PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

LOLA v1.1 – An Upgrade in Hardware and Software Design for Dynamic Multi-Contact Locomotion

Philipp Seiwald[1,3], Shun-Cheng Wu[2], Felix Sygulla[1], Tobias F. C. Berninger[1], Nora-Sophie Staufenberg[1], Moritz F. Sattler[1], Nicolas Neuburger[1], Daniel Rixen[1] and Federico Tombari[2]

***Abstract*— This paper presents recent and ongoing hardware and software upgrades to our humanoid robot LOLA. The purpose of these modifications is to achieve dynamic multi-contact locomotion, i. e., fast bipedal walking with additional hand-environment support for increased stability and robustness against unforeseen disturbances. The upper body of LOLA has been completely redesigned with an enhanced lightweight torso frame and more robust arms with additional degrees of freedom, which extend the reachable workspace. The mechanical structure of the torso is optimized for stiffness with the help of an experimental modal analysis performed on the real hardware, while the new arm topology is the result of kinematic optimization for typical use-cases in a multi-contact setting. We also propose extensive changes to our software framework, which include a complete redesign of the onboard, real-time perception and navigation module. Although the hardware upgrade is finished and the overall software design is complete, the implementation of various modules is still work in progress.**

I. INTRODUCTION

Despite the steady progress made in academic and industrial research, full-sized humanoid robots have not made it yet to our everyday lives. Although various prototypes have demonstrated impressive skills, these are mostly restricted to specific tasks in a controlled environment. It seems that even recent high-end systems still do not display the versatility, robustness and reliability needed for real-world applications.

Nevertheless, even though the steps appear small (from a global perspective), continuous progress is being made. In 2004, the Chair of Applied Mechanics at TUM began developing the high-end prototype LOLA, to be used as a research platform for investigating the electromechanical design and real-time control of fast bipedal walking machines. Since recently, our focus has been more and more on dynamic multi-contact locomotion - by which we mean fast bipedal walking where additional hand-environment contacts take place. Our main intention is to improve stability and robustness against disturbances in environments typically made for humans, pushing research prototypes further into the direction of real-world applications. Figure 1 shows examples of our envisaged target scenarios.

Note that in the field of humanoid robotics, the term *multi-contact* is used ambiguously. In contrast to complex patterns such as climbing or crawling, we focus on *dynamic* locomotion, which preserves the main characteristics of bipedal gait. By restricting ourselves to gaited motion, we search for effective methods that maintain real-time performance.

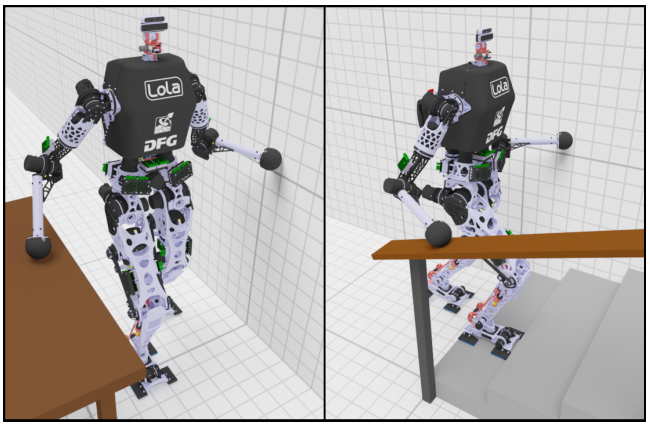

Fig. 1. Target scenarios in a multi-contact setting. When walking, the robot should be able to support itself against walls, tables and handrails to increase its stability margins. Renderings show the new version of LOLA.

In this paper, we present recent and ongoing upgrades to LOLA, which are intended to ready it for the kind of multi-contact scenario described above. The presented work is part of a collaboration between the Chair of Applied Mechanics at TUM, which focuses on hardware design, motion planning and control, and the Chair for Computer Aided Medical Procedures & Augmented Reality at TUM, which investigates suitable algorithms for computer vision. The contribution of this paper is a concise presentation of:

- realized and planned changes in hardware and software,
- tools used for redesigning critical elements, and
- the development process as a guide to future designs.

The work presented covers several disciplines, which can not all be presented here in full detail. We therefore restrict ourselves to the redesign of the hardware, which is now complete (and for which a live demonstration is planned at the *2020 IEEE-RAS International Conference on Humanoid Robots*). Moreover, we propose an extension to our software framework, which incorporates extensive changes and should be regarded as work in progress.

The following section presents related work clustered into the topics hardware design, perception and motion planning. Section III provides an overview of the limitations of the previous system along with the requirements of a multi-contact scenario. The redesigned hard- and software are presented in Sections IV and V, respectively. Finally, preliminary results and general conclusions are given in Sections VI and VII.

[1]Technical University of Munich, Chair of Applied Mechanics, Boltzmannstraße 15, 85748 Garching, Germany
[2]Technical University of Munich, Chair for Computer Aided Medical Procedures & Augmented Reality, Boltzmannstraße 3, 85748 Garching, Germany
[3]Corresponding author; e-mail: philipp.seiwald@tum.de

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

II. Related Work

A. Upper Body Design for Humanoids

Since the design of LOLA's hip, pelvis, and leg [1] has proven to perform very well in versatile and dynamic walking, cf. [2], [3], we decided to keep it unchanged and only focus on modifications to the upper body. The main purpose of the torso is to connect the limbs and to integrate core components. Despite its central role, there is only little literature devoted to investigating its mechanical design. This may be because, in most cases, the torso has no active or passive *Degrees of Freedom (DoF)*. However, it has been shown in [4] that the structural dynamics of a stiff torso can be critical to the performance of the low-level controller and need to be designed carefully. We will return to this issue in Section III. Examples of torso designs with active spines driven by a parallel mechanism are CAUTO [5] and CHARLIE [6]. In our case, the moderate increase in suppleness does not justify the greater complexity for which reason we stick to a stiff torso. However, we keep the two existing DoFs located in LOLA's pelvis, which significantly increase the kinematic capabilities for performing challenging maneuvers.

