

ViGWaM — An Emulation Environment for a Vision Guided Virtual Walking Machine

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

Institute of Automatic Control Engineering, Technische Universität München, D-80290 Munich, Germany
ViGWaM@lsr.ei.tum.de
<http://www.lsr.ei.tum.de>

Abstract. Goal-oriented vision-based walking is a fundamental high-level problem in legged locomotion. In this article, a biologically inspired cascaded control structure to realize smooth machine walking is introduced being applicable to various types of walking machines. An innovative strategy for walking pattern synthesis for adaptive biped locomotion is presented in combination with a step sequence planning algorithm. A guidance system controls the walking process based on environment perception using stereo image processing and gaze control. In the emulation environment, a set of stereo cameras mounted on a pan-tilt-head can be moved over a prototypical scenario synchronized in a feedback loop with a simulated virtual walking machine to achieve realistic perception conditions. The behavior of the controlled virtual walking machine is visualized in an augmented reality display layed over the real image of the experimental setup taken by an external camera.

1 Introduction

Robots are not exclusively used in industrial environments anymore. They are in the process to become an integrated part of the society of the future. Humanoid robots represent a significant step in this evolution towards artificial “creatures”. They are designed to perform tasks in indoor and outdoor environments, typically structured to suit human locomotion requirements. Humanoid robots can be employed in servicing tasks, where close interaction between human and robot is crucial. In addition, they can be applied to various other tasks like inspection in hazardous environments, e.g. in nuclear facilities, disaster areas, or in outer space. In many situations the application of walking machines, in particular biped humanoid robots, seems to be much more appropriate compared to wheeled robotic systems.

Research activities in the past have mainly focused on the design, simulation, stabilization and construction of walking machines [5, 6, 8, 16, 20–22]. Walking machine prototypes were confined to the laboratories until Honda presented the humanoid robot P2 [13, 14], and the new version P3. These humanoid robots show mobility features similar to human locomotion and have the capability to avoid or surmount simple obstacles or to climb stairs. The impressive results in a live demonstration by the Honda robots made many researchers think that some of the major problems of biped walking were solved leaving little room for further research. The approach taken by Honda of recording human walking patterns and using these as input to control the locomotion is innovative, but it does not explain why a given walking pattern is adequate nor does it give any insight into how various behavior primitives can be selected, concatenated, and mixed to realize locomotion across difficult terrain [11]. In addition to these issues related to step and step trace (behavior) planning, another challenging and widely open area of research is the development of (visual) perception methods to achieve goal-oriented smooth and reliable behavior.

Typical walking machines still lack this ability of perception based intentional goal-oriented locomotion behavior when compared to biological or human walking. Even P2, which is equipped with a set of cameras, does not seem to rely on visual information for autonomous navigation, guidance, obstacle detection, and avoidance, or the selection and adaptation of situation-dependent walking patterns. It is noteworthy that past robotics research has shown only minor interest in studying the important relationship between locomotion and vision, although it seems to be a key issue in the development of an intelligent autonomous walking robot. Some ideas in this direction have been reported by [15, 17, 26, 30].

Our research is focused on the relationship between vision and biped walking. Clearly the majority of the results obtained for a biped robot, as an interesting and challenging special case of legged locomotion, can also be transferred to other types of machines. For validation purposes a typical scenario has been considered, including various obstacles, stairs, and a step trace to reflect indoor and outdoor conditions of daily life. Figure 1 shows a vision-based walking robot in such an environment. To support studies on vision-based goal-oriented walking, it is not necessary to work with physical walking machine hardware. We propose a novel emulation environment that allows the extrapolation and validation of our vision-based control approach for biped walking. It also facilitates

the generalization to other, non-biped, walking machines. In this rapid prototyping environment, head and thereby resulting camera motions of a computer simulated walking machine are emulated in a real environment. Data from real visual perception and image processing are coupled with this vision-guided virtual walking machine — ViGWaM. Its operation is visualized in an augmented reality display, where an image of an external camera observing the environment is overlaid with a virtual 3D graphic visualization of ViGWaM.

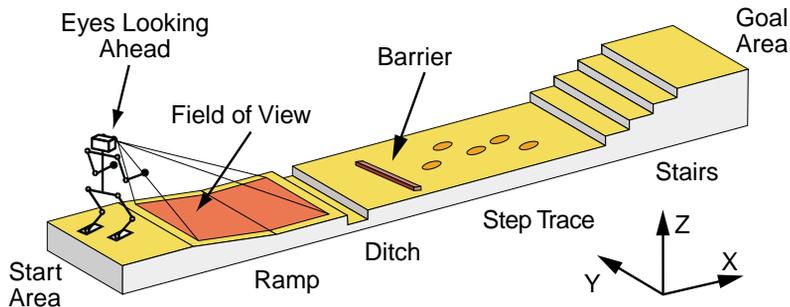


Fig. 1. Vision-based walking robot in prototypical test scenario.

For the organization of the article: Section 2 discusses major research directions and the proposed cascaded control architecture with an emphasis on the biological inspiration of this work. An approach to walking pattern planning for smooth, goal-oriented locomotion is presented in Section 3. In Section 4 an image processing and gaze control approach is proposed. The emulation of the simulated walking motion for hardware-in-the-loop experiments is presented in Section 5. Experimental results in Section 6 demonstrate the advances towards vision-based goal-oriented walking by combining the novel methods presented.

2 Research Directions and Control Approach

Biped walking is a complex problem that has been studied for many years, but the combination with other important biologically inspired planning and perception capabilities, as vision, remains a widely open research field as mentioned above. The study of the relationship between walking and vision is particularly challenging. It is rather complicated to find a starting point for this research when denying the superb example of a human being. Consequently, investigations originating from biology and biomechanics on human vision and walking have been chosen as the basis for our research. However, solutions offered by nature are not always satisfactory from a technical point of view. “Natura non facit saltus”: the evolution must be considered as a very interesting but slow process, whose results are often partial solutions but not optimal [10]. Research must be based on the knowledge of natural facts but its scope must be wider. In our research human nature is not closely imitated but rather taken as an inspiration.

The objective of our research is the development of a vision-guided humanoid walking robot with the capability to walk safely and smoothly through a prototypical environment as shown in Figure 1. This scenario, representing e.g. a sidewalk, contains the most important typical obstacles to be faced when walking in indoor and outdoor environments. For this purpose, a cascaded control structure as shown in Figure 2 is proposed. The two cascaded control loops for navigation (1) and guidance (2) supervise two inner control loops for the walking machine (3) and the view direction (4). The general goal derived from the locomotion mission (5) is decomposed by the global path planner (7) into several subgoals based on *a priori* knowledge stored in a global map (6). The result is a global COM (center of mass) path for the walking machine to reach the goal. The path is then transformed into a step sequence (8) also considering a local vicinity map (9), regularly updated by the vision for action module (10). The planned step sequence is to guarantee the safe motion over various obstacles. The step sequence planner makes use of a data base (11), where the parameters of the different walking patterns and their corresponding joint space trajectories are stored. The step parameters are then fed to the controlled walking machine (12) and the desired steps are executed by means of the control loop for stabilization (3), which affects the COM and the ZMP (zero moment point) and must deal with the disturbances due to the interaction of the walking machine with the environment.

The motion of the walking machine in the environment results in a movement of the head (13). Camera motions caused by head motions can be considered as a disturbance for the visual perception system. That is why, the control loop for view direction stabilization (4) becomes necessary in order to simplify the task of visual perception and the image processing algorithms. The results of the algorithms for scene reconstruction are the decisive information for updating the vicinity map. The information stored in this map (9) is used for step sequence planning (8) and to

control the view direction (14). The view direction controller determines the location for the cameras to point at. Thus, both the information coming from the image processing and the actual intention (5), e.g. self-localization or search for obstacles, should be taken into consideration.

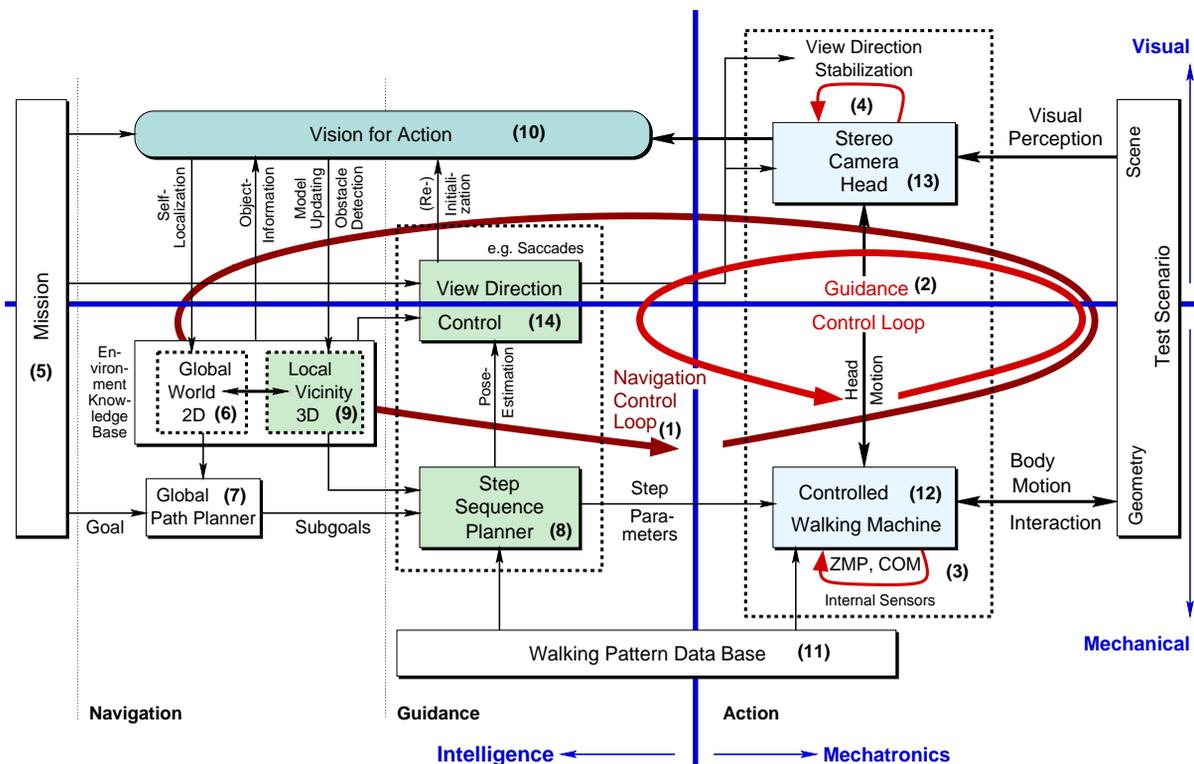


Fig. 2. Biologically inspired control architecture for vision-based walking.

