

Vision Guided Biped Walking: Step Sequence Planning, Visual Perception and Feedback

J DENK, R CUPEC, J F SEARA, O LORCH, and G SCHMIDT

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

ABSTRACT

This paper presents key ideas of our approach to perception guided robot walking. Results of computer vision are employed for reactive step sequence adaptation. Our approach and the guidance system were tested and evaluated in experiments with the biped robot BART-UH [4] developed at IRT of the Universität Hannover.

1 INTRODUCTION

The research of our group is directed to vision guided autonomous biped locomotion. For this purpose we investigated the task of enabling a biped robot to reach a goal area in a scenario with prototypical obstacles [3, 9, 12]. Only little related research has been conducted so far. Experiments in perception-guided biped walking have been described in [14] using vision and in [6] using an ultrasonic range sensor for environmental perception. Newer important investigations are simulations of biped walking in scenarios with obstacles assuming a known environmental model [7] or a virtual sensor [15]. Recently, videos have been presented, which show the robot H7 walking around obstacles guided by stereo cameras [5].

To support our studies, a stand-alone vision-based guidance system for walking machines has been developed. Information about obstacles in the walking trail is gathered using a line-based stereo vision algorithm for visual feedback. This information enables a reactive step sequence planner to determine the parameters of an adequate step sequence to stride over the obstacles. In recent experiments we validated our approach by utilizing the guidance system as supervisory controller for BART-UH, a biped robot developed at IRT of the Universität Hannover, Germany [1, 4].

2 CONCEPT OF VISION-BASED GUIDANCE

For performing a given locomotion task, the guidance system of a walking machine must provide the robot with the parameters of an appropriate step sequence. The sequence must allow

the robot to reach a goal position taking into account obstacles in the walking trail. To select appropriate steps for overcoming obstacles, the reactive *step sequence planner* needs sufficiently accurate information about obstacle locations relative to the walking machine as well as obstacle dimensions.

In the guidance strategy described in this article, obstacle information is supplied by *visual feedback*. The architecture of the visual-based guidance system is shown in Fig. 1.

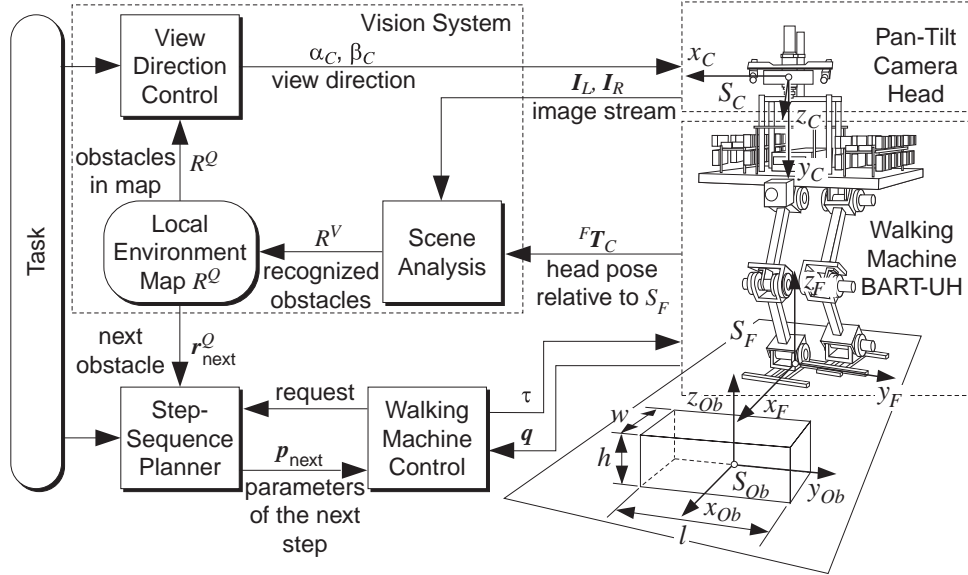


Figure 1. Architecture of the vision-based guidance system.

The environment of the walking machine is perceived by a stereo-camera pair mounted on a pan-tilt head. The resulting stream of camera images I_L and I_R represents the input to a *scene analysis algorithm*. The algorithm recognizes rectangular objects laying on the walking trail and estimates their size and pose relative to the coordinate system S_F which is fixed in the center of the current standing foot of the walking machine. Object recognition and visual estimation is performed once per step using the pose of the camera reference frame S_C relative to S_F . The necessary transformation ${}^F T_C$ between these two coordinate systems is obtained by direct kinematics and depends on the joint angles of the robot.