In contrast, there are numerous studies on the electromechanical design of humanoid arms. Most of these share the objective of replicating the versatility, flexibility and strength of the human arm. We make specific reference to [7] as a preliminary study, as it presents a simplified kinematic model of the human arm and the corresponding workspace. The proposed topology features a 2 DoF *inner shoulder*, representing the sternoclavicular joint, a 3 DoF *outer shoulder*, representing the glenohumeral and acromioclavicular joint, and a single DoF for elbow flexion/extension. Unlike complex biomechanical models, this model draws no explicit connections to the human anatomy but instead identifies the kinematics from experimental data obtained from visually tracked human subjects. It is thus well-suited to the design of robotic arms which have similar kinematics but typically apply other principles of actuation. The influence of the inner shoulder joint is discussed in [8], while the model is extended by an additional translational DoF between the inner and outer shoulder joint in [9].

The majority of humanoid arm designs is similar to those found in classic robotic manipulators but with focus on low weight and compliance. Examples are the upper-body designs of ARMAR-III and ARMAR-4 presented in [10] and [11], respectively. While ARMAR-III features a classical 7-DoF arm, an inner shoulder joint is added in ARMAR-4. Further development lead to the KIT DUAL ARM SYSTEM [12], which was specifically designed for two-handed manipulation tasks. Another example following the 'classical' approach is JUSTIN [13], which combines two DLR-LWR-III manipulators for dexterous manipulation. A study devoted to the optimum arrangement of shoulder joints in humanoid robots is presented in [14]. There are also a number of arm designs with alternative principles of actuation, such as the tendon-driven DAVID [15] and BLADE [16], which attempt to replicate the dexterity and flexibility of the human arm and shoulder, respectively. Although close to human anatomy, such systems are rather complex. Finally, there are various hydraulically and pneumatically actuated systems, which also include hardware driven by artificial muscles. Unfortunately, these typically suffer from high response times, a high mass due to the fluid supply, and modeling difficulties, especially with braided artificial muscles, cf. [17].

B. Visual Perception for Humanoids

So that it can autonomously navigate its environment, a robot has to maintain a meaningful model of its surroundings. One of the first methods of dense real-time scene reconstruction based on active depth sensors to be proposed was KINECTFUSION [18]. It is based on a voxel representation with a *Truncated Signed Distance Field (TSDF)* for integrating depth information into the scene. Each input depth map is treated as a surface, for which frame-to-model tracking and mapping is performed. The drawbacks of this method are its high memory consumption and lacking of relocalization. NIESSNER et al. [19] propose a voxel hash table to improve memory usage. They divide the 3D space into small chunks and only assign memory to a chunk if a point is observed in it. This means that sparsity can be dealt with by hashing. ELASTICFUSION [20] adopts the 3D data representation introduced in Keller et al. [21] replacing voxels with surfels, which enables the same frame-to-model tracking method to be used and a dense surface to be reconstructed with much lower memory consumption. They also propose a relocalization method based on a deformation graph to reduce the error via non-rigid optimization. INFINITAM [22] is a framework which uses KINECTFUSION with voxel hashing as in [19]. Also, as does ELASTICFUSION, it uses randomized ferns [23] to resume the pose when tracking is lost.

Aiming in another direction, HORNUNG et al. [24] propose mapping 3D space by octree-based occupancy fusion to reduce memory consumption. However, unlike TSDF-based approaches, occupancy fusion is not suitable for providing fine scene reconstruction details, i. e., at sub-voxel resolution. Also, the boundaries of occupancy maps are not clearly defined. VESPA et al. [25] modify this approach by choosing a standard deviation value in the occupancy distance function that is proportional to the measured depth distance for more realism with respect to the depth camera noise model.

STASSE et al. [26] published one of the first implementations of a *Simultaneous Localization and Mapping (SLAM)* system for humanoid robots. Their method requires the scene (which is represented by a sparse point map) to be static. SCONA et al. [27] use ELASTICFUSION to fuse motion data obtained from the humanoid VALKYRIE. Their method is able to construct a semi-dense map and handle slightly dynamic environments. ZHANG et al. [28] propose a method to handle image blur from robot motion, dynamic scene elements and tracking failures, based on a surfel fusion system [20]. Dynamic scene parts are assumed to come from human motion. GROTZ et al. [29] propose a method of simultaneously reconstructing and reasoning about a scene. Their method predicts object classes by combining

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

a primitive detection approach based on a scene graph [30] with an off-the-shelf object detector [31]. However, their approach does not consider completing the object geometry, and their perception system is not able to handle rapid viewpoint changes and lost tracking.

C. Motion Planning for Multi-Contact Locomotion

Again, we wish to emphasize that the term *multi-contact* is used ambiguously in relation to bipeds. It can refer to multiple contacts per foot [32], additional hand-environment contacts for gaited patterns like stair climbing [33], or highly complex non-gaited motions such as climbing a ladder or crawling through a tunnel [34]. The different meanings relate mainly to different target applications. Note that the presented modifications to LOLA are aimed at the second type of applications, i. e., *gaited* bipedal walking augmented by (unilateral) hand support.

A core aspect of multi-contact planning is the selection of an appropriate stability criterion. BRETL et al. [35] proposed an algorithm for computing a support region for static equilibrium. This is applied for example in [36], where contact-consistent elastic strips are used to detect a contact sequence and the motion in-between. Note that the popular *Zero-Moment Point (ZMP)* [37] for dynamic walking applies solely to co-planar contacts. However, (rather complex) extensions exist for the multi-contact case, cf. [38].

In [39], a method of planning contact sequences is proposed. Visual input is used to assign affordances to individual surfaces in the robot's surrounding. The choice of contact partners is based on a model learned from observed human motion. The computation time is in the range of few seconds; however, neither the transition between poses, nor dynamics are considered. WERNER et al. [33] propose a scheme to find quasi-stable configurations by a quadratic program with appropriate constraints. Moreover, the transitions between configurations are computed using a constrained, bidirectional, rapidly-exploring random tree. This method is fast enough for real-time execution, but contact points for hand support have to be chosen manually, and the generated motion is quasi-static. A common approach for planning complex non-gaited motions is to formulate a comprehensive optimization problem. As a very generic approach, it can be applied to various scenarios and tasks. However, the computation time is high, and care must be taken to avoid local minima. A framework based on this approach is presented in [40].

