On the Dualities Between Grasping and Whole-Body Loco-Manipulation Tasks

Tamim Asfour, Júlia Borràs, Christian Mandery, Peter Kaiser, Eren Erdal Aksoy and Markus Grotz

1 Introduction

While efficient solutions have been found for walking in different scenarios [16, 25], including rough terrain and going up/down stairs, humanoid robots are still not able to robustly use their arms to gain stability, robustness and safety while executing locomotion tasks. The ability of reaching for supports can be crucial to increase robustness in tasks that require balance like walking or general locomotion, but also for increasing dexterity and maneuverability in complex manipulation tasks. Nevertheless, to execute such tasks in an autonomous way, we need to better understand the principles of whole-body coordination in humans, the variety of supporting whole-body postures available and how to transition between them.

Kinematically, a humanoid balancing with multi-contacts is equivalent to a closed kinematic chain mechanism, where the closed chains are formed through the contacts with the environment. Contacts can be modeled as joints that can range from 0 DoFs (a planar contact without friction) to 5 DoFs (point contact with friction) [34]. Closed kinematic chain mechanisms constitute a big family that includes parallel robots, cable driven robots, cooperative robotic arms, multi-legged robots and multi-fingered hands among others. Dynamically, when a humanoid uses its body to gain stability through contacts with its environment, the dynamic equations to achieve equilibrium are the same as those of closed kinematic chain mechanisms where the chains are closed through contacts with an object or with the environment. Although these parallelisms in kinematics and dynamics have been acknowledged by many authors [5, 27, 36, 37, 39], fewer works try to find connections and transfer of techniques between those fields of robotics [7, 8, 15, 42]. We are interested in using techniques developed in the context of grasping with multi-fingered hands to apply them to the study of whole-body postures with multi-contacts, where the body plays

T. Asfour (✉) · J. Borràs · C. Mandery · P. Kaiser · E.E. Aksoy · M. Grotz
Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology,
Karlsruhe, Germany
e-mail: asfour@kit.edu

a double role: the role of the hand and the role of the manipulated object that can be moved through the contact reaction forces with the environment.

In this context, this paper presents our works exploring several different aspects of grasping that can be transferred to whole-body balance, such as grasping taxonomies [9], analysis of human motion data to classify and analyze grasps [31] and grasping affordances [23].

The main tools to understand how the hand can hold an object are the grasp taxonomies [4, 13, 19, 26]. Grasp taxonomies have been proven to be useful in many contexts: to provide a benchmark to test the abilities of a new robotic hand, to simplify grasp synthesis, to guide autonomous grasp choices, or to inspire hand design, among others. In our work in [9], we propose a classification of whole-body poses for balance exploring similar criteria as these used for taxonomies in grasping. Most grasping taxonomies define two main categories: precision and power grasping. In addition, Cutkosky classifies grasps according to object shape and tasks [13]. Kamakura according to task and the hand areas of contact [4, 26] and Feix et al. according to number/type of contacts and the configuration of the thumb [18]. We can directly use the idea of precision versus power grasping for whole-body poses: poses that use contact with the torso versus poses where contacts are realized only using the body end-effectors. But there is also an important difference since almost all whole-body poses are non-prehensile grasps, i.e., grasps that use the gravity to hold the object. In [9], we explore in detail criteria used in grasping to define a taxonomy of whole-body poses. A full taxonomy of whole-body poses can have many interesting applications and uses, such as a tool for autonomous decision making, a guide to design complex motions combining different whole-body poses, a way to simplify the control complexity, a benchmark to test abilities for humanoid robots, and a way to improve recognition of body poses and transitions between them.

Robotics in general, but particularly humanoid robotics, has always been inspired by biological human experience and the anatomy of the human body. In particular, grasping has collected human motion data to generate grasp synergies [11, 44] and classify and analyze most common grasps [17, 19], among other applications. However, human motions involving support contacts have almost not been studied [38] and even less how healthy subjects choose to make use of contacts with support surfaces. Our works like [31] explore the transitions between the whole-body poses of the proposed taxonomy by analyzing human motion data. While in classic locomotion actions such as walking and running the transitions between double and single support poses are very well-understood [29, 52], such transitions can become much more complex when e.g. the possibility of leaning against a surface with the hands is considered. In our work, we are interested in identifying balance poses during motions to be able to automatically perform a segmentation based on support poses. To study pose transition in [31], we analyze real human motion data captured with a marker-based motion capture system and post-processed using our unifying Master Motor Map (MMM) framework [2, 32, 33, 46], to gain information about the poses that are used while executing different locomotion and manipulation tasks like those shown in Fig. 1.
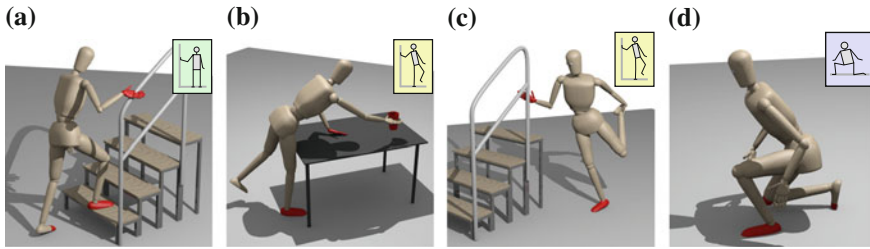
Fig. 1 When performing locomotion (**a**), manipulation (**b**), balancing (**c**) or kneeling (**d**) tasks, the human body can use a great variety of support poses to enhance its stability. Automatically detecting such support contacts allows for an automatic identification of the visited support poses and their transitions

Finally, we want to revisit the works where we explore the transfer of grasp affordances to the whole-body. The concept of affordances was originally introduced by Gibson [20] in the field of cognitive psychology for describing the perception of action possibilities. It states that objects suggest actions to the agent due to the object's shape and the agent's capabilities. A chair for example affords *sitting*, a cup *drinking* and a staircase *climbing*. Various works apply the concept of affordances to the field of grasping and manipulation, primarily for learning grasp affordances, e.g. by initial visual perception and subsequent interaction with an object [14, 41] or by focusing on either haptic [6] or visual [40] perception. In our previous work, we introduced the concept of Object-Action Complexes (OACs) to formalize affordances and link objects and actions into co-joint representations of sensorimotor behaviors in the robotic context [28]. In [23, 24], whole-body affordance hypothesis are defined as an association of a whole-body stable action to a perceived primitive of the environment. The concept of affordances is applied to actions related to the whole body of a humanoid agent, particularly actions for stabilization and combinations of whole-body locomotion and manipulation actions, i.e. loco-manipulation actions. Based on previous approaches like [6, 41], these works aim at deriving, refining and utilizing whole-body affordances like holding, leaning, stepping-on or supporting in unknown environments.

