Online Environment Reconstruction for Biped Navigation

Philipp Michel†, Joel Chestnutt†, Satoshi Kagami‡, Koichi Nishiwaki‡, James Kuffner†‡ and Takeo Kanade†‡

†The Robotics Institute
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
{pmichel,chestnutt,kuffner,kanade}@cs.cmu.edu

‡Digital Human Research Center
National Institute of Advanced Industrial Science and Technology
2-41-6, Aomi, Koto-ku, Tokyo
135-0064, Japan
{s.kagami,k.nishiwaki}@aist.go.jp

***Abstract*— As navigation autonomy becomes an increasingly important research topic for biped humanoid robots, efficient approaches to perception and mapping that are suited to the unique characteristics of humanoids and their typical operating environments will be required. This paper presents a system for online environment reconstruction that utilizes both external sensors for global localization, and on-body sensors for detailed local mapping. An external optical motion capture system is used to accurately localize on-board sensors that integrate successive 2D views of a calibrated camera and range measurements from a SwissRanger SR-2 time-of-flight sensor to construct global environment maps in real-time. Environment obstacle geometry is encoded in 2D occupancy grids and 2.5D height maps for navigation planning. We present an on-body implementation for the HRP-2 humanoid robot that, combined with a footstep planner, enables the robot to autonomously traverse dynamic environments containing unpredictably moving obstacles.**

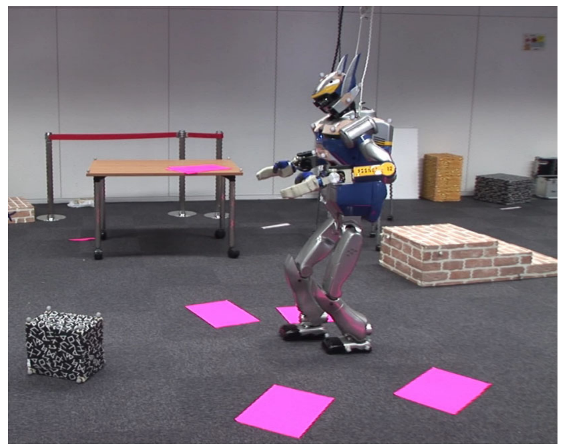


Fig. 1. The HRP-2 humanoid navigating autonomously through a set of obstacles during vision-guided walking.

I. Introduction

Significant progress has been made towards stable robotic bipedal walking, leading to an increased research interest in developing autonomous navigation strategies tailored specifically to humanoid robots. The ability of legged robots to step not only around but also over and onto some obstacles makes them particularly well suited for environments designed for humans, which often contain a wide variety of objects and obstacles such as furniture, stairs, doors, and uneven ground. A key requirement for humanoid navigation autonomy is the ability to build reliable environment representations from sensor data upon which a navigation planner can operate. To efficiently plan walking paths leading to a particular goal location, global maps of the robot's surroundings must be generated. However, on-body perception is inherently limited in range and direction. Some traditional approaches to perception have thus focused on recovering local, robot-centric environment information, typically allowing only short-term navigation strategies such as reactive obstacle avoidance to be implemented. Others have used off-body visual sensing to compute global maps, sacrificing some degree of robot autonomy as well as being susceptible to sensing limitations such as occlusions or parallax induced by a fixed view of the scene.

In this paper, we present an approach to environment reconstruction for the purposes of biped robot navigation that uses on-body sensors localized globally in an indoor environment. Given a calibrated camera, such localization allows for a 2D occupancy grid of the floor area in terms of obstacles and free space to be recovered from the video stream using image warping and segmentation. When the same method is applied to localize a range sensor, distance measurements can be converted into 2.5D height maps of the robot's surroundings. In both cases, maps are constructed by integrating sensor data accumulated over time as the robot moves through the environment. While each sensor measurement thus only reconstructs a part of the environment seen from a partial view of the scene, accurate sensor localization allows successive measurements to be correctly placed in a global map of the robot environment, which can then be used for path planning.