III. PREVIOUS LIMITATIONS AND REQUIREMENTS

We omit a detailed description of LOLA's previous hard- and software setup. Instead we refer to various publications that cover its core components: [1] and [41] (electro-mechanical hardware), [42] (communication), [43] (computer vision), [44] (navigation), [45] (motion planning), [46] (stabilization), [47] (collision avoidance), [48] (disturbance rejection) and [49] (simulation). In this section, we will focus on the *limitations* of the previous setup with respect to the multi-contact scenario, i. e., the reason for the upgrade.

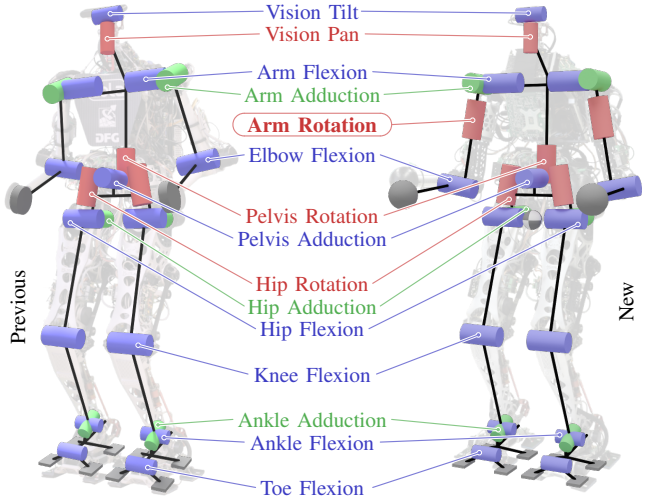

Fig. 2. Previous (left, 24 DoF) and new (right, 26 DoF) kinematic topology of LOLA. The new upper body design adds "Arm Rotation Right/Left".

A. Hardware

LOLA's previous topology compromised a total of 24 DoFs[1], see Figure 2 left. The design placed the focus on sophisticated legs, leaving the upper body somewhat rudimentary. In particular, each arm featured only three joints, which were used to compensate leg dynamics during fast walking. However, in consideration of the additional hand contacts, the reachable workspace of the hands was rather poor, especially in the robot's lateral proximity, see Figure 3. Moreover, the 'hands' were simple dummy weights, which lacked any sensor or dedicated contact surface.

In addition to the humble arm design, the torso also had its issues. the stiffness of the backbone and shoulders - a T-shaped, riveted tubular scaffolding - was too low. As a result, the first critical eigenmode (visible to the low-level controller) occurred at 9.7 Hz already, see [4] for a full *Experimental Modal Analysis (EMA)* of the previous setup. Unfortunately, bipedal walking excites modes in such low frequency domains, which limits the performance of the low-level controller, [4]. Finally, it was clear that the previous upper body design would not have been able to withstand the increased loads in a multi-contact setting.

B. Software

The previous computer vision system used the input of a single depth sensor to classify the environment into floor, platforms and obstacles, see [43] for details. However, a multi-contact scenario requires a more sophisticated environment model. In particular, the new vision system has to be able to distinguish between possible contact partners, such as a solid wall and surfaces that are unsafe for interaction but are still considered obstacles, such as a swivel chair.

The environment model is then transferred to the navigation module. In the past, path planning was restricted

[1]Note that the very first version of LOLA had 25 DoFs, due to its additional DoF for camera vergence, [1]. However, as the stereo camera system was replaced by an RGB-D sensor, only pan and tilt remained.

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

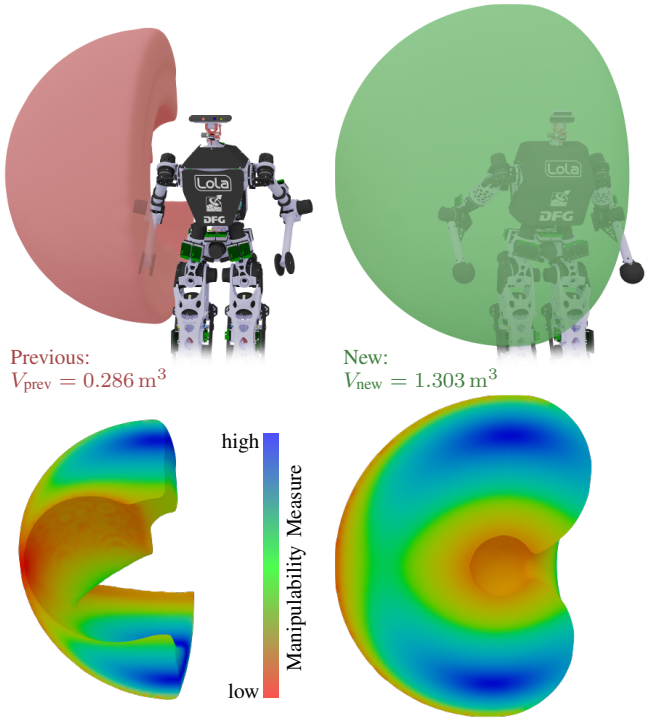

Fig. 3. The hand's reachable task space with previous (left) and new (right) arm kinematics. Here, the task space is defined as the Cartesian position of the right hand relative to the (fixed) torso. The volume is discretized by voxels of 1 cm size and sampled with $1.1 \cdot 10^8$ (previous, 3 DoF) and $1.5 \cdot 10^{11}$ (new, 4 DoF) configurations. For each sample, the actual joint limits are respected, but collisions of the hand with the torso or lower body are not considered. As a post-processing step, the surface of the volume is extracted using the *Marching Cubes* algorithm [50] and subsequently smoothed. Top: the robot's reachable volume. Bottom: section view of the (per-voxel maximum) manipulability measure as defined in [51].

to simple bipedal walking. Unfortunately, a naive extension of the search space to incorporate additional hand-support would violate our real-time constraints. Another issue was the robot's poor localization which was based solely on odometry and suffered from significant drift due to slipping contacts. Since the high-performance *Inertial Measurement Unit (IMU)* located in the torso of LOLA is also subject to drift, the only (fully autonomous) solution turns out to be localization by visual input. Finally, up to now, the hand motion was not accounted for explicitly during the planning, but was merely a result of a nullspace optimization to reduce the angular momentum of the robot, [47]. However, the multi-contact scenario requires a transition of the hand position from nullspace to task space and vice versa.

