A Bimanual Manipulation Taxonomy

Franziska Krebs and Tamim Asfour

***Abstract*—The ability of humans to bimanually manipulate objects is unprecedented and has not been matched by robots yet. The goal-oriented coordination requires the consideration of both temporal and spatial constraints between the hands. Within this work, we propose a taxonomy which differentiates between the different coordination patterns observed in human bimanual manipulation. The taxonomy rests on the key aspects of coordination and interaction of the hands, hand roles and symmetry in the execution of bimanual tasks. To validate the taxonomy, we propose a contact- and graph-based representation of the task, which combined with a rule-based classification provides the category of bimanual actions.**

***Index Terms*—Bimanual Manipulation, Learning from Demonstration, Human and Humanoid Motion Analysis and Synthesis**

I. Introduction

BEHAVIORAL studies provide evidence that bimanual tasks are more than the simple sum of unimanual tasks [1] as they have to consider spatial and temporal coordination as well as the interactions between both hands. When performing skillful manipulation tasks, bimanual coordination is essential for achieving the task goal and poses a challenge for the sensorimotor control system of a human or a humanoid robot. Understanding the underlying concepts of bimanual manipulation is not only essential for the task execution on bimanual robots such as humanoids but also for rehabilitation e. g. after unilateral stroke [2]. In robotic grasping, taxonomies are a common technique to address the complexity of hand design and grasp synthesis [3], [4], [5], [6]. While some previous works also employ taxonomies for more general manipulation scenarios [7], [8], within this work we will explicitly consider bimanuality. Such a categorization of bimanual patterns can be leveraged e. g. for learning task models from human demonstration, effective human-robot collaboration, action recognition as well as to derive constraints for the coordination and execution of robot bimanual manipulation tasks.

In this work, we aim at creating a comprehensive taxonomy for bimanual manipulations (see Figure 2) that incorporates previous knowledge about bimanuality in neuroscience and robotics but with emphasis on the usability for analysis and synthesis of bimanual robotic tasks. In addition to the conceptual definition of the taxonomy, we provide an analysis of the taxonomy using the KIT Bimanual Manipulation Dataset [9] demonstrating that different categories of bimanual actions can be distinguished using a rule-based classification system. Finally, we discuss and show how the taxonomy can be used for the segmentation of bimanual tasks.

Manuscript received: February, 23, 2022; Revised May, 16, 2022; Accepted June, 15, 2022. This paper was recommended for publication by Editor Dana Kulić upon evaluation of the Associate Editor and Reviewers' comments. This work has been supported by the German Federal Ministry of Education and Research (BMBF) under the project CATCH-HEMI (01KU2012) and by the Carl Zeiss Foundation under the JuBot project.

The authors are with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Germany {franziska.krebs,asfour}@kit.edu

Digital Object Identifier (DOI): see top of this page.

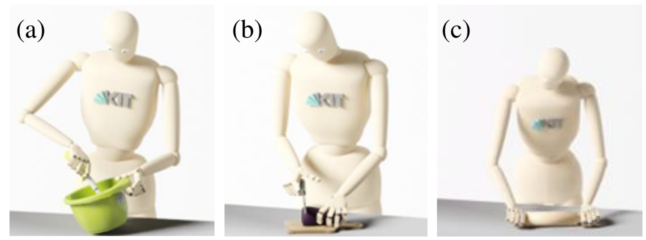

Fig. 1. Examples of bimanual actions: Asymmetric such as stir (a) and cut (b), and symmetric such as rolling (c).

II. Related Work

We review and discuss previous studies on bimanual coordination in neuroscience and rehabilitation as well as on bimanual manipulation in robotics.

A. Bimanual Coordination in Humans

Understanding motor control and motor learning in bimanual manipulation and coordination have a long history in the research areas of neuroscience, neuro-rehabilitation and clinical assessment of dysfunctional execution of everyday activities. Bimanual manipulation and coordination in humans is a complex process that is learned during childhood [10] and that can be disrupted by neurodegenerative diseases and brain pathologies ([11], [12]). As discussed in [13], two main theoretical frameworks for studying bimanual motor control exist: i) the information-processing and ii) the dynamic pattern perspective. The information-processing perspective considers bimanual movements as a task that faces structural interference due to limited neural resources resulting in neural leakage during bimanual movements, which can be overcome by training. The dynamic pattern perspective describes biological systems that are composed of different subsystems in terms of time-dependent changes, in which behaviors emerge in a self-organized manner.