In this work, we present a summary review of these works to visualize successful transfer of knowledge from grasping to whole-body motion analysis with multi-contacts, and we point out directions of current and future work following the same idea of research. The paper is organized as follows. Section 2 summarizes the work where the taxonomy of whole-body poses is defined, summarizing the used criteria for classification. Section 3 revises the results of the analysis of motion data, and the automatic generation of a graph of pose transitions that is compared to the taxonomy of the previous section. Section 4 revisits the works where we extract whole-body affordances from unknown scenes and we validate the approach with an experiment executed with ARMAR-IIIa [1]. Finally, Sect. 5 gives a summary and provides ideas for current and future research.
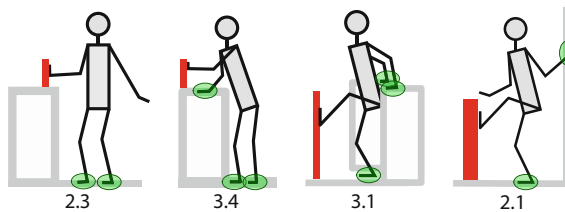
2 Taxonomy of Whole-Body Support Poses

When considering the whole body interacting with the environment, there is a wide range of different postures that the robot can adopt. In [9], we were interested in those poses that use contacts for balancing. Then, the limb end-effectors that are not used for balancing can be used to perform other manipulation tasks. This way, we provide a framework for loco-manipulation poses. In other words, contacts with environmental elements that do not provide support are not considered for the taxonomy classification. For instance, in Fig. 2, green marked contacts define the support pose, while the rest are contacts intended to manipulate the environment that do not affect the support pose definition.

In Table 1, the taxonomy proposed in [9] is shown. It contains a total of 46 classes, divided into three main categories: standing, kneeling and resting. Each row corresponds to different number of supports, and in each row, different columns correspond to different contact types (see contact type legend at the bottom left corner of Table 1). In addition, colors differentiate type of leg supports and poses under the gray area use line contacts (with arms or legs). The lines between boxes indicate hypothesis of pose transitions between poses assuming only one change of support at a time.

The criteria considered for the definition of the taxonomy are

1. **Number of contacts**: One of the first relevant characteristics that greatly modifies the complexity of a motion is the number of contacts and supports with the environment. Kinematically, each support creates a new closed kinematic loop, and therefore, reduces by one the dimension of the feasible configuration space [3, 22]. Dynamically, planning of complex motions tested on humanoid robots report higher execution times per higher number of supports [43].
2. **Type of contacts**: From the control point of view, the nature of the contact used to provide the support [12, 35] and the part of the body that performs the support are relevant and important, because the resultant kinematics of the robot changes accordingly. A fingertip contact is usually modelled as a point contact with friction, the foot support as plane contacts and arm leaning can be modelled using line with friction model [34]. Figure 3 shows the types of support that we considered for the taxonomy with the legs and arms. To keep our taxonomy simple, we consider only 5 types: hold, palm, arm, feet, and knee support. These lead to the


Fig. 2 The support poses to perform the task of hit an object are defined by the contacts highlighted in *green*. The numbers under the sketches refer to the id number of the support class in Table 1

**Table 1** Taxonomy of whole-body support poses
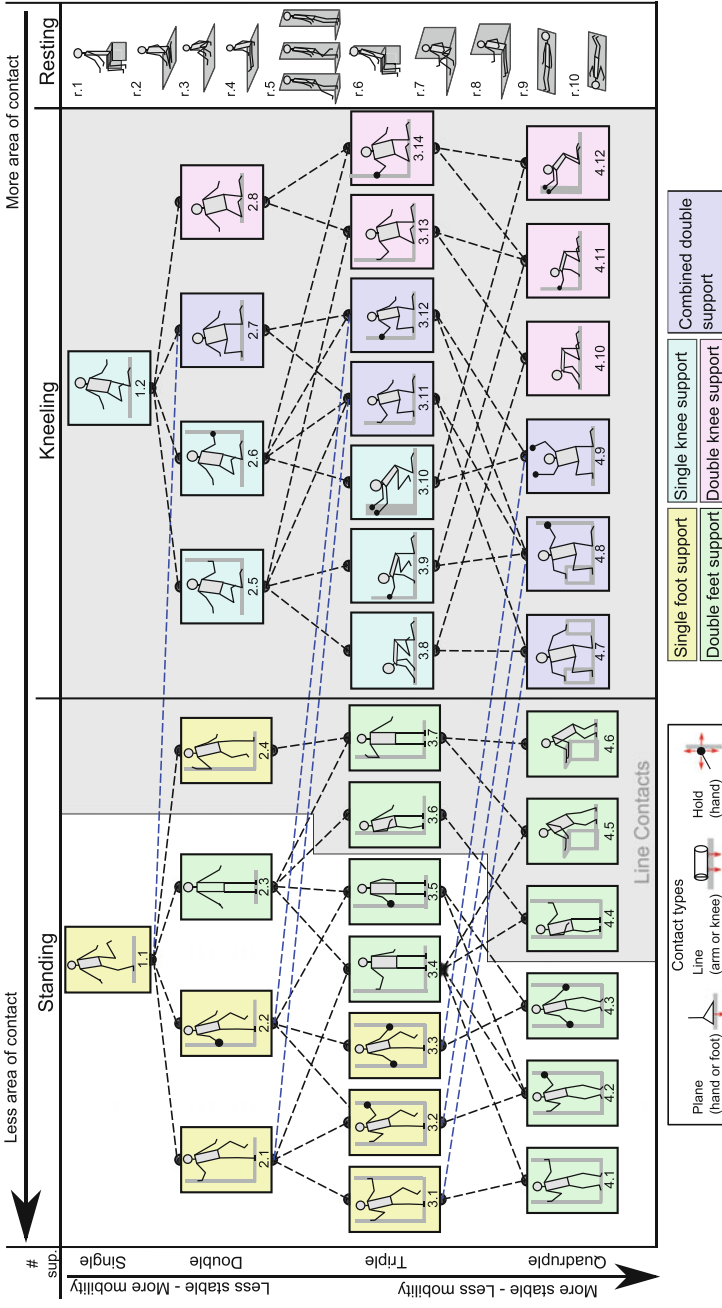
The proposed taxonomy has 46 classes, including 18 standing poses, 18 kneeling poses and 10 resting configurations. Sketches represent all the ranges of poses with the same number of supports and type of contact. Each class includes the symmetric cases when applicable. The lines provide hypothetical road maps to transition from one pose to another, assuming one contact change at a time. Lines also provide a hierarchy among the poses. Blue lines represent transitions between different categories (from standing to kneeling).