The remainder of this paper is organized as follows: Section II presents some related research. In Section III, we describe how images from a calibrated and localized camera can be warped to reconstruct the ground plane. The generation of height maps using range data from a SwissRanger SR-2 time-of-flight sensor is detailed in Section IV. Section V presents results from our implementation on an HRP-2 humanoid, shown in Figure 1, which combines the generated environment maps with a footstep planner to accomplish autonomous biped locomotion. Finally, Section VI gives a summary, discussion and pointers to future work.

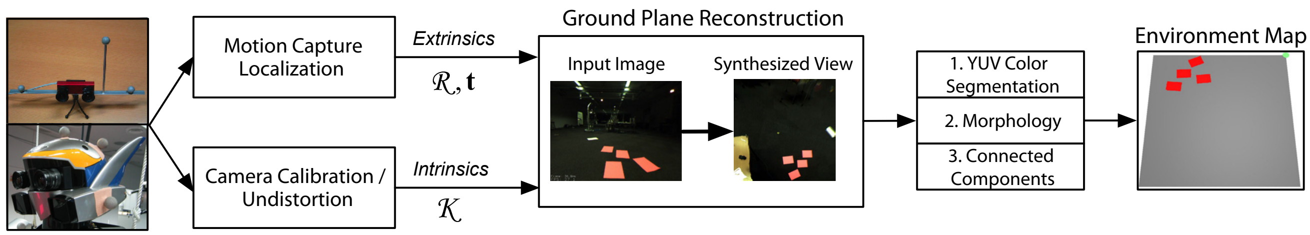


Fig. 2. Overview of the image-based reconstruction process. An environment map of obstacles on the floor is constructed from a calibrated, moving camera localized using motion capture.

II. Related Work

Efficient and accurate perception plays a crucial role in any complete approach to robotic navigation. For biped humanoids with many degrees of freedom, fast movement, and sensor instability induced by walking, environment reconstruction is particularly challenging.

Full estimation of the 3D scene geometry and camera ego-motion by image feature tracking and structure from motion has been extensively studied (e.g. [1]). However, its computational complexity has prevented efficient online robotic implementations. Instead, recent research has focused on combining stereo correspondence with robot-centric maps [2] or with localization using dead-reckoning or visual odometry, leading to several successful on-body implementations for humanoids [3], [4]. Drawbacks to such approaches include the still considerable processing required for stereo, reliance on highly textured scenes, and the accumulation of mapping error over time that is characteristic of visual odometry. Until online SLAM techniques become viable in terms of accuracy and efficiency, accurate global localization will likely require some instrumentation of the environment. For example, techniques utilizing known landmarks, indoor GPS systems, or optical motion capture systems offer localization information free of cumulative errors and have reliable accuracy within the instrumented area. Our experiments have shown that accurate global localization, when combined with a range sensor that is compact enough for humanoid applications and less computationally demanding, enables real-time 3D scene mapping with sufficient performance and accuracy for autonomous navigation. In our system, we utilize optical motion capture data to perform accurate global localization of the on-board sensors, which is important for computing reliable maps.

For the purposes of biped locomotion, full 3D mapping may often not be necessary. Several previous systems have focused on representing the robot environment as 2D occupancy grids [5] or 2.5D height maps [6], using standard vision techniques as well as stereo [7]. Given knowledge about the environment, many biped navigation approaches thus far have implemented reactive perception-based obstacle avoidance [8]–[11]. Such methods either do not account for the global configuration of obstacles between the robot and the desired goal, or do so in a limited way, which makes locomotion in truly cluttered environments challenging. Others accomplish obstacle avoidance using approximate planning techniques [12] that do not fully exploit the unique stepping capabilities of bipeds. To compute efficient navigation strategies for humanoids, we have adopted a footstep planning approach as proposed by Kuffner et al. [13], [14] and Chestnutt et al. [15] that uses global environment representations, incrementally built as the robot navigates its environment.