The approaches proposed here are related to various methods in the literature on machine walking and computer vision. As already mentioned, our research directions focus on the development of methods for the combination of perception and action, applicable for goal-oriented vision-guided locomotion, and not on construction of walking machine hardware. The proposed vision-based biologically inspired cascaded control architecture determines several basic research directions, cf. Figure 2:

- (i) The development of strategies to synthesize walking patterns for smooth variable locomotion. Many researchers have concentrated on walking machines with simplified dynamics and specially on 2D walking [3, 18, 27]. The approach here deals with 3D walking of a biped with high mobility.
- (ii) Online step sequence planning allowing the robot to reach and stride over obstacles without collisions.
- (iii) Biologically motivated studies on intention dependent gaze control to increase visual information in a task and goal dependent way, and predictive feature (re-)allocation reducing the computational effort of image processing.
- (iv) Algorithms for robust image processing and obstacle reconstruction under the influences of head motion in 6 DOFs during walking in a distributed computer architecture.

3 Walking Pattern Synthesis

In order to move in an only partly known environment, a walking machine must be capable of adapting its steps to overcome obstacles. The need of smooth, stable motion requires predictive, real-time path and step planning algorithms to avoid major gait pattern modifications. In addition, the environment perception sensor should not be disturbed by the walking motion.

Usually the problem of executing a defined step with a biped while simultaneously ensuring its stability is handled by preplanning a suitable trajectory. 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 single walking primitives for level steps with different step lengths, for stair climbing and for striding over obstacles. The walking primitives are computed in such a way that ViGWaM is statically stable at each point of the path. Dynamic stability is assumed by executing the statically stable walking primitives with low speed. During online execution the step sequence planner appropriately selects and concatenates the walking primitives to a walking pattern using the information about the type of the next prototypical obstacle and its relative distance.

Related work to our approach is the inverted pendulum method [19] allowing the analytical calculation of dynamically stable trajectories for the 2D case, i.e. with the side stability of the biped ensured e.g. by mechanical design. For the 3D case the articles by Nishiwaki [25] and Nagasaka [24] provide methods to plan dynamically stable trajectories in Cartesian space by optimization.

After a brief description of the ViGWaM model in Section 3.1, the conditions for stability are stated in Section 3.2. The planning method for walking primitives is discussed in Section 3.3. Considerations on walking pattern concatenation for level walking are outlined in Section 3.4 and the algorithm for step sequence planning is presented in Section 3.5.

3.1 Model of ViGWaM

The kinematic structure of ViGWaM is illustrated in Figure 3. The 15 bodies are modeled as mass points; the total height (ground to cameras) is 1.70 m. For walking primitive planning the left foot is assumed to be fixed in the origin of a local coordinate system; the posture of the biped is expressed by the joint angles $\mathbf{q} = [q_1 \dots q_{10}]^T$. According to [7] the 10 degrees of freedom enable the robot to move without restrictions in the sagittal plane in the scenario shown in Figure 1. The position $\boldsymbol{\xi} = [x \ y \ z]^T$ and orientation given by roll-pitch-yaw angles $\boldsymbol{\Phi} = [\gamma \ \beta \ \alpha]^T$ of the head of ViGWaM, and right foot are $[\boldsymbol{\xi}_H, \boldsymbol{\Phi}_H]$, $[\boldsymbol{\xi}_R, \boldsymbol{\Phi}_R]$, respectively.

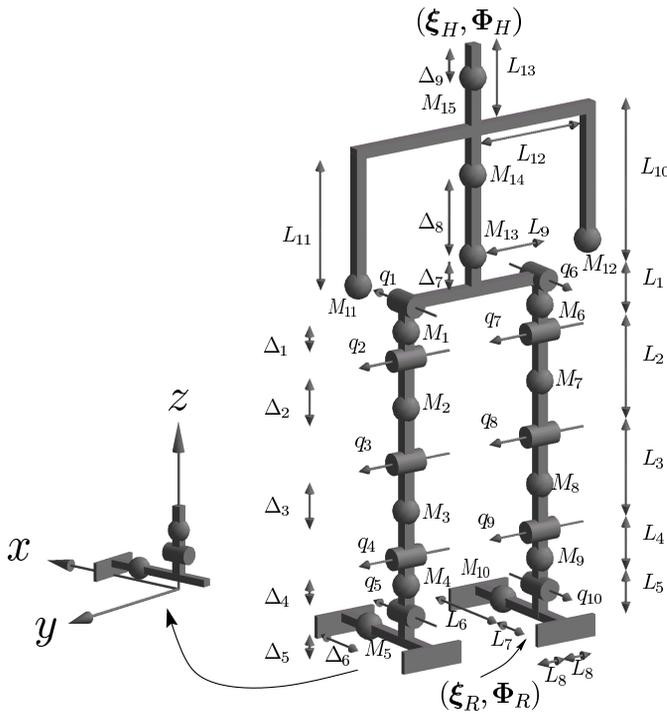


Fig. 3. ViGWaM kinematics from the back side ($\mathbf{q} = \mathbf{0}$).

Table 1. ViGWaM parameters.

$M_1 = M_6$	2.57 kg	L_8	0.05 m
$M_2 = M_7$	3.44 kg	L_9	0.12 m
$M_3 = M_8$	2.84 kg	L_{10}	0.49 m
$M_4 = M_9$	0.04 kg	L_{11}	0.45 m
$M_5 = M_{10}$	0.69 kg	L_{12}	0.20 m
$M_{11} = M_{12}$	2.00 kg	L_{13}	0.24 m
M_{13}	1.19 kg	Δ_1	0.00 m
M_{14}	6.40 kg	Δ_2	0.22 m
M_{15}	4.00 kg	Δ_3	0.12 m
L_1	0.00 m	Δ_4	0.00 m
L_2	0.44 m	Δ_5	0.03 m
L_3	0.44 m	Δ_6	0.06 m
L_4	0.00 m	Δ_7	0.10 m
L_5	0.09 m	Δ_8	0.25 m
L_6	0.17 m	Δ_9	0.06 m
L_7	0.05 m		

Table 2. Limits of ViGWaM joint-angles.

$q_{1/6,min}$	-0.436 rad	$q_{1/6,max}$	0.349 rad
$q_{2/7,min}$	-1.222 rad	$q_{2/7,max}$	0.436 rad
$q_{3/8,min}$	0.087 rad	$q_{3/8,max}$	1.571 rad
$q_{4/9,min}$	-0.873 rad	$q_{4/9,max}$	0.349 rad
$q_{5/10,min}$	-0.175 rad	$q_{5/10,max}$	0.262 rad

3.2 Walking Stability Conditions

A widely used stability criterion for biped walking is the zero moment point — ZMP condition, see e.g. [29]. The ZMP $\mathbf{p} = [x_p \ y_p \ 0]^T$ is characterized as the point on the contact surface, where the x - and y -component of the

moment generated by the reaction forces and the reaction torques vanishes. Using the principle of D'Alembert the ZMP can be calculated as

$$x_p = \frac{\sum_{i=1}^M M_i [(\ddot{z}_i + g)x_i - z_i \ddot{x}_i]}{\sum_{i=1}^M M_i (\ddot{z}_i + g)} \quad y_p = \frac{\sum_{i=1}^M M_i [(\ddot{z}_i + g)y_i - z_i \ddot{y}_i]}{\sum_{i=1}^M M_i (\ddot{z}_i + g)} \quad (1)$$

where $M = 15$ is the number of mass points, see Figure 3, and M_i their mass. Their Cartesian position, acceleration vector is given by $\xi_i(\mathbf{q}) = [x_i \ y_i \ z_i]^T$, $\ddot{\xi}_i(\mathbf{q}) = [\ddot{x}_i \ \ddot{y}_i \ \ddot{z}_i]^T$, respectively; g is the gravitational acceleration in negative z -direction. The physical parameters of ViGWaM are given in Table 1.

The ZMP stability condition holds, if \mathbf{p} remains inside the supported area of the stance foot during single-support and the polygon formed by both feet during double-support, see Figure 4. For $\ddot{\xi}_i = \mathbf{0}$, $i = 1, \dots, M$ the ZMP simplifies from (1) to

$$x_c = \frac{\sum_{i=1}^M M_i x_i}{\sum_{i=1}^M M_i} \quad y_c = \frac{\sum_{i=1}^M M_i y_i}{\sum_{i=1}^M M_i} . \quad (2)$$

In this case the ZMP coincides with the projection $\mathbf{c} = [x_c \ y_c]^T$ of the center of mass (COM) to the floor, also referred to as the projected COM (PCOM). A *statically stable path* for ViGWaM is characterized by the PCOM residing inside the supported area for all points along the path.

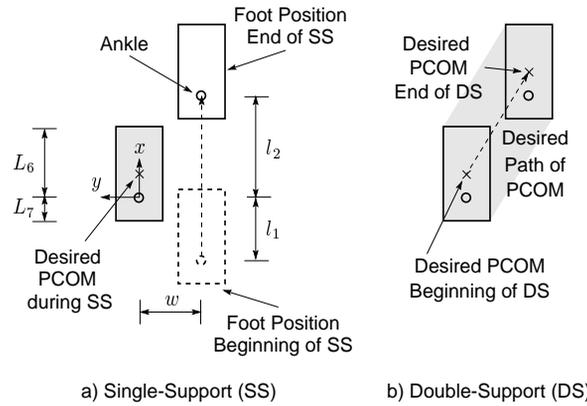


Fig. 4. Walking primitive and supported area (gray) - view on the contact-surface.

Under the assumption of small accelerations $\ddot{\xi}_i$ the ZMP will not significantly differ from the PCOM. If we further assume that the path has a sufficiently large stability margin — PCOM near the center of the supported area — the biped can move slowly along the path without violating the ZMP stability condition.

3.3 Walking Primitive Planning

A statically stable walking primitive is defined in the configuration space by the path $\mathbf{q}(s)$, $s \in [0, 1]$ with the left foot fixed in the origin of a local coordinate system. The path of the right foot is given by $\xi_R(s)$, $s \in [0, 1]$.

The *first phase* of the walking primitive is a single-support phase, see Figure 4(a). The initial configuration of ViGWaM is given by $\mathbf{q}(0)$ with the right foot just having lifted off the ground at $\xi_R(0) = [-l_1 \ -w \ -h_1]^T$. Following the desired path $\xi_{R,d}(s)$, $s \in [0, 0.5[$ the right foot moves towards $\xi_R(0.5) = [l_2 \ -w \ h_2]^T$. During the whole phase the PCOM resides in its desired position $\mathbf{c}_d(s) = [0.5(L_6 - L_7) \ 0]^T$, $s \in [0, 0.5[$ in the center of the left foot.

The *next phase* is a double-support phase, starting in the configuration $\mathbf{q}(0.5)$ with the right foot just touching the ground at $\xi_R(0.5) = [l_2 \ -w \ h_2]^T$. Then the PCOM is shifted along

$$\mathbf{c}_d(s) = \begin{bmatrix} x_{c,d} \\ y_{c,d} \end{bmatrix} = \begin{bmatrix} (2s - 1)l_2 + 0.5(L_6 - L_7) \\ -(2s - 1)w \end{bmatrix}, \quad s \in [0.5, 1] \quad (3)$$

from the center of the left foot towards the center of the right foot within the supported area by both feet, see Figure 4(b). After completion of the walking primitive the final configuration of ViGWaM is $\mathbf{q}(1)$.