In bimanual manipulation research, there is an imbalance regarding the type of tasks studied as most studies are concerned with cyclic bimanual coordination while object-oriented and goal-directed bimanual tasks have been often neglected [14]. In [15] it is investigated which factors effect the choice between unimanual, self-handover and symmetric bimanual actions in transport tasks. Hereby, self-handover is mainly used to transport an object between the right and left hemisphere while bimanual transport, compared to unimanual, provides increased stability but at the cost of a higher effort.

Several classifications of bimanual manipulation have been presented in the past. Guiard analyzes in [16] bimanual manipulation by considering the roles of the two hands. For symmetrical movements, both hands take the same role. An example of this is the transport of a large box. In contrast, the hands have different roles for asymmetric movements. For example, one hand stabilizes an object while the other acts on it. In literature such movements are also called *role-differentiated bimanual manipulations (RDBMs)* [17]. Within such manipulations, the dominant and non-dominant hands often assume specific roles. According to [16], the non-dominant hand provides a spatial frame of reference within which the dominant hand moves. Furthermore, there is a contrast in the spatio-temporal scale of the motions of the hands. Sainburg [18] introduced the dynamic dominance hypothesis, stating that the dominant hand is more effective in adapting to novel task dynamics. The work in [19] compares different definitions, assessment methods, and robotic devices for therapy. The goal is to standardize the terms and methods in robotic and sensor-based assessment and to establish a common language for communication and collaboration between clinicians, neuroscientists, and engineers conducting research on interlimb coordination. The authors further present a taxonomy of interlimb activities. This classification has no hierarchical structure but defines terms for the description of one limb independent of the other (e. g. periodic) and relative to each other (e. g. in-phase). In [2], Kantak et al. present a classification of bimanual tasks to study bimanual coordination in rehabilitation. According to the proposed classification, bimanual tasks can be characterized by the symmetry of arm movements, the task goal, and the necessity of cooperative interaction. These insights from bimanual coordination in humans can play an important role in the execution of goal-directed bimanual tasks on a robot.

B. Bimanual Manipulation in Robotics

Bimanual manipulation in robotics is still an underdeveloped research area with large potential [20]. In [21], the authors provide a survey on dual-arm manipulation in robotics addressing scientific problems related to control, planning and execution. The work in [22] addresses the problem of extracting constraints in asymmetric bimanual tasks from human demonstrations where the tasks are either autonomously executed by a robot or in interaction with a human. A master-slave relationship between both end-effectors is assumed and the respective interaction forces are explicitly considered. In [23], the authors extract the hand dominance based on the force-motion relation between the hands as a part of the intent to extract relevant task constraints. [24] presents a framework for intuitive bimanual telemanipulation. Next to a dual-arm teleoperation mode, for symmetric motions, a control strategy is implemented where both robot arms are controlled by a single human arm. Experiments for box transport tasks show that in two out of three tasks an improved performance and smaller variance is achieved for the symmetric control strategy.

A manipulation motion taxonomy is presented in [25] that prioritizes motion and contact features. The authors consider the following aspects in their taxonomy: *contact type*, *engagement type* (rigid/soft), *trajectory type*, *contact duration* (discontinuous/continuous) and *manual operation* (unimanual/bimanual). In [26], a hand pose taxonomy for high-precision, bimanual fine-manipulation tasks common in watchmaking is presented. The taxonomy is based on the analysis of virtual fingers in relation to force/torque demands. While the taxonomy is applied to each hand individually, the authors show how the hand pose matrix can be used to describe and visualize functional distributions across both hands. A sparse matrix indicates a low variance in hand pose combinations, while a concentration of entries in the upper or lower diagonals indicates the handedness. This representation is less suited to classify individual motion segments but rather to analyze hand pose selection over a full task.