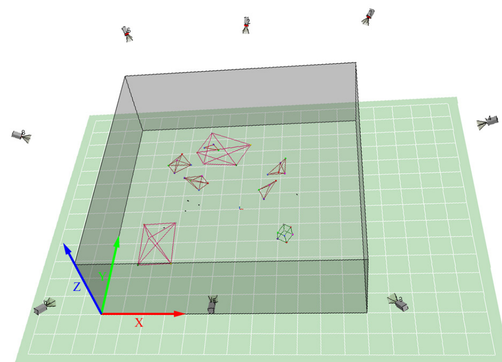


Fig. 3. View of the floor area covered by motion capture, showing world coordinate system and various registered objects being tracked.

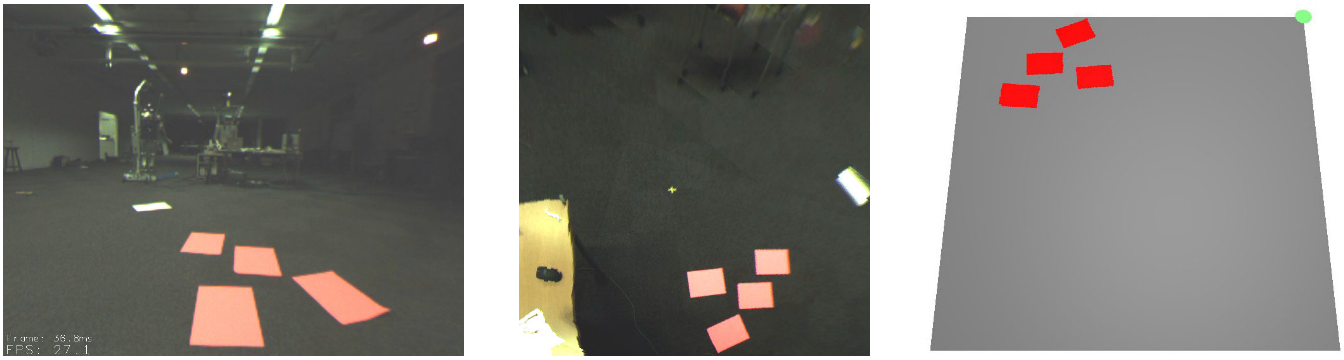
III. Reconstruction by Image Warping

Using images from a calibrated on-body camera, we reconstruct the robot's surroundings as if viewed from a virtual camera mounted overhead and observing the scene in its entirety. By accurately tracking the camera's position using motion capture, we are able to recover the full projection matrix. This enables a 2D collineation, or homography, between the floor and the image plane to be established, allowing incoming camera images to be warped onto the ground plane. A step of obstacle segmentation then produces a 2D occupancy grid of the floor. All processing described is done in real-time or faster on commodity hardware. Figure 2 gives an overview of the reconstruction process.

A. Localization using Motion Capture

Our operating environment consists of a laboratory room outfitted with a Motion Analysis high speed motion capture system [16] comprised of 8 digital cameras operating at up to 240Hz and covering a $16m^2$ floor area. The room provides an indoor operating environment spacious enough for encompassing a wide range of potential humanoid activities. Figure 3 contains a diagram of the setup.

For image-based reconstruction, we use a firewire camera mounted on the robot head or a separate metal frame, both equipped with retroreflective markers and shown in Figure 4.



(a) (b) (c)

Fig. 5. Example camera image *(a)*, taken from the frame-mounted camera. Synthesized top-down view *(b)* of the ground plane. Corresponding environment map generated *(c)*.

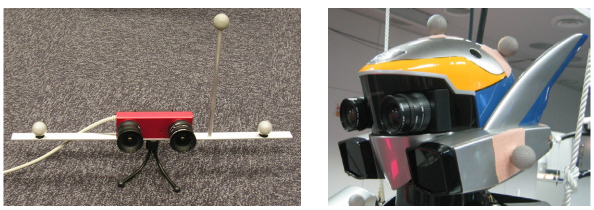


Fig. 4. Marker equipped camera bodies used for localization: metal frame *(left)* and humanoid head *(right)*. Only one camera from the stereo pair is used.

The motion capture system tracks and outputs the 3D positions of the markers on the camera body in real-time, allowing its location and orientation to be established. The transform defined by the known arrangement of markers on the body can then be used to recover a translation vector $\mathbf{t}$ corresponding to the world coordinates of the camera's optical center as well as a 3×3 rotation matrix $\mathcal{R}$ representing the direction of the optical axis. $\mathcal{R}$ and $\mathbf{t}$ thus define the rigid transform representing the camera's extrinsic parameters.