Each object recognized in the field of view of the biped is represented by a vector $\mathbf{r}_i^V = [c_i^V \ x_i^V \ y_i^V \ \alpha_i^V \ w_i^V \ h_i^V \ l_i^V]^T$, where c_i^V denotes the class of object; x_i^V , y_i^V and α_i^V represent the position and orientation of the object reference frame S_{Ob} relative to S_F and w_i^V , h_i^V and l_i^V are object dimensions, cf. Fig. 1.

Since the field of view is limited, the camera system should be directed in such a way, that the currently most relevant objects in the walking trail are visible. This is obtained by a *gaze controller*, which selects the pan and tilt angle α_C and β_C of the camera system corresponding to the view direction with the maximal visual information content [12].

The set of objects $R^V = \{r_1^V, r_2^V, \dots, r_n^V\}$ recognized in one pair of images I_L and I_R is used for updating the *local environment map* R^Q . This map represents the set of all obstacles r_j^Q appearing in the walking trail during one experiment.

The vector r_{next}^Q containing the information about the next obstacle in the walking trail is provided as input to the step sequence planner. Since the space of steps which a robot can execute is restricted by dynamics and mechanical design, only particular step combinations can be executed. This fact has to be regarded by the planning algorithm, when selecting an appropriate step sequence to overcome the obstacle. Since uncertainty in visual estimation decreases as the robot approaches the obstacle, the step sequence is replanned in each step using current visual information. This strategy also allows to react to changes in the obstacle situation. The parameters p_{next} of the next step to be executed are sent to the robot locomotion controller [4].

3 VISUAL FEEDBACK

Different approaches to scene interpretation for walking machines have been reported including depth-map generation [5] and model-based stair recognition [2, 11]. In our approach a model-based object recognition algorithm is adopted, which recognizes separate rectangular objects using line segments extracted from a *single image* together with the information about the orientation of the camera system. Size and position of objects relative to the walking machine are obtained by *stereo vision*.

3.1 Recognition of rectangular objects

Recognition of rectangular objects is based on 2D straight line segments obtained by edge detection and segmentation of edge contours. Hypothetically, each detected 2D line segment $\overline{m_i m_j}$ can represent a projection of a 3D object edge $\overline{x_i x_j}$, cf. Fig. 2. The relation between an

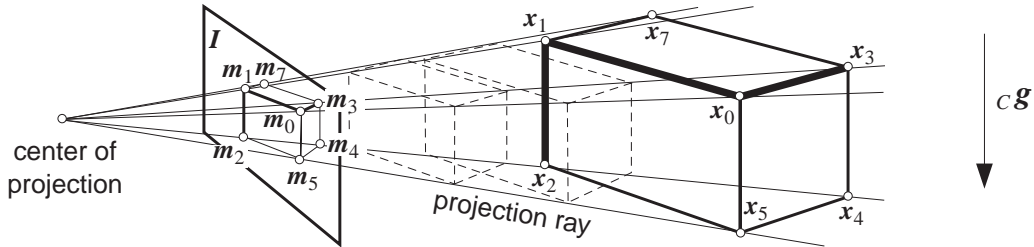


Figure 2. Projection of a 3D object to a camera image I . Groups of orthogonal edges (bold lines) are used as indicators for rectangular objects.

endpoint m_i of a 2D line and the corresponding endpoint ${}_F x_i$ of a 3D edge represented in the foot frame S_F is given by

$$[m_i^T \ 1 \ s_i]^T = s_i \mathbf{P}({}^F T_C)^{-1} [{}_F x_i^T \ 1]^T, \quad (1)$$

where $\mathbf{P} \in \mathbb{R}^{4 \times 4}$ is the homogeneous projection matrix of the camera obtained by camera calibration, s_i is a scaling factor defining the position of ${}_F x_i$ on the projection ray. The trans-

formation ${}^F T_C$ from S_C to S_F is obtained by direct kinematics using the information from the joint encoders.

For each detected 2D line two hypotheses are generated. One hypothesis assumes a 2D line to represent a projection of a *vertical* 3D edge of an object and the other assumes the line to represent a projection of a *horizontal* 3D edge.