Fig. 3 Types of support contacts with arms and legs

hold   palm   arm   feet   knee
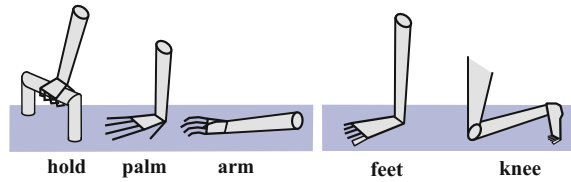
Fig. 4 Different body shapes for the pose 2.3
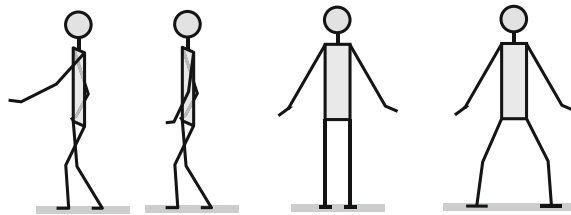
consideration of 51 combinations from which we have selected 36 (corresponding to the standing and kneeling poses). This choice has been done assuming that some combinations, while feasible, are not common.

3. **Shape of the environment**: Many grasping taxonomies include the shape of the object as a criteria for grasp selection. Indeed, object shape and size have a great influence on the ability for grasping and manipulation. However, there is a fundamental difference between hand grasping and whole-body poses: the need of gravity to reach force closure. A hand grasp will always start with no contacts at all, and after grasping, it may or may not start a manipulation motion that can be maintaining constant contacts (in-grasp manipulation) or performing re-grasps [30]. On the contrary, a whole-body grasp is always part of a motion sequence of re-grasps that will always start with at least one contact with the environment (even if one of the phases has no contact as in a running locomotion or jumping). For this reason, we believe that whole-body grasp choice will not depend as much on the shape of the environment, but on the task/motion the pose occurs. Therefore, our taxonomy does not include this criteria of classification.

4. **Shape of the body**: While we believe shape of the environment is not relevant in our case, the shape of the body performing the pose is relevant because it depends on the task and can influence the transitions between different poses. For instance, the shape of the body on the pose 3.4 when walking with a handrail will be different than going upstairs with a handrail. Also, if when performing a locomotion, the shape of the body in a double foot support pose (pose 2.3) contains a hand reaching further, like in the left figure of Fig. 4, it is very probable that the following pose will be a one with hand contact. However, the number of shapes that each pose can adopt is very large and the size of the taxonomy grows exponentially. Therefore, we will classify shape poses in a different hierarchy of classification, that is left for future work.

5. **Stability**: The taxonomy in Table 1 is organized so that the less stable poses, with less number of supports lie on the upper left side, while the most stable ones on the lower right side, assuming that the more number of contacts and the larger the surfaces of contact, the more stable the robot is. Works like [21] show that there is a trade-off between stability and maneuverability during a goal-directed whole-body movements. In the taxonomy, we observe a similar trade-off with mobility versus stability. However, it has to be noted that inside any class it is possible to obtain different levels of stability depending on the support region [10] (that greatly depends on the body shape) and the sum of the contact wrenches [50].

6. **Power grasps versus resting poses**: In addition to the standing and kneeling poses we have added 10 extra classes where there is contact with the torso. We call them resting poses and they are the equivalent to power grasps where there is contact of the object with the palm. Poses from r.1 to r.4 are poses where still balance needs to be achieved, but the inclination of the torso needs to be controlled. Poses from r.5 to r.6 are stable provided that the areas of contact are flat and with friction. Finally, using poses from r.7 to r.10 the robot is unlikely to lose balance and can be considered safe and completely in rest, but with very limited mobility.

   At this stage of work, no transitions are shown between resting poses and the rest of the table. Such transitions are more complex and require further analysis that will be left for future work.

7. **Pose transitions and motions**: In the next section, we will show how we have studied support pose transitions by analyzing human motion capture data. However, the taxonomy provides preliminary hypothesis of possible pose transitions using lines connecting poses in the taxonomy. Physically, a transition between two classes can happen by first imposing the constraints of the current and destination class, and then shifting to only the constraints of the destination class. This induces the definition of two types of motions:

   a. **Inside class motion**: A purely manipulation action will happen inside a single class. It includes other manipulation motions and therefore, extra contacts with objects, always with the objective of manipulation. As a manipulation motion, it can be semantically segmented and interpreted as done in [48].

   b. **Transition class motion**: motions that define a transition between poses. The motion still occurs inside a class, but the motion consists in the shifting towards a destination class, as part of a locomotion. For instance, a double feet support motion that shifts towards a right foot support ($2.3 \rightarrow 1.1$). These are the kind of transitions that are studied in [31] and summarized in the next section.

Note that both motions happen always inside the same support class, but in the second case, the destination class is relevant for the motion definition.

3   Detection of Whole-Body Poses and Segmentation

The framework presented in Sect. 2 introduces a set of segmentation criteria for a given motion that, provided that we can differentiate support contacts and manipulation contacts, subdivides a motion into pieces that can be related to types of actions. For actions identified as manipulation (inside pose motion), further segmentation based on the manipulation contacts can be performed [48], providing a hierarchy of segments distinguishing between the locomotion and the manipulation parts of an action. In the work [31], we proposed a method to detect support contacts that allows us to automatically segment motion data based on the support poses. This allows us to analyze support pose transitions during 121 loco-manipulation motions recorded using an optical marker-based Vicon MX motion capture system. This motion analysis can provide a better semantic understanding of complex locomotion and manipulation actions for imitation learning and autonomous decision making applications. Our motions recordings contain also information of the location and movement of the environmental elements, such as manipulated objects or objects to provide support. The KIT Whole-Body Human Motion Database [32] contains a large set of motions, providing raw motion capture data, corresponding time-synchronized video recordings and processed motions. The motions recorded for the work [31] can be found in the KIT Whole-Body Human Motion Database.[1]

Finally, the motions are post-processed using the The Master Motor Map (MMM) framework [33, 46]. This provides an open-source framework for capturing, representing and processing human motion. It includes a unifying reference model of the human body for the capturing and analysis of motion from different human subjects. The kinematic properties of this MMM reference model are based on existing biomechanical analysis by Winter [51] and allow the representation of whole-body motions using 58 degrees of freedom (DoF): 6 for the root pose and 52 for the body torso, extremities and head.