B. Ground Plane Reconstruction

To recover the intrinsic parameters and the radial distortion coefficients, we initially perform camera calibration offline based on Tsai's camera model [17]. During operation, incoming camera images are then rectified on-the-fly.

By combining the 3×3 upper triangular intrinsic parameter matrix $\mathcal{K}$ with the matrix of extrinsics established previously using motion capture, we are able to recover the full 3×4 camera projection matrix $\mathcal{M}$ as:

$$\mathcal{M} = \mathcal{K} \left[\begin{array}{cc} \mathcal{R} & \mathbf{t} \end{array} \right] \tag{1}$$

$\mathcal{M}$ uniquely maps a scene point (X, Y, Z) to a point on the image plane (u, v). Absent any scene constraints, however, the inverse mapping does not allow a unique scene point to be recovered for a given image point, but rather yields the equation of a ray along which the scene point must lie. For the purpose of building a 2D occupancy grid of the environment for biped navigation, however, we can assume that all scene points of interest lie in the $Z = 0$ plane. This assumption implies that scene elements significantly violating the planar constraint will appear distorted in the synthesized top-down view of the ground plane.

Scene planarity then allows ground plane points $\mathbf{q} = (X, Y, 1)^T$ in homogeneous coordinates to be related to points $\mathbf{p} = (u, v, 1)^T$ in the image plane via a 3×3 ground-image homography matrix $\mathcal{H}$ as $\mathbf{p} \equiv \mathcal{H}\,\mathbf{q}$. $\mathcal{H}$ can be constructed from the projection matrix $\mathcal{M}$ by considering the full camera projection equation

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ m_1 & m_2 & m_3 & m_4 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2}$$

and realizing that the constraint $Z = 0$ cancels the contribution of column m_3. $\mathcal{H}$ is thus simply composed of columns m_1, m_2 and m_4, yielding the desired 3×3 planar homography, defined up to scale with 8 degrees of freedom.

The recovered homography matrix is square and hence easily inverted. $\mathcal{H}^{-1}$ can then be used to warp incoming camera images onto the ground plane and thus accumulate an output image, resembling a synthetic top-down view of the floor area. In practice, applying $\mathcal{H}$ to points on the ground plane and then sampling pixel values (with optional sub-pixel interpolation) from their corresponding points in the image plane produces an improved result. The ground plane image is thus incrementally constructed in real-time as the camera moves through the scene, selectively yielding information about the environment. Updates to the output image proceed by overwriting previously stored data. The recovered view implicitly supports varying level of detail, achieved by moving the camera closer to the scene area of which a higher resolution reconstruction is desired. Figure 5(a) shows a camera image from a typical sequence and Figure 5(b) the corresponding synthesized floor view.

C. Obstacle Segmentation

We employ colored markers to denote the planar obstacles on the floor used for image-based reconstruction. To build an

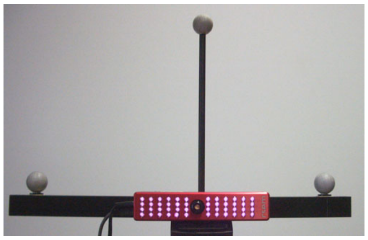


Fig. 6. The SwissRanger SR-2 time-of-flight sensor mounted on marker-equipped metal frame.

occupancy grid from the synthesized top-down floor view, a step of color segmentation is performed in YUV space [18], which provides robustness to color intensity changes due to variability in environment lighting. We define segmentation thresholds by sampling pixel values offline for obstacles placed in a variety of locations on the floor. Simple color thresholding then produces a binary mask indicating whether each pixel corresponds to an obstacle or to the floor. To eliminate noise, a pass of erosion/dilation using a rectangular structuring element is performed, followed by a step of connected component labeling that groups pixels belonging to the same obstacle. Calculating the moments of each obstacle blob then provides information about its centroid, area, major/minor axes and orientation on the floor. Given the shapes and locations of the obstacles and the known world dimensions of the floor area, it is straightforward to generate binary environment maps of arbitrary grid size for the purposes of robot navigation. Figure 5(c) gives an example.