Planning of a walking primitive is accomplished by solving a static optimization problem. For this purpose each of the continuous paths $\mathbf{q}(s)$, $\xi_R(s)$ and $\mathbf{c}(s)$, $s \in [0, 1]$ is represented by $N + 1$, $N = 2k$, $k \in \mathbb{N}$ discrete vectors $\mathbf{q}^{(j)}$, $\xi_R^{(j)}$ and $\mathbf{c}^{(j)}$, $j = 0, \dots, N$, equally spaced along the path. The necessary mapping is exemplified for the path $\mathbf{q}(s)$, $s \in [0, 1]$ of the joint angles, i.e.

$$\mathbf{q}^{(j)} = [q_1^{(j)} \dots q_{10}^{(j)}]^T = [q_1(s = \frac{j}{N}) \dots q_{10}(s = \frac{j}{N})]^T, \quad j = 0, \dots, N$$

yielding $\mathbf{q}^{(0)}, \dots, \mathbf{q}^{(N/2-1)}$ for the description of the single-support phase and $\mathbf{q}^{(N/2)}, \dots, \mathbf{q}^{(N)}$ for the double-support phase. The total path is then described by the augmented vector

$$\mathbf{Q} = [\mathbf{q}^{(0)}, \dots, \mathbf{q}^{(N)}].$$

A statically stable path is obtained by optimization, with the performance index as the sum over the squared Cartesian distances of the actual $\mathbf{c}^{(j)}$ from the desired PCOM $\mathbf{c}_d^{(j)}$ and the squared deviation of the direction of the z -unit vector $\mathbf{e}_{z,H}$ of the head frame from the vertical z -axis of the local coordinate system

$$\min_{\mathbf{Q}} = \sum_{j=0}^N \left[\left(x_c(\mathbf{q}^{(j)}) - x_{c,d}^{(j)} \right)^2 + \left(y_c(\mathbf{q}^{(j)}) - y_{c,d}^{(j)} \right)^2 + \left([0 \ 0 \ 1] \mathbf{e}_{z,H}(\mathbf{q}^{(j)}) - 1 \right)^2 \right]. \quad (4)$$

subject to

$$\mathbf{Q}_{min} \leq \mathbf{Q} \leq \mathbf{Q}_{max} \quad (5)$$

$$\xi_R(\mathbf{q}^{(j)}) - \xi_{R,d}^{(j)} = \mathbf{0}, \quad j = 0 \dots N \quad (6)$$

$$[0 \ 0 \ 1] \mathbf{e}_{z,R}^{(j)} = 1, \quad j = 0 \dots N \quad (7)$$

$$\mathbf{g}(\mathbf{q}^{(0)}, \mathbf{q}^{(N)}) = \mathbf{0} \quad (8)$$

where (5) are upper and lower bounds for the joint angles, cf. Table 2; (6) are equality constraints to ensure that the right foot $\xi_R^{(j)}$ remains on its prespecified desired path $\xi_{R,d}^{(j)}$; (7) ensures that the unit vector $\mathbf{e}_{z,R}$ points exactly upwards resulting in the right foot being parallel to the ground; the equality constraints (8) on the initial and final configuration depend on the type of the walking primitive.

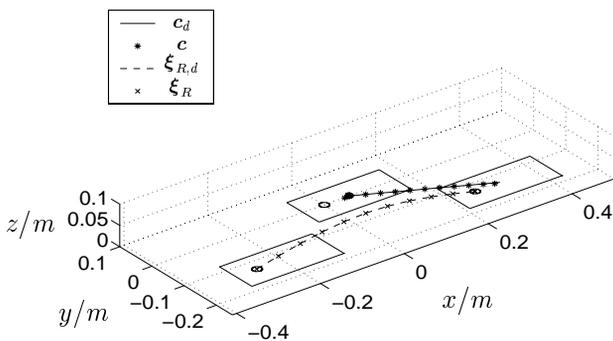


Fig. 5. Walking primitive after optimization process.

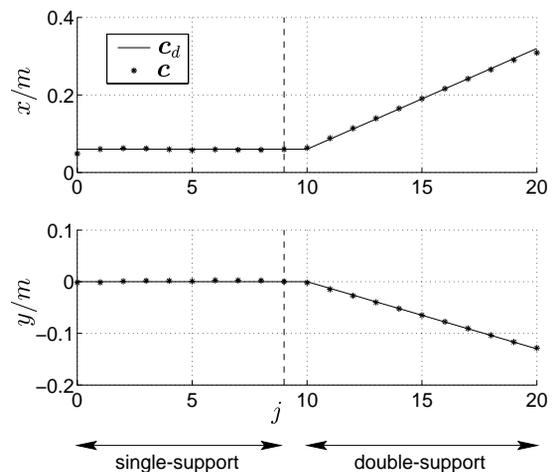


Fig. 6. x/y-path of desired and resulting PCOM.

As an example, the results for the synthesis of a level walking primitive ($h_1 = h_2 = 0.0 \text{ m}$) with a step length of $l_1 = l_2 = 0.26 \text{ m}$ and $N = 20$ are presented. The desired path $\xi_{R,d}(s)$ of the swinging leg during the

single-support phase is given by

$$\begin{bmatrix} x_{R,d}(s) \\ y_{R,d}(s) \\ z_{R,d}(s) \end{bmatrix} = \begin{cases} \begin{bmatrix} x_{R,d} \\ -w \\ u \sin\left(\pi \frac{x_{R,d} + l_1}{l_1 + l_2}\right) \end{bmatrix} & \text{for } 2s = \frac{\int_{-l_1}^{x_{R,d}} \sqrt{1 + \left(\frac{u\pi}{l_1 + l_2} \cos\left(\pi \frac{x + l_1}{l_1 + l_2}\right)\right)^2} dx}{\int_{-l_1}^{l_2} \sqrt{1 + \left(\frac{u\pi}{l_1 + l_2} \cos\left(\pi \frac{x + l_1}{l_1 + l_2}\right)\right)^2} dx}, s \in [0, 0.5[\end{cases} \quad (9)$$

with $u = 0.04 m$ and $w = 0.13 m$ resulting in the points $\xi_{R,d}^{(j)}$, $j = 0 \dots N/2 - 1$ equidistantly spaced over the length of the curve

$$z_{R,d}(x_{R,d}) = u \sin\left(\pi \frac{x_{R,d} + l_1}{l_1 + l_2}\right), \quad y_{R,d} = -w, \quad x_{R,d} \in [-l_1, l_2].$$

The width $w = 0.13 m$ between both feet is held relatively small in order to reduce the lateral swinging motion necessary to keep the PCOM over the supported area and thus to decrease the head motion disturbing the image-processing algorithms. The condition on the initial and final configuration of ViGWaM follows by the mapping (11) as

$$\mathbf{g}(\mathbf{q}^{(0)}, \mathbf{q}^{(N)}) = [q_6^{(N)} \dots q_{10}^{(N)} q_1^{(N)} \dots q_5^{(N)}]^T - [q_1^{(0)} \dots q_{10}^{(0)}]^T = \mathbf{0}.$$

This optimization problem (4)-(8) is solved by application of the MATLAB SQP-routine `fmincon`. The gradient of the performance index is supplied in analytical form by a symbolic computation tool. The walking primitive is computed by three subsequent optimization runs with $N = 10, 15, 20$. For the first optimization an approximate starting solution is supplied. For $N = 10, 15$ the starting solution is gained by a linear interpolation of the previous result. Convergence is fast; the final solution for a walking primitive is usually found in $\leq 1 min$ computation time on a UltraSparc II 400, although convergence towards a steady path is not necessarily guaranteed by the problem formulation. Figure 5 depicts the desired and resulting path of the PCOM and the right foot. Figure 6 shows the corresponding x - and y -coordinates of the PCOM and Figure 7 the paths of the joint angles. The path of the head $\xi_H = [x_H \ y_H \ z_H]^T$, $\Phi_H = [\gamma_H \ \beta_H \ \alpha_H]^T$ is illustrated in Figure 8. It is obvious that the resulting Cartesian head motion is still large despite all efforts.

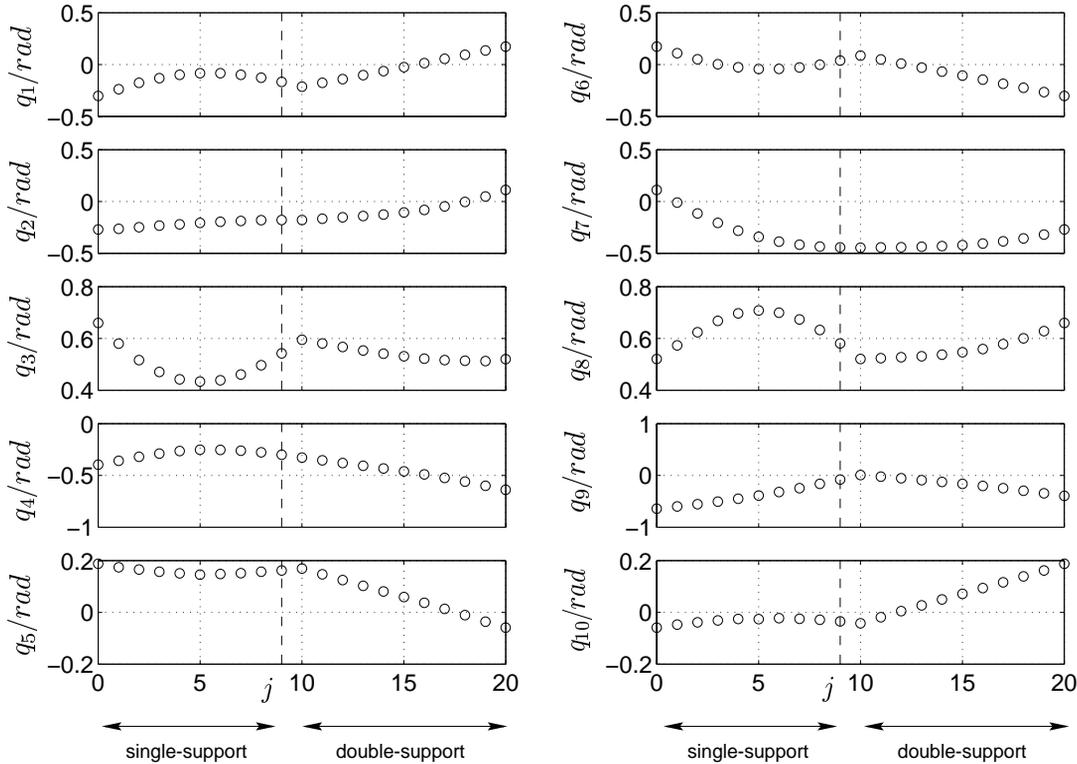


Fig. 7. Resulting paths of the joint angles.