There are already some approaches where a categorization of bimanual manipulations is formulated and later used for control purposes. A classification of dual-arm manipulation is presented in [27] and used for task segmentation in the context of robot programming by demonstration. Here, bimanual tasks are divided into two categories: *uncoordinated* and *coordinated*. Coordinated tasks are further subdivided into *symmetric* and *asymmetric* tasks. In symmetric coordinated tasks, both hands grasp the same object while in asymmetric tasks they manipulate different objects. Further, the authors present a control architecture that allows switching between different bimanual modes. In [28], bimanual manipulation tasks are divided into *non-coordinated* and *coordinated* tasks. The authors distinguish between *goal-coordinated* and *bimanual* operations that can be *symmetric/asymmetric* and *congruent/non-congruent*. The goal was to design an impedance controller for contact-based bimanual operations. Similar to [27], the work considers a decomposition of a task into single arm and bimanual actions. Another application is the intuitive control of a prosthetic [29]. Movements are classified depending on movement onset and movement direction into the categories of *unimanual*, *bimanual synchronous* and *bimanual asynchronous*. Upon the recognition of a bimanual category, the wrist rotation of the prosthesis is automatically controlled accordingly. In [30] a bimanual action vocabulary is proposed to improve the performance of a dual-arm teleoperation system. The vocabulary includes: *fixed offset*, *one hand fixed*, *self-handover* and *one hand seeking*. The authors in [31] leverage a taxonomy for subsymbolic motion representation and introduce the Extended-Cooperative-Task Space (ECTS) for coordinated motions of two end-effectors. The two ECTS coefficients are used to split the end-effector motions into absolute and relative parts. According to the values of the coefficients, the motions are categorized as uncoordinated or coordinated with the subcategories of parallel, blended and serial. The work in [32] focuses on the online recognition of bimanual coordination modes for teleoperated robots. This includes differentiating between different symmetry types and relative movement directions.

Our goal is to propose a taxonomy with clearly differentiable categories and special focus on applications in bimanual robotic manipulation. As discussed in Section II-A, clear categorizations of bimanual manipulation exist in neurorehabilitation. However, within this field, less attention is paid to deriving methods and tools for automatic recognition of different bimanual actions but to the assessment of therapy progress in the context of bimanual motor coordination. In particular, Kantak's taxonomy [2] contains important criteria. While symmetry is of great importance in neurology due to the characteristic muscle activation or the way of interhemispheric communication, the focus in robotics is rather on the fulfillment of or dependencies on task goals to be achieved. In contrast, previous approaches to robotics focus less on the precise definition and classification of different categories but on providing a schema that supports the development of control and planning strategies for dual-arm manipulation tasks. Accordingly, some consider a limited set of bimanual categories [22], [29]. In other cases the presented taxonomies are very promising but their precise definition is lacking since the focus of the respective publication was purely on the technical implementation [28]. Taking into account previous work in neuroscience (in particular [2]) and robotics, we propose a bimanual manipulation taxonomy that is not primarily tailored to the evaluation of therapy progress in rehabilitation, but rather dedicated to the representation of bimanual robotic manipulation tasks. This includes learning such representations from human motion data and making use of this knowledge to improve the execution of bimanual manipulation tasks in humanoid robotics.

III. Key Aspects of Bimanual Manipulations

In the following, we discuss and elaborate on key aspects observed in bimanual manipulation and which must be considered in a new taxonomy for bimanual manipulation tasks in robotics. These key aspects are concerned with (i) coordination between both hands, (ii) physical interaction between both hands, (iii) the role of each hand, and (iv) symmetry in the task.