IV. Reconstruction from Range Data

In order to fully exploit the walking capabilities of biped humanoids, we would like to map complex environments containing arbitrary non-planar obstacles that the robot can step over, around, or onto during autonomous locomotion. To build 2.5D height maps of the robot's environment, we use a CSEM SwissRanger SR-2 time-of-flight (TOF) range sensor [19], mounted on a marker-equipped frame as shown in Figure 6 and tracked by motion capture in the same manner as described above. Being able to accurately and globally localize the sensor allows us to convert distance measurements into clouds of 3D points in world coordinates, from which environment height maps can be cumulatively constructed. The reconstruction process runs in real-time.

A. Sensor Calibration

The SwissRanger SR-2 emits radio-frequency modulated light in the near-infrared spectrum, which is backscattered by the scene and detected by a 124×160 pixel CMOS CCD. The distance of each sensor element to the reflecting scene object is calculated from the time-of-flight of the backscattered light, by measuring the phase delay between the emitted and detected signal at each pixel. The sensor has a reported operating range of up to 7.5m and sub-centimeter depth resolution. In contrast to stereo, a highly textured scene is not required. Also, TOF-based depth recovery requires significantly less computational resources than disparity estimation and can be done on-camera by an FPGA. Most importantly, the sensor's weight and dimensions enable flexible on-body placement on a humanoid robot.

The sensor ships in an uncalibrated state, with lens distortion unaccounted for and distance measurements not reflecting the true distance to scene objects. We first recover the intrinsic parameters and radial distortion coefficients using Zhang's camera model [20] during a standard checkerboard-based calibration step. This method works despite the sensor's comparatively low resolution and allows subsequent distance images to be rectified.

To obtain accurate range data, the raw distance measured at each sensor element must be corrected by a per-pixel offset. These offsets, as well as the exposure time, are specific to the interval within which range measurements are to be made. We compute the 124×160 offset mask and optimal exposure time by placing the sensor in front of a featureless plane (e.g. a wall surface), at a known distance that lies in the desired measurement interval, and comparing the raw measurements to ground truth.

B. Building Height Maps

Having corrected the range measurements with the offset mask and estimated the sensor's focal length during calibration, we are able to convert per-pixel distances (u, v, d) into camera-centric 3D coordinates $\mathbf{P}_C = (X, Y, Z)^T$ of the measured scene points. The extrinsic parameters $\mathcal{R}$ and $\mathbf{t}$ localizing the sensor and recovered using motion capture allow us to construct the 4×4 transform matrix converting between the world frame and the camera frame in homogeneous coordinates as:

$$\mathcal{T} = \begin{bmatrix} & \mathcal{R} & & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$\mathcal{T}$ is easily inverted and can then be used to reconstruct each measured scene point in world coordinates via

$$\mathbf{P}_W \equiv \mathcal{T}^{-1}\mathbf{P}_C \tag{4}$$

The maximum Z value of the measured scene points over a given position on the floor can then be recorded to build 2.5D height maps of the environment, which ignore vertical concavity but are particularly suitable representations for biped navigation planning. We found that reconstruction was significantly improved by a step of median filtering of the height values, leading to more accurate recovery of objects with smooth horizontal surfaces such as boxes, tables, chairs and stairs. We also imposed a distance cutoff that effectively ignores scene points whose measured distance to the sensor lies outside of the range interval it was calibrated for. Figure 7 shows an example reconstruction.

V. Implementation & Results

We have implemented our reconstruction approach on an HRP-2 humanoid robot and combined it with a path planner that operates at the level of footsteps, allowing the biped's unique locomotion capabilities to be specifically exploited during autonomous navigation.

