

Electronic version of an article published as
International Journal of Humanoid Robotics
Vol. 12, No. 3 (2015) 155031
DOI: 10.1142/S0219843615500310
© World Scientific Publishing Company

EXTRACTION OF WHOLE-BODY AFFORDANCES FOR LOCO-MANIPULATION TASKS

PETER KAISER, NIKOLAUS VAHRENKAMP
FABIAN SCHÜLTJE, JÚLIA BORRÀS and TAMIM ASFOUR

*Institute for Anthropomatics and Robotics (IAR)
Karlsruhe Institute of Technology (KIT)
Adenauerring 2, 76131 Karlsruhe, Germany
peter.kaiser@kit.edu*

Received 26 April 2015
Revised 7 May 2015
Accepted 10 June 2015

Humanoid robots that have to operate in cluttered and unstructured environments, such as man-made and natural disaster scenarios, require sophisticated sensorimotor capabilities. A crucial prerequisite for the successful execution of whole-body locomotion and manipulation tasks in such environments is the perception of the environment and the extraction of associated environmental affordances, i.e. the action possibilities of the robot in the environment. We believe that such a coupling between perception and action could be a key to substantially increase the flexibility of humanoid robots.

In this paper, we approach the affordance-based generation of whole-body actions for stable locomotion and manipulation. We incorporate a rule-based system to assign affordance hypotheses to visually perceived environmental primitives in the scene. These hypotheses are then filtered using extended reachability maps that carry stability information, for identifying reachable affordance hypotheses. We then formulate the hypotheses in terms of a constrained inverse kinematics problem in order to find whole-body configurations that utilize a chosen set of hypotheses.

The proposed methods are implemented and tested in simulated environments based on RGB-D scans as well as on a real robotic platform.

Keywords: Whole-Body; Affordances; Locomotion; Manipulation; Humanoids.

1. Introduction

One of the most fundamental questions in robotics research is how to enable robots to autonomously interact with unknown environments. This problem has been partially addressed by numerous works that try to bridge the gap between low-level control and high-level abstract reasoning.^{1,2,3,4} Most of these publications focus on manipulation tasks with single robotic arms or upper body humanoids with mobile platforms. Bipedal humanoid robots add more complexity to the problem, in terms

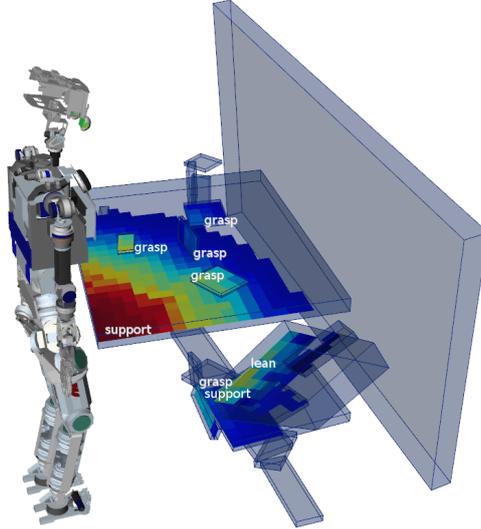


Fig. 1: The proposed methods allow the identification of affordance hypotheses from active vision and inertial sensor data.

of their kinematic structure as well as in terms of the possible ways of interacting with the environment. Particularly, constraints on balance have to be satisfied, including the utilization of the environment to enhance stability.

This paper extends our previous work⁵ in which we presented a first step towards enabling humanoids to interact with unknown environments. We improved state-of-the-art techniques to fuse inertial and visual information to a depth model which is then segmented and categorized into geometric primitives. We assume that the environment is unknown and that a safe navigation in unstructured environments requires the ability to utilize walls and other objects for stabilization. We employ active vision in order to detect environmental primitives that the robot can possibly use for stabilizing interactions (see Fig. 1).

Due to the known structures of human-centered environments and the humanoid kinematics of the robot, we can assume some prior knowledge on the scene in order to infer affordance hypotheses based on shapes, sizes or orientations of the detected primitives. For instance, we can assume that vertical, large planes are probably walls that can afford to lean on them.

The concept of affordances was first proposed by JJ. Gibson⁶ in the context of ecological psychology. Since then, it has been applied to several fields of research, from cognitive science and neuropsychology to human-computer interaction and autonomous robotics. In the original psychological context, the main idea behind the concept of affordance was that perception is economical, i.e. instead of modeling the whole world, only the relevant environmental information is perceived.

In the context of autonomous robots, affordances have been used to simplify

complex tasks, such as grasp planning³. Works, e.g. by ten Pas and Platt⁷, show that grasp affordances have the potential to completely avoid complex grasp planning, by associating unknown objects with known geometries for which the robot has predefined grasps. Similarly, Bierbaum et al.² proposed practical and efficient solutions to grasp planning for unknown objects based on affordances and potential fields. We think that the research on whole-body motion with contacts, e.g. Sentis et al.^{8,9} or Lengagne et al.¹⁰ can greatly benefit from the use of whole-body affordances to break down the problem in separate parts that can be organized by a high level reasoning process.

In this work, we define a whole-body affordance hypothesis as an association of a whole-body stable action to a perceived primitive of the environment. Based on previous approaches^{2,11}, we aim at deriving, refining and utilizing whole-body affordances like holding, leaning, stepping-on or supporting in unknown environments.

For representation and execution of whole-body actions, we will rely on the concept of *Object-Action Complexes* (OACs)¹. The concept of OACs states that the execution of an action is tightly related to the object that the action involves. This viewpoint of objects and actions being coupled is related to the concept of affordances. One could think of affordances as preconditions for the instantiation of OACs.

To generate utilizable affordance hypotheses, we rely on the extension of manipulability maps¹² to whole-body stability maps¹³. Such maps are discrete representations of the robot's workspace. For each end effector pose, the extended reachability maps contain the best possible stability rating among the whole-body configurations that realize the respective end effector poses. We use stability maps for detecting affordances in reach and for computing feasible end effector poses for the utilization of affordances.