A *vertical* 3D edge $\overline{x_i x_j}$ is parallel to the gravity axis represented by the unit vector \mathbf{g} . In that case, the following equation is satisfied

$${}^F \mathbf{x}_j(s_j) - {}^F \mathbf{x}_i(s_i) = \lambda \left\| {}^F \mathbf{x}_j(s_j) - {}^F \mathbf{x}_i(s_i) \right\| {}^F \mathbf{g}, \quad \lambda \in \{-1, 1\}. \quad (2)$$

Since the current standing foot is assumed to reside flat on the ground, ${}^F \mathbf{g}$ is parallel to the z_F -axis of the frame S_F , i. e. ${}^F \mathbf{g} = [0 \ 0 \ -1]^T$.

A *horizontal* 3D edge $\overline{x_i x_j}$ is perpendicular to the gravity axis, i. e.

$$\left({}^F \mathbf{x}_j(s_j) - {}^F \mathbf{x}_i(s_i) \right)^T {}^F \mathbf{g} = 0. \quad (3)$$

The equations (1), (2), and (3) are used to reduce the number of false hypotheses about *vertical* edges [8] and to determine the orientation of *horizontal* edges relative to the frame S_F .

The orientations of the hypothetical 3D edges are used for edge grouping. Chains of three mutually orthogonal edges, cf. Fig. 2, as well as groups of two or three parallel edges are used as indicators for rectangular objects. For each edge group an endpoint ${}^F \mathbf{x}_0(s_0)$ is selected as a reference point and coordinates of all other corner points of the hypothetical object are expressed as functions of s_0 . Hence, the hypothetical object defined by the scaled corner points ${}^F \mathbf{x}_i(s_0)$ actually represents a family of objects corresponding to the same projection, as illustrated in Fig. 2. Verification of the object hypotheses is performed by matching the projections of the edges of a hypothetical object to the 2D lines detected in image I .

3.2 Object reconstruction using stereo vision

The scaling factor s_0 needed for full reconstruction of the object is obtained using stereo vision. The stereo vision algorithm determines the correspondences between lines $\overline{m_i^L m_j^L}$ detected in the left camera image I_L and lines $\overline{m_k^R m_l^R}$ detected in the right camera image I_R . Two corresponding lines represent the projections of a 3D edge to I_L and I_R . The coordinates of the endpoints of the detected 3D edges relative to the frame S_F are obtained by triangulation.

The scaling factor s_0 is estimated by minimizing the sum of squared distances between the corner points ${}^F \mathbf{x}_i(s_0)$ provided by the recognition algorithm and the corresponding points ${}^F \hat{\mathbf{x}}_i$ estimated using triangulation, i. e.

$$\hat{s}_0 = \arg \min_{s_0} \sum_{i=1}^N \left\| {}^F \hat{\mathbf{x}}_i - {}^F \mathbf{x}_i(s_0) \right\|^2, \quad (4)$$

where N is the number of corner points of the object confirmed by stereo vision.

3.3 Recognition of footprints

Besides the rectangular obstacles, the scene analysis algorithm is capable of recognizing elliptical footprints on the ground in which the robot should step. The recognition is performed

by least-squares fitting of ellipses to the contours obtained by edge detection. In the current implementation, the position of a footprint relative to the S_F is determined by intersecting the projection ray corresponding to the center of the footprint with the ground plane.

3.4 Practical considerations

In case of walking robots reliable detection of obstacles and precise estimation of their size and pose relative to the robot's standing foot is essential for correct step sequence planning. Unlike a wheeled robot, however, a walking robot represents a rather long kinematic chain connecting the camera system to the ground. Complex kinematics of a walking robot can cause a substantial uncertainty in determining the transformation ${}^F T_C$. This uncertainty increased by movements of the camera system during walking can degrade the precision of visual estimation.

Furthermore, collisions between the robot's feet and the ground can generate vibrations of the cameras resulting in blurred images. In order to reduce this effect, our vision system triggers the image acquisition at an instant during the step when camera movement is slow. The state during execution of a step corresponding to the slowest head motion is defined by the robot's trajectory planning algorithm [10].