Support poses of the human subject are detected by analyzing the relation of the MMM reference model to the floor and environmental elements. For this purpose, we only consider objects which exhibit low movement during the recorded motion as suitable environmental elements to provide support. For every motion frame, we use the forward kinematics of the reference model to calculate the poses of the model segments that we consider for providing supports. These model segments represent the hands, feet, elbows and knees of the human body.

A segment s of the reference model is recognized as a support if two criteria are fulfilled. First, the distance of s to an environmental element must be lower than a threshold $\delta_{dist}(s)$. Distances to environmental elements are computed as the distances between pairs of closest points from the respective models with triangle-level accuracy using Simox [47]. Additionally, the speed of segment s, computed from smoothed velocity vectors, has to remain below a threshold $\delta_{vel}(s)$ for a certain

---

[1] See https://motion-database.humanoids.kit.edu/details/motions/<ID>/ with ID ∈ {383, 385, 410, 412, 415, 456, 460, 463, 515, 516, 517, 520, 521, 523, 527, 529, 530, 531, 597, 598, 599, 600, 601, 604, 606, 607}.

number of frames, starting with the frame where the support is first recognized. The thresholds are chosen empirically: $\delta_{vel} = 200 \frac{mm}{s}$, $\delta_{dist}(Feet) = \delta_{dist}(Hands) = 15\,mm$, $\delta_{dist}(Knees) = 35\,mm$ and $\delta_{dist}(Elbows) = 30\,mm$. The support pose is defined by the contacts that are providing support to the subject. We ignore parts of the motion where the human body is not supported at all as an empty support pose, e.g. during running. Also, some practical assumptions are used, such as that a knee support also implies a foot support. We have manually validated the segmentation method error by exploring frame by frame the detected support segments. Results can be seen in [31]. They show that about 4.5% of the poses are missed, but the missed poses are always double foot supports (with or without hand). Only 2.1% of the poses are incorrectly detected.

3.1 Data Driven Analysis of Transitions Between Whole-Body Support Poses

Without taking into account kneeling motions, we have recorded and analyzed 110 motions including locomotion, loco-manipulation and balancing tasks. In the following, we present some analysis on the most common pose transitions. We ignore kneeling motions because we do not have enough data yet to get significant results. In every motion, both the initial and the final pose are double foot supports and the time spent on these poses is arbitrary. Therefore, they have been ignored for the statistical analysis. Without counting them, we have automatically identified a total of 1323 pose transitions lasting a total time of 541.48 s (9.02 min). In Table 2, each cell represents the transition going from the pose indicated by the row name to the pose indicted by the column name. In each cell, we show first the percentage of

Table 2 Percentages of appearances and time spent for each transition (%appearance, %time)

	1Foot	1Foot-1Hand	2Feet	2Feet-1Hand	2Feet-2Hands	1Foot-2Hands	Totals × pose
1Foot	4.38, 5.69%	9.30, 7.90%	22.90, 25.56%	0.15, 0.26%	–	0.08, 0.04%	36.81, 39.44%
1Foot-1Hand	9.15, 3.64%	1.81, 2.26%	0.08, 0.03%	12.24, 16.59%	0.08, 0.02%	0.15, 0.02%	23.51, 32.57%
2Feet	16.02, 10.05%	0.15, 0.04%	×	3.48, 2.23%	0.08, 0.06%	–	19.73, 12.38%
2Feet-1Hand	0.23, 0.07%	11.72, 4.38%	4.61, 5.31%	×	0.98, 0.15%	–	17.54, 9.92%
2Feet-2Hands	–	–	–	0.83, 1.22%	×	0.68, 0.75%	1.51, 1.97%
1Foot-2Hands	–	0.53, 1.27%	–	–	0.38, 2.45%	×	0.91, 3.72%

occurrence of the transition with respect to the total number of transitions detected, and secondly the percentage of time spent on the origin pose before reaching the destination pose, with respect to the total time of all motions. The last column is the accumulation of percentages per each pose, and the rows are sorted from the most to the least common pose. According to Table 2, the most common transitions are 1Foot→2Feet (22.90% of appearance) and 2Feet→1Foot (16.02% of appearance). These are the same transitions of walking that have been widely studied. Although all motions contain some steps of normal walking, they also involve hand supports, and therefore, these transitions show different behaviours because they are part of a more complex set of transitions. It must be noted that the loop transitions 1Foot→1Foot, and 1Foot-1Hand→1Foot-1Hand are mostly missed double foot supports and we will not include them in the analysis.

Figure 5 shows the automatically generated transition graph, considering also the start and end poses of each motion and all the kneeling motions. Each edge
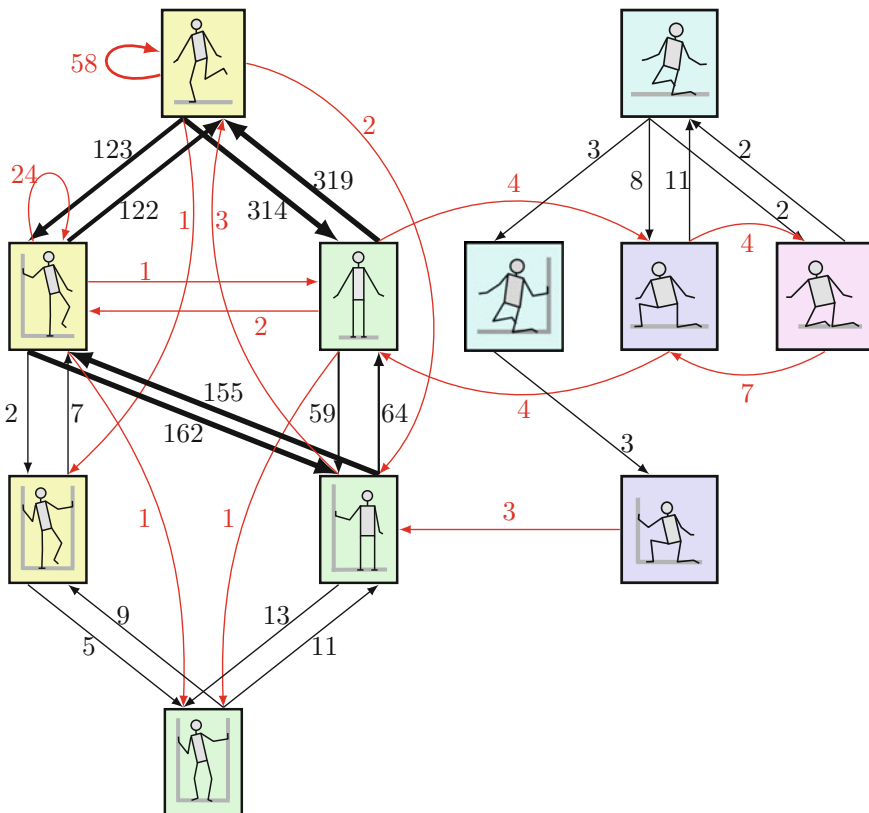
Fig. 5 Transition graph of whole-body pose transitions automatically generated from the analyzed motions. Labels on edges indicate the number of transitions found of each type