Based on these results we show that by formulating the problem of affordance utilization as a constrained inverse kinematics problem, we can compute whole-body robot configurations that realize a chosen set of affordances. Due to perceptual errors however, an additional step of whole-body control needs to be incorporated in order to actually establish stable contact with the environment. This step is currently left for future work.

We have implemented the perceptual pipeline and the affordance generation methods in our robot framework ArmarX¹⁴ and evaluated them in different unknown scenarios involving small and big objects, stairs and walls. We additionally implemented a simple verification strategy and let the humanoid robot ARMAR-III¹⁵ perceive and verify environmental affordances.

In the remainder of the paper, Section 2 describes the extension of reachability maps and Section 3 explains how affordances are assigned to detected environmental primitives, incorporating information from extended reachability maps. Section 4 discusses the generation of whole-body robot configurations based on a chosen set of affordances to utilize and Section 5 demonstrates a simple affordance verification

strategy on the humanoid robot ARMAR-III. Finally, Section 6 discusses the results and outlines our ideas for future work.

2. Extended Reachability Maps

In our previous work¹³ we proposed an extension of classic reachability maps^{16,17} in order to capture additional quality indices like end effector manipulability or whole-body stability, besides the raw reachability information. An extended reachability map for an end effector e is denoted as \mathcal{R}_e and the quality value stored in \mathcal{R}_e for the end effector pose $\mathbf{p} \in SE(3)$ is accessed as:

$$\mathcal{R}_e(\mathbf{p}) \in [0, 1] \quad (1)$$

Depending on the type of quality information stored with \mathcal{R}_e , it can also be referred to as \mathcal{S}_e , in case of stability information, or as \mathcal{M}_e , in case of manipulability information. A 3D visualization of exemplary reachability data is shown in Fig. 2.

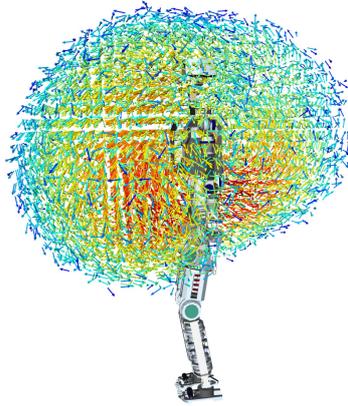


Fig. 2: A visualization of an extended reachability map containing whole-body support information for the right hand of the humanoid robot ARMAR-4¹⁸.

Several quality measures can be used together with extended reachability maps, such as stability, manipulability, end effector visibility, human-likeness of posture, or energy efficiency. We investigated two quality measures which are of high interest in humanoid robotics: *manipulability* and *stability*.

Manipulability measure

The manipulability index of an end effector pose expresses the robot's capability to adjust the pose in workspace. This capability can be important since it gives a whole-body controller the freedom to react on inaccuracies in perception and actuation as

well as on external disturbances while maintaining a desired end effector pose. To compute the manipulability index of a given robot configuration we use the extended manipulability formulation¹⁹ in order to consider joint limits and the robot's self-distance. An exemplary extended reachability map with manipulability information is depicted in Fig. 2.

Stability measure

Stability information allows to identify end effector poses which cannot be reached with a stable whole-body configuration. Even poses that are just reachable with a low stability value should be avoided since in such cases even small disturbances could make the robot fall. The quality measure $stability(\mathbf{c})$ is computed by projecting the center of mass (CoM) $\mathbf{x}_{com}(\mathbf{c}) \in \mathbb{R}^3$ of the current robot configuration $\mathbf{c} \in \mathbb{R}^n$ to the ground plane, resulting in $\mathbf{x}'_{com}(\mathbf{c}) \in \mathbb{R}^2$. An exemplary configuration of ARMAR-4 together with the support polygon and the CoM projection $\mathbf{x}'_{com}(\mathbf{c})$ is depicted in Fig. 3.

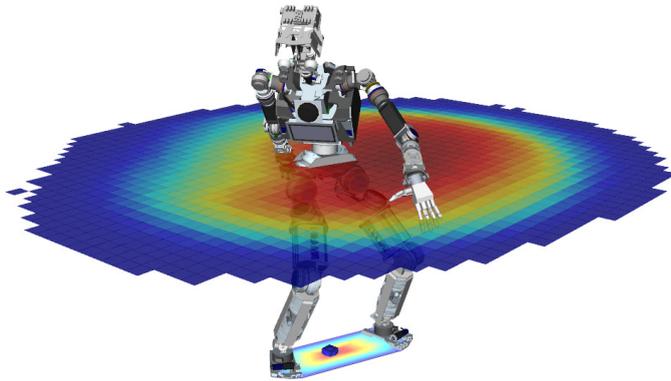


Fig. 3: Cut through the whole-body stability map for the left hand of the simulated robot ARMAR-4. The stability rating depends on the distance of the projected center of mass $\mathbf{x}'_{com}(\mathbf{c})$ (blue box) to the border of the support polygon s_c .

If $\mathbf{x}'_{com}(\mathbf{c})$ lies outside the support polygon s_c that is spanned by the contact points between feet and ground, the robot is not in a statically-stable configuration and the resulting quality value is set to zero. Otherwise, the distance of $\mathbf{x}'_{com}(\mathbf{c})$ to the support polygon's border ∂s_c is put in relation to the distance of the support polygon's center $\mathbf{x}_{center}(\mathbf{c})$ to ∂s_c (see Eq. 2).

$$stability(\mathbf{c}) = \frac{\min \{ \|\mathbf{x}'_{com}(\mathbf{c}) - \mathbf{y}\|, \forall \mathbf{y} \in \partial s_c \}}{\min \{ \|\mathbf{x}_{center}(\mathbf{c}) - \mathbf{y}\|, \forall \mathbf{y} \in \partial s_c \}} \in [0, 1] \quad (2)$$

2.1. Fusion of extended reachability maps

Individual extended reachability maps provide one type of workspace quality information. In order to combine different types of quality information, fusion operations on extended reachability maps become necessary. In this work we fuse a set of extended reachability maps $\mathcal{R}_1, \dots, \mathcal{R}_N$ by multiplying the individual quality values:

$$\mathcal{R}(\mathbf{p}) = \prod_{i=1}^N \mathcal{R}_i(\mathbf{p}) \quad (3)$$

However, different operations are possible, e.g. linear combinations for weighting the contribution of individual maps \mathcal{R}_i . Fig. 4 shows the map $\mathcal{R}_{Left\ Hand}$ resulting from the fusion of the extended reachability map $\mathcal{S}_{Left\ Hand}$ and $\mathcal{M}_{Left\ Hand}$ storing stability and manipulability information, respectively.