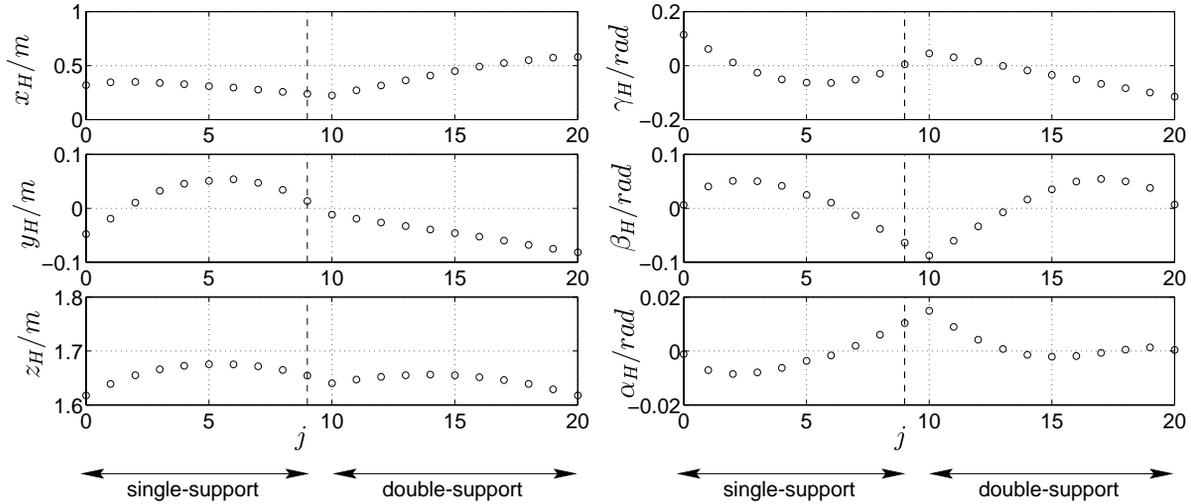


Fig. 8. Path of the walking machine head.

3.4 Walking Primitive Concatenation

To concatenate several walking primitives into a walking pattern further considerations are necessary. For clarity the discussion is restricted to level walking resulting from a concatenation of walking primitives characterized by $l_1, l_2, w, h_1 = h_2 = 0 m$ and referred to as $\mathbf{Q}_{l_1 \rightsquigarrow l_2}$.

Because of the symmetry in the kinematic model of ViGWaM a walking primitive $\tilde{\mathbf{Q}}_{l_1 \rightsquigarrow l_2}$ for the right foot supporting the biped is easily obtained from a walking primitive $\mathbf{Q}_{l_1 \rightsquigarrow l_2}$ for its left foot by the mapping

$$\mathbf{q}_{l_1 \rightsquigarrow l_2}^{(j)} = [q_1^{(j)} \dots q_{10}^{(j)}]^T \rightarrow \tilde{\mathbf{q}}_{l_1 \rightsquigarrow l_2}^{(j)} = [q_6^{(j)} \dots q_{10}^{(j)} \quad q_1^{(j)} \dots q_5^{(j)}]^T, \quad j = 0, \dots, N. \quad (10)$$

In order to enable cyclic walking with a given step length $l_1 = l_2 = l$ a walking primitive for a single step fulfilling the conditions

$$\mathbf{g}(\mathbf{q}_{l \rightsquigarrow l}^{(0)}, \mathbf{q}_{l \rightsquigarrow l}^{(N)}) = \mathbf{q}_{l \rightsquigarrow l}^{(N)} - \tilde{\mathbf{q}}_{l \rightsquigarrow l}^{(0)} = \mathbf{0} \quad (11)$$

needs to be planned. Based on this assumption the primitives $\mathbf{Q}_{l \rightsquigarrow l}, \tilde{\mathbf{Q}}_{l \rightsquigarrow l}, \mathbf{Q}_{l \rightsquigarrow l}, \tilde{\mathbf{Q}}_{l \rightsquigarrow l}, \dots$ can be executed subsequently.

For the concatenation of two cyclic walking primitives $\mathbf{Q}_{a \rightsquigarrow a}, \mathbf{Q}_{b \rightsquigarrow b}$ with different step lengths $a \neq b$, a transition walking primitive $\mathbf{Q}_{a \rightsquigarrow b}$ is defined, see Figure 9. This transition primitive has to satisfy

$$\mathbf{g}_1(\mathbf{q}_{a \rightsquigarrow b}^{(0)}) = \mathbf{q}_{a \rightsquigarrow b}^{(0)} - \tilde{\mathbf{q}}_{a \rightsquigarrow a}^{(N)} = \mathbf{0}, \quad \mathbf{g}_2(\mathbf{q}_{a \rightsquigarrow b}^{(N)}) = \mathbf{q}_{a \rightsquigarrow b}^{(N)} - \tilde{\mathbf{q}}_{b \rightsquigarrow b}^{(0)} = \mathbf{0}. \quad (12)$$

When the transition primitive is mapped according to (10) the primitive $\tilde{\mathbf{Q}}_{a \rightsquigarrow b}$ shown in the center of Figure 9 is obtained. This primitive can be concatenated with the two cyclic primitives yielding the complete transition. The resulting walking pattern is depicted on the right in Figure 9 allowing the biped, starting with a initial step length (leg distance) of a , to slowly execute three steps with the lengths a, b, b .

In order to enable walking with n different step lengths, a database of n cyclic walking primitives with different step lengths l is computed beforehand. Considering possible transitions between the set of cyclic primitives $p = n(n-1)$ transition primitives are necessary. In order to allow ViGWaM to walk with step lengths ranging from l_{min} to l_{max} discretized with Δl , $n = 1 + (l_{max} - l_{min})/\Delta l$ cyclic walking primitives and $p = n(n-1)$ transition primitives are to be calculated. Conditions on the values of l_{min} and l_{max} result from the strategy for step sequence planning as explained in the next section.

The level walking primitives already allow ViGWaM to cope with obstacles such as ditch and step trace. If one ensures that the last step before a bar or stairway has a fixed step length, striding over a bar requires a concatenation of two walking primitives with a modified path of the swinging leg. Stair climbing requires one cyclic primitive with $h_1 = h_2 = 0.15 m$ corresponding to the stair height for cyclic stair walking, one primitive for climbing the first stair with $h_1 = 0 m, h_2 = 0.15 m$ and another one for leaving the last stair with $h_1 = 0.15 m, h_2 = 0 m$.

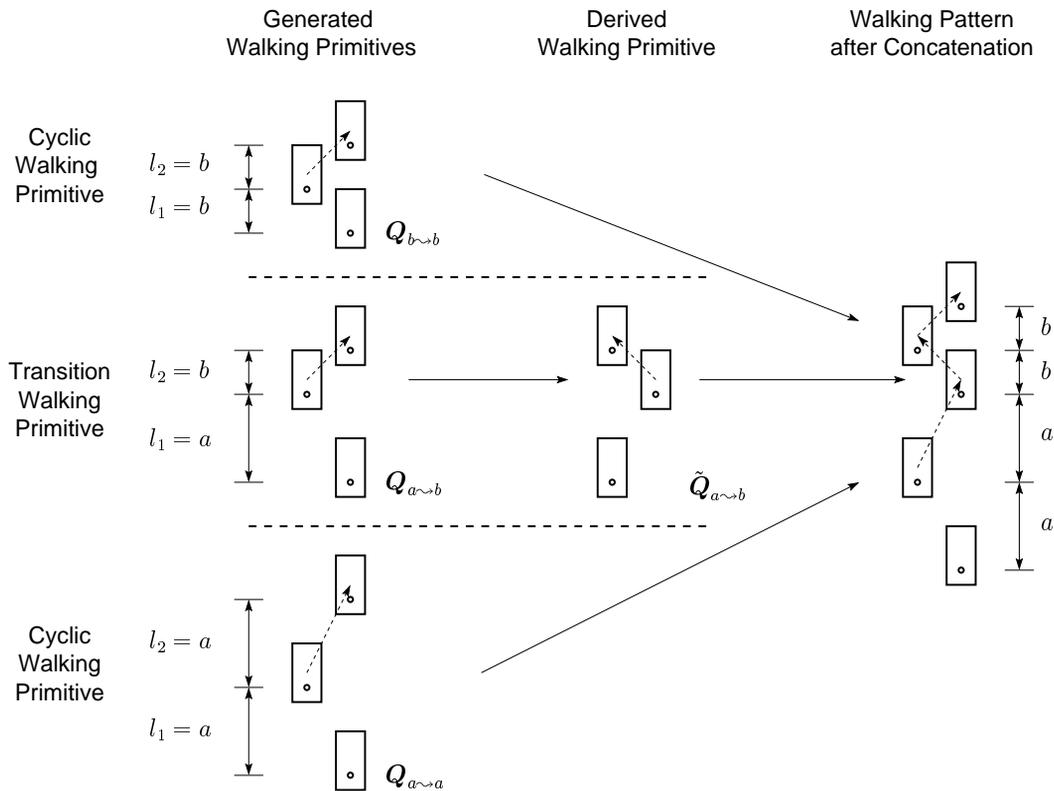


Fig. 9. Changing step lengths by a transition primitive.

3.5 Step Sequence Planner

The task of the step sequence planner is to select and concatenate the appropriate walking primitives in order to position ViGWaM in the desired location relative to an obstacle. As long as the distance l_d between the estimated position of the next supporting foot and the desired location relative to the next obstacle is bigger than $3l_n$ (3-steps-ahead-strategy) ViGWaM walks with a fixed norm step length $l_n = 0.26 m$. When the step sequence is adapted the distance l_d is $\leq 3l_n$. Depending on the obstacle type and in order to realize smooth walking, the step sequence is planned according to one of the following three rules:

- rule A:** Obstacle should be reached with the next supporting foot \Rightarrow execute 2 steps with $l_d/2$.
- rule B:** Obstacle should be reached with the next swinging foot \Rightarrow execute 3 steps with $l_d/3$.
- rule C:** Obstacle should be reached with one of both feet \Rightarrow execute 2 steps with $l_d/2$ or 3 steps with $l_d/3$ depending on which step length is closer to l_n .

Considering the worst cases it is easy to identify the interval of step lengths that the biped has to support as $[\frac{2}{3}l_n, \frac{3}{2}l_n]$. This is ensured by the ViGWaM database providing 676 walking primitives for walking with the step lengths $l = 0.15 m, 0.16 m, \dots, 0.40 m$.

Figure 10 illustrates a situation where the algorithm for step sequence planning is used in order to step into the first footprint of a step trace. The upper part shows the situation from the side, the lower part the view from above. ViGWaM enters the scene from the left hand side walking cyclic by the norm step length l_n and receives information on the relative position to the first footprint by vision for action as soon as it is detected. During execution of the second illustrated cyclic step it detects that the position of the next supporting foot, in this case the left one, will be closer than $3l_n$ to the first footprint. As the first footprint has to be reached with the right foot, rule B is applied and three steps with length $l_d/3$ are planned. After that first one step with the step length l_{t1} , corresponding to the distance between the 1. and 2. footprint, and then one step with l_{t2} , corresponding to the distance between the 2. and 3. footprint as measured by vision, is executed.

