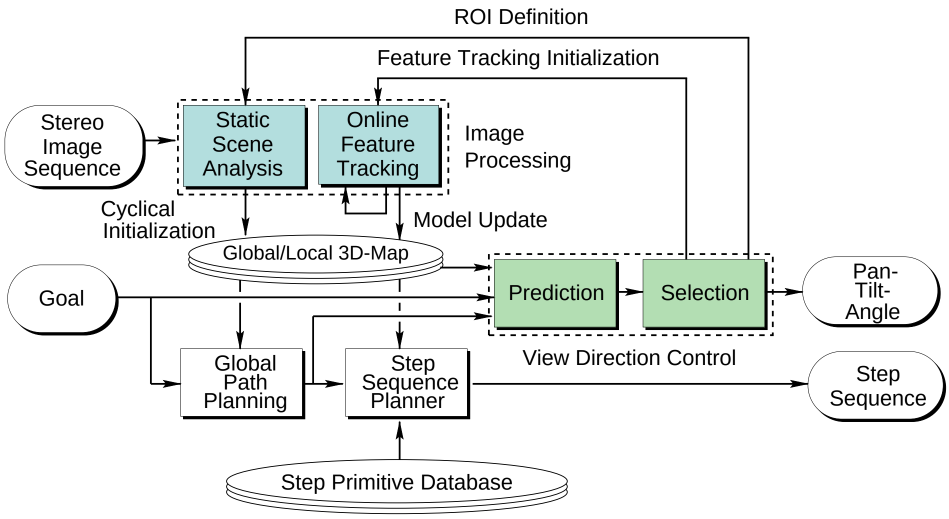

Figure 2: Cascaded vision architecture.

For safe walking, it suffices to know the pose of the obstacle relative to the walking machine, given that information on the type and size of the obstacle is provided by the map. Thus, the representation of a set of obstacles considered in the image stream can be reduced to a single feature representing the individual object in the image plane, e.g. a single border line at the front of a barrier. Such features are tracked online in the image stream. Their 3D-reconstruction can be performed with common reconstruction algorithms in camera frame rate which is explained in more detail in [4]. With this architecture the time-consuming scene interpretation and feature tracking can be performed simultaneously in the distributed computer network. The corresponding results are fused via the 3D-map.

Furthermore, goal-oriented locomotion causes variations in the field of view. The result is that new regions are mapped on the vision sensor while already analyzed features will disappear. This makes the two processes, image analysis and tracking, necessary. To compensate the unintended variations in the field of view and for maximization of visual information during walking, the sensor gaze is adapted. A gaze controller changes the view direction depending on the current goal of the walking machine and accomplishes the task to determine "where to look next?". This gaze controller manages the determination

of new appearing regions and the reallocation of already tracked features temporarily not detectable (**R**egion-**O**f-**I**nterest-Definition), see Figure 2.
Analysis of human gaze behaviour during walking has shown that there is a predictive component in the way visual information is used in locomotion [1]. With this biological inspiration the gaze controller makes predictive use of the available information about the scene accumulated in the 3D-map, see Figure 2. To predict the optimal orientation of the visual sensor, the current walking machine location and actual motion parameters are used. Among several predicted sensor orientations, the optimal is selected.

4 Emulation of ViGWaM

In lack of physical walking machine hardware we proposed a novel alternative for experimental validation of our approaches [3]. The upgraded hardware-in-the-loop emulation environment (a new perpendicularly mounted 3 DOF SCARA-type robot arm) with real vision components and the vision guided virtual walking machine — ViGWaM, is shown in Figure 3. This configuration allows the emulation of the head motion in 6 DOFs, see Figure 1. The emulator architecture comprises the vision controller, step planner, ViGWaM simulation, and the augmented reality visualization subsystem.