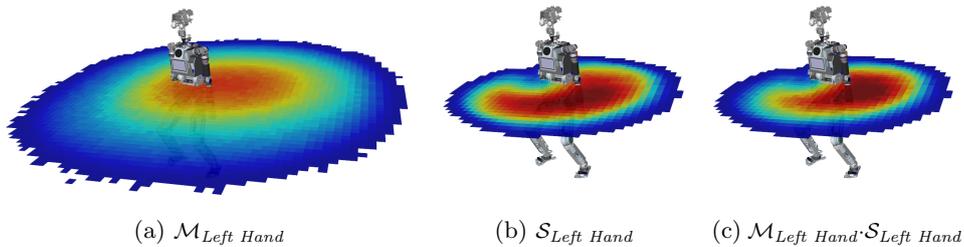


Fig. 4: Fusion of extended reachability maps with manipulability and stability information for the left hand $\mathcal{M}_{Left\ Hand}$ and $\mathcal{S}_{Left\ Hand}$ into a combined map $\mathcal{M}_{Left\ Hand} \cdot \mathcal{S}_{Left\ Hand}$. For clearer visualization, the map values are scaled to the full range of $[0, 1]$.

2.2. Evaluation of the map creation process

While querying an extended reachability map is efficient, the actual generation is a time-consuming process which is intended to be carried out offline. Fig. 5 displays the generation progress over time for an extended reachability map for ARMAR-4's right hand, containing stability information.

In this case the map contained about 40 million cells that have been filled to up to 12% with stability information after a generation time of 22 hours. It can be clearly seen, that a large part of the 6D workspace is not statically reachable due to kinematic limitations and stability constraints, i.e. for a 3D position only a small subset of the possible end effector orientations are reachable for the robot.

Since the map generation procedure is probabilistic, it can be speeded up with regard to multicore systems by running N generation processes in parallel. The resulting maps need to be merged afterwards in order to obtain a single result.

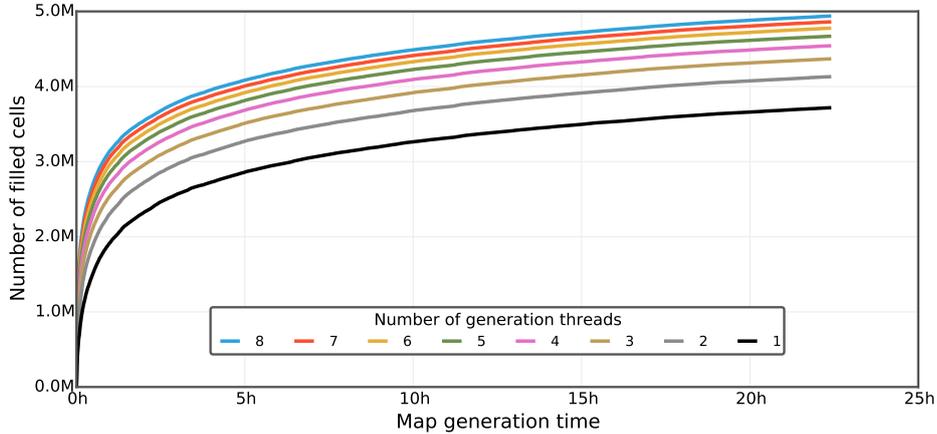


Fig. 5: Plots of the filling levels of an extended reachability map during its generation processes with 1 (black) up to 8 (blue) generation threads. In this example, an extended reachability map $\mathcal{S}_{Right\ Hand}$ was built, containing stability information for the right hand. The captured workspace contained a total of about 40,000,000 cells.

Fig. 5 shows the number of filled cells for different values of N , ranging from 1 to 8. The plots show the bounded nature of the map generation process. They also show that running multiple generation processes in parallel significantly speeds up the overall generation process. For example, while a single-threaded process is able to fill about 3.7 million cells in 22 hours, the combination of eight individual generation processes can fill the same amount of cells in about 2.6 hours. The benefit of employing additional generation processes decreases with the number of already running processes.

3. Extraction of Affordance Hypotheses

In this sections, strategies are proposed for suggesting affordance hypotheses based on visually perceived environmental primitives. Our preliminary experiments focus on affordances related to planar surfaces, although there is no principle limitation to these. Extension to curved surfaces, like cylindric or spherical ones, or volumetric primitives is possible and initial experiments have been conducted.

The methods for visual perception⁵ allow the detection and approximation of environmental surfaces based on RGB-D camera images. Fig. 6 shows the depth image of an exemplary scene (Fig. 6a) and the set of primitives resulting from the perception process (Fig. 6b).