corresponds to a transition, and their labels to the number of times we have found it. Edges plotted in red correspond to transitions where two simultaneous changes of contacts occur. In the taxonomy of Table 1, we assumed that only one change of support should be allowed per transition. While this is still desirable for robotics, it is also obvious that some human transitions involve two contact changes. For instance, in push recovery motions, humans usually lean on the wall using both arms at the same time to increase stability. Many of the red edge transitions in Fig. 5 occur in balancing tasks. In the transition graph shown in Fig. 5, we can quickly see that red edges are of significantly lower frequency than the black ones, except the loop edges in the 1Foot and 1Hand-1Foot poses, that are caused by either jumps or missed double foot supports. They correspond to the 4.5% transitions missed by our segmentation method reported before. This data-driven transition graph is influenced by the type of motions we have analyzed, using only one handle or one hand support. Only balancing poses reach the four support poses. In future work, we will analyze walking motions with handles on both sides.

Most of the hypothetical transitions in Table 1 are correct, except 1Foot → 1Foot1Knee that does not appear in real data. This is because the subject uses support with the tip of the foot until contact is reached with the knee, and this is detected as a foot support that corresponds to a tip-toe support. Therefore, all the edges in red between double foot support (with and without hand) and kneeling are correctly detected and should be corrected in our taxonomy in Table 1. In the future, we will study if we should distinguish between tip-toe and sole support. Figure 6 shows an example of segmentation result, corresponding to the time line of a motion where the subject goes upstairs using a handle on his right side. In blue, we show the long locomotion transitions. The supporting pose for these transitions alternates between 1Foot-1Hand, used to swing forward the foot not in contact, and 1Foot, used to swing forward both the handle hand and the foot not in contact. This is because we only provide one handle. Another interesting thing to notice is that the short locomotion transitions appear in clusters, composed by a sequence of two transitions. We have observed this in many of the motions and we have observed that the order of the transitions inside these clusters does not matter, just the start and end poses. We believe that each cluster could be considered as a composite transition where several contact changes occur. As future work, we want to detect and model these clusters to
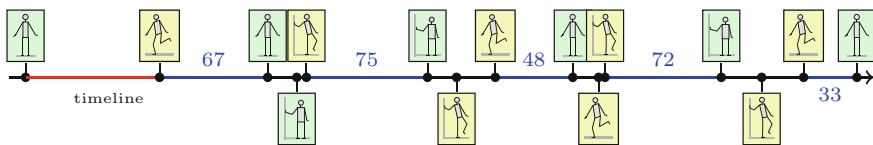
Fig. 6 Result of the segmentation for one of the motions upstairs with handle. The segment shown in *red* represents the initial pose transition, that has an arbitrary length. *Blue* segments represent transitions where the foot swings. *Blue* labels/numbers indicate transition durations. We can see that the human alternates between single foot support swing and 1Foot-1Hand support swing using the handle

identify rules that allow us to automatically generate sequences of feasible transitions according to extremities available for contacts.

4 Whole-Body Affordances

Grasp affordances rely on perception methods, either visual or haptic, to perceive the geometry of the object, and then associate grasp strategies according to the recognized geometric shapes [6]. Similarly, in [23, 24] we relay on a visual perception system with active cameras that can collect point clouds, register them and then extract geometric primitives from an unknown scene. In [23] we proposed to assign hypotheses for whole-body affordances like *support*, *lean*, *grasp* or *hold* to environmental primitives based on shape, size and orientation. Large vertical planes for instance are assumed to indicate *lean*-affordances. These kind of affordances are of basic importance for whole-body stabilization. However, further possible whole-body affordances exist and are of special interest when manipulating large, and possibly heavy, objects, for instance for removing debris from a blocked pathway. Examples for whole-body affordances indicating manipulability of objects are *pushability* and *liftability*, which are experimentally evaluated in [24]. The association of affordances is based on a set of rules shown in Table 3. While an exhaustive evaluation of the available types of whole-body affordances still remains to be done, *pushability* and *liftability* are certainly essential. The work show that we can integrate and evaluate the processes of affordance perception, validation and utilization on a real-world robotic system considering all the affordance types.

The constants λ_i from Table 3 are currently application specific. However, we think that there is a fixed set of affordance extraction parameters that will work reasonably well for our scenarios due to the following reasons:

- Research shows that agents infer affordances based on a body-scaled metric, i.e. with respect to the proportions of their bodies [49].
- We primarily focus our studies to disaster scenarios that contain at least partly intact man-made elements like doors, handrails or staircases. These elements usually have standardized dimensions known beforehand.

Figure 7 visualize the environmental primitives and their associated affordances from point cloud example corresponding to a staircase scenario. The primitives are assigned meaningful whole-body affordances based on the rules from Table 3. The proposed framework successfully identifies the existing cylindrical and planar primitives. More examples of different scenes can be found in [24]. The strategies for affordance extraction are purely based on visual perception and are therefore only *affordance hypotheses* subject to further investigation and validation by the robot. In [23], precomputed reachability maps help to discard non utilizable affordances. In [24] there is no reliable mechanism for verifying the existence of affordanceswithout

Table 3 Example of a set of rules for affordance derivation and possible validation strategies. The operator $\uparrow$ tells if two vectors point into the same direction. The values λ_i are implementation-specific constants