IV. HARDWARE REDESIGN

A. Kinematic Optimization

A complete redesign of the upper body allows us to overcome previous barriers; but it also poses a broad range of questions. One of these is the optimal kinematic topology of the arms. Although LOLA's new arm design is inspired by humans, we do not attempt to replicate it. This is largely because we do not consider grasping or manipulation, preferring simplicity, stiffness and dynamic performance over

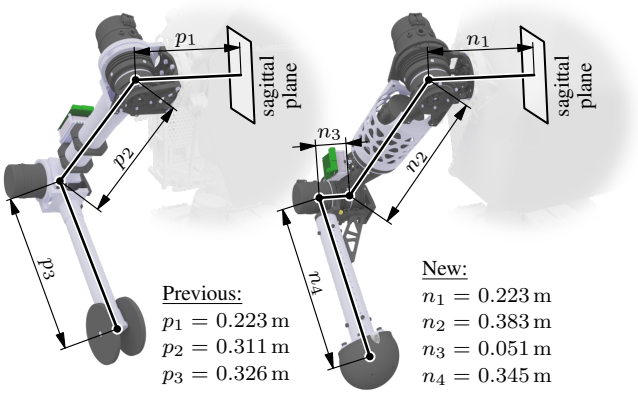

Fig. 4. Segment proportions in the previous (left, 3 DoF) and new (right, 4 DoF) arm designs.

dexterity. Neither do we consider an inner shoulder joint. Instead, we performed a kinematic optimization to find an optimal choice of joint arrangement and segment proportions for certain test cases in a multi-contact setting, [52]. We omit a detailed explanation of the optimization process, however, its steps can be summarized as follows:

- Evaluate the task space of the topology candidate with respect to various metrics on a per-voxel basis.
- Sum up the (weighted) per-voxel metrics over a task-specific volume to obtain a global cost value.
- Choose a new topology candidate and repeat.

Note that the 'task space' can be defined in different ways, depending on the action under consideration. In general, however, we define it as the position and orientation of the hand alias *Tool Center Point (TCP)*. A full description of the procedure is given in [52]. We would like to point out that a generalized version of the subroutine for evaluating the task space of an arbitrary kinematic chain is available in our open-source C++ library `broccoli`[2], [53]. The published code was used to perform the comparison of the task space presented in Figure 3. Since it is always difficult to represent a volume on a two-dimensional medium, we recommend viewing the animated task space visualization in the accompanying video: https://youtu.be/mpDqMFppT68.

A comparison of the segment proportions for the previous and new arm design is shown in Figure 4. The new topology is the combined result of the kinematic optimization together with 'general' design considerations. An example of the latter is the additional offset n_3, which was introduced to minimize the risk of collision in typical multi-contact scenarios.

B. Actuation and Sensing

From a pure kinematic perspective, since all three axes of the shoulder joint have a common point of intersection, the specific order of the DoFs does not matter. However, if we additionally consider joint limits or varying drive types, then the arrangement has a significant effect. In particular, the choice of 'sufficient' actuator torques has proven to be

[2]See `TaskSpaceEvaluator` in the `analysis` module.

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

TABLE I

DRIVE SPECIFICATION OF THE NEW 4 DOF ARM DESIGN.

Drive	τ_{perm}	τ_{peak}	τ_{human}	$\Delta\varphi$	Mass
Arm Flexion	49 Nm	147 Nm	61 Nm	360°	1.14 kg
Arm Adduction	31 Nm	99 Nm	30 Nm	180°	0.97 kg
Arm Rotation	31 Nm	99 Nm	34 Nm	360°	1.05 kg
Elbow Flexion	31 Nm	99 Nm	49 Nm	182°	0.97 kg

difficult, as there is no general design rule. Ideally, LOLA should be able to perform supporting motions similar to humans. Thus, we chose a pragmatic solution and designed the maximum joint torques to be similar to an adult human. For this, we used common maximum interaction forces[3] from the military and NASA standards, [54] and [55], respectively.

The joint drives have a custom, lightweight design and consist of a brushless DC motor, a Harmonic Drive gear, a motor-side incremental encoder, and a link-side absolute encoder packed into a custom housing, see [1] for details. Note that, with the exception of 'Vision Pan/Tilt', LOLA's drive system is organized into modules, [1] . In particular, we reuse modules D and E for arm flexion and rotation and the module F for arm adduction and elbow flexion, respectively. The drive specification is presented in Table I and shows the maximum permanent torque τ_{perm}, the maximum peak torque τ_{peak}, the equivalent maximum human torque[4] τ_{human}, and the mass of the drive. Note that the available joint range $\Delta\varphi$ exceeds human capabilities.

Finally, we restrict ourselves to unilateral contacts, i. e., pushing interactions. Thus, we omit a dedicated wrist and hand design. This significantly reduces the complexity, mass and cost of the design. However, we did add the commercial six-axis force-torque sensors Schunk FTE-Axia80 to the 'wrist' to enable the measurement of interface forces in a multi-contact scenario. These sensors feature a compact, fully integrated design and are directly connected to LOLA's main EtherCAT communication bus running at 4 kHz.

C. Torso Design

The main drawbacks of the previous torso design were the insufficient stiffness of its backbone and that it would not have been able to withstand the increased loads. However, there were also other issues, such as the very limited extensibility and complex power bus system. Thus, we decided to redesign the complete torso from scratch. The mechanical structure was changed from a T-shaped, tubular scaffolding to a more cube-shaped framework, see Figure 5 top. This significantly increases the second moment of area, especially in directions identified as 'critical' in the EMA of the previous hardware. Moreover, instead of connections through lightweight rivets, we switched to a (slightly heavier) screwed tube system, which is more robust against vibrations and simplifies future extension. As with the rest of LOLA, most parts are made by milling and turning using high-strength aluminum alloys.