In the following, the proposed process of affordance suggestion is explained. First, we pursue a rule-based assignment of affordance hypotheses to environmental

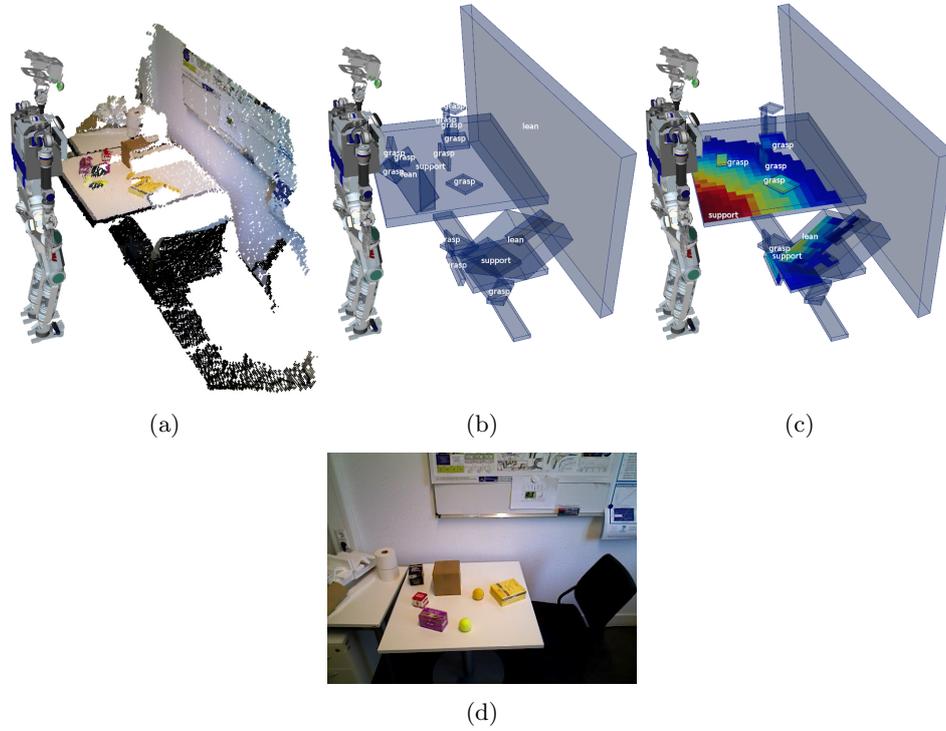


Fig. 6: Processing of an exemplary depth image (a) into a set of environmental primitives with assigned affordance hypotheses (b). Finally extended reachability maps are employed for determining reachable affordance hypotheses (c). The scene contains a table with several objects on it, a chair next to the table and a wall behind the table (d).

surfaces based on parameters like extent or orientation. Then, stability information from extended reachability maps is used for limiting the amount of found hypotheses to directly usable ones.

3.1. Suggestion of affordance hypotheses

Affordance hypotheses are suggested based on rules that incorporate parameters of the perceived primitives like orientation or extent. This approach eventually results in a set of rules that link geometric primitives to affordance hypotheses, similar to Varadarajan and Vicze²⁰. An exemplary set of such rules is given in Table 1.

For example, a planar surface that is sufficiently large and oriented horizontally, e.g. a table, suggests the affordance *support*. A long curved surface of a certain radius, e.g. a handrail, suggests the affordance *hold*. The last column of Table 1 describes the preferred end effector pose when utilizing the respective affordance,

Table 1: Example of a set of rules for affordance derivation. See Fig. 7 for \mathbf{x}_{eef} , \mathbf{y}_{eef} and \mathbf{z}_{eef} . The operator \uparrow tells if two vectors point into the same direction^a. The λ_i are implementation-specific constants.

Affordance	Surface	Parameters	Conditions	EEF
Support	Planar	Normal \mathbf{n} Area a	$\mathbf{n} \uparrow \mathbf{z}_{world}$ $a \geq \lambda_1$	$\mathbf{z}_{eef} \uparrow \mathbf{n}$
Lean	Planar	Normal \mathbf{n} Area a	$\mathbf{n} \perp \mathbf{z}_{world}$ $a \geq \lambda_2$	
Grasp	Planar	Normal \mathbf{n} Area a	$a \in [\lambda_3, \lambda_4]$	
	Curved	Radius r Direction \mathbf{d}	$r \in [\lambda_5, \lambda_6]$ $\ \mathbf{d}\ \leq \lambda_7$	$\mathbf{y}_{eef} \uparrow \mathbf{d}$
Hold	Curved	Radius r Direction \mathbf{d}	$r \in [\lambda_8, \lambda_9]$ $\ \mathbf{d}\ \geq \lambda_{10}$	

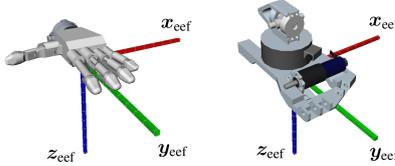


Fig. 7: The TCP coordinate systems for the left hand (left) and the left foot (right) of ARMAR-4.

(see also Fig. 7). This will be of interest in Section 3.2.

Using the rules outlined in Table 1, the system can identify several affordance hypotheses in the exemplary scene (see Fig. 6b). In the next steps, the resulting hypotheses are filtered according to their stable reachability, based on precomputed extended reachability maps.

3.2. Determination of reachable hypotheses

The previous sections show that, based on depth models obtained from active cameras, a robot can identify plenty of primitives p_i in a scene and is able to assign affordance hypotheses h_i to these primitives:

$$\mathcal{H} = \{(p_1, h_1), \dots, (p_k, h_k)\}. \quad (4)$$

For planning purposes it is important to identify \mathcal{H}_R , the subset of hypotheses that are directly reachable for the robot, either for utilization or for verification.

$$^a \mathbf{v} \uparrow \mathbf{w} \leftrightarrow \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|} \approx 1$$

For each affordance h_i , Table 1 constraints the set of possible end effector poses by fixing one axis of the end effector’s local coordinate system (see Fig. 7). The resulting constrained space of orientations will be denoted as $\Omega_{(p_i, h_i)}$. The geometric shapes of the primitives together with the suitable end effector poses allow us to assign a stability value to each point $\mathbf{x} \in \partial p_i$ on the surface of the primitive p_i :

$$stability_{(p_i, h_i)}(\mathbf{x}) = \max \{ \mathcal{S}_e(\mathbf{x}, \mathbf{R}) : \mathbf{R} \in \Omega_{(p_i, h_i)} \}. \quad (5)$$

This value tells how stable the robot would be when reaching for the different points on the primitive’s surfaces while maintaining the preferred end effector orientation.