4 REACTIVE STEP SEQUENCE PLANNING

We assume a symmetric biped with a robot control system allowing execution of steps from a discrete set of M feasible walking primitives, cf. Fig. 3, as basis for situation dependent adaptive biped locomotion [3, 7, 9, 15]. For purposes of step sequence planning and to permit fast and efficient collision checking, walking primitive properties are characterized in terms of the vector $\mathbf{p}^\zeta = [l_I^\zeta \ l_{II}^\zeta \ c^\zeta]^T$, $\zeta \in \{1, \dots, M\}$, with l_I and l_{II} denoting step length of the preceding and current step, respectively. Step clearance c and the parameters $\Delta_{b/a}$, which are assumed to be identical for all walking primitives, define the collision free space needed for stepping over an obstacle. For the sake of clarity, the discussion is restricted to level walking in the context of this paper. Furthermore, the mapping from a parameter vector \mathbf{p} to a walking primitive ζ is considered to be one-to-one.

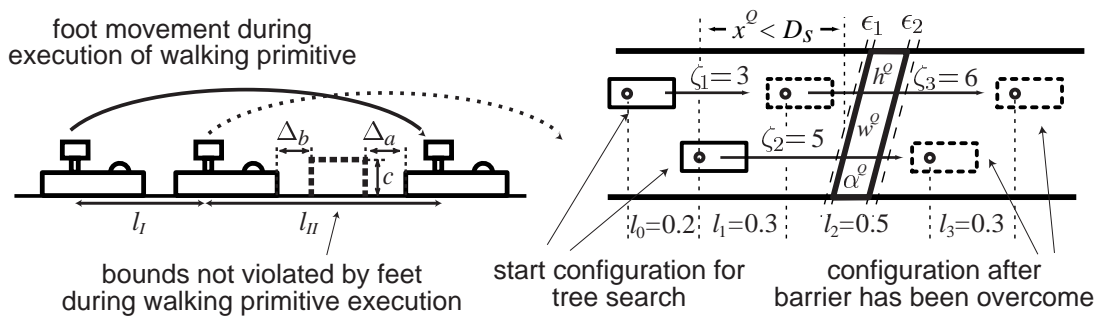


Figure 3. First step over a barrier by execution of a walking primitive defined by step lengths l_I, l_{II} , clearance c and the parameters Δ_b, Δ_a (left). Situation before a barrier, with distance x^Q , height h^Q , width w^Q , angle α^Q , and security regions given by ϵ_1, ϵ_2 . Feasible step sequence $\zeta_1, \zeta_2, \zeta_3$ dashed (right).

During locomotion the set of walking primitives needs to be searched for a step combination allowing to traverse the next obstacle, cp. [7,15], e.g. a barrier as illustrated in Fig. 3. The distance x^Q to the obstacle and other relevant obstacle information are provided by the vision system. Uncertainty in these parameters resulting from limited precision of image processing and dead-reckoning are regarded by introducing security areas ϵ_1, ϵ_2 . Thereby, the obstacle boundaries are virtually enlarged.

For structured representation and a systematic search, the parameters of all walking primitives and the information on their possible combinations are represented by a directed graph. A simplified example is shown in Fig. 4 (left). A node represents the state of the robot after execution of a walking primitive and is labeled with the corresponding step length l . The walking primitives themselves are represented by the edges labeled with the corresponding parameters p^ζ , $\zeta \in \{1, \dots, M\}$ and costs k^ζ as a measure for the efficiency of the step.

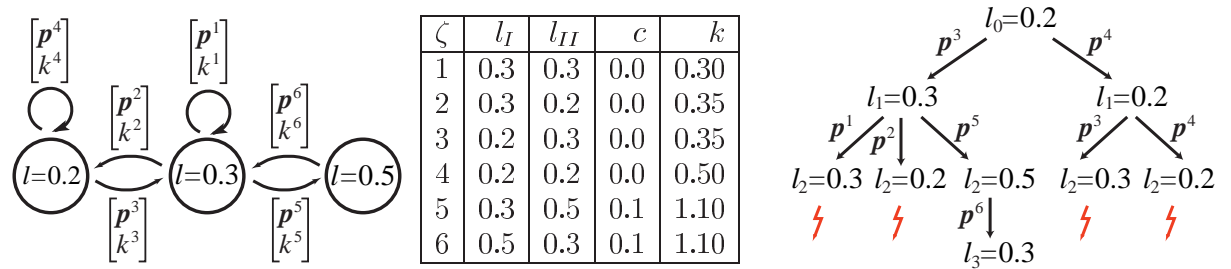


Figure 4. Graph structure defining feasible walking primitives $p^\zeta = [l_I^\zeta \ l_{II}^\zeta \ c^\zeta]^T$, their costs k^ζ and their possible combinations (left). Search tree resulting from depth first search for situation before a barrier illustrated in Fig. 3. Flashes indicate collisions (right).