A. Coordination

This criterion is concerned with the question of whether there is any kind of spatial or temporal coordination between the hands defined by spatial or temporal task constraints. Uncoordinated movements are in principle simultaneously executed unimanual actions. This means an action is uncoordinated if the same result can be achieved by executing the action consecutively in arbitrary order on a single arm. Both arms are neither spatially nor temporally coordinated and do not serve directly connected goals. For example, one hand holds a coffee cup while the other hand takes notes. Both hands must fulfill the task-specific constraints, but spatial coordination is limited to avoiding collisions and a temporal coupling does not exist in such a situation. In humans, actions performed by the individual hands in an uncoordinated manner cannot be arbitrarily complex. In the example above, the one hand can only fulfill the demanding action of writing because the holding of the cup is so highly automated that it hardly takes up any cognitive resources. In practice, it is often difficult to distinguish whether a movement is coordinated or uncoordinated since the relationship often becomes apparent not at a single point in time but only when viewed over a period of time. If one hand closes the lid of a chest while the other hand holds an object, the connection is unclear. But if it can be seen how one hand opened the chest so that the other could take something out, they are probably coordinated. Another challenging category is formed by actions including support poses. In [8], the authors present an approach of differentiating between support poses and manipulation by analyzing the transitions in a whole-body support pose taxonomy. In the case of leaning on a table with one hand in order to grasp a distant object with the other hand, there is a dependency despite the seemingly different activities of the hands. Without the support, the person would not be able to grasp the distant object. In contrast, in the context of support poses, there are also situations in which both hands are independent. For example, when one hand is holding a cup of coffee, while the other hand is holding a handrail while climbing a staircase. It is obvious that some cases are, even for humans, difficult to recognize from pure observation and are manifested only if one repeats the action execution in another way.

B. Interaction

Independent of potentially high-level coordination imposed by spatial or temporal constraints, physical interaction between the hands might or might not occur when executing bimanual tasks. Here, we refer to interaction in situations, in which forces are transmitted between both hands, either directly, via a common object or several objects. This includes holding a large object with both hands, but also when one hand holds the object and the other hand applies force to it using a tool, as when holding a bottle with one hand and unscrewing the lid with the other hand. Such interactions are crucial from a robotics control point of view, especially for bimanual dexterous manipulation tasks, in which the physical interactions with objects and potentially between the hands are essential for successful task completion.

C. Hand Roles

In bimanual manipulation tasks, the hands might have distinct roles. In [16], Guiard presents an abstract view on symmetry in bimanual manipulation as it does not refer to the movement but to the abstract roles of the hands. In symmetric manipulation tasks, both hands take the same role e. g. when both hands hold and transport a large box. In such a case, a fixed transformation between the hands exists and both arms, together with the box, form a closed kinematic chain. In contrast, the hands take different roles in asymmetric movements, e. g. when one hand stabilizes an object while the other performs an action on it, such as stirring in a cup while holding it. In humans, the hand roles in asymmetric tasks are highly correlated with the hand dominance. The non-dominant hand is usually the stabilizing hand providing the reference frame for the dominant hand. Furthermore, there are differences in the temporal-spatial frame of the movements with the dominant hand usually having a higher frequency [16]. However, these roles are not statically assigned to the right or left hand but might be changed to optimally incorporate an action into a complete task. For example, we might perform the finer manipulation of closing the lid of a bottle with the left hand since the right hand already holds the bottle from drinking.

D. Symmetry

One characteristic of bimanual tasks is the symmetry of both arm movements observed in tasks such as bimanual transporting of a tray, reaching for a large object, or conducting an orchestra in which the motion pattern of both arms share similarities. The importance of symmetry is stressed in several human behavior studies in neurology and neuroscience (see [2]) as well as in the assessment of bimanual coordination in rehabilitation. Information about symmetry can play an important role in programming and executing bimanual tasks on a dual-arm system as reference motion trajectories of one arm can be generated by "mirroring" movements of the other arm.

IV. Taxonomy for Bimanual Manipulation

Based on the discussion of the key aspects in bimanual manipulation in Section III, we propose a taxonomy that classifies bimanual manipulation tasks from a robotic-centered perspective (see Figure 2). In particular, our goal is to propose a taxonomy that supports learning tasks models for bimanual manipulations from human observation and the execution of such task models on bimanual robot systems such as humanoid robots.

On the first level, we differentiate between coordinated and uncoordinated bimanual actions based on whether there are any spatial and/or temporal constraints that are important for the execution of the task. Uncoordinated actions can be seen as simultaneously executed unimanual actions. Thus, unimanual actions are considered to be a subcategory of uncoordinated bimanual motions, where one hand just fulfills no explicit task. Within the category of coordinated actions, the degree of interdependence between hands varies. We define *loose* coupling as dependencies that impose constraints on the actions of the hands but only in the sense of common via points (spatial) or synchronization points (temporal) where certain relations between the hands have to hold but a permanent dependency of the trajectories does not exist. Self object hand-overs are a common example of a combination of temporal and spatial coupling, which is essential particularly in the pre-handover phase. During the handover, additional constraints resulting from physical interaction of the hands through the object must be fulfilled.