Based on the stability value defined in Eq. 5, the set \mathcal{H}_R of reachable affordance hypotheses can be determined by omitting hypotheses for which the stability value lies below a threshold σ :

$$\mathcal{H}_R = \left\{ (p, h) \in \mathcal{H} : \exists \mathbf{x} \in \partial p : stability_{(p, h)}(\mathbf{x}) > \sigma \right\} \quad (6)$$

Fig. 6c depicts the result of the affordance assignment process. It shows only those affordances whose stability rating lies above a threshold σ . Furthermore, the affordance labels are attached to the points with the highest stability ratings.

3.3. Identification of end effector poses for affordance utilization

Extended reachability maps have already been used in order to reduce the full set of inferred affordance hypotheses \mathcal{H} to the set of reachable hypotheses \mathcal{H}_R . Based on the rules in Table 1, we can infer the set of possible end effector poses for each point on the surface of an environmental primitive. Each of the resulting end effector poses can then be examined in terms of the incorporated extended reachability map \mathcal{R} in order to find the most promising poses in terms of whole-body stability^b. Fig. 8 depicts the process of generating the six best grasps for each detected environmental primitive.

A similar method can be employed to find promising foot poses for utilizing *support* affordances. In terms of the above example, we try to find poses for the right foot that allow the robot to climb the first step of the stairs. Fig. 9 shows the projection of the fused map $\mathcal{M}_{Right\ Foot} \cdot \mathcal{S}_{Right\ Foot}$ to the detected primitives as well as the most promising foot poses for each *support* hypothesis.

4. Towards Finding Affordance-Utilizing Robot Configurations

In this section we present an approach for computing initial whole-body configurations that utilize a chosen set of affordances. As stated above, the perceptual and

^bOther measures as well as the combination of different measures are possible¹³

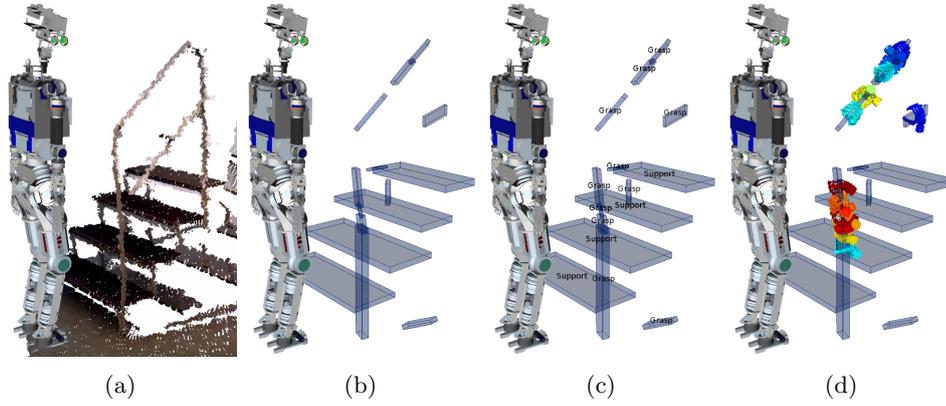


Fig. 8: An exemplary depth image of a staircase with a handrail (a) is segmented into environmental primitives (b). Affordance hypotheses based on the rules in Table 1 are generated (c). Finally, the proposed target poses for the right hand are computed based on a fusion of reachability and stability information (d). The color indicates the corresponding quality information value from blue (low) to red (high).

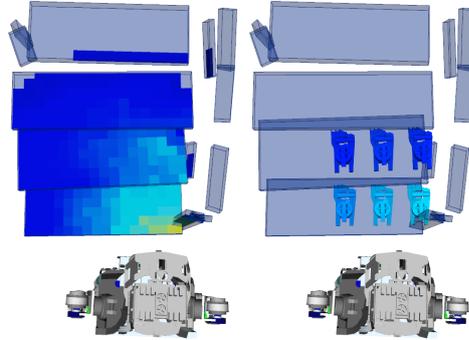


Fig. 9: $\mathcal{M}_{RightFoot} \cdot \mathcal{S}_{RightFoot}$ projected to the detected primitives (left). The six best proposed target poses for the right foot, based on the fusion of reachability and stability information (right).

interpretational errors in the process make an additional control strategy necessary that actually establishes contact with the environment. However, for initiating this control procedure it is crucial to find a stable start configuration that is already close to satisfying all considered constraints. We formulate the problem as a constrained inverse kinematics problem^{21,22,23} and then use state-of-the art tools²⁴ for solving it.

Task space regions

Constraints on end effector pose are formulated in terms of *Task Space Regions (TSRs)*²⁵. A TSR is a six-dimensional interval \mathbf{B}^w defined in the coordinates of a frame w , in our case the environmental primitives' local frames. The idea of TSRs additionally includes an end effector offset \mathbf{T}_e^w . An exemplary TSR for an environmental primitive p with extents $(2l_x, 2l_y, 2l_z)$ having a *support* affordance can look as follows:

$$\mathbf{B}_{support}^p = \begin{bmatrix} x_{\min} & x_{\max} \\ y_{\min} & y_{\max} \\ z_{\min} & z_{\max} \\ \psi_{\min} & \psi_{\max} \\ \theta_{\min} & \theta_{\max} \\ \phi_{\min} & \phi_{\max} \end{bmatrix} = \begin{bmatrix} -l_x & l_x \\ -l_y & l_y \\ -l_z & l_z \\ 0 & 0 \\ 0 & 0 \\ -\pi & \pi \end{bmatrix} \quad (7)$$

The TSR $\mathbf{B}_{support}^p$ fixes two out of three orientational dimensions, while giving the end effector the freedom to rotate around the support surface's normal. As described by Berenson et al.²⁵, TSRs can easily be formulated as kinematic constraints.