To determine an appropriate step sequence for overcoming the next obstacle, a search tree is constructed recursively by depth first search based on the information about feasible step combinations stored in the graph, see Fig. 4 (right). The starting node corresponds to the state of the robot at the beginning of the step sequence. Each branch in the tree represents a possible sequence of walking primitives $s = [\zeta_1 \ \zeta_2 \ \dots]^T$. A branch is searched until the obtained step sequence violates the current obstacle situation or until a feasible solution allowing to pass the obstacle has been found.

Usually there exist multiple feasible solutions $s^j = [\zeta_1^j \ \zeta_2^j \ \dots \ \zeta_{N_j}^j]^T$, $j = 1 \dots F$. From this set the solution is chosen, which fits the requirements of the actual task best. In the current implementation the step sequence minimizing the summarized costs k relative to the traveled distance is selected as

$$\nu = \arg \min_{j \in \{1, \dots, F\}} \frac{\sum_{i=1}^{N_j} k_i^j}{\sum_{i=1}^{N_j} \frac{1}{2}(l_{i-1}^j + l_i^j)} \quad , \quad (5)$$

where $l_0^j = l_0$, $j = 1 \dots F$ is the step length corresponding to the node from which the search was started. The costs k of a walking primitive are heuristically defined in such a way, that the execution of walking-primitives with step lengths close to a nominal step length l_n is favored as inspired by human locomotion [9], e.g. by

$$k = \frac{l_I + l_{II}}{2 - \kappa |l_I - l_n| - \kappa |l_{II} - l_n|} \quad , \quad (6)$$

with κ a heuristically chosen constant. Alternatively, costs k based on the electrical energy needed for the execution of a walking primitive might be used to reduce the power consumption of the system.

As long as there is no obstacle in a distance x^Q closer than $D_s \approx 3 l_n$ the biped walks with a nominal step length l_n . Otherwise, walking primitives obtained by a tree search are executed until the obstacle has been passed. A search is conducted once per step using current obstacle information provided by vision. From the resulting step sequence s^ν the parameters $\mathbf{p}_{next} = \mathbf{p}^{\zeta_1^\nu}$ of the first planned primitive are sent to the robot locomotion controller while the parameters of $\zeta_2^\nu, \zeta_3^\nu, \dots, \zeta_{N_\nu}^\nu$ are discarded. This continuous replanning takes into account that results of visual perception are getting more precise the shorter the distance to the obstacle. In addition it allows to react to changes in the obstacle situation.

To achieve continuous and secure walking, the parameters \mathbf{p}_{next} of the next step to be executed must be known before the current step ends. Therefore a search must be aborted, if it is not completed when new parameters \mathbf{p}_{next} are requested by the robot locomotion controller. If feasible solutions have been found, the currently best solution is transmitted. Otherwise, the biped is stopped and the search is restarted allowing a longer search period.

The biped is also stopped if a search is completed without finding a step sequence for overcoming the next obstacle e.g. if the obstacle is too high. The search is restarted periodically and availability of a solution will indicate that the obstacle has been removed. In such a case the biped resumes locomotion demonstrating its reactive behaviour.

5 EXPERIMENTAL RESULTS

To test the suitability and efficiency of our approach the stereo camera head developed at our laboratory was mounted on top of the robot BART-UH. The stereo-vision system comprises two cameras with view-angles of 55° in horizontal and 42° in vertical direction and a stereo baseline of 240 mm. Image resolution is 640×480 pixels and 256 gray values. Details about the robot BART-UH, the communication structure and the experimental setup can be found in [4, 10].

Due to kinematic constraints resulting from overlapping feet, the range of step lengths which BART-UH can execute [1] is decomposed into two continuous intervals $l \in [0.08 \text{ m}, 0.12 \text{ m}]$ and $l \in [0.29 \text{ m}, 0.38 \text{ m}]$, cf. Fig. 5. To adapt the step sequence planner to the requirements of BART-UH, the set of walking primitives used for step sequence planning was obtained by discretization of these intervals with $\Delta l = 0.01 \text{ m}$. The nominal step length was $l_n = 0.35 \text{ m}$.