4 Image Processing and Gaze Control

Smooth walking in an environment with obstacles requires a reliable and predictive classification and pose estimation of all obstacles in the walking trail. Furthermore, online perception capabilities are essential for performing

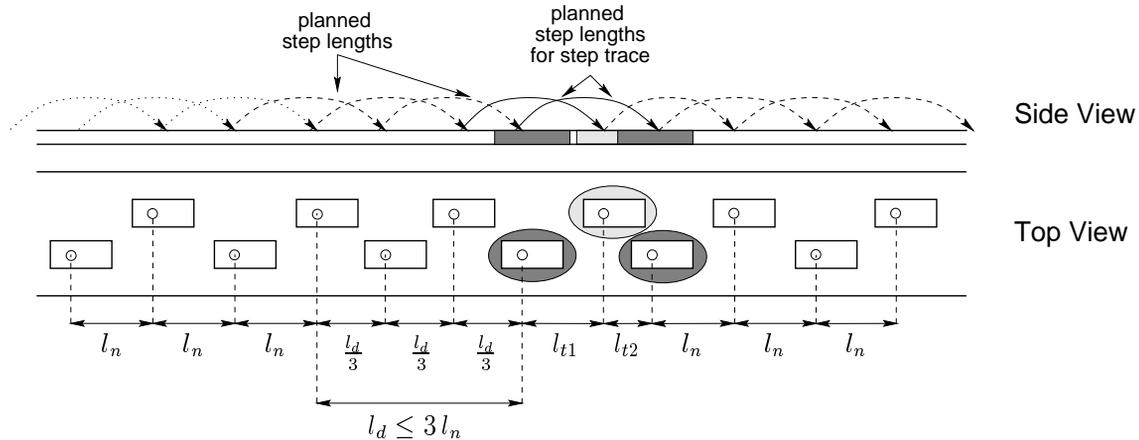


Fig. 10. Combining walking primitives by step sequence planning stepping into a step trace.

the step sequence adaptation. Thereby, head and resulting camera motion during walking in general may cause problems for image processing algorithms. Especially, orientation changes around the optical axis prove to be most critical. In addition, the variety of obstacles indicated in Figure 1 requires the extraction of specific online trackable features, e.g. blobs (binary large objects), edges, and corners, describing the obstacle in the image stream. For this purpose, elementary modules for robust real-time image processing software are employed. Since visual resources are always limited, it is not possible to track all objects and to analyze all aspects of a scenario simultaneously. This situation is analogous in biological examples [1, 2]. Therefore, some means of controlling the view direction of the visual stereo sensors is required in order to focus on currently relevant objects and to maximize information acquisition.

4.1 Vision Architecture

Image processing time is a critical factor for smooth walking in nearly unknown environments. 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 interpretation potential for object recognition. To merge the advantages of both approaches, a cascaded image processing architecture for “vision for action” has been developed. Figure 11 shows the proposed architecture with the two image processing blocks highlighted. In the outer loop for image analysis, online capability is not required. In fact, static images are analyzed slower than camera frame rate and the results are integrated into a 3D-map of the environment. All relevant data of the objects are stored in this map: type and dimension of the obstacle, pose in the world, visibility, and estimated accuracy.

For safe walking, it suffices to know the pose of the obstacle relative to the walking machine, as 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 using the XVISION-library [12]. With this architecture the time-consuming scene interpretation and feature tracking can be performed simultaneously and the results are fused via the 3D-map.

Furthermore, goal-oriented locomotion causes variations in the field of view, with the result that new regions are mapped on the vision sensor and already analyzed features disappear, thus making the two processes, image analysis and tracking, necessary. To influence the unintended variations in the field of view, and for maximization of visual information during walking, a variable gaze is employed. A gaze controller, proposed in Section 4.2, changes the view direction depending on the current goal of the walking machine. The view direction is selected from predicted future perceptions of the scenario. This gaze controller manages the determination of new appearing regions and the reallocation of already tracked features temporarily not detectable.

4.2 Gaze Control

A gaze controller accomplishes the task to determine “where to look next?”. This is an intention/mission problem and in order to build an intelligent visual system, it must be considered a fact that perception is intimately related to the physiology and psychology of the perceiver and the tasks that it performs. A gaze controller is only applicable together with an active vision system. An active vision system is defined to be able to react in order to control

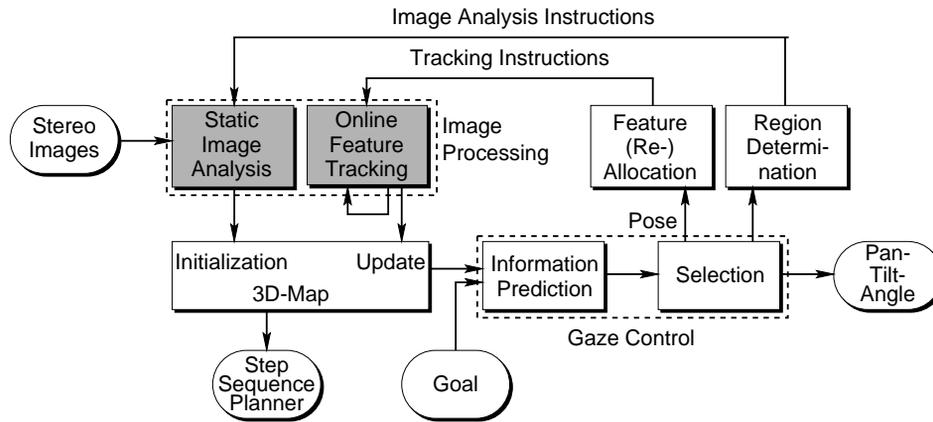


Fig. 11. Cascaded image processing architecture.

the geometry of its sensory apparatus. An active vision system is a system able to adjust its visual parameters in a controlled manner in order to extract useful data about the scene in space and time that allow it to best perform a set of tasks [4].

Since in biologically inspired architectures the field of view is always limited, a means of directing the fixation of the visual sensors is required. To accomplish this, a view direction (gaze) control algorithm must be employed to determine how to position the sensor so that the interesting objects of the scene are within its field of view. A gaze controller is to position the visual sensors (overt attention in biological systems) and is not a mechanism to determine the high saliency regions in a given scene (covert attention) [9].

These facts result from considering the visual sensors of the walking machine as a system to perform purposive vision, i.e. vision is not considered in isolation, but as part of a complex system that interacts in specific ways with the environment. The relationship of the visual system to the world consists of perceptual capabilities and supporting actions. This relationship does not require an elaborate categorical representation of the world (qualitative vision in opposition to purposive vision).

Very few articles have been presented on the topic of view direction control. The “where to look next?”-problem has been treated so far in a different way. The goal was to determine the most conspicuous point or feature within an image in order to direct attention towards it, i.e. shifts on covert attention.

4.3 ViGWaM Gaze Controller

In the experiment presented in Section 6, the mission will be constrained to walking straight ahead in the prototypical scenario avoiding collision with typical obstacles. In this case the objective of gaze control is to ensure that objects of interest remain within the field of view as long as possible. Other tasks, such as self-localization or the generation of tracking instructions for image processing, cf. Figure 11, are irrelevant in this situation.

Obviously view direction depends heavily on the type of locomotion task, such as global navigation or local guidance. For example, high quality information for local guidance, i.e. the exact pose of upcoming obstacles, can be considered of little significance for planning of a global mission-dependent walking path.

Without gaze control, which adapts the view direction depending on the requirements of the current situation, it would not be possible to achieve the required obstacle perception accuracy and the object would disappear from the field of view much earlier than desired. In our setup it is required that the object remains visible until a distance to the object $l_d \leq 0.78 m$. This corresponds to the 3-steps-ahead-strategy of the step sequence planner described in Section 3.5. Figure 12 shows the dependency of the field of view in plane x -direction on the camera tilt-angle β for an view-angle of $0.786 rad$; C_y is the y -coordinate in image coordinates.

In our experimental setup, the camera pair is mounted on a pan-tilt-head. This means that the gaze controller disposes of two variables, pan-angle (α) and tilt-angle (β), to control the perception process. The operation can be explained by considering the mapping of a 3D-point ${}_{WM} \mathbf{x} = [x \ y \ z]^T$ (represented in the provisional walking machine frame) into a 2D-point $\mathbf{m} = [u \ v]^T$ (represented in the image coordinate frame). The x/y -plane of the provisional walking machine frame coincides with the ground plane. Its x -axis is the walking direction, and the position of the origin is the vertical projection of the camera reference frame origin on the ground, translated in y -direction to the center between both feet. This coordinate system is computationally decoupled with respect to changes of camera orientation, height of cameras above the ground, and camera movements in y -direction. Thus,

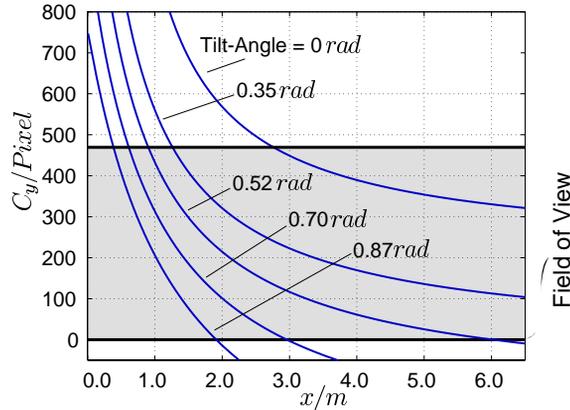


Fig. 12. Field of view dependence in x -direction on the tilt-angle.

this coordinate system is appropriate for step sequence planning. This mapping can be represented by:

$$s[\mathbf{m}^T \ 1 \ k]^T = \mathbf{P}^C \mathbf{T}_{WM} [\mathbf{w}_M \mathbf{x}^T \ 1]^T \quad (13)$$

where s, k are two arbitrary scalars and \mathbf{P} the homogeneous projection matrix of the camera, which maps a 3D-point given in camera reference frame into a 2D-point in image reference frame. The matrix ${}^C\mathbf{T}_{WM}$ represents the homogeneous transformation between the provisional walking machine frame and the camera reference frame. In this transformation matrix the pan and tilt-angle variables determine how the visual perception is performed by selecting a specific view direction.

Analysis of human gaze behaviour during walking has shown that there is a predictive component in the way visual information is used in locomotion [1,2]. The human brain uses pre-stored information for navigation and guidance. This information, for example the approximate position of an oncoming obstacle on the walking trail, helps to direct the visual sensor. This is the biological basis of the predictive component incorporated in the ViGWaM gaze controller.

Using the available information about the scene (i.e. 3D-map, see Figure 11), the current walking machine location and actual movement parameters, the gaze controller predicts the optimal orientation of the visual sensor, in this case tilt-angle, for a future location. The optimal orientation is found by choosing the state of maximum information among a series of predicted states (PS). A predicted state is specified by the future location of the head and the variable value of the tilt-angle, that determines the field of view, see Figure 13. We predict the future camera view for a range of tilt-angles at the end of the present step because at this point of time the foot location is known and the head movement in x -direction is minimum. The tilt-angle is changed from 0.35 rad , field of view reaching in x -direction approximately from 1.45 m to ∞ , to the maximum permitted by the used hardware 0.87 rad (0.60 to 1.85 m), see Figure 12. With this tilt-angle range, the walking trail remains in a reasonable view.

The information content of a given state is quantified, depending on the current goal, taking into account the number of objects in view, their distance to the walking machine and their relative position to the center of the camera view:

$$I_{PS}(\beta) = \sum_{i=1}^N \left(\frac{f(d_i)}{p_i^2(\beta)} \right) \quad (14)$$

where N is the number of obstacles in the field of view, d_i is the distance from the walking machine to the obstacle i , and p_i is the distance of the obstacle to the center of the image in the camera view. The function f is defined as follows:

$$f(d_i) = \frac{C}{d_i^2}, \quad \begin{cases} C = 100 & \text{if } 0.8 \text{ m} > d_i \\ C = 10 & \text{if } 1.3 \text{ m} > d_i \geq 0.8 \text{ m} \\ C = 1 & \text{if } d_i \geq 1.3 \text{ m} \end{cases} \quad (15)$$

The value of I_{PS} decreases with increasing d_i because the accuracy of the stereo-based 3D-reconstruction decreases with distance. It decreases with increasing p_i because an object projected in the outer region of the image cannot be analyzed with the same accuracy as an object projecting in the image center due to lens distortions. The optimal tilt-angle β^* is obtained by maximizing (14).