Static stability

As we compute the start configuration as a basis for further investigation if the assumed affordance hypotheses actually prevail, we cannot rely on a stabilizing contact to the considered environmental primitive p . Hence, we need to make sure that the computed start configuration is statically stable, leaving the robot enough tolerance to maneuver for an affordance verification procedure. Our kinematic constraint for static stability bases on *COG Jacobians*²⁶. Assuming the robot to be composed of N links l_i with their respective masses m_i and centers of gravity $\mathbf{x}_{\text{cog}_i}$, the COG Jacobian \mathbf{J}_{COG} is calculated as:

$$\mathbf{J}_{\text{COG}} = \frac{1}{\sum_{i=1}^N m_i} \sum_{i=1}^N m_i \cdot \mathbf{J}_{l_i}(\mathbf{x}_{\text{cog}_i}) \quad (8)$$

Finding suitable start configurations for approaching affordance utilization requires at least the discussed constraints on whole-body stability and end effector pose. However, employment of further constraints, e.g. joint torque minimization, human-likeness of posture, is possible as long as these constraints can be formulated in terms of the above problem. Taking additional constraints into account makes the overall problem harder and will eventually make more sophisticated solvers necessary, e.g. Kaiser et al.²⁷. Fig. 10 depicts two whole-body configurations computed based on a chosen set of perceived environmental primitives.

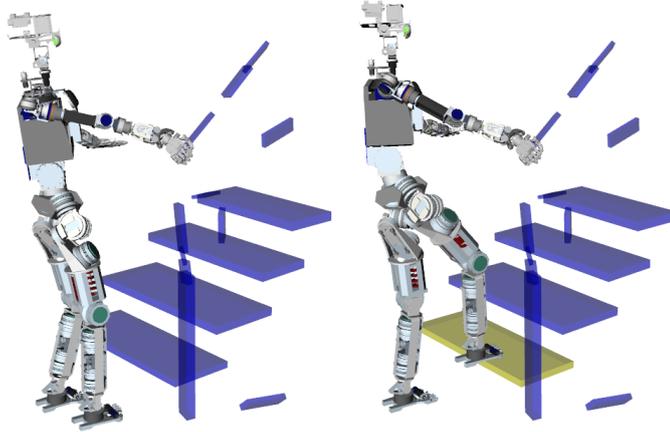


Fig. 10: Two statically stable whole-body configurations that approach the utilization of a *grasp* affordance at the handrail (left) or the combination of a *grasp* affordance at the handrail and a *support* affordance at the first step (right).

5. Verification of Whole-Body Affordances

The extraction of environmental primitives and the consecutive derivation of whole-body affordance hypotheses as outlined in the previous sections does not produce results that are directly ready for use on a real robot platform, due to two main reasons:

- (1) The perceptual process relies on the fusion of RGB-D camera images with inertial sensor data. This procedure introduces noise that results in a certain amount of error in the extraction of environmental primitives. Position and orientation of the resulting primitives are therefore only approximately known.
- (2) We assign affordance hypotheses based on pure geometric attributes of the extracted environmental primitives. There is no further perceptual step that estimates the level of robustness of an extracted primitive. Detected supporting structures can therefore easily collapse when the robot tries to utilize assumed affordances.

Due to the level of uncertainty in the perception-based process of affordance extraction, an additional step has to be carried out by the robot in order to verify affordance hypotheses, resulting in actual affordances. The schematic process of affordance verification is outlined in Fig. 11.

As a first step towards verification of whole-body affordances, we have implemented the full process of primitive extraction and affordance derivation in the robot framework ArmarX¹⁴ and employed a simple force-based verification strategy outlined in Fig. 12 on the humanoid robot ARMAR-III.

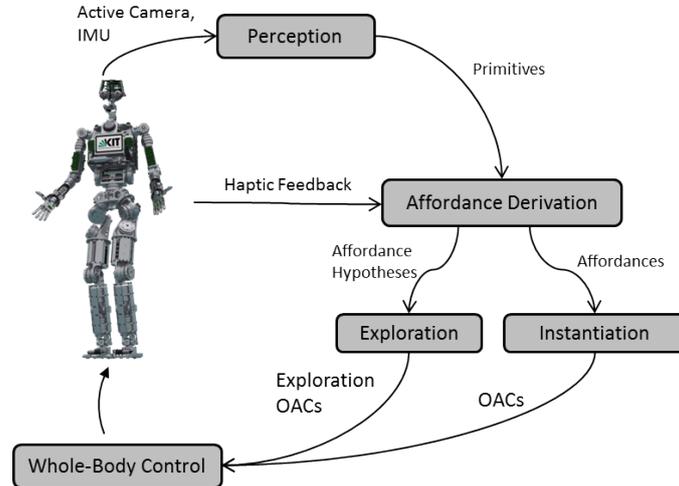


Fig. 11: The process of detection and exploration of whole-body affordances: Based on sensory information from active cameras or IMUs, the perceptual component produces an abstract representation of the environment. The resulting primitives are used for deriving affordance hypotheses. One possible choice is then to trust a derived hypothesis, in which case it directly results in an OAC instance that can be executed. The other choice is to start an exploration process to estimate the affordance’s reliability and the execution parameters. In this case, exploration OACs are executed and the sensed feedback again contributes to the affordance assignment step.

6. Conclusions

This paper presents our approach to the detection of whole-body affordance hypotheses based on the fusion of visually perceived environmental primitives. This incorporates a predefined set of rules that links symbolic affordances to properties of the extracted primitives like orientation or extent. The extracted affordance hypotheses are fused with an extended reachability map covering stability information in order to determine affordance hypotheses that are reachable for the robot in a statically stable manner.

In a second step we propose a method for deriving promising robot configurations that utilize a chosen set of affordances. This method is based on the idea that affordances act as constraints in a constraint-based inverse kinematics problem. The proposed methods on affordance extraction and configuration generation have been implemented and evaluated in simulation based on real RGB-D data.