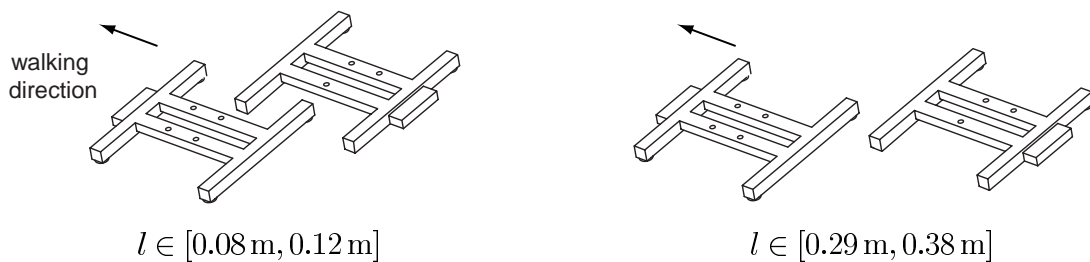


Figure 5. Overlapping feet of BART-UH and intervals of possible step lengths.

The capability of the biped to react to changes in the environment was validated by putting a ribbon in front of the robot during walking, as shown in Fig. 6 (a). When the obstacle came inside the view of the cameras (b), the ribbon was identified as a rectangular obstacle on the ground. The obstacle location and dimensions were estimated and the local environment map was updated. Consequently, the step sequence was adapted allowing BART-UH to successfully overcome the obstacle (c-f).

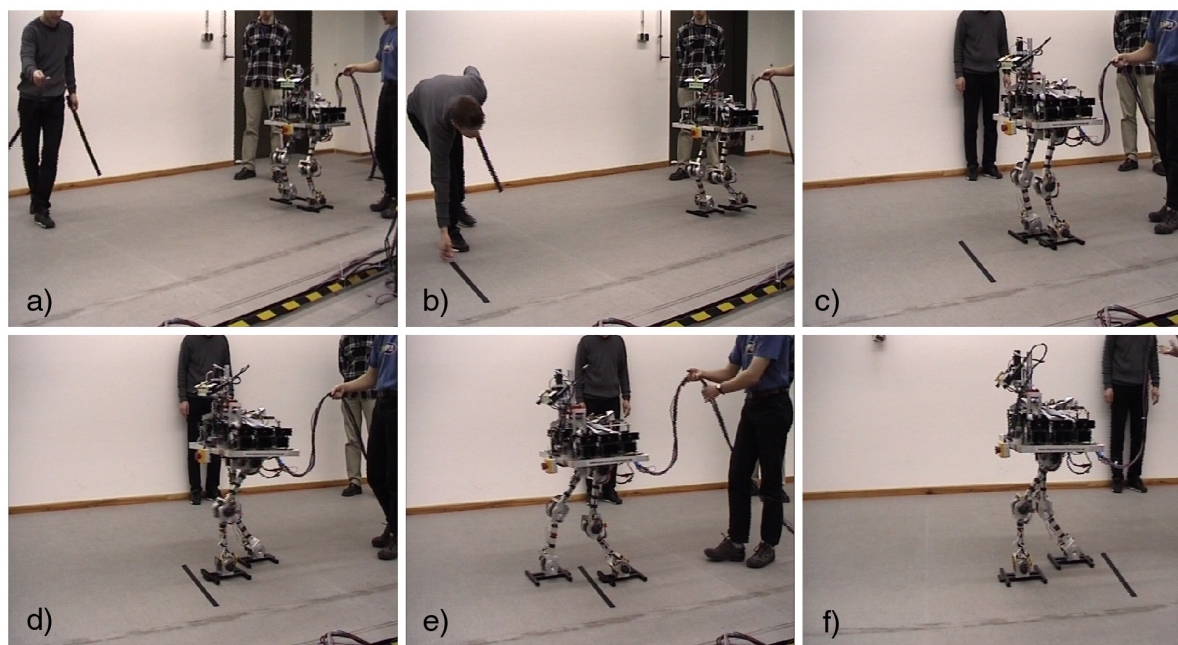


Figure 6. A ribbon-type obstacle is inserted in the walking trail during operation of BART-UH. The obstacle is recognized and localized and the robot adapts its steps overcoming the ribbon in a smooth and stable way.