| Affordance	Shape	Parameters	Conditions	Valid.
Support (S)	Planar	Normal $\boldsymbol{n}$	$\boldsymbol{n} \uparrow z_{world}$	(1a)
		Area a	$a \geq \lambda_1$	
Lean (Ln)	Planar	Normal $\boldsymbol{n}$	$\boldsymbol{n} \perp z_{world}$	(1a)
		Area a	$a \geq \lambda_2$	
Grasp (G)	Planar	Normal $\boldsymbol{n}$	$a \in [\lambda_3, \lambda_4]$	(3)
		Area a		
	Cylindrical	Radius r	$r \in [\lambda_5, \lambda_6]$	
		Direction $\boldsymbol{d}$	$\|\boldsymbol{d}\| \leq \lambda_7$	
	Spherical	Radius r	$r \in [\lambda_8, \lambda_9]$	
Hold (H)	Cylindrical	Radius r	$r \in [\lambda_{10}, \lambda_{11}]$	(2a)
		Direction $\boldsymbol{d}$	$\|\boldsymbol{d}\| \geq \lambda_{12}$	
Push (P)	Planar	Normal $\boldsymbol{n}$	$\boldsymbol{n} \perp z_{world}$	(1b)
		Area a	$a \leq \lambda_{13}$	
Lift (Lf)	Planar	Normal $\boldsymbol{n}$	$a \leq \lambda_{15}$	(2b)
		Area a		
	Cylindrical	Radius r	$r \leq \lambda_{15}$	
		Direction $\boldsymbol{d}$	$\|\boldsymbol{d}\| \leq \lambda_{16}$	
	Spherical	Radius r	$r \leq \lambda_{17}$	
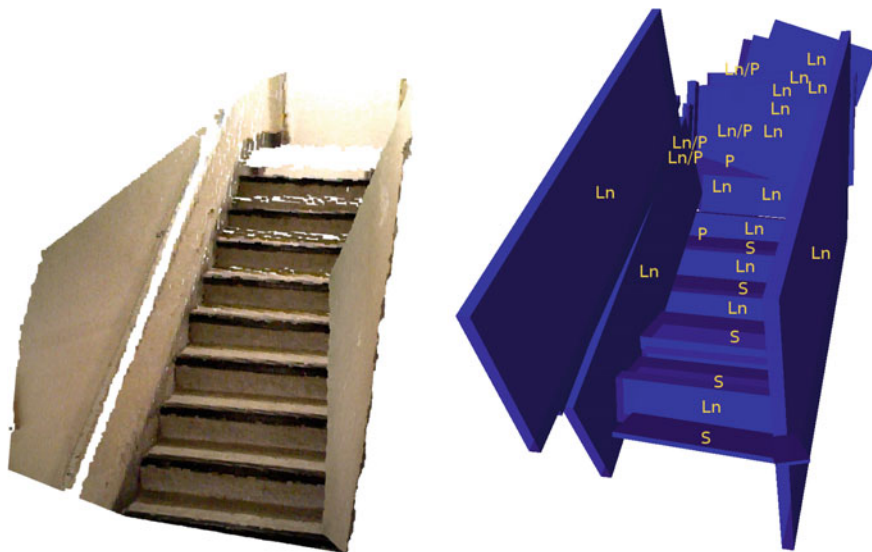
Fig. 7 Example of the results of the affordance extraction process (*right*) from a segmented point cloud (*left*). The example scenario is staircase. The affordance tags **S**, **Ln**, **G**, **P** and **Lf** refer to Table 3. For more examples see [24]

establishing contact to the underlying primitives. Referring to Table 3, different force-based validation strategies exist based on the affordance hypothesis to investigate:

1. Exert a force along the primitive's normal **n** and compare the resistance force against a minimum ϑ_1 (1a) or a maximum ϑ_2 (1b).
2. Grasp the primitive and exert forces perpendicular to the primitive's direction **d**. Compare the resistance force against a minimum ϑ_3 (2a) or a maximum ϑ_4 (2b).
3. Push the primitive and perceive the caused effect.

Considering further sensor modalities apart from contact forces is of interest and can lead to more sophisticated and accurate validation strategies. Validating the *pushability* of a very light object for instance might not result in a reliable resistance force feedback. Possible solutions for cases like this include tactile feedback or the comparison of RGB-D images before and after the push, similar to [45]. These strategies were validated with an experiment carried out on the humanoid robot

(a) Perception   (b) Validation   (c) Execution
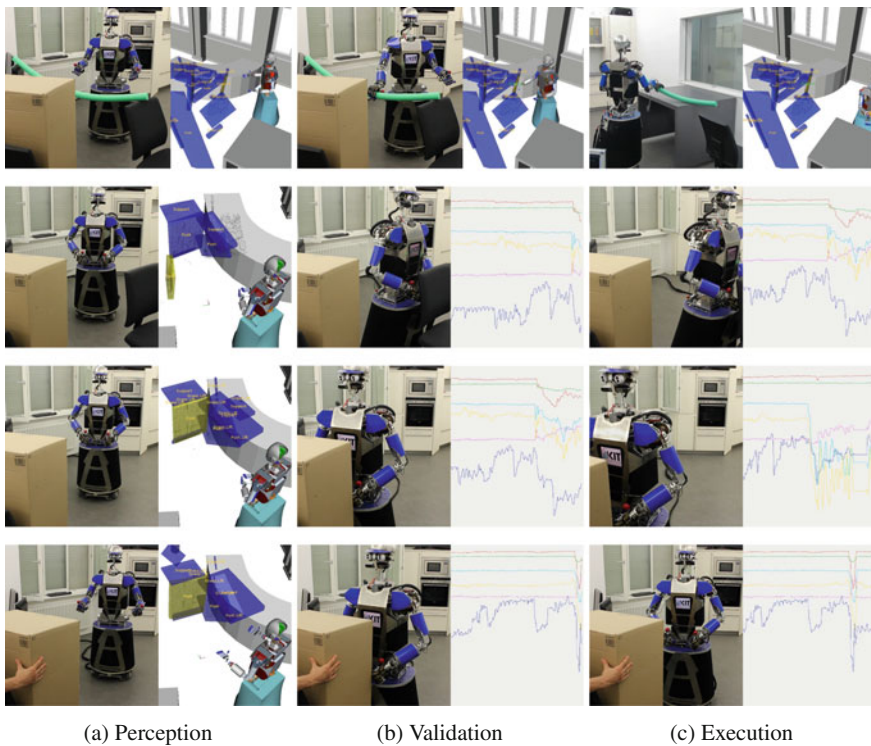
Fig. 8 The three stages of perception, validation and execution of whole-body affordances in four different scenarios: A pipe that can be grasped and lifted (*first row*), a chair that can be pushed (*second row*), a box that can be pushed (*third row*) and a box that is fixed and cannot be pushed (*fourth row*). The plots visualize the force amplitudes (y-axis) measured in the robot's left wrist over time (x-axis), while the *blue* curve represents the force in pushing direction

ARMAR-III, demonstrating the perception and validation of affordance hypotheses for *pushability* and *liftability*. In the experiment ARMAR-III is facing a cluttered arrangement of different obstacles, i.e. debris, that block its way: A chair, a box and a pipe (see Fig. 8, top left corner). The robot has no prior knowledge on the types or locations of the employed obstacles, the only information it gets results from the perceptual pipeline. Figure 8 displays snapshots of different stages of the experiment: perception (first column), validation (second column) and execution (third column). The perception stage displays the initial obstacle arrangement and its representation after the perceptual pipeline in terms of primitives and affordance hypotheses. The validation stage includes the establishment of contact with the selected primitive and the affordance validation based on the obstacle's resistant force. In the execution phase, the robot has validated the affordance in question and starts pushing or lifting the obstacle, respectively. The robot successfully identifies all three obstacles and starts by validating the liftability of the pipe (Fig. 8, first row). The validated liftability is then exploited for moving the obstacle away. In the next steps the robot identifies the chair and the box as pushable obstacles and validates these affordances accordingly (Fig. 8, second row, third row). The last row of Fig. 8 displays a repetition of the previous scene with a fixed box. The robot again assigns a pushability hypothesis to the box, but fails to validate this hypothesis. Hence, the corresponding push cannot be executed. A detailed description of the whole process is given in [24].