In addition to the simulated results, we have implemented and tested a first strategy for validating *support* and *lean* hypotheses that have before been assumed purely based on visual perception. The affordance hypotheses are validated by ob-

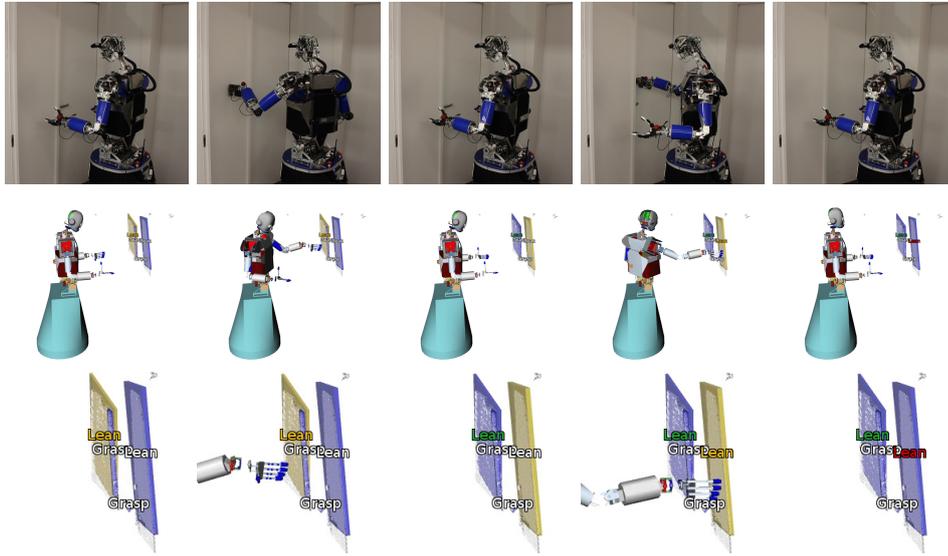


Fig. 12: A simple affordance verification strategy based on measuring the resistance force of perceived environmental primitives that presumably afford leaning. The robot perceives the two sides of a swing door, one of which being locked. It perceives two vertical planar primitives and starts verifying the assigned *lean* affordance hypotheses by exerting a force to each of the sides of the door. Based on the measured resistance force, the left side of the door is detected as a verified lean affordance while the right side of the door is rejected as it does not resist the exerted force.

serving the reaction forces measured in the robot’s wrist while touching the corresponding environmental primitives. The validation strategy has been implemented in the robot framework ArmarX on the humanoid robot ARMAR-III.

Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme under grant agreement no. 611832 (WALK-MAN).

References

1. N. Krüger, C. Geib, J. Piater, R. Petrick, M. Steedman, F. Wörgötter, A. Ude, T. Asfour, D. Kraft, D. Omrčen, A. Agostini, and R. Dillmann, “Object-Action Complexes: Grounded Abstractions of Sensorimotor Processes,” *Robotics and Autonomous Systems*, vol. 59, pp. 740–757, 2011.
2. A. Bierbaum, M. Rambow, T. Asfour, and R. Dillmann, “Grasp Affordances from Multi-Fingered Tactile Exploration using Dynamic Potential Fields,” in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 168–174, 2009.

3. E. Şahin, M. Çakmak, M. R. Doğar, E. Uğur, and G. Üçoluk, “To afford or not to afford: A new formalization of affordances toward affordance-based robot control,” *Adaptive Behavior*, vol. 15, no. 4, pp. 447–472, 2007.
4. D. Leidner, A. Dietrich, F. Schmidt, C. Borst, and A. Albu-Schäffer, “Object-Centered Hybrid Reasoning for Whole-Body Mobile Manipulation,” pp. 1828–1835.
5. P. Kaiser, D. Gonzalez-Aguirre, F. Schültje, J. Borràs, N. Vahrenkamp, and T. Asfour, “Extracting Whole-Body Affordances from Multimodal Exploration,” in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 1036–1043, 2014.
6. J. J. Gibson, *The Ecological Approach to Visual Perception*. Psychology Press, 1979.
7. A. ten Pas and R. Platt, “Localizing Grasp Affordances in 3-D Points Clouds Using Taubin Quadric Fitting,” in *International Symposium on Experimental Robotics (ISER)*, 2014.
8. L. Sentis, J. Park, and O. Khatib, “Compliant Control of Multicontact and Center-of-Mass Behaviors in Humanoid Robots,” *IEEE Transactions on Robotics*, vol. 26, no. 3, pp. 483–501, 2010.
9. L. Sentis, “Compliant Control of Whole-Body Multi-Contact Behaviors in Humanoid Robots,” in *Motion Planning for Humanoid Robots*, pp. 29–66, Springer, 2010.
10. S. Lengagne, J. Vaillant, E. Yoshida, and A. Kheddar, “Generation of whole-body optimal dynamic multi-contact motions,” *The International Journal of Robotics Research*, vol. 32, no. 9, pp. 1104–1119, 2013.
11. M. Popović, D. Kraft, L. Bodenhausen, E. Başeski, N. Pugeault, D. Kragic, T. Asfour, and N. Krüger, “A strategy for grasping unknown objects based on co-planarity and colour information,” *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 551–565, 2010.
12. N. Vahrenkamp, T. Asfour, and R. Dillmann, “Efficient inverse kinematics computation based on reachability analysis,” *International Journal of Humanoid Robotics*, vol. 9, no. 04, 2012.
13. N. Vahrenkamp, P. Kaiser, D. Gonzalez-Aguirre, and T. Asfour, “Extended Reachability Maps for Grasping, Manipulation and Locomotion,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (submitted)*, 2015.
14. N. Vahrenkamp, M. Wächter, M. Kröhnert, K. Welke, and T. Asfour, “The ArmarX Framework - Supporting high level robot programming through state disclosure,” *Information Technology*, vol. 57, no. 2, pp. 99–111.
15. T. Asfour, K. Regenstein, P. Azad, J. Schröder, N. Vahrenkamp, and R. Dillmann, “ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control,” in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 169–175, 2006.
16. F. Zacharias, C. Borst, and G. Hirzinger, “Capturing Robot Workspace Structure: Representing Robot Capabilities,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3229–3236, 2007.
17. R. Diankov, *Automated Construction of Robotic Manipulation Programs*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, 2010.
18. T. Asfour, J. Schill, H. Peters, C. Klas, J. Bücken, C. Sander, S. Schulz, A. Kargov, T. Werner, and V. Bartenbach, “ARMAR-4: A 63 DOF Torque Controlled Humanoid Robot,” in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, pp. 390–396, 2013.
19. N. Vahrenkamp, T. Asfour, G. Metta, G. Sandini, and R. Dillmann, “Manipulability Analysis,” in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pp. 568–573, 2012.
20. K. M. Varadarajan and M. Vincze, “AfNet: The Affordance Network,” in *Proceedings*