Another type of obstacle considered in the experiments is a step trace , which is used to examine the capability of the guidance system to cope with combinations of obstacles, see Fig. 7. The task of the robot is to step into the elliptical footprints with subsequent steps. As seen in the image sequence BART-UH reached the first footprint correctly with its left foot (b) and followed the step trace by adapting the next two step lengths appropriately (c,d).

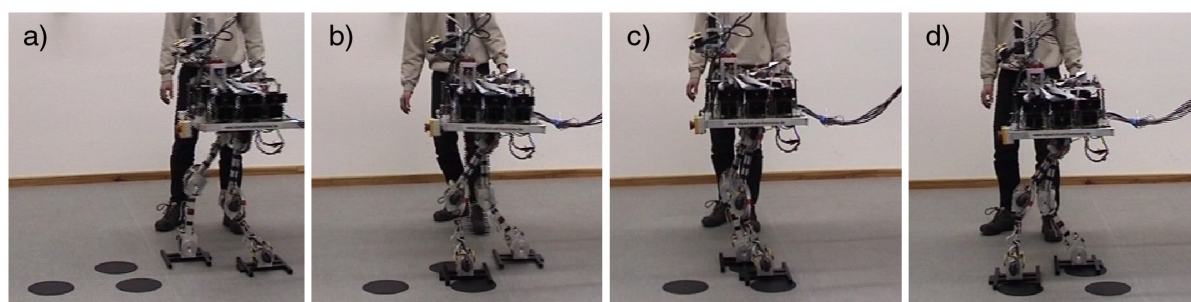


Figure 7. BART approaches a step-trace and places it's feet correctly into the footprints.

The capability of the system to distinguish between obstacles which can be overcome and those which cannot, was demonstrated by placing a box in the walking trail, see Fig. 8. This obstacle was too high for the robot to step over it. When the vision system recognized the box and estimated its dimensions, the step sequence planner could not find a step sequence allowing to overcome the box and the robot stopped (b). As soon as the obstacle was removed from the walking trail, the vision system updated the local environment map. Accordingly, a step sequence guiding the robot towards the goal position was found and BART-UH continued walking (c).

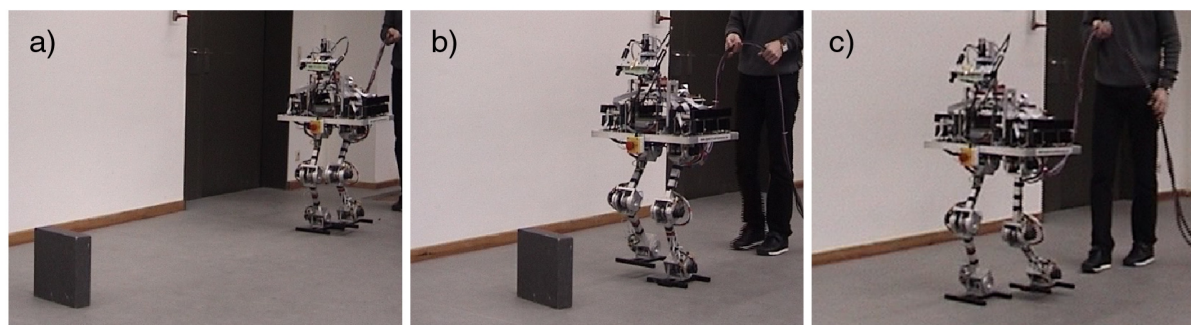


Figure 8. BART-UH approaches a box, which cannot be overcome. The robot stops and resumes locomotion as soon as the obstacle is removed.

6 CONCLUSIONS

In this article an approach to vision-guided robot walking is described. A line-based stereo vision algorithm is applied to provide information about rectangular obstacles. This information is used by a reactive step sequence planner to select appropriate steps allowing the robot to stride over the obstacles in the walking trail.

The presented experiments with the biped robot BART-UH of Universität Hannover were conducted to examine the performance of the developed guidance system. Obstacle recognition and precision of visual estimation were not degraded significantly by the effects of camera movements caused by walking as images were acquired at an instant during step execution when camera movement is slow. The dynamic updating of the local environment map and continuous step sequence planning allowed the robot to successfully overcome obstacles in the walking trail and to react to changes in the environment.