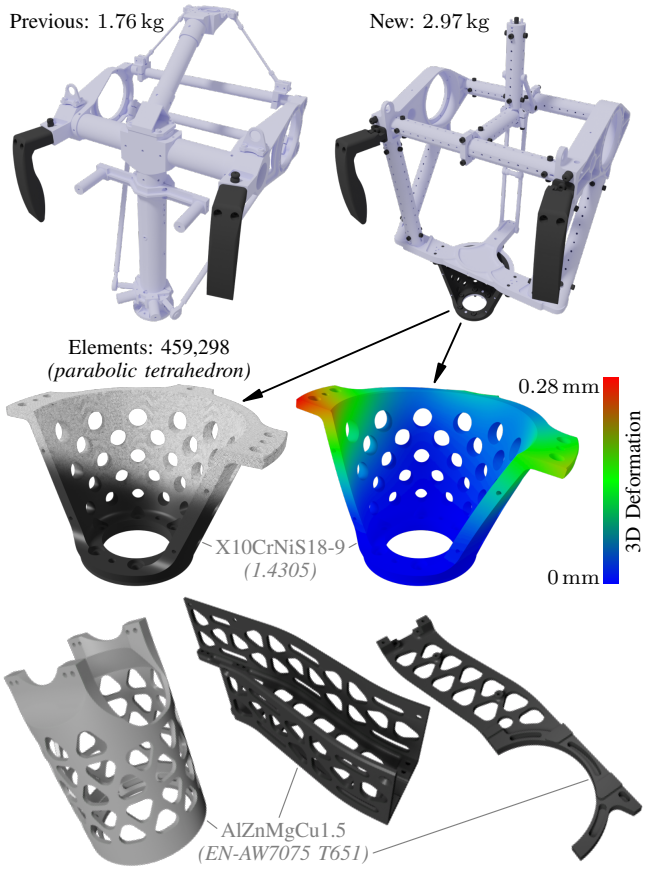

Fig. 5. Top: rear view of previous (left) and new (right) torso frame. Center: FEM analysis of central mounting flange (stainless steel) coupling the lower and upper body of LOLA. The visible deformation of the mesh (right) is scaled by a factor of 50. Bottom: lightweight structures (aluminum alloy) of the new arm design milled from the solid.

To enable the new design to withstand the increased loads of a multi-contact scenario, critical parts were identified and analyzed using the *Finite Element Method (FEM)*, see Figure 5 center. To do this, we defined a load case in which interaction forces of 100 N are applied to both hands in all three axes simultaneously. Note that an FEM analysis requires the definition of reasonable loads, which is difficult if the part under investigation has many mechanical interfaces. Thus, most parts of the new upper body (including the arms) are designed in accordance with conventional engineering practice for low mass and high stiffness, omitting a dedicated FEM analysis, see Figure 5 bottom.

In order to simplify the power supply system, each electronic component of the robot is now connected directly to an input voltage of either 80 V (motors) or 24 V (rest). Note that LOLA still does not have an onboard power supply, although this might be a feature of a future revision. As part of the overall upgrade, we enhanced the onboard computing power to feature two identical Mini-ITX industrial boards, each with an Intel Core i7-8700 (6x3.2 GHz) CPU and 32GB DDR4 RAM. One of these PCs runs QNX Neutrino

[3]In these standards, the maximum permitted interaction force refers mainly to manipulation. However, we assume that typical supporting motions cause similar loads.

[4]According to [54, p.80,330] the specified (maximum) torques τ_{human} can be applied by 95% of male subjects in the US General Forces.

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

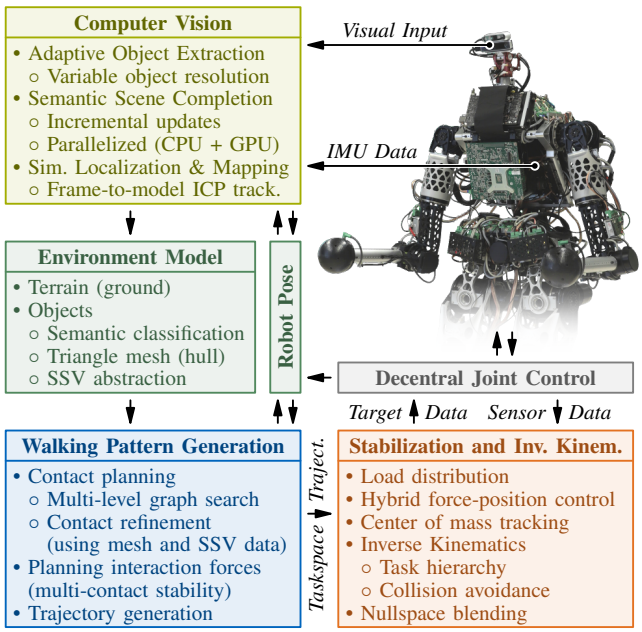

Fig. 6. Extension of LOLA's software framework for dynamic multi-contact locomotion. The redesign includes major upgrades of the existing pipeline for bipedal walking. Modifications of the modules *Computer Vision (CV)*, *Walking Pattern Generation (WPG)*, and *Stabilization and Inverse Kinematics (SIK)* are currently in development.

RTOS 7.0 (64bit) and is used for our planning and control algorithms, while the other is dedicated to computer vision tasks and runs Ubuntu 20.04 (64bit). The vision PC has an additional Nvidia Quadro P2000 GPU accelerator and communicates with the control PC via Gigabit Ethernet. Although the head DoFs alias 'Vision Pan/Tilt' remain the same, they have been slightly modified to accommodate the new Intel RealSense D435 and T265 vision sensors.

V. SOFTWARE REDESIGN

Figure 6 contains an overview of the planned software framework. Note that the design of its architecture is complete, however, the underlying algorithms are still in development. Since the proposed modifications are extensive, we only focus on the core components of each module.

A. Computer Vision (CV)

The CV system processes visual input to build an *Environment Model*, including a rough reconstruction of the terrain and a more detailed representation of possible collision and/or contact partners. For each such object, a corresponding semantic classification, a triangle mesh describing its surface, and an abstract geometric representation through *Swept Sphere Volumes (SSV)*, cf. [43], is generated.