Therefore, we define an additional category of *tightly coupled* coordinated actions, which are characterized not only by spatial and temporal constraints but also by force-based constraints resulting from contact-rich interaction of the hands and a dependency of the hands on trajectory-level. Within this category obviously, the question arises how exactly the trajectories depend on one another. In this context, we refer to Guiard's work [16] regarding the roles of the hands where the hands have distinct roles in the so-called *asymmetrical* activities. According to this, the non-dominant hand provides a reference frame for the dominant hand so that the trajectory of the dominant hand can be formulated dependent on the non-dominant hand. In general, we need to determine the hand dominance not only for different persons but also within different tasks. We denote bimanual activities, in which both hands have the same role as *symmetric*. Even though this primarily refers to the roles of the hands and not to their motion trajectories, this commonly coincides. We define symmetric actions as the ones where both hands are grasping and manipulating the same object and thus also move in a symmetrical way. In such a case, the dependency is even stronger since there is a fixed transformation between the hands.

TABLE I
Bimanual action categories

Bimanual Category	Abbreviation
No action	*no_action*
Unimanual left	*uni_left*
Unimanual right	*uni_right*
Loosely coupled & uncoordinated bimanual	*loosely*
Tightly coupled asymmetrical left dominant	*tightly_asym_left*
Tightly coupled asymmetrical right dominant	*tightly_asym_right*
Tightly coupled symmetrical	*tightly_sym*

Despite its enormous importance in neuroscience, we consider geometric motion symmetry to be less important for robotics. In *loosely coupled* actions, motion symmetry takes on a functional character rather in exceptional cases. While it is not necessary for bimanual reaching it might be relevant in gestures such as conducting an orchestra. Therefore, in our taxonomy, we consider symmetry only for tightly coupled actions, focusing on the symmetry of the distribution of roles of the hands rather than on strict geometric symmetry.

V. Analysis of the Taxonomy

We provide an analysis of the proposed taxonomy using recordings of human bimanual activities provided in the KIT Bimanual Manipulation Dataset [9] and show which categories of bimanual actions in the dataset can be identified. We consider *loosely coupled* and *uncoordinated bimanual* actions in the same category since they cannot be distinguished based on pure motion features and without a semantic understanding of the entire task. However, these two categories are clearly distinct, and the fact that they cannot be distinguished is a limitation of the current classification approach. In our analysis, we also add the category *no_action* to indicate an idle state and support a correct segmentation of the entire task. This results in the categories of actions listed in Table I.

Our goal is to determine the characteristics of bimanual manipulation actions during the execution of daily household tasks. This would provide a data-driven validation of the proposed taxonomy. To achieve this goal, we introduce a pipeline for data processing, feature extraction and classification of bimanual actions into different bimanual manipulation categories (see Figure 3). The 6D pose and configuration of the hand, 6D poses of all objects in the scene as well as the 3D models of these objects and the human hand are used as input. In the first step, we construct a contact graph, in which objects and hands are the nodes and contact relations between objects and hands are edges (Section V-B). In the second step, we extract for each hand and each object, i. e. for each node of the graph, motion features consisting of poses and velocities that are used in the third step by a rule-based classification