- of the 11th Asian Conference on Computer Vision, pp. 512–523, 2012.
21. B. Siciliano and J.-J. E. Slotine, “A General Framework for Managing Multiple Tasks in Highly Redundant Robotic Systems,” in *IEEE International Conference on Advanced Robotics (ICAR)*, pp. 1211–1216, 1991.
 22. L. Sentis and O. Khatib, “Synthesis of Whole-Body Behaviors Through Hierarchical Control of Behavioral Primitives,” *International Journal of Humanoid Robotics*, vol. 2, no. 4, pp. 505–518, 2005.
 23. A. Escande, N. Mansard, and P.-B. Wieber, “Fast Resolution of Hierarchized Inverse Kinematics with Inequality Constraints,” pp. 3733–3738, 2010.
 24. N. Vahrenkamp, M. Kröhnert, S. Ulbrich, T. Asfour, G. Metta, R. Dillmann, and G. Sandini, “Simox: A Robotics Toolbox for Simulation, Motion and Grasp Planning,” in *International Conference on Intelligent Autonomous Systems (IAS)*, pp. 585–594, 2012.
 25. D. Berenson, S. Srinivasa, and J. Kuffner, “Task Space Regions: A Framework for Pose-Constrained Manipulation Planning,” *International Journal of Robotics Research (IJRR)*, vol. 30, pp. 1435 – 1460, October 2011.
 26. T. Sugihara, Y. Nakamura, and H. Inoue, “Realtime Humanoid Motion Generation through ZMP Manipulation based on Inverted Pendulum Control,” in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1404–1409, 2002.
 27. P. Kaiser, D. Berenson, N. Vahrenkamp, T. Asfour, R. Dillmann, and S. Srinivasa, “Constellation - An Algorithm for Finding Robot Configurations that Satisfy Multiple Constraints,” in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 436–443, 2012.



Peter Kaiser received his diploma degree from Karlsruhe Institute of Technology (KIT) in 2014. He was a visiting researcher at Carnegie Mellon University in 2010 and at University of Edinburgh in 2013. He is currently a Ph.D. student at the Institute for Anthropomatics and Robotics at KIT where he is a member of the High-Performance Humanoid Technologies group. His major research interests are

planning and learning of whole-body actions, in particular actions related to locomotion and manipulation, based on the concept of affordances.



Nikolaus Vahrenkamp received his Diploma and Ph.D. degrees from the Karlsruhe Institute of Technology (KIT), in 2005 and 2011, respectively. Currently, he is a postdoctoral researcher at the Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT) and works on software development, grasping and mobile manipulation for the humanoid robots of the ARMAR family.

From 2011 to 2012, he was postdoctoral researcher at the Cognitive Humanoids Lab of the Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology (IIT), where he worked on grasp and motion planning for the humanoid robot iCub. Nikolaus Vahrenkamp is the author of over 40 technical publications, proceedings and book chapters. His research interests include humanoid robots,

motion planning, grasping and sensor-based motion execution.



Fabian Schültje received the M.Sc. degree in mechanical and process engineering from the Technical University of Darmstadt (TUD), Germany in 2014. During his studies he worked at the Institute for Mechatronic Systems in Mechanical Engineering (IMS) in the field of prosthetics. He is currently a Ph.D. student at the Karlsruhe Institute of Technology (KIT) in the High-Performance Humanoid Technologies group.

His major interests are in the field of modeling and control, in particular whole-body control algorithms related to locomotion and manipulation including hardware development and control for the humanoid robots of the ARMAR family.



Júlia Borràs received the M.Sc. degrees in mathematics and computer science from the Technical University of Catalonia, Barcelona, Spain, in 2004 and the Open University of Catalonia, Barcelona, in 2006, respectively. From 2004 to 2007, she worked with several companies as a programmer. In 2007, she joined the Institut de Robòtica i Informàtica Industrial, where she completed her Ph.D. degree in advanced automation and robotics from the

Technical University of Catalonia, and the Spanish Scientific Research Council, Institut de Robòtica i Informàtica Industrial, Barcelona, in 2011. From November 2011 to July 2013, she was a Postdoctoral Associate at the Grasping and Manipulation, Rehabilitation Robotics and Biomechanics (GRAB) Laboratory, Yale University (USA). In February 2014, she joined the High Performance Humanoid Technologies (H2T) team at KIT (Germany) as a postdoctoral researcher. She has worked in kinematics and singularities of parallel robots, and applied later these frameworks to grasping and whole-body humanoid poses.



Tamim Asfour is full Professor at the Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT). He is chair of Humanoid Robotics Systems and head of the High Performance Humanoid Technologies Lab (H²T). His current research interest is high performance humanoid robotics. He is developer and leader of the development team of the ARMAR humanoid robot family. He has been active in the field of Humanoid Robotics

for the last 14 years resulting in about 150 peer-reviewed publications with focus on engineering complete humanoid robot systems including humanoid mechatronics and mechano-informatics, grasping and dexterous manipulation, action learning from human observation, goal-directed imitation learning, active vision and active touch, whole-body motion planning, system integration, robot software and hardware control architecture. He received his diploma degree in Electrical Engineering in 1994 and his PhD in Computer Science in 2003 from the University of Karlsruhe.