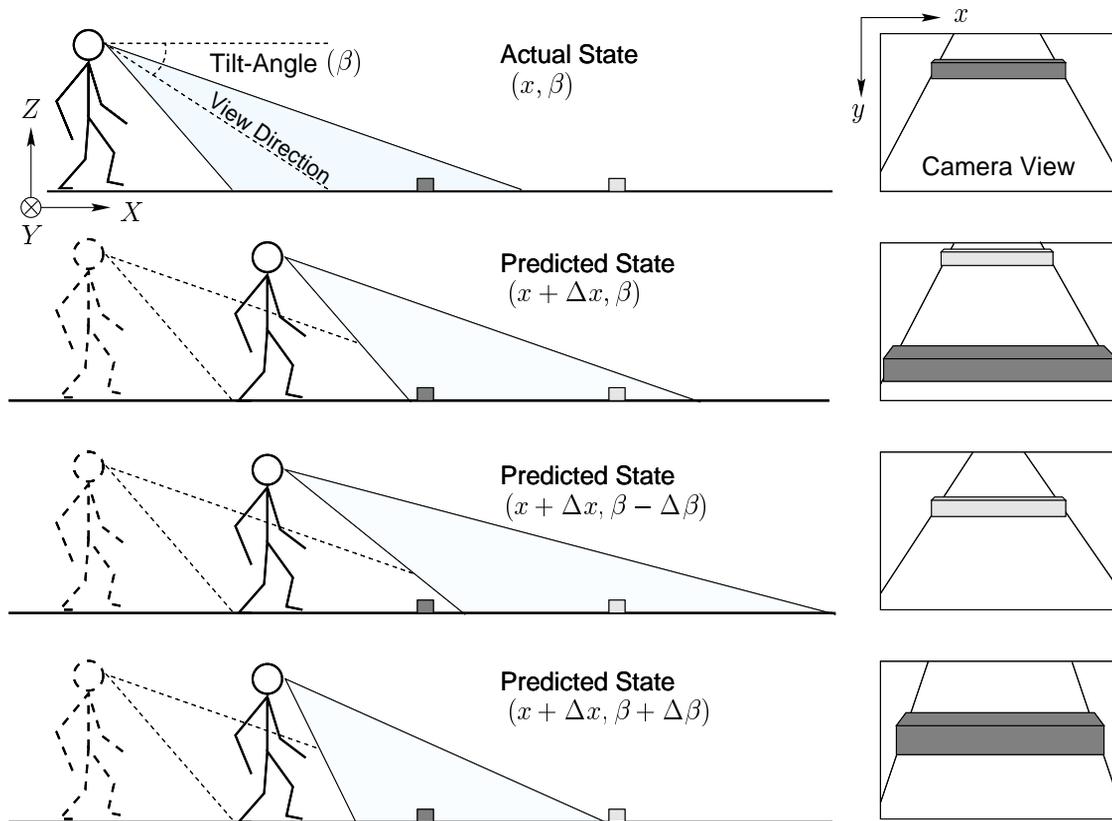


Fig. 13. Predictive component in ViGWaM gaze control strategy.

4.4 Line-Based Object Reconstruction

In this section, we will focus on the image based online reconstruction of line-based objects, such as the barrier and the staircase in the walking scenario. We assume the dimensions of the barrier and staircase to be known from the 3D-map and the lines representing the transition from ground plane to obstacle to be tracked when visible in the image stream of one of the two cameras. The allocation of features is achieved by the gaze controller. Let \mathbf{P} be the homogeneous projection matrix of the camera, which maps the 3D-line (represented in the camera reference frame) into a 2D-line (represented in the image coordinate frame) and $\mathbf{m}_1 = [u_1 \ v_1]^T$ and $\mathbf{m}_2 = [u_2 \ v_2]^T$ two arbitrary points of the 2D-line. Then three 3D-points (${}^C\mathbf{x}_1$ to ${}^C\mathbf{x}_3$) of the plane, in which the perceived 3D-line is located can be computed with

$$[{}^C\mathbf{x}_1^T \ 1]^T = \mathbf{P}^{-1}[\mathbf{m}_1 \ 1 \ 1]^T, [{}^C\mathbf{x}_2^T \ 1]^T = \mathbf{P}^{-1}[\mathbf{m}_2 \ 1 \ 1]^T, s[{}^C\mathbf{x}_3^T \ 1]^T = \mathbf{P}^{-1}[\mathbf{m}_2 \ 1 \ s]^T \quad (16)$$

where $s \neq 1$ is an arbitrary scalar. By inverse kinematics, these points are transformed with $({}^C\mathbf{T}_{WM})^{-1}$ into the ground fixed coordinate system and ${}_{WM}\mathbf{x}_1 \dots \mathbf{x}_3$ are obtained.

With three points the normal representation of the plane E_1 can be written as

$$E_1 : \mathbf{n}_1 \mathbf{x} = d_1 \quad \text{with} \quad \mathbf{n}_1 = \frac{({}_{WM}\mathbf{x}_2 - {}_{WM}\mathbf{x}_1) \times ({}_{WM}\mathbf{x}_3 - {}_{WM}\mathbf{x}_1)}{|({}_{WM}\mathbf{x}_2 - {}_{WM}\mathbf{x}_1) \times ({}_{WM}\mathbf{x}_3 - {}_{WM}\mathbf{x}_1)|} \quad \text{and} \quad d_1 = \mathbf{n}_1^T {}_{WM}\mathbf{x}_1. \quad (17)$$

We assume that the considered line representing the obstacle is in the x/y -plane E_0 of the ground plane. With

$$E_0 : \mathbf{n}_0 \mathbf{x} = d_0 \quad \text{with} \quad \mathbf{n}_0 = [0 \ 0 \ 1]^T \quad \text{and} \quad d_0 = 0 \quad (18)$$

the 3D-line G_1 can be computed by

$$G_1 : \mathbf{x} = \mathbf{a}_1 + t\mathbf{b}_1 \quad \text{with} \quad \mathbf{b}_1 = \mathbf{n}_0 \times \mathbf{n}_1, \quad \mathbf{a}_1 = d_0\mathbf{n}_0 + \frac{d_1 - d_0\mathbf{n}_0^T\mathbf{n}_1}{\mathbf{n}_1^T(\mathbf{b}_1 \times \mathbf{n}_0)}(\mathbf{b}_1 \times \mathbf{n}_0) \quad \text{and} \quad t \in \mathbb{R}. \quad (19)$$

Other line-based obstacles can be detected and located similarly. Experiments based on an implementation of this algorithm are presented in Section 6.

4.5 Point-Based Object Reconstruction

In case of point based objects, represented as blobs (binary large objects) in the image stream, e.g. the footprints in our walking scenario, 3D-reconstruction is easily realized as follows. Given two points in correspondence $\mathbf{m} = [u \ v]^T$ and $\mathbf{m}' = [u' \ v']^T$ — the correspondence problem can be solved with aid of the 3D-map and gaze controller. Let ${}_C\tilde{\mathbf{x}} = [x \ y \ z]^T$ be the corresponding 3D-point in space with respect to the projective basis of the stereo-camera-pair, with \mathbf{P} and \mathbf{P}' the projection matrices we have:

$$s [\mathbf{m}^T \ 1 \ k]^T = \mathbf{P} [{}_C\tilde{\mathbf{x}}^T \ t]^T, \quad (20)$$

$$s' [\mathbf{m}'^T \ 1 \ k']^T = \mathbf{P}' [{}_C\tilde{\mathbf{x}}^T \ t]^T, \quad (21)$$

where s , s' , k , k' , and t are five arbitrary scalars. Let \mathbf{p}_i and \mathbf{p}'_i be the vectors corresponding to the i th row of \mathbf{P} and \mathbf{P}' , respectively. The two scalars can be computed with $k=k'=t=1$ as $s = \mathbf{p}_3^T [{}_C\tilde{\mathbf{x}}^T \ 1]^T$ and $s' = \mathbf{p}'_3^T [{}_C\tilde{\mathbf{x}}^T \ 1]^T$. Eliminating s and s' from (20) and (21) yields the following equation:

$$\mathbf{A} [{}_C\tilde{\mathbf{x}}^T \ 1]^T = \mathbf{0}, \quad (22)$$

where \mathbf{A} is the matrix

$$\mathbf{A} = [\mathbf{p}_1 - u\mathbf{p}_3, \mathbf{p}_2 - v\mathbf{p}_3, \mathbf{p}'_1 - u'\mathbf{p}'_3, \mathbf{p}'_2 - v'\mathbf{p}'_3]. \quad (23)$$

Solving this linear equation leads to the corresponding 3D-point [28]. As the projective coordinates ${}_C\tilde{\mathbf{x}}$ are defined up to a scale factor t , we can impose $\| [{}_C\tilde{\mathbf{x}}^T \ 1]^T \| = 1$, then the solution is well known to be the eigenvector of the matrix $\mathbf{A}^T\mathbf{A}$ associated with the smallest eigenvalue.

As the type of the obstacles are initially known, planning a smooth step sequence over these obstacles is performed by the step sequence scheme described in Section 3. Hence, image processing and therewith step sequence planning can be performed at camera frame rate with satisfactory 3D-reconstruction results.

5 Emulation of ViGWaM

An implementation of the guidance control scheme, shown in Figure 2, on a real walking machine would be the most suitable way to demonstrate overall performance. In our research we chose an alternative in lack of physical walking machine hardware. We propose a hardware-in-the-loop emulation environment approach with real vision components and the vision guided virtual walking machine — ViGWaM presented below.

As part of the walking machine emulator an ideal simulation of the controlled walking machine including all dynamical effects of the walking machine mechanism and its interaction with the environment — such as contact situations between feet and objects — would be required. This simulation would deliver realistic head/camera motions of the walking machine as a result of the controlled walking process. However, hardware-in-the-loop experiments require online simulation of the controlled walking machine, which currently seems to be unfeasible for non-simplistic mechanical models due to the necessary computational effort required for numerical integration. In our emulation approach, the controlled walking machine behavior is analyzed offline. The walking patterns to move the body of the walking machine are generated by an optimization process as described in Section 3. As a result of the interaction of the walking machine with its environment head motion, see Figure 8, occurs. This head motion is the input for the motion emulator. This somewhat ideal (but reasonable) assumption means that all controllers work perfectly and that there is no uncontrolled, unknown or undesired head motion.

Currently a wheeled mobile platform is used to carry a perpendicularly mounted SCARA-type robot arm. The pan-tilt head (PTU-46-17.5, Directed Perception Inc.) driven by stepper motors and a pair of NTSC-cameras (XC-999, Sony) is mounted on the robot arm, see Figure 14 [23]. With the mobile platform in longitudinal direction, the robot arm for movements in y -, z -direction and orientation changes around the x -axis, and the pan-tilt head for gaze control — orientation around the y - and z -axis — the cameras can travel over the scenario and emulate the ViGWaM head motion in 6 DOFs.