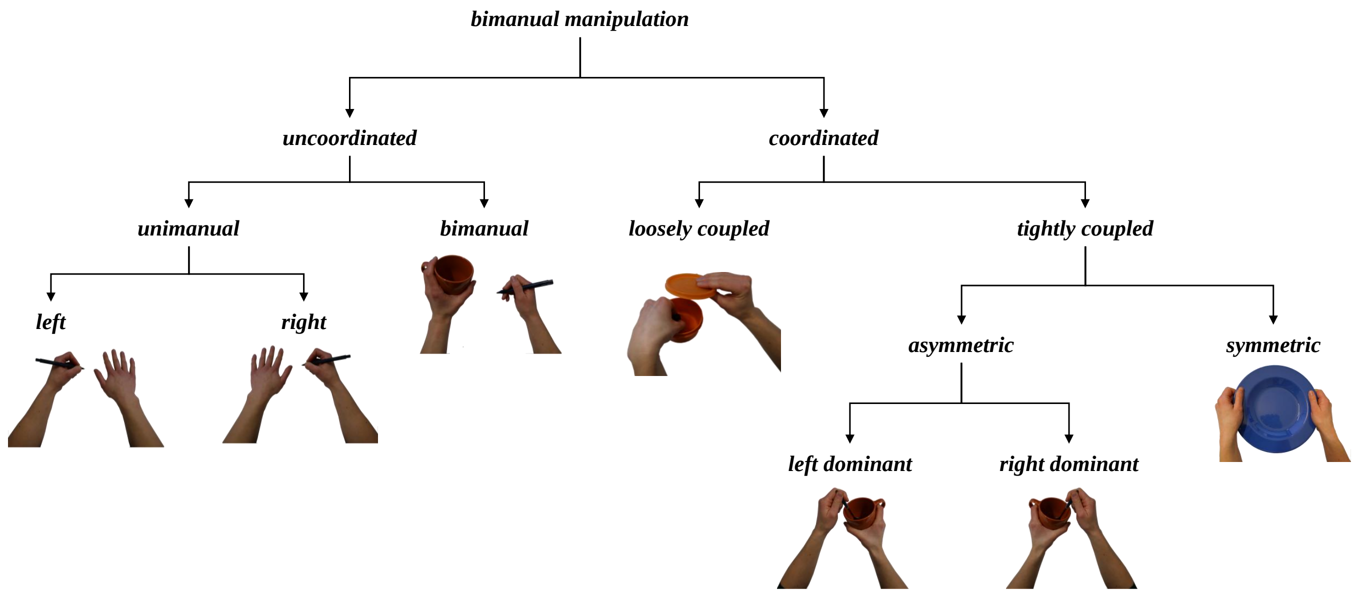

Fig. 2. Bimanual manipulation taxonomy. Tasks are classified based on the aspects *coordination*, *interaction*, *hand role* and *symmetry*.

system that is inspired by the proposed taxonomy to determine the characteristics of bimanuality.

A. Dataset

For our analysis, we use 120 recordings from the KIT Bimanual Manipulation Dataset [9]. This dataset includes multimodal recordings of several bimanual household activities. Each recorded motion starts and ends with both hands on the table and contains one bimanual manipulation task. The dataset is publicly available on our KIT Whole-Body Human Motion Database[1]. For this analysis, we use the whole-body human and object motion data, recorded with a marker-based Vicon motion capture system at 100 Hz and the hand motion that is recorded by 18 DoF data gloves at 90 Hz. Therefore, precise whole-body motions, including the hand's configurations as well as the 6D pose of all involved objects, are available. We use the open-source Master Motor Map (MMM) framework[2] that provides a reference model of the human body, unifying data formats and tools for capturing, representing as well as analyzing human motions. The details of the MMM with the marker set on the human body and its procedures for the reconstruction of human motion from different input data are described in [33]. The result of the motion reconstruction are the joint angles and 6D pose of the human body including hand pose and configuration for every frame of the captured motion. With the MMM model, we can also derive the 6D pose of the end-effectors based on a virtual tool center point located in the palm.

B. Contact Graph Construction and Feature Extraction

Our goal is to investigate whether and which of the bimanual categories proposed by our taxonomy appear in

[1]https://motion-database.humanoids.kit.edu/
[2]https://mmm.humanoids.kit.edu

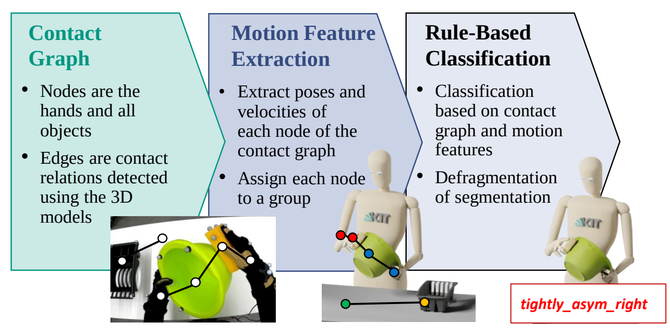

Fig. 3. Pipeline for extracting bimanual categories. The figure shows also an example of the contact graph with nodes associated to the rightGroup (red), leftGroup (blue), background (green) and scene (yellow) as well as the detected bimanual category as tightly coupled, asymmetric with a dominant right hand.