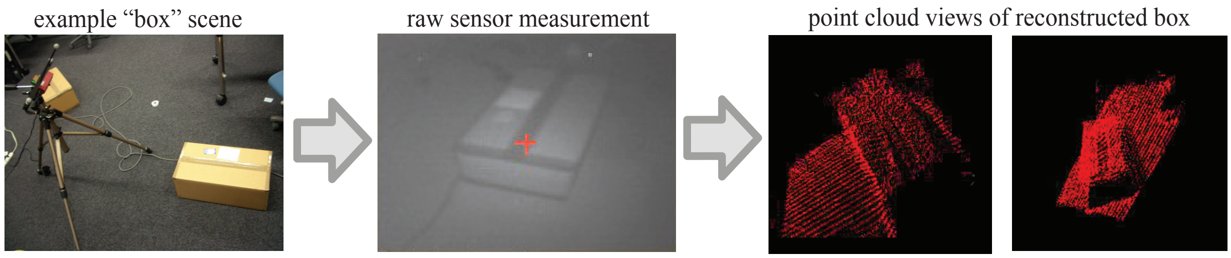


Fig. 7. Swiss Ranger viewing an obstacle box placed on the floor *(left)*. Raw measurements recorded by the sensor *(center)*. Two views of the 3D point cloud reconstruction of the box *(right)*.

We base our planner on the one described by Chestnutt et al. [21] and Michel et al. [5]. It computes an optimal sequence of footsteps that allow the robot to reach a particular goal location while avoiding obstacles. In particular, this also includes solutions that involve stepping over obstacles. For the experiments in this paper, the planner was configured to operate on 2D binary occupancy grids of the floor area, recovered in real-time using image warping. During the walking sequence, the path to the goal is dynamically adjusted by re-planning based on updated environment information gathered from sensing. The planner is given a time-limited horizon not exceeding one step cycle, with an updated plan being generated every time the robot's stance foot changes while walking. Planning does not proceed from scratch each time, however. When replanning, those parts of the previous footstep path still valid given the current environment are used to seed the search queue. Efficient plan reuse coupled with real-time environment reconstruction allows the robot to navigate dynamic environments containing unpredictably moving obstacles.

Figure 8(a) shows HRP-2 walking towards a goal location while the surrounding obstacles are being moved by a human. The experimenter was asked to intentionally block off the currently computed path to the goal. The planner quickly adjusts for the sudden occlusion, computing a sequence of footsteps that allows the robot to reach the goal. By having an experimenter additionally reposition the goal location, HRP-2 can traverse the room continuously without stepping on obstacles.

We have also combined our 2D environment reconstruction approach with an adaptive joystick-based steering method that allows for intuitive user control of the robot while still selecting safe alternate foot placements in the presence of obstacles [22]. Figure 8(b) shows HRP-2 under such joystick control. Given a simple 'forward' directional command, a detected obstacle is avoided by a sidestepping motion automatically computed using information derived online from the environment map.

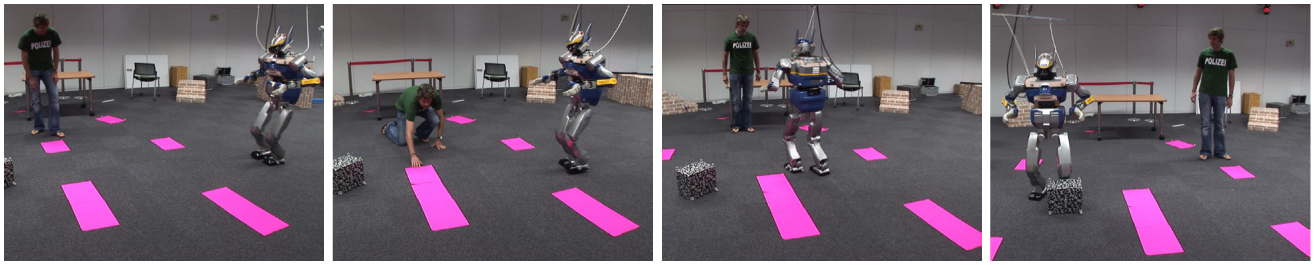
Being able to accurately establish the optical center and axis of the camera from the motion capture markers placed on the robot head is crucial for recovery using image warping. To optimize the marker-to-camera transform, one planar obstacle was outfitted with retroreflective markers. Its ground truth position on the floor was then gathered from motion capture and used to adjust the transform accordingly. This ensures precise alignment of the computed 2D occupancy grids with the real-world obstacles throughout the walking sequence, allowing the robot to step within less than 5cm of the obstacles during walking while still avoiding them. Residual errors in the reconstruction do not accumulate over time, a significant advantage of using motion capture to perform sensor localization when compared to visual odometry approaches, which rely on the tracking of scene features in order to recover camera motion and are prone to drift.