To achieve this, we propose a framework consisting of two pipelines: *Adaptive Object Extraction (AOE)* and *Semantic Scene Completion (SSC)*. The framework is built on top of INFINITAM [22] and uses an occupancy distance function from [25]. The AOE pipeline extracts objects from the reconstructed scene at an adaptive resolution, to accelerate the conversion of the occupancy probability to representative SSV objects. The resolution is determined by the object type and the size of its visible surface area. The SSC pipeline, [56], uses a data-driven approach to incrementally predict a completion for geometric shapes in the reconstructed scene, while reasoning about its semantic information. In order to maintain real-time performance, we trigger the SSC pipeline asynchronously and update regions incrementally. Moreover, the completion network is accelerated through parallel processing on the onboard CPU and GPU.

Besides the environment model, a shared 6D robot pose will also be maintained by the CV and WPG module. This refers to a localization of the robot which is continuously updated and corrected by visual input, IMU data, and odometry obtained from the planned motion. Note that for the new torso design, we reused the high-precision IMU and placed it (again) in the center of the torso scaffold, see Figure 6. The visual localization is run by a SLAM algorithm, which uses the frame-to-model *Iterative Closest Point (ICP)* tracking method of [18] for localization in the reconstructed scene.

B. Walking Pattern Generation (WPG)

The environment model and the localized robot pose are then used by the WPG module for contact planning, i. e., to find a sequence of foothold and (optional) hand-contacts. Because of the additional hand-environment contacts, this involves a high-dimensional search space which turns out to be infeasible for real-time execution. However, an internal pre-study showed that a hierarchical graph search drastically reduces the total computation time and thus meets our real-time requirements. In particular, we split the problem into a coarse-grained first-level search and a subsequent fine-grained second-level search. The latter is augmented by the optimum 'path' found in the first level. Once a feasible contact sequence is found, the result can be refined in a postprocessing step, using shape information from the environment model to find optimal contact positions.

To generate feasible task space trajectories, we retain our ZMP based approach, see [45]. In order to incorporate the dynamic effects of the additional hand-environment contacts, we consider them as (planned) disturbances. These forces are heuristically determined from semantic information relating to the contact partner, which we obtain from our environment model. The disturbance is then fed into the equations of motion of a five-mass model (approximating the robot dynamics) used to determine the *Center of Mass (CoM)* motion, see [45]. We are aware that the traditional ZMP concept is restricted to planar contacts, and that a heuristically defined interaction force does not provide the full potential of multi-contact locomotion. However, as stated in the introduction, we are focusing on (augmented) *bipedal* gait and try to find a real-time solution.

C. Stabilization and Inverse Kinematics (SIK)

The SIK module stabilizes the feasible trajectories from the WPG module (CoM and foot/hand motion, desired CoM wrench) by means of feedback control. A balance controller

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

modifies the desired CoM torques on the basis of IMU inclination data. The resulting wrench is realized through vertical acceleration of the CoM and contact force control, see [46]. Currently, the total wrench is distributed between the end effectors in contact using a heuristic, see [49]. A more general solution for arbitrary hand and foot contacts is subject to ongoing research. For the contact controllers, which modify the end effector trajectories in task space, separate instances are planned for each end effector with an explicit contact model. The final trajectories will then be processed with an *Inverse Kinematics (IK)* solution capable of blending between different task space representations. The motion of the arm can then either be defined in the nullspace by secondary tasks, e. g. angular momentum minimization and collision avoidance, or is made part of the task space for deliberately making contact with the environment. The resulting joint-space trajectories are sent to the robot's hardware control architecture [42] and tracked by decentralized joint position controllers.

VI. Preliminary Results

The hardware upgrade of LOLA is complete. The new upper body is shown in Figure 6, but we also recommend watching the video https://youtu.be/mpDqMFppT68, which shows the assembly of the new upper body and visualizes the FEM- and task space analysis in 3D. All in all, the upgrade increased the DoF count to 26 (+2), the total height to $1.763\,\mathrm{m}$ (+1.21 %), and the total mass to $68.2\,\mathrm{kg}$ (+7.74 %).

A. Experimental Modal Analysis (EMA)

In order to validate if the new torso design overcomes the issues mentioned in Section III-A, we repeated the EMA of [4] with the new hardware. Our measurements show that the torso now behaves much more rigidly, see Figure 7, even for higher frequency mode shapes. The dynamic behavior of the mechanical structure is now dominated by deformations in the arms and legs, while the torso and hip stay quite rigid relative to each other (up to 22 Hz). This is a significant improvement, which allows us to further push the bandwidth of our controllers. See [57] for details on the new EMA.

B. Initial Operation

We successfully performed initial tests to validate and demonstrate the basic functionality of the new hardware, see https://youtu.be/JCYmq6u0EEc. The video shows basic stabilization during standing and simple bipedal walking.

C. Computer Vision

Although the new software framework for LOLA is still under development, some components of the CV module have already been published, e. g. SCFusion, see [56] and the related video https://youtu.be/LtTFgASSLWc.

VII. Conclusions and Next Steps

The development of complex systems such as a humanoid robot is always a big challenge. During design, manufacturing, assembly, and initial operation, various problems emerged which (luckily) could be fixed without undue effort.

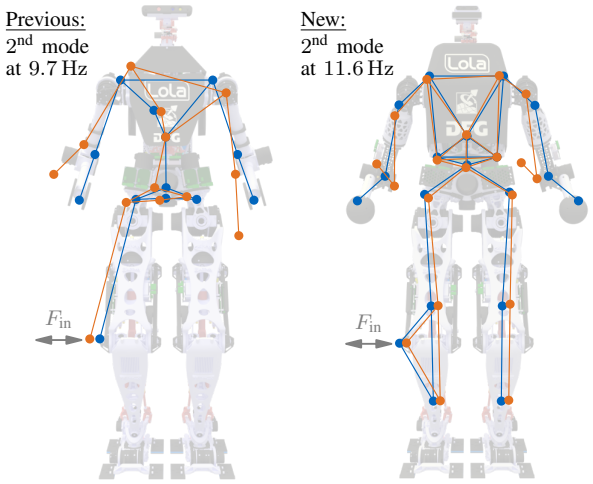

Fig. 7. Experimental modal analysis: The critical second mode shape (orange) of LOLA's mechanical structure with the undeformed sensor positions (blue) are shown. The previous (left) and new (right) system have both been excited at the right knee, cf. F_{in}. See [57] for details.