5 Conclusions

This work revisits several works from our previous work that explore transfer of techniques from grasping to whole-body loco-manipulation tasks. In this context, we have proposed a taxonomy of whole-body balancing poses containing 46 classes, divided into three main categories, considering number and type of support and possible transitions between support poses. We have analyzed known grasping criteria used to classify robot grasps, but focusing on the demands of whole-body poses. As opposed to grasping, we have given less relevance to environment shape and more to the type of contact the body uses to provide a support pose. We have also presented an analysis of support poses of more than 100 motion recordings showing different locomotion and manipulation tasks. Our method allowed us to retrieve the sequence of used support poses and the time spent in each of them, providing segmented representations of multi-contact whole-body motions. Although the most common pose transitions are the ones involved in walking, we have shown that the 1Foot-1Hand and the 2Foot-1Hand poses also play a crucial role in multi-contacts motions. The data-driven generated graph of transitions validates the transitions proposed in our taxonomy. We believe that our motion segmentation by support poses and time spent per transition provides a meaningful semantic representation of a motion. Finally, we have shown how the concept of grasp affordances can also be applied to whole-body affordances. Using common sense knowledge of a perceived unknown scene,

whole-body affordances are assigned to geometries and are then validated through physical interaction with the scene.

This work opens the door to many exciting future directions. First, each class of poses in the taxonomy corresponds to an infinite number of possible body configurations depending on location and orientation of contacts and the body shape. Future work directions include finding the most relevant *whole-body eigen-grasps* based on the collected human motion data, that is, we will apply dimensionality reduction to deal with the large space of whole-body configuration and to determine whole-body eigen-grasps associated with support poses. Secondly, we are interested in analyzing our motion representations to find semantic rules that can help define new motions for different situations, with the objective of building a grammar of motion poses based on the introduced taxonomy. Storing each transition as a motion primitive, we are also interested in performing path planning at a semantic level based on support poses. Finally, we plan to use the extracted and validated affordances to generate sequences of whole-body poses that generate locomotions with multi-contacts, utilizing perceived location of the possible contacts and the learned motion primitives for each pose transition. In conclusion, our works present a step further in the comprehension of how humans can utilize their bodies to enhance stability for locomotion and manipulation tasks. We believe the proposed ideas have a lot of potential to be used in many areas of humanoid robotics.

References

1. Asfour, T., Regenstein, K., Azad, P., Schröder, J., Vahrenkamp, N., Dillmann, R.: ARMAR-III: an integrated humanoid platform for sensory-motor control. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids)
2. Azad, P., Asfour, T., Dillmann, R.: Toward an unified representation for imitation of human motion on humanoids. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 2558–2563 (2007)
3. Berenson, D., Srinivasa, S.S., Kuffner, J.: Task space regions: a framework for pose-constrained manipulation planning. Int. J. Robot. Res. **30**(12), 1435–1460 (2011)
4. Bernardin, K., Ogawara, K., Ikeuchi, K., Dillmann, R.: A sensor fusion approach for recognizing continuous human grasping sequences using hidden markov models. IEEE Trans. Robot. **21**(1), 47–57 (2005). doi:10.1109/TRO.2004.833816
5. Bicchi, A., Melchiorri, C., Balluchi, D.: On the mobility and manipulability of general multiple limb robots. IEEE Trans. Robot. Autom. **11**(2), 215–228 (1995)
6. Bierbaum, A., Rambow, M., Asfour, T., Dillmann, R.: Grasp affordances from multi-fingered tactile exploration using dynamic potential fields. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids), pp. 168–174 (2009)
7. Borràs, J., Dollar, A.M.: Analyzing dexterous hands using a parallel robots framework. Auton. Robots **36**(1–2), 169–180 (2014)
8. Borràs, J., Dollar, A.M.: Dimensional synthesis of three-fingered robot hands for maximal precision manipulation workspace. Int. J. Robot. Res. **34**(14), 1731–1746 (2015)