VI. Discussion

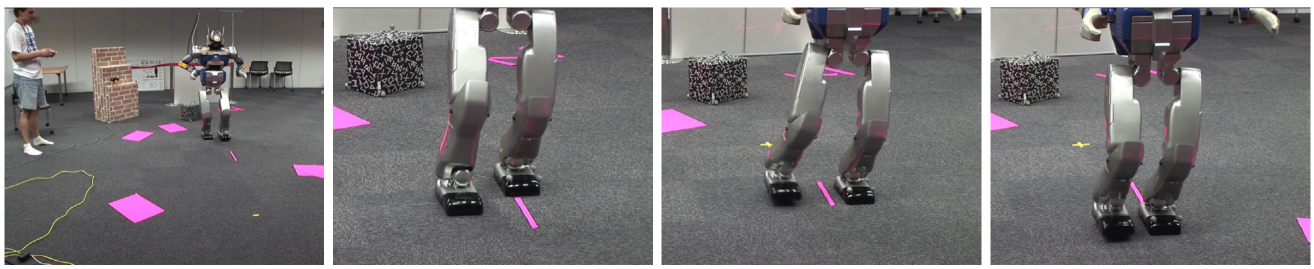
We have presented a method of environment reconstruction tailored to humanoid robots that uses motion capture to localize a calibrated camera or a range sensor in order to construct global representations of the robot's surroundings in real-time during locomotion. By combining the recovered environment representations with a planner operating at the level of footsteps and supporting efficient plan reuse, we have enabled an HRP-2 humanoid robot to autonomously navigate unknown, obstacle-filled environments in real-time.

In our experiments, we found that tight temporal synchronization of the sensor data with the positioning information from motion capture was crucial for accurate reconstruction. Ideally, one would like the motion capture to return positioning data for precisely the point in time when a sensor measurement was acquired. Delays in acquisition, processing and communications make this difficult and are responsible for a temporal discrepancy between motion capture and sensor data. As a remedy, we keep a manually adjusted offset into a history of the last 100 motion capture frames and use the measurement that visually minimizes error and jitter in the reconstruction. We are currently investigating more sophisticated methods, including improved cameras with global shutter and better ways of sequencing and comparing both sensor and motion capture data.

We are also presently working on integrating the SwissRanger SR-2 on HRP-2's body, by mounting it on the robot's arm for added mobility. Because the sensor needs to be calibrated for a certain overall measurement range, only those parts of the scene lying in the calibrated interval can be properly reconstructed, making large-scale footstep planning with height maps difficult. As an alternative, we are focusing on using the range finder to accurately reconstruct the 3D



(a)



(b)

Fig. 8. Examples of online environment reconstruction used during biped locomotion: HRP-2 navigating autonomously towards a goal location (box) with unpredictably moving obstacles *(a)*. Avoiding a thin, elongated obstacle during a joystick controlled walking sequence *(b)*.

shape of obstacles in close proximity to the robot. In particular, we intend to use the reconstruction method outlined in this paper to accurately estimate environment obstacle geometry, such as the height and spacing of steps in front of the robot for the purposes of stair and obstacle climbing. These additional capabilities will allow the robot to more freely navigate human environments of increasing complexity.

References

[1] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *Int. J. of Computer Vision*, vol. 9, no. 2, pp. 137–154, November 1992.

[2] K. Sabe, M. Fukuchi, J.-S. Gutmann, T. Ohashi, K. Kawamoto, and T. Yoshigahara, "Obstacle avoidance and path planning for humanoid robots using stereo vision," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'04)*, New Orleans, LA, April 2004.