The walking scenario similar to Figure 1 has been physically built including prototypical obstacles. It comprises several obstacle modules, which can be arranged arbitrarily into various configurations. To keep this experimental setup manageable in our laboratory, all components, such as the height of ViGWaM, stair height, stair depth etc., are scaled down by a ratio 3 : 2. This means that the ViGWaM stereo camera pair is located approximately 1.13 m above the ground compared to the 1.7 m of human eyes.

To visualize the operation of ViGWaM, an image of an external scene camera observing the scenario is overlaid with a 3D-animation of ViGWaM. The right side of Figure 15 shows the prototypical environment including

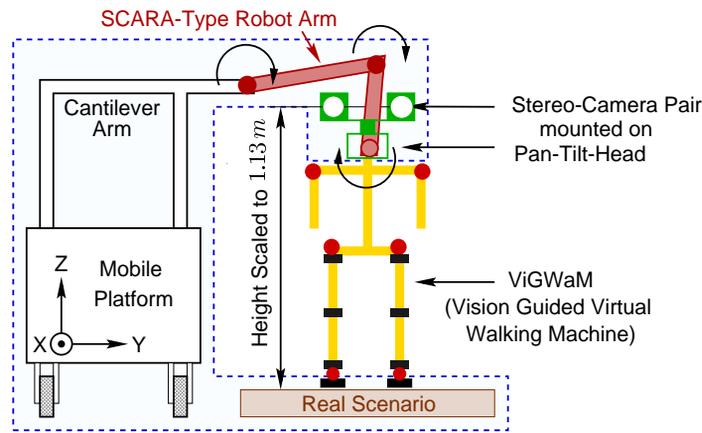


Fig. 14. Emulation environment with real hardware-in-the-loop components (dashed box).

the ViGWaM graphics overlay. During an experiment the movement of the real camera head above the walking trail and the gait of ViGWaM overlaid on the image of the external camera on a monitor can be observed simultaneously. As a result head motion of ViGWaM corresponds with the camera motion performed by the mobile platform, the robot arm, and the pan-tilt head. The emulator architecture, shown in Figure 15 comprises the vision controller, step planner, ViGWaM simulation part, and the augmented reality visualization part.

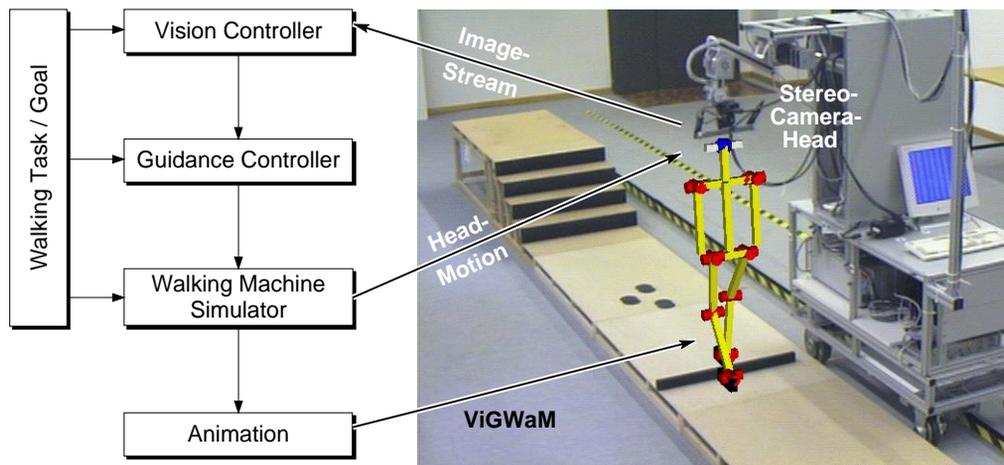


Fig. 15. Emulator architecture and experimental setup with ViGWaM overlaid.

6 Experimental Results

To illustrate main features of our approach, we describe an experiment with three prototypical obstacles, a barrier, three footprints with the task to step inside, and a stairway. The type and dimensions of the obstacles are known *a priori* and specified in a map. The distances of the obstacles from the starting point of the left foot: $d_b = 1.482\text{ m}$ of the barrier, $d_{f_1} = 2.525\text{ m}$, $d_{f_2} = 2.640\text{ m}$, and $d_{f_3} = 2.792\text{ m}$ of the footprints, and $d_s = 3.930\text{ m}$ of the first stair are assumed to be known *a priori* with an error of $\pm 0.05\text{ m}$.

Figure 16 shows the configuration of the obstacles and the resulting view direction from the gaze controller during walking. The variable tilt-angle is shown as a continuous line.

Figure 17 shows the unfiltered error e of visual distance measurement during walking. As an example for the footprints only the error to the first print is shown. The roughness of the error results from the fact, that the features are not continuously tracked in the image stream, during rapid tilt motion. Reallocation of the features, describing the obstacle in the image stream, is performed when the walking primitives are switched. It is noteworthy that

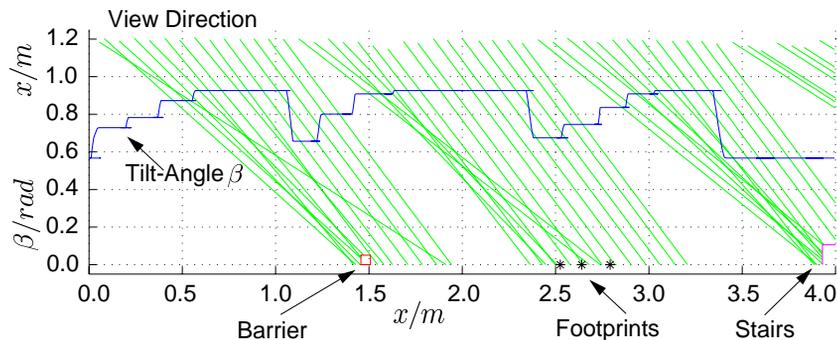


Fig. 16. View direction during walking.

line-based 3D-reconstruction is more accurate than point-based reconstruction. However, in the phase where the step sequence is adapted in a distance $l_d \leq 3l_n$ in front of the obstacle, the error $|e|$ is ≤ 0.015 m, which proved to be sufficient for vision based walking.

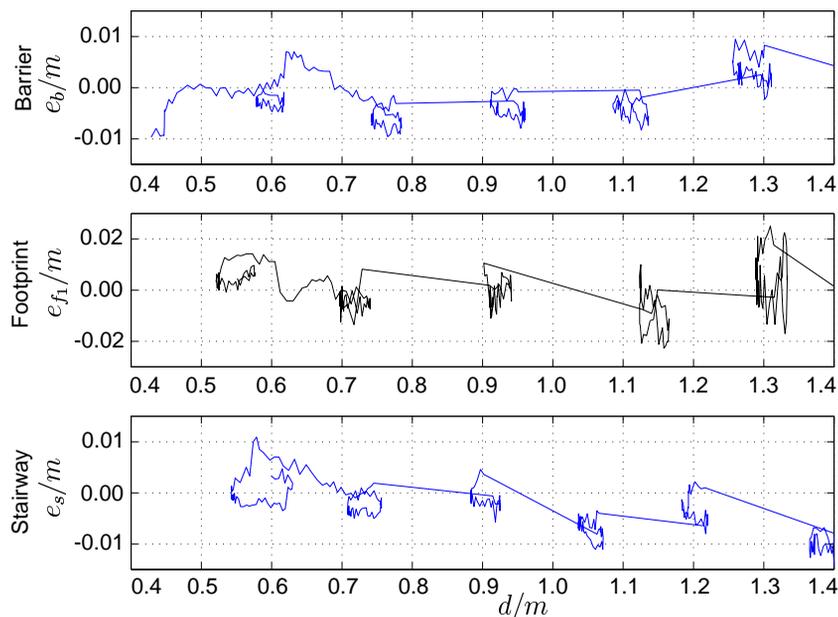


Fig. 17. Distance measurement errors of obstacles during walking.

In the upper part of Figure 18 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 not have been possible.

Figure 19 indicates, how ViGWaM strides safely over the barrier, into the footprints, and climbs the stairway employing the proposed vision based step sequence adaptation method. Figure 19(a,b) shows the approach of ViGWaM to the front of the barrier with different step lengths, which is surmounted with high clearance steps in (c). After the barrier (d) the steps for the three footprints are planned. Figure 19 (e-g) depicts the locomotion on the step trace of the footprints. The stairway approach is shown in Figure 19 (h,i) and Figure 19 (j-l) demonstrates the successful stair climbing of ViGWaM; see also www.lsr.ei.tum.de/~vigwam for video sequences of ViGWaM.

7 Conclusions

In this article we have presented a novel control architecture for smooth goal-oriented vision-based guidance and control of walking machines. The cooperation and information exchange between the individual modules of this

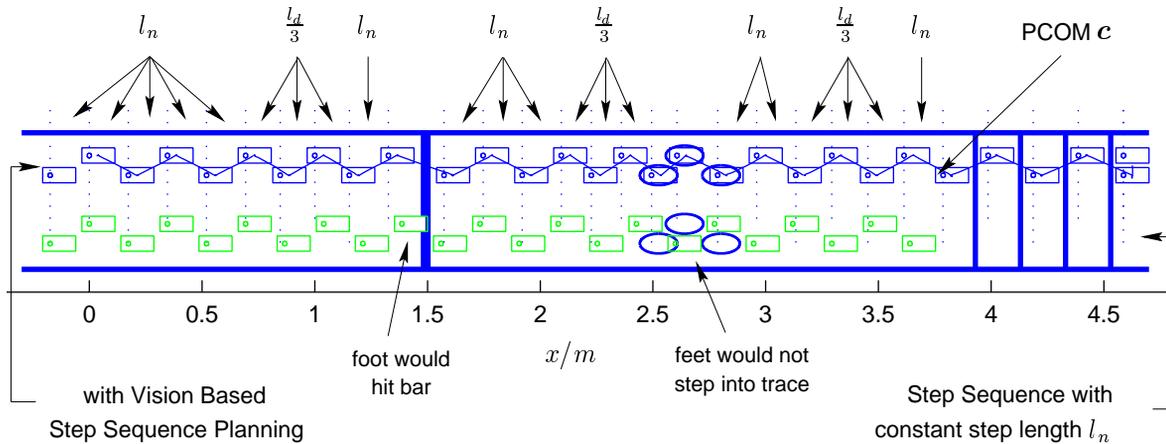


Fig. 18. Resulting step sequence with and without vision-based step length adaptation and progression of PCOM.

cascaded control approach has been biologically inspired. On the basis of this general approach to perception-based walking the key research directions of walking pattern synthesis, step sequence planning, gaze control strategies, and development of robust image processing algorithms have been identified. An emulation environment is used to validate the vision-based walking pattern and step sequence planning strategies to guide the humanoid walking machine ViGWaM in a modular prototypical scenario structured similarly to indoor and outdoor environments.