Future research will focus on improvement of the robustness of visual estimation using EKF techniques and advanced gaze control strategies [13]. The algorithms for reactive step sequence planning will be extended to curve walking along a local path and will be tested with walking primitive databases synthesized by optimal control methods [3].

ACKNOWLEDGMENT

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) within the “Autonomous Walking” Priority Research Program.

REFERENCES

1. **Albert, A. and Gerth, W.** (2001) New Path Planning Algorithms for Higher Gait Stability of a Bipedal Robot. Proceedings of the International Conference on Climbing and Walking Robots (CLAWAR 2001), Karlsruhe, Germany, pp. 521–528.
2. **Cupec, R., Lorch, O. and Schmidt, G.** (2001) Object Recognition and Estimation of Camera Orientation for Walking Machines. Tagungsband zum 17. Fachgespräch Autonome Mobile Systeme, Stuttgart, Germany, Springer, pp. 1–10.
3. **Denk, J. and Schmidt, G.** (2001) Synthesis of a Walking Primitive Database for a Humanoid Robot using Optimal Control Techniques. Proceedings of the IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS 2001), Tokyo, Japan, pp. 319–326.
4. **Gerecke, M., Albert, A. and Gerth, W.** (2002) Vision Guided Biped Walking: Trajectories and Communication. International Conference on Climbing and Walking Robots (CLAWAR 2002), Paris, France, accepted.
5. **Kagami, S., Nishiwaki, K., Kuffner, J. J. Jr., Kuniyoshi, Y., Inaba, M. and Inoue, H.** (2001) Design and Implementation of Software Research Platform for Humanoid Robotics: H7. Proceedings of the IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS 2001), Tokyo, Japan, pp. 253–258.
6. **Kajita, S. and Tani, K.** (1996) Adaptive Gait Control of a Biped Robot based on Real-time Sensing of the Ground Profile. Proceedings of the IEEE International Conference on Robotics and Automation, Minneapolis, Minnesota, pp. 570–577.
7. **Kuffner, J. J. Jr., Nishiwaki, K., Kagami, S., Inaba, M. and Inoue, H.** (2001) Footstep Planning Among Obstacles for Biped Robots. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, Hawaii, USA, 500–505.
8. **Lebègue, X. and Aggarwal, J. K.** (1993) Significant Line Segments for an Indoor Mobile Robot. IEEE Trans. on Robotics and Automation, Vol. 9, No. 6, Dec. 1993, pp. 801–815.
9. **Lorch, O., Denk, J., Seara, J. F., Buss, M. and Schmidt, G.** (2000) Coordination of Perception and Locomotion Planning for Goal-Oriented Walking. Proc. of the Int. Conference on Climbing and Walking Robots (CLAWAR 2000), Madrid, Spain, pp. 183–192.
10. **Lorch, O., Albert, A., Denk, J., Gerecke, M., Cupec, R., Seara, J. F., Gerth, W. and Schmidt, G.** (2002) Experiments in Vision-Guided Biped Walking. IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, accepted.
11. **Pack, D. J.** (1996) Perception-Based Control for a Quadruped Walking Robot. Proceedings of the IEEE International Conference on Robotics and Automation, Minneapolis, Minnesota, pp. 2994–3001.
12. **Seara, J. F., Schmidt, G. and Lorch, O.** (2001) ViGWaM Active Vision System—Gaze Control for Goal-Oriented Walking. Proceedings of the International Conference on Climbing and Walking Robots (CLAWAR 2001), Karlsruhe, Germany, pp. 265–272.
13. **Seara, J. F., Strobl, K. H. and Schmidt, G.** (2002) Information Management for Gaze Control in Vision Guided Biped Walking. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Lausanne, Switzerland, accepted.
14. **Shin, D. et al.** (1989) Realization of Obstacle Avoidance by Biped Walking Robot Equipped with Vision System. Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Tsukuba, Japan, pp. 268–275.
15. **Yagi, M. and Lumelsky, V.** (1999) Biped Robot Locomotion in Scenes with Unknown Obstacles. Proceedings of the IEEE International Conference on Robotics and Automation, Detroit, Michigan, pp. 375–380.