9. Borràs, J., Asfour, T.: A whole-body pose taxonomy for loco-manipulation tasks. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1578–1585 (2015)
10. Bretl, T., Lall, S.: Testing static equilibrium for legged robots. IEEE Trans. Robot. **24**(4), 794–807 (2008)
11. Ciocarlie, M., Goldfeder, C., Allen, P.: Dimensionality reduction for hand-independent dexterous robotic grasping. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3270–3275 (2007)
12. Collette, C., Micaelli, A., Andriot, C., Lemerle, P.: Robust balance optimization control of humanoid robots with multiple non coplanar grasps and frictional contacts. In: ICRA 2008. IEEE International Conference on Robotics and Automation, pp. 3187–3193. IEEE (2008)
13. Cutkosky, M.R.: On grasp choice, grasp models, and the design of hands for manufacturing tasks. IEEE Trans. Robot. Autom. **5**(3), 269279 (1989)
14. Detry, R., Kraft, D., Kroemer, O., Bodenhagen, L., Peters, J., Krüger, N., Piater, J.: Learning grasp affordance densities. Paladyn J. Behav. Robot. **2**(1), 1–17 (2011)
15. Ebert-Uphoff, I., Voglewede, P., et al.: On the connections between cable-driven robots, parallel manipulators and grasping. IEEE Int. Conf. Robot. Autom. **5**, 4521–4526 (2004)
16. Englsberger, J., Ott, C., Albu-Schaffer, A.: Three-dimensional bipedal walking control using divergent component of motion. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2600–2607 (2013)
17. Feix, T., Bullock, I.M., Dollar, A.M.: Analysis of human grasping behavior: object characteristics and grasp type. IEEE Trans. Haptics **7**(3), 311–323 (2014)
18. Feix, T., Pawlik, R., Schmiedmayer, H.B., Romero, J., Kragic, D.: A comprehensive grasp taxonomy. In: Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation (2009)
19. Feix, T., Romero, J., Schmiedmayer, H.B., Dollar, A.M., Kragic, D.: The grasp taxonomy of human grasp types. IEEE Trans. Hum.-Mach. Syst. (2015). In press
20. Gibson, J.J.: The Ecological Approach to Visual Perception. Psychology Press, Hove (1979)
21. Huang, H.J., Ahmed, A.A.: Tradeoff between stability and maneuverability during whole-body movements. PLoS One **6**(7), e21,815 (2011)
22. Jaillet, L., Porta, J.M.: Path planning under kinematic constraints by rapidly exploring manifolds. IEEE Trans. Robot. **29**(1), 105–117 (2013)
23. Kaiser, P., Gonzalez-Aguirre, D., Schültje, F., Sol, J.B., Vahrenkamp, N., Asfour, T.: Extracting whole-body affordances from multimodal exploration. In: Proceedings of the IEEE-RAS International Conference on Humanoid Robots (2014)
24. Kaiser, P., Grotz, M., Aksoy, E.E., Do, M., Vahrenkamp, N., Asfour, T.: Validation of whole-body loco-manipulation affordances for pushability and liftability. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids) (2015)
25. Kajita, S., Kanehiro, F., Kaneko, K., Fujiwara, K., Harada, K., Yokoi, K., Hirukawa, H.: Biped walking pattern generation by using preview control of zero-moment point. In: Proceedings of the IEEE International Conference on Robotics and Automation, vol. 2, p. 16201626 (2003)
26. Kamakura, N.: Te no katachi Te no ugoki. Ishiyaku, Tokyo (1989)
27. Kerr, J., Roth, B.: Analysis of multifingered hands. Int. J. Robot. Res. **4**(4), 3–17 (1986)
28. Krüger, N., Geib, C., Piater, J., Petrick, R., Steedman, M., Wörgötter, F., Ude, A., Asfour, T., Kraft, D., Omrčen, D., Agostini, A., Dillmann, R.: Object-action complexes: grounded abstractions of sensorimotor processes. Robot. Auton. Syst. **59**, 740–757 (2011)
29. Kwon, T., Shin, S.Y.: Motion modeling for on-line locomotion synthesis. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 29–38. ACM (2005)
30. Ma, R.R., Dollar, A.M.: On dexterity and dexterous manipulation. In: 2011 15th International Conference on Advanced Robotics (ICAR), pp. 1–7. IEEE (2011)
31. Mandery, C., Borràs, J., Jochner, M., Asfour, T.: Analyzing whole-body pose transitions in multi-contact motions. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids), pp. 5411–5418 (2015)

32. Mandery, C., Terlemez, O., Do, M., Vahrenkamp, N., Asfour, T.: The KIT whole-body human motion database. In: International Conference on Advanced Robotics (ICAR), pp. 329–336 (2015)
33. Mandery, C., Terlemez, O., Do, M., Vahrenkamp, N., Asfour, T.: Unifying representations and large-scale whole-body motion databases for studying human motion. IEEE Trans. Robot. **32**(4), 796–809 (2016)
34. Mason, M.T.: Chapter 4.3 Kinematic models of contact. Mechanics of Robotic Manipulation, pp. 86–88. The MIT Press, Cambridge (2001)
35. Mason, M.T.: Mechanics of Robotic Manipulation. MIT Press, Cambridge (2001)
36. Nakamura, Y.: Grasp and manipulation. J. Meas. Control **29**(3), 206–212 (1990) (Japanese)
37. Nakamura, Y., Nagai, K., Yoshikawa, T.: Dynamics and stability in coordination of multiple robotic mechanisms. Int. J. Robot. Res. **8**(2), 44–61 (1989)
38. Nori, F., Peters, J., Padois, V., Babic, J., Mistry, M., Ivaldi, S.: Whole-body motion in humans and humanoids. In: Workshop on New Research Frontiers for Intelligent Autonomous Systems (2014)
39. Orin, D., Oh, S.: Control of force distribution in robotic mechanisms containing closed kinematic chains. J. Dyn. Syst. Meas. Control **103**(2), 134–141 (1981)
40. Pas, A., Platt, R.: Localizing grasp affordances in 3-D points clouds using taubin quadric fitting. In: International Symposium on Experimental Robotics (ISER) (2014)
41. Popovi, M., Kraft, D., Bodenhagen, L., Baeski, E., Pugeault, N., Kragic, D., Asfour, T., Krüger, N.: A strategy for grasping unknown objects based on co-planarity and colour information. Robot. Auton. Syst. **58**(5), 551–565 (2010)
42. Porta, J.M., Ros, L., Bohigas, O., Manubens, M., Rosales, C., Jaillet, L.: The cuik suite: analyzing the motion closed-chain multibody systems. IEEE Robot. Autom. Mag. **21**(3), 105–114 (2014)
43. Saab, L., Ramos, O.E., Keith, F., Mansard, N., Soueres, P., Fourquet, J.Y.: Dynamic whole-body motion generation under rigid contacts and other unilateral constraints. IEEE Trans. Robot. **29**(2), 346–362 (2013)
44. Santello, M., Flanders, M., Soechting, J.F.: Postural hand synergies for tool use. J. Neurosci. **18**(23), 10105–10115 (1998)
45. Schiebener, D., Ude, A., Asfour, T.: Physical interaction for segmentation of unknown textured and non-textured rigid objects. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 4959–4966 (2014)
46. Terlemez, O., Ulbrich, S., Mandery, C., Do, M., Vahrenkamp, N., Asfour, T.: Master Motor Map (MMM) - framework and toolkit for capturing, representing, and reproducing human motion on humanoid robots. In: IEEE-RAS International Conference on Humanoid Robots (Humanoids), pp. 894–901 (2014)
47. Vahrenkamp, N., Kröhnert, M., Ulbrich, S., Asfour, T., Metta, G., Dillmann, R., Sandini, G.: Simox: A robotics toolbox for simulation, motion and grasp planning. In: International Conference on Intelligent Autonomous Systems (IAS), pp. 585–594 (2012)
48. Wächter, M., Schulz, S., Asfour, T., Aksoy, E., Wörgötter, F., Dillmann, R.: Action sequence reproduction based on automatic segmentation and object-action complexes. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids), pp. 189–195 (2013)
49. Warren, W.H.: Perceiving affordances: visual guidance of stair climbing. J. Expe. Psychol. **10**(5), 683–703 (1984)
50. Wieber, P.B.: On the stability of walking systems. In: Proceedings of the International Workshop on Humanoid and Human Friendly Robotics (2002)
51. Winter, D.A.: Biomechanics and Motor Control of Human Movement, 4th edn. Wiley, Hoboken (2009)
52. Yin, K., Loken, K., van de Panne, M.: Simbicon: simple biped locomotion control. ACM Trans. Gr. **26**(3), 105–1–105–10 (2007)