We attribute this to our structured development process, which we have attempted to outline in this paper. For similar projects, we recommend avoiding over-engineering in favor of pragmatic and robust solutions that can be refined in subsequent iterations.

Initial experiments with the new hardware show promising results; the main issues of the previous system seem to have been resolved and our extensions are working as expected. Although the hardware upgrade is a major milestone, more work is currently performed to accomplish the presented extensions to the software framework.

Acknowledgment

This work is supported by the *German Research Foundation* (DFG, project number 407378162). We would like to give special thanks to Simon Gerer, Georg König, Andreas Köstler and Georg Mayr for their invaluable support during manufacturing, assembling, and wiring the new system. Moreover, we thank our students Daniel Pölzleitner and Reinhold Poscher for their assistance in the CAD design.

References

[1] S. Lohmeier, "Design and Realization of a Humanoid Robot for Fast and Autonomous Bipedal Locomotion," Dissertation, TUM, 2010. [Online]. Available: https://mediatum.ub.tum.de/980754

[2] T. Buschmann, V. Favot, S. Lohmeier, *et al.*, "Experiments in Fast Biped Walking," in *IEEE ICM*, 2011.

[3] A.-C. Hildebrandt, R. Wittmann, F. Sygulla, *et al.*, "Versatile and robust bipedal walking in unknown environments: real-time collision avoidance and disturbance rejection," *Autonomous Robots*, 2019.

[4] T. F. C. Berninger, F. Sygulla, S. Fuderer, *et al.*, "Experimental Analysis of Structural Vibration Problems of a Biped Walking Robot," in *IEEE ICRA*, 2020.

[5] D. Cafolla and M. Ceccarelli, "An Experimental Validation of a Novel Humanoid Torso," *Robotics and Autonomous Systems*, vol. 91, 2017.

[6] D. Kühn, A. Dettmann, and F. Kirchner, "Analysis of Using an Active Artificial Spine in a Quadruped Robot," in *IEEE ICCAR*, 2018.

[7] J. Lenarčič and A. Umek, "Simple Model of Human Arm Reachable Workspace," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 8, 1994.

PREPRINT - Final published paper: https://doi.org/10.1109/HUMANOIDS47582.2021.9555790

[8] J. Lenarčič and N. Klopčar, "Positional kinematics of humanoid arms," *Robotica*, vol. 24, no. 1, 2006.

[9] N. Klopčar and J. Lenarčič, "Kinematic Model for Determination of Human Arm Reachable Workspace," *Meccanica*, vol. 40, 2005.

[10] A. Albers, S. Brudniok, J. Ottnad, *et al.*, "Upper Body of a new Humanoid Robot - the Design of ARMAR III," in *IEEE-RAS Humanoids*, 2006.

[11] T. Asfour, J. Schill, H. Peters, *et al.*, "ARMAR-4: A 63 DOF Torque Controlled Humanoid Robot," in *IEEE-RAS Humanoids*, 2013.

[12] S. Rader, L. Kaul, H. Fischbach, *et al.*, "Design of a High-Performance Humanoid Dual Arm System with Inner Shoulder Joints," in *IEEE-RAS Humanoids*, 2016.

[13] C. Ott, O. Eiberger, W. Friedl, *et al.*, "A Humanoid Two-Arm System for Dexterous Manipulation," in *IEEE-RAS Humanoids*, 2006.

[14] M. Bagheri, A. Ajoudani, J. Lee, *et al.*, "Kinematic Analysis and Design Considerations for Optimal Base Frame Arrangement of Humanoid Shoulders," in *IEEE ICRA*, 2015.

[15] M. Grebenstein, A. Albu-Schäffer, T. Bahls, *et al.*, "The DLR Hand Arm System," in *IEEE ICRA*, 2011.

[16] Y. Sodeyama, I. Mizuuchi, T. Yoshikai, *et al.*, "A Shoulder Structure of Muscle-Driven Humanoid with Shoulder Blades," in *IEEE/RSJ IROS*, 2005.

[17] B. Tondu, S. Ippolito, J. Guiochet, *et al.*, "A Seven-degrees-of-freedom Robot-arm Driven by Pneumatic Artificial Muscles for Humanoid Robots," *The International Journal of Robotics Research*, vol. 24, no. 4, 2005.

[18] R. A. Newcombe, S. Izadi, O. Hilliges, *et al.*, "KinectFusion: Real-time dense surface mapping and tracking," in *IEEE ISMAR*, 2011.

[19] M. Nießner, M. Zollhöfer, S. Izadi, *et al.*, "Real-time 3D Reconstruction at Scale Using Voxel Hashing," *ACM Transactions on Graphics*, vol. 32, no. 6, 2013.

[20] T. Whelan, S. Leutenegger, R. Salas-Moreno, *et al.*, "ElasticFusion: Dense SLAM without a pose graph," in *Robotics: Science and Systems*, 2015.

[21] M. Keller, D. Lefloch, M. Lambers, *et al.*, "Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion," in *IEEE 3DV*, 2013.

[22] V. A. Prisacariu, O. Kähler, S. Golodetz, *et al.*, "InfiniTAM v3: A Framework for Large-Scale 3D Reconstruction with Loop Closure," *arXiv preprint: 1708.00783*, 2017.

[23] B. Glocker, J. Shotton, A. Criminisi, *et al.*, "Real-Time RGB-D Camera Relocalization via Randomized Ferns for Keyframe Encoding," *IEEE TVCG*, vol. 21, no. 5, 2015.

[24] A. Hornung, K. M. Wurm, M. Bennewitz, *et al.*, "OctoMap: an efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, 2013.

[25] E. Vespa, N. Nikolov, M. Grimm, *et al.*, "Efficient Octree-Based Volumetric SLAM Supporting Signed-Distance and Occupancy Mapping," *IEEE RAL*, vol. 3, no. 2, 2018.

[26] O. Stasse, A. J. Davison, R. Sellaouti, *et al.*, "Real-Time 3D SLAM for Humanoid Robot considering Pattern Generator Information," in *IEEE/RSJ IROS*, 2006.