human demonstrations of daily tasks. Thus, it is important to determine the semantic spatial and temporal relations between hands and objects as well as motion characteristics of both hands during the task execution. Given a scene with 1) a set of objects represented by their geometric 3D models and 6D poses and 2) the human represented by the 3D geometric model of the MMM with its root pose in the scene, body, and hand configuration as well as 6D end-effectors poses, we construct a contact graph of the scene whose nodes represent the objects and the hands and whose edges indicate contact between objects and hands. To detect contact relations based on 3D object models, we rely on our previous work on extracting semantic relations [34] and segmentation of human demonstrations [35]. The contact graph is constructed frame-wise and represents contact relations and their changes during the execution of a task. The bimanual category is determined based on the topological structure of the graph. In addition, motion features of the hands and objects are important. Thus, each node of

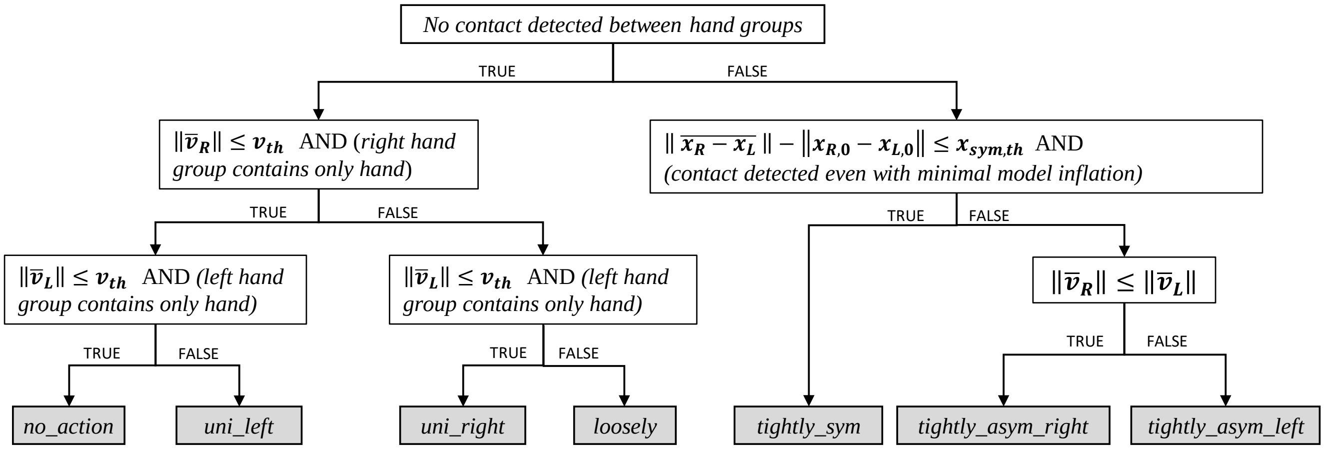

Fig. 4. Decision tree for the rule-based classification

the graph is enriched by information about 1) the global pose and velocity of the node 2) an object ID as unique identifier of the node, e. g. the object names *rolling_pin*, *sponge_small*, etc. and 3) a group ID for *rightGroup*, *leftGoup*, *background* and *scene*.

The contact graph is constructed by first evaluating the contact relations in the scene and then assigning objects to different groups. We manually assign objects in the scene that are fixed and will not be manipulated such as a table or a wall to the *background* group. This is needed in order to eliminate contacts detected between the hands and these objects. The *rightGroup* and *leftGroup* contain the hands themselves and all objects which are indirectly or directly in contact with the respective hand. The right hand and left hand are assigned to the corresponding groups *rightGroup* and *leftGroup*, which are continuously updated based on changes of the topological structure of the contact graph. Objects in contact with members of the *rightGroup* or *leftGroup*) are added to this group unless they are already part of another group. All remaining nodes are assigned to the *scene* group. The pose of each node (hands and objects) is provided by the marker-based VICON system used in the recorded task demonstrations and the velocity of each node is approximated using numerical differentiation.