For locomotion over different prototypical obstacles, an optimization approach to generate statically stable walking patterns has been proposed. The walking patterns are computed offline for various step lengths, clearance heights during the steps and stored in a walking pattern database for online application. On the basis of a distance measurement of the walking machine to obstacles, a step sequence planning strategy combines the walking patterns such that obstacles can be surmounted or the feet can be placed into certain desired locations of a step trace.

The developed strategy for gaze control uses an evaluation function to estimate the information content of visible, tracked features in the camera images. The index sums over visible features and decreases with increasing distance of a feature from the walking machine and from the center of the camera image. This overall information content estimate depends on the pan-tilt angles of the stereo camera head. In the prototypical scenario straight walking along a line is considered. Hence, again by optimization the optimal tilt angle to maximize the visual information content can be determined. According to this angle the pan-tilt head is pointed in this “best” view direction.

Robust image processing algorithms for feature recognition and tracking are advantageous in general and especially important in walking machines because of possible rotations of the cameras around the optical axes. In this article, we proposed line- and point-based object recognition algorithms combined with online tracking with fast and robust tracking algorithms from the XVISION-library. Image processing results have shown that feature distances can be measured with an accuracy of about 0.015 m . This accuracy is more than sufficient for adequate step sequence planning to surmount the obstacles of the prototypical environment.

The experimental emulation setup combines head motion emulation of the statically stable ViGWaM using a wheeled mobile platform, a SCARA-type robot arm, and a pan-tilt-head with a stereo camera pair. The motion behavior of ViGWaM is visualized in an augmented reality using 3D-graphics layed over an image of an external camera observing the prototypical scenario. The presented ViGWaM emulation approach provides head motion of different types of legged and non-legged machines and is an effective rapid prototyping platform for the analysis and understanding of vision-based guidance towards goal-oriented locomotion.

Future work will focus on the approximative online simulation of a dynamically stable walking machine, to achieve smoother and faster machine walking, and extension of the image processing algorithms as well as the gaze (view direction) control behavior to more general environments.

8 Acknowledgments

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

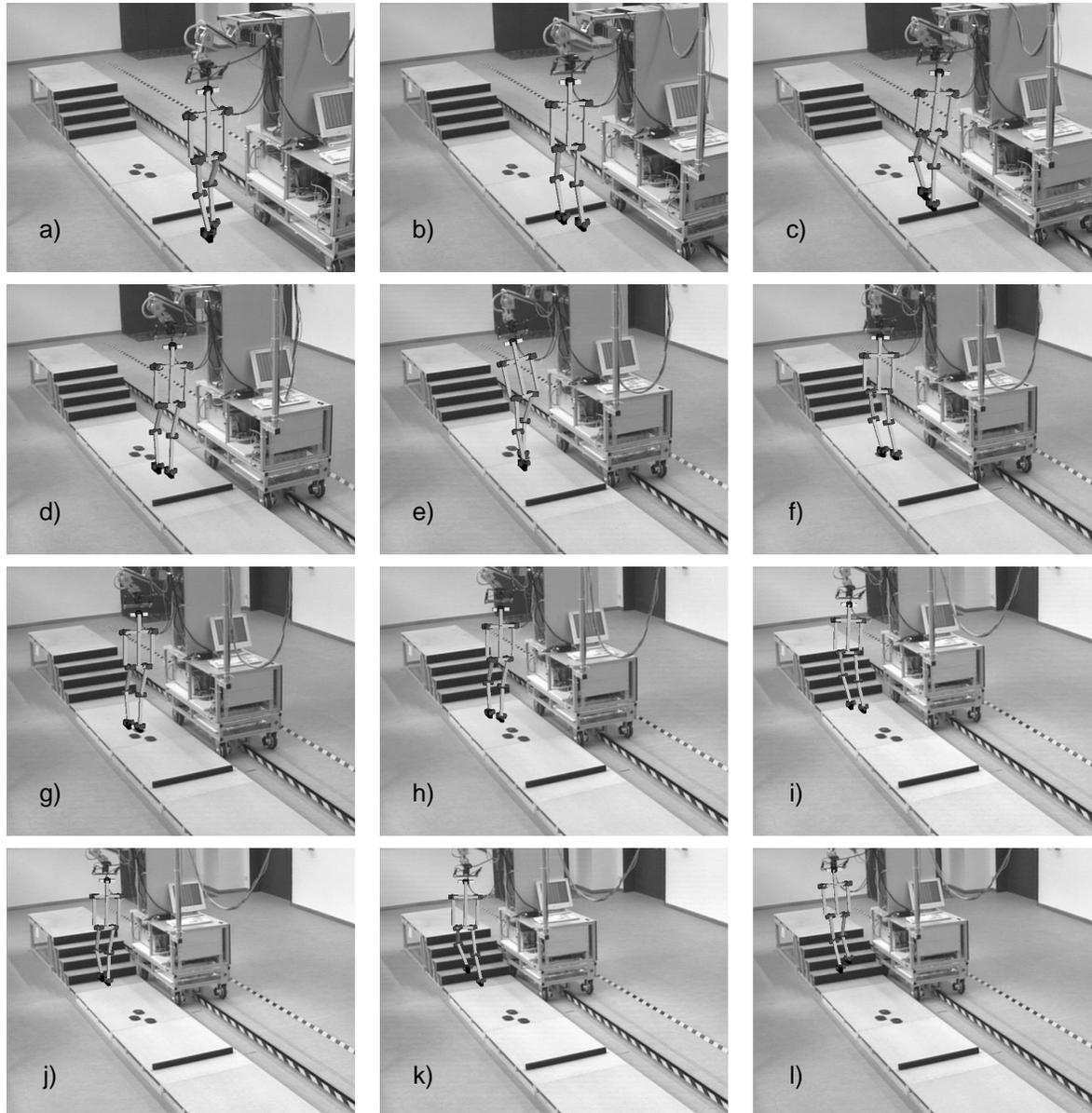


Fig. 19. ViGWaM walking over the scenario with three basic types of obstacles, a barrier, footprints (step trace), and stairs.

References

- [1] Aftab E. Patla. Understanding of roles of vision in the control of human locomotion. *Gait and Posture*, 5:54–69, 1997.
- [2] Aftab E. Patla and Joan N. Vickers. Where and when do we look as we approach and step over an obstacle in the travel plath. *NeuroReport*, 8(17):3661–3665, December 1997.
- [3] A. Albert and J. Hofschulte. Trajektorienplanung für Bewegungen eines autonomen, zweibeinigen Roboters im menschlichen Lebensraum. *at—Automatisierungstechnik*, 48(6):296–304, 2000.
- [4] Y. Aloimonos. Introduction: Active Vision Revisited. Technical report, University of Maryland, 1999.
- [5] T. Arakawa and T. Fukuda. Natural Motion Generation of Biped Locomotion Robot using Hierarchical Trajectory Generation Method Consisting of GA, EP Layers. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 211–216, Albuquerque, New Mexico, 1997.
- [6] C. Chevallereau, A. Formal'sky, and B. Perrin. Control of a Walking Robot with Feet Following a Reference Trajectory Derived from Ballistic Motion. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1094–1099, Albuquerque, New Mexico, 1997.
- [7] Fred R. Sias Jr. and Yuan F. Zheng. How Many Degrees-of-Freedom Does a Biped Need. In *Proceedings of the International Workshop on Intelligent Robots and Systems IROS*, pages 297–302, Tsuchiura, Ibaraki, Japan, 1990.

- [8] Y. Fujimoto and A. Kawamura. Simulation of an Autonomous Biped Walking Robot Including Environmental Force Interaction. *IEEE Robotics and Automation Magazine*, 5(2):33–42, June 1998.
- [9] Gal Sela and Martin D. Levine. Real-Time Attention for Robotic Vision. *Real Time Imaging*, 3:173–194, 1997.
- [10] Gero von Randow. *Roboter: unsere naechsten Verwandten*. Rowohlt, 1997.
- [11] Gill A. Pratt. Legged Robots at MIT: What is new since Raibert? *Climbing and Walking Robots: CLAWAR News*, 4:4–7, February 2000.
- [12] G. D. Hager and K. Toyama. The XVision System: A General-Purpose Substrate for Portable Real-Time Vision Applications. *Computer Vision and Image Understanding*, 69(1):23–37, January 1998.
- [13] K. Hirai. Current and Future Perspective of Honda Humanoid Robot. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS*, pages 500–508, Grenoble, France, 1997.
- [14] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka. The Development of Honda Humanoid Robot. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1321–1326, Leuven, Belgium, 1998.
- [15] K. Hosoda, M. Kamado, and M. Asada. Vision-Based Servoing Control for Legged Robots. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3154–3159, Albuquerque, New Mexico, 1997.
- [16] M. Inaba, T. Igarashi, S. Kagami, and H. Inoue. A 35 DOF Humanoid that can Coordinate Arms and Legs in Standing up, Reaching and Grasping an Objekt. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS*, pages 29–36, Osaka, Japan, 1996.
- [17] S. Kagami, K. Okada, M. Kabasawa, Y. Matsumoto, A. Konno, M. Inaba, and H. Inoue. A Vision Legged Robot as a Research Platform. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS*, pages 235–240, Victoria, B.C., Canada, 1998.
- [18] S. Kajita and K. Tani. Study of Dynamic Biped Locomotion on Rugged Terrain. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1405–1411, Sacramento, California, 1991.
- [19] S. Kajita and K. Tani. Adaptive Gait Control of a Biped Robot based on Realtime Sensing of the Ground Profile. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 570–577, Minneapolis, Minnesota, 1996.
- [20] F. Kanehiro, M. Inaba, and H. Inoue. Development of a Two-Armed Bipedal Robot That Can Walk and Carry. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS*, pages 23–28, Osaka, Japan, 1996.
- [21] J. Laci, H. Hooshang, and C. Bradley. A Control Strategy for Adaptive Bipedal Locomotion. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 563–569, Minneapolis, Minnesota, 1996.
- [22] K.-P. Lee, T.-W. Koo, and Y.-S. Yoon. Real-Time Dynamic Simulation of Quadruped Using Modified Velocity Transformation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1701–1706, Leuven, Belgium, 1998.
- [23] O. Lorch, M. Buss, and G. Schmidt. An Emulation Environment for the Development of a Vision-Based Virtual Walking Machine. *International Journal Automation Austria*, 7(2):57–66, 1999.
- [24] K. Nagasaka, H. Inoue, and M. Inaba. Dynamic Walking Pattern Generation for a Humanoid Robot Based on Optimal Gradient Method. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pages 908–913, Tokyo, Japan, 1999.
- [25] K. Nishiwaki, K. Nagasaka, M. Inaba, and H. Inoue. Generation of Reactive Stepping Motion for a Humanoid by Dynamically Stable Mixture of Pre-designed Motions. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pages 902–907, Tokyo, Japan, 1999.
- [26] D. 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.
- [27] J. H. Park and K. D. Kim. Biped Robot Walking Using Gravity-Compensated Inverted Pendulum Mode and Computed Torque Control. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 3528–3533, Leuven, Belgium, 1998.
- [28] T. Viéville. *A few steps towards 3D Active Vision*, volume 33. Springer Series in Information Sciences, 1997.
- [29] M. Vukobratović and M. Kirčanski. *Kinematics and Trajectory Synthesis of Manipulation Robots*. Springer-Verlag, Berlin, Germany, 1 edition, 1986.
- [30] 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.