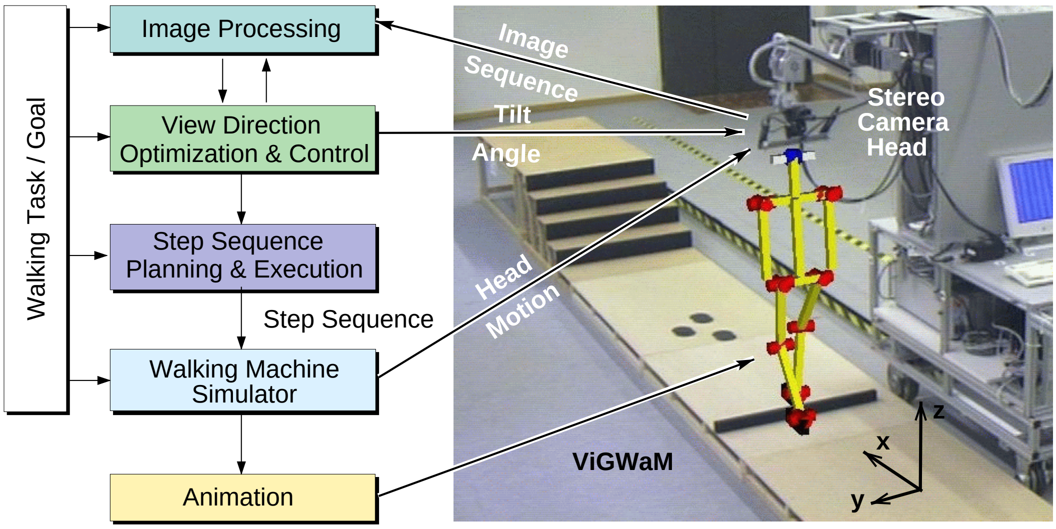

Figure 3: Emulator architecture and experimental set-up with ViGWaM overlayed.

5 Experimental Results

To illustrate some of the benefits of our approach, we describe an experiment comprising a sequence of 3 prototypical obstacles, a barrier, a set of three footprints to be stepped on, and a stairway. The type and dimensions of the obstacles are known *a priori* and specified in a map. The distances between the obstacles and the starting point of the left foot are assumed to be known *a priori* with an error of $\pm 0.05\,m$.
In the upper part of Figure 4 the step sequence resulting from the experiment is compared to a nominal step sequence with a constant step-length l_n in the lower part. One can clearly see, that without step sequence planning striding over the bar, stepping into the footprints and climbing the stairs would have failed.

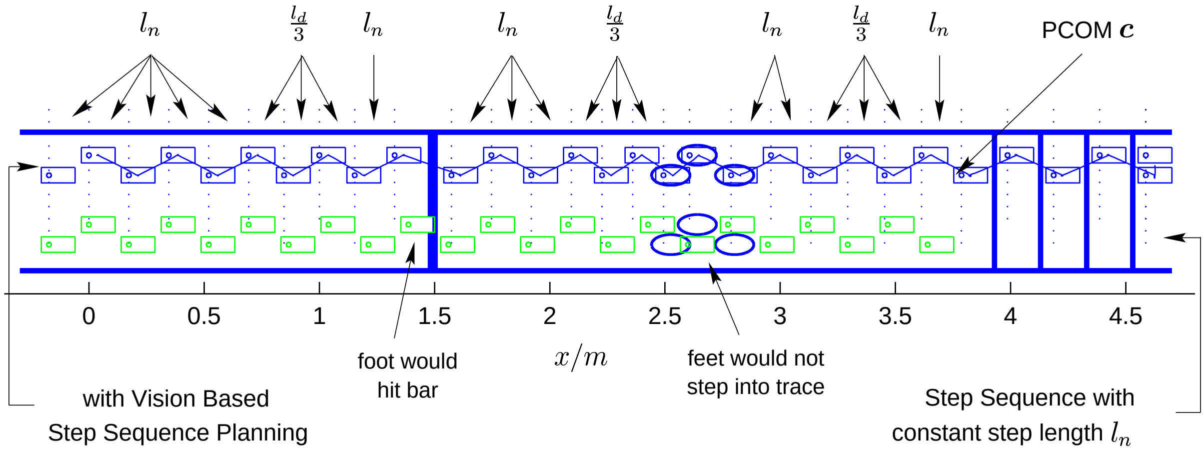

Figure 4: Resulting step sequence with and without vision-based step length adaptation and progression of the projected center of mass $\boldsymbol{c}$.

Figure 5 indicates, how ViGWaM strides safely over the barrier, into the footprints, and climbes the stairway employing the vision based step sequence adaptation method. Figure 5 (a,b) shows the approach of ViGWaM to the front of the barrier with different step lengths. In (c) the barrier is surmounted with high clearance steps. After that the steps for the three footprints are planned (d). Figure 5 (e-g) depicts the locomotion on the step trace of the footprints. The stairway approach is shown in Figure 5 (h,i) and Figure 5 (j-l) demonstrates the successful stair climbing of ViGWaM.

6 Conclusions

This article presents an approach to goal-oriented vision-based walking. Key research directions such as walking pattern synthesis, step sequence planning, gaze control strategies, and development of robust image processing algorithms are identified and described. For solving the walking task, the main ideas on each research direction have been pointed out without describing further details. The use of preplanned variable joint trajectories for walking gained by solving an optimization problem and online concatenating primitives to a step sequence, have been outlined. Furthermore the cascaded vision control architecture to support online perception of the environment with use of a predictive gaze controller has been described.
An emulation environment is employed to validate the vision-based step sequence planning strategies, based on reliable image processing algorithms, and walking pattern concatenation to guide ViGWaM in a modular prototypical scenario including three different types of obstacles. The walk of ViGWaM over the path demonstrates the performance of the approach chosen.
Future work will focus on dynamically stable walking to achieve smoother and faster machine walking, and extensions of the image processing algorithms as well as the gaze (view direction) control behavior to more general environmental situations.