[3] Y. Takaoka, Y. Kida, S. Kagami, H. Mizoguchi, and T. Kanade, "3D map building for a humanoid robot by using visual odometry," in *Proc. of the IEEE Int. Conf. on Systems, Man & Cybernetics (SMC'04)*, October 2004, pp. 4444–4449.

[4] S. Kagami, Y. Takaoka, Y. Kida, K. Nishiwaki, and T. Kanade, "Online dense local 3D world reconstruction from stereo image sequences," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'05)*, August 2005, pp. 2999–3004.

[5] P. Michel, J. Chestnutt, J. Kuffner, and T. Kanade, "Vision-guided humanoid footstep planning for dynamic environments," in *Proc. of the IEEE-RAS/RSJ Int. Conf. on Humanoid Robots (Humanoids'05)*, December 2005, pp. 13–18.

[6] J.-S. Gutmann, M. Fukuchi, and M. Fujita, "A floor and obstacle height map for 3D navigation of a humanoid robot," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'05)*, Barcelona, Spain, April 2005.

[7] S. Kagami, K. Nishiwaki, J. Kuffner, K. Okada, M. Inaba, and H. Inoue, "Vision-based 2.5D terrain modeling for humanoid locomotion," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'03)*, September 2003, pp. 2141–2146.

[8] M. Yagi and V. Lumelsky, "Local on-line planning in biped robot locomotion amongst unknown obstacles," *Robotica*, vol. 18, no. 4, pp. 389–402, 2000.

[9] M. Yagi and V. J. Lumelsky, "Biped robot locomotion in scenes with unknown obstacles," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'99)*, Detroit, MI, May 1999, pp. 375–380.

[10] O. Lorch, J. Denk, J. F. Seara, M. Buss, F. Freyberger, and G. Schmidt, "ViGWaM - an emulation environment for a vision guided virtual walking machine," in *Proc. of the IEEE-RAS/RSJ Int. Conf. on Humanoid Robotics (Humanoids 2000)*, 2000.

[11] R. Cupec, O. Lorch, and G. Schmidt, "Vision-guided humanoid walking - concepts and experiments," in *Proc. of the 12th Int. Workshop on Robotics in Alpe-Adria-Danube Region (RAAD'03)*, Cassino, Italy, May 2003.

[12] J.-S. Gutmann, M. Fukuchi, and M. Fujita, "Real-time path planning for humanoid robot navigation," in *Proc. of the Int. Joint Conf. on Artificial Intelligence (IJCAI'05)*, Edinburgh, UK, July 2005, pp. 1232–1237.

[13] J. Kuffner, K. Nishiwaki, S. Kagami, M. Inaba, and H. Inoue, "Footstep planning among obstacles for biped robots," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'01)*, 2001, pp. 500–505.

[14] J. Kuffner, K. Nishiwaki, S. Kagami, Y. Kuniyoshi, M. Inaba, and H. Inoue, "Online footstep planning for humanoid robots," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'03)*, 2003.

[15] J. Chestnutt, J. Kuffner, K. Nishiwaki, and S. Kagami, "Planning biped navigation strategies in complex environments," in *Proc. of the IEEE-RAS/RSJ Int. Conf. on Humanoid Robots (Humanoids'03)*, Munich, Germany, October 2003.

[16] Motion Analysis Corporation, "Motion Analysis Eagle Digital System," Web: http://www.motionanalysis.com.

[17] R. Y. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Miami Beach, FL, 1986, pp. 364–374.

[18] J. Bruce, T. Balch, and M. Veloso, "Fast and inexpensive color image segmentation for interactive robots," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS-2000)*, Japan, October 2000.

[19] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger)," Available online at: http://www.swissranger.ch, 2004.

[20] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. of the Int. Conf. on Computer Vision (ICCV '99)*, Corfu, Greece, September 1999, pp. 666–673.

[21] J. Chestnutt, M. Lau, G. Cheung, J. Kuffner, J. Hodgins, and T. Kanade, "Footstep planning for the Honda ASIMO humanoid," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'05)*, April 2005.

[22] J. Chestnutt, P. Michel, K. Nishiwaki, J. Kuffner, and S. Kagami, "An intelligent joystick for biped control," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA'06)*, May 2006.