[27] R. Scona, S. Nobili, Y. R. Petillot, *et al.*, "Direct Visual SLAM Fusing Proprioception for a Humanoid Robot," in *IEEE/RSJ IROS*, 2017.

[28] T. Zhang, E. Uchiyama, and Y. Nakamura, "Dense RGB-D SLAM for Humanoid Robots in the Dynamic Humans Environment," in *IEEE-RAS Humanoids*, 2018.

[29] M. Grotz, P. Kaiser, E. E. Aksoy, *et al.*, "Graph-Based Visual Semantic Perception for Humanoid Robots," in *IEEE-RAS Humanoids*, 2017.

[30] R. Schnabel, R. Wessel, R. Wahl, *et al.*, "Shape Recognition in 3D Point-Clouds," in *WSCG*, 2008.

[31] J. Redmon, S. Divvala, R. Girshick, *et al.*, "You only look once: Unified, real-time object detection," in *IEEE CVPR*, 2016.

[32] J. Lack, M. J. Powell, and A. D. Ames, "Planar Multi-Contact Bipedal Walking Using Hybrid Zero Dynamics," in *IEEE ICRA*, 2014.

[33] A. Werner, B. Henze, D. A. Rodriguez, *et al.*, "Multi-Contact Planning and Control for a Torque-Controlled Humanoid Robot," in *IEEE/RSJ IROS*, 2016.

[34] A. Escande, A. Kheddar, and S. Miossec, "Planning contact points for humanoid robots," *Robotics and Autonomous Systems*, vol. 61, no. 5, 2013.

[35] T. Bretl and S. Lall, "Testing Static Equilibrium for Legged Robots," *IEEE T-RO*, vol. 24, no. 4, 2008.

[36] S.-Y. Chung and O. Khatib, "Contact-Consistent Elastic Strips for Multi-Contact Locomotion Planning of Humanoid Robots," in *IEEE ICRA*, 2015.

[37] M. Vukobratović and B. Borovac, "Zero-Moment Point – Thirty Five Years of its Life," *International Journal of Humanoid Robotics*, vol. 1, no. 1, 2004.

[38] S. Caron, Q.-C. Pham, and Y. Nakamura, "ZMP Support Areas for Multicontact Mobility Under Frictional Constraints," *IEEE T-RO*, vol. 33, no. 1, 2017.

[39] P. Kaiser, C. Mandery, A. Boltres, *et al.*, "Affordance-Based Multi-Contact Whole-Body Pose Sequence Planning for Humanoid Robots in Unknown Environments," in *IEEE ICRA*, 2018.

[40] K. Bouyarmane, S. Caron, A. Escande, *et al.*, *Multi-contact Motion Planning and Control*. Springer Netherlands, 2019.

[41] V. Favot, "Hierarchical Joint Control of Humanoid Robots," Dissertation, TUM, 2016. [Online]. Available: https://mediatum.ub.tum.de/?id=1294180

[42] F. Sygulla, R. Wittmann, P. Seiwald, *et al.*, "An EtherCAT-Based Real-Time Control System Architecture for Humanoid Robots," in *IEEE CASE*, 2018.

[43] D. Wahrmann, A.-C. Hildebrandt, T. Bates, *et al.*, "Vision-Based 3D Modeling of Unknown Dynamic Environments for Real-Time Humanoid Navigation," *International Journal of Humanoid Robotics*, vol. 16, no. 01, 2019.

[44] A.-C. Hildebrandt, M. Klischat, D. Wahrmann, *et al.*, "Real-Time Path Planning in Unknown Environments for Bipedal Robots," *IEEE RA-L*, vol. 2, no. 4, 2017.

[45] P. Seiwald, F. Sygulla, N.-S. Staufenberg, *et al.*, "Quintic Spline Collocation for Real-Time Biped Walking-Pattern Generation with variable Torso Height," in *IEEE-RAS Humanoids*, 2019.

[46] F. Sygulla and D. Rixen, "A force-control scheme for biped robots to walk over uneven terrain including partial footholds," *International Journal of Advanced Robotic Systems*, vol. 17, no. 1, 2020.

[47] M. Schwienbacher, "Efficient Algorithms for Biped Robots – Simulation, Collision Avoidance and Angular Momentum Tracking," Dissertation, TUM, 2013. [Online]. Available: https://mediatum.ub.tum.de/1175522

[48] R. Wittmann, A.-C. Hildebrandt, D. Wahrmann, *et al.*, "Model-Based Predictive Bipedal Walking Stabilization," in *IEEE-RAS Humanoids*, 2016.

[49] T. Buschmann, "Simulation and Control of Biped Walking Robots," Dissertation, TUM, 2010. [Online]. Available: https://mediatum.ub.tum.de/997204

[50] W. E. Lorensen and H. E. Cline, "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," *Computer Graphics*, vol. 21, no. 4, 1987.

[51] T. Yoshikawa, "Manipulability of Robotic Mechanisms," *The International Journal of Robotics Research*, vol. 4, no. 2, 1985.

[52] N. Neuburger, "Kinematic Structure Optimization for Humanoid Robots," Master's thesis, TUM, 2019. [Online]. Available: https://mediatum.ub.tum.de/1523789

[53] P. Seiwald and F. Sygulla, "broccoli: Beautiful Robot C++ Code Library," 2020. [Online]. Available: https://gitlab.lrz.de/AM/broccoli

[54] United States of America - Department of Defense, *Design Criteria Standard Human Engineering (MIL-STD-1472G)*. CreateSpace, 2012.

[55] National Aeronautics and Space Administration, *Man-Systems Integration Standards (NASA-STD-3000 Volume I)*. NASA, 1995. [Online]. Available: https://msis.jsc.nasa.gov/Volume1.htm

[56] S.-C. Wu, K. Tateno, N. Navab, *et al.*, "SCFusion: Real-time Incremental Scene Reconstruction with Semantic Completion," in *IEEE 3DV*, 2020, accepted - preprint: arXiv:2010.13662.

[57] T. F. C. Berninger, P. Seiwald, F. Sygulla, *et al.*, "Evaluating the Mechanical Redesign of a Biped Walking Robot Using Experimental Modal Analysis," in *IMAC*, 2021, (accepted).