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

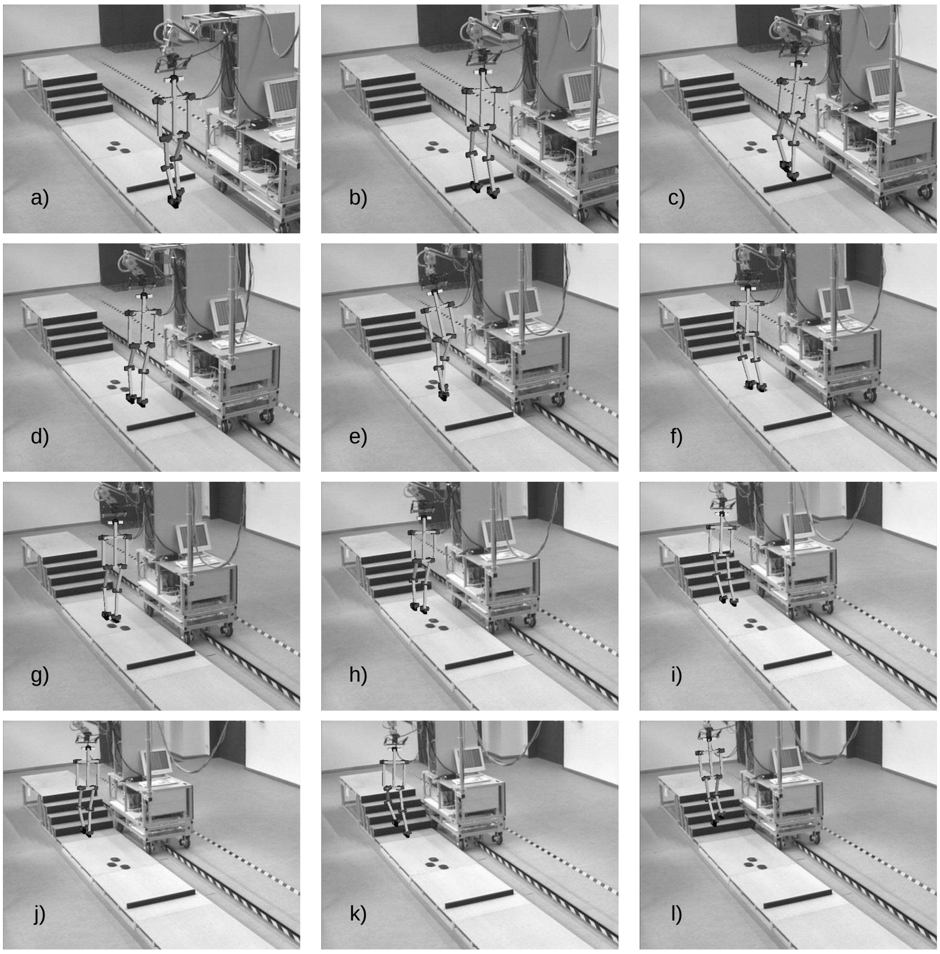

Figure 5: ViGWaM walking over the scenario with three basic types of obstacles, a barrier, footprints (step trace), and stairs.

References

[1] Aftab E. Patla. Understanding of roles of vision in the control of human locomotion. *Gait and Posture*, 5:54–69, 1997. 4

[2] Daniel Pack. Perception–Based Control for a Quadruped Walking Robot. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2994–3001, Minneapolis, Minnesota, 1996. 1

[3] O. Lorch, M. Buss, and G. Schmidt. An Emulation Environment for the Development of a Vision-Based Walking Robot. In *Proceedings of the 8th International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD'99)*, pages 63–68, München, Germany, June 1999. 1, 3, 4

[4] O. Lorch, J. Denk, J. F. Seara, M. Buss, F. Freyberger, and G. Schmidt. ViGWaM — An Emulation Environment for a Vision Guided Virtual Walking Machine. In *Proceedings of the IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Cambridge, Massachusetts, USA, September 2000. 2, 3

[5] M. Vukobratović and M. Kirćanski. *Kinematics and Trajectory Synthesis of Manipulation Robots*. Springer-Verlag, Berlin, Germany, 1 edition, 1986. 2

[6] M. Yagi and V. Lumelsky. Biped Robot Locomotion in Scenes with Unknown Obstacles. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 375–380, Detroit, Michigan, 1999. 1