While we rely on object information for the contact graph construction, the approach is object agnostic since the bimanual category is only determined based on the topological structure of the graph. Since force information is not available in the dataset, in our analysis we only consider contact information between the hands and objects in the scene.

C. Classification and Segmentation

To determine the category of bimanual actions, we apply a rule-based classification by which the decision is made based on predefined and interpretable *if-else* rules. The decision tree and rules are given in Figure 4. The starting point is the contact graph described above, which is constructed for every frame of the motion to be analyzed. For motion classification, we use a sliding window approach with a window size of 10 frames. We provide a publicly available implementation of our method[3].

We start at the root node of the decision tree and check whether there is contact between elements the *rightGroup* and *leftGroup*, i. e. whether there is an edge in the contact graph between the two groups of graph nodes. If there is no contact and thus no physical interaction between the hands, the average vector norms of the hand velocities $\|\bar{v}_L\|$ and $\|\bar{v}_R\|$ are used for decisions on the next level of the tree. If the velocity for one of the hands falls below a certain threshold v_{TR} (here 20 $\frac{\text{mm}}{\text{s}}$), the respective hand is considered to be not moving. In case the hand is also not in contact with any object, apart from background objects, the hand is considered not active. If only one hand is active, the leaf node of the tree will have the label *uni_right* or *uni_left* depending on the moving hand. If no hand is active, the motion segment is classified as *no_action*, and *loosely* if both hands are active.

In case of contact between the hand groups, the average distance $\|\overline{x_R - x_L}\|$ between the hands during the motion execution is compared against the distance at the beginning of the respective window $\|x_{R,0} - x_{L,0}\|$ (right sub-tree in Figure 4). If the difference is below a certain threshold $x_{sym,th}$ (here $1\ mm$), the motion segment is classified as *tightly_sym*. In the initial contact evaluation, 3D models are inflated (by $15\ mm$) and contact is detected if the inflated models overlap. At this point we test whether the contact persist even with minimal model inflation ($3\ mm$). This is an additional necessary requirement for motions of the class *tightly_sym*. If the motion is not classified as *tightly_sym*, the velocities of the hands are compared. According to Guiard [16], the hand with the higher mean velocity is considered the dominant hand. Therefore, the classification results in either *tightly_asym_right* or *tightly_asym_left* as label of the leaf node of the tree.

To apply the procedures described above to longer actions or sequences, a sliding window with a window size of 10 frames is used, where each window provides a category as result

[3]https://gitlab.com/h2t/software/bimanualcategoryclassification

of the classification. Adjacent windows of the same category form a common segment. An additional defragmentation step is applied to eliminate resulting small segments. In summary, with this approach, we obtain not only a classification of bimanual actions into the categories defined by the taxonomy but also a segmentation of the demonstration.

D. Results and Discussion

In total, 120 recordings of two subjects (60 for each) performing bimanual daily activities were used for the analysis. Because the recordings are limited to isolated, specifically bimanual manipulations performed by right-handed individuals, the reference data do not include all of the bimanual categories. Therefore, *uncoordinated* bimanual tasks are not contained in the dataset. In case of occurrence, they would be classified as *loosely* coupled. A distinction between *uncoordinated* and *loosely* coupled actions could take place in an additional step, in which a larger temporal section of the motion is considered. Considering the number of frames in the manually labeled reference data, *loosely* occurs most frequently (46.02 %), followed by *tightly_asym_right* (28.19 %), *tightly_sym* (15.81 %), *no_action* (7.14 %), and *uni_left* (2.84 %). The two categories *unimanual_right* and *tightly_asym_left* do not occur at all.

The manually segmented reference data and the automatically generated labels are compared frame-wise. The micro, macro and weighted F_1 score are computed using the tools provided by the Python library scikit-learn. Figure II shows the results for each action, each subject and all motions combined. The best results are obtained independent of the subject for the actions *Stir* and *Wipe*. The lowest scores were obtained for the *Roll* and *Cut* actions. For *Cutting*, this is due to inaccurate modeling of vegetables in the dataset and the lack of relative motion at the turning point during the cutting motion. For *Roll*, the symmetry criterion is not always met exactly. One option to tackle such issues in the future would be to not assign a bimanual category to each frame but instead to each action segment. Since the majority of each segment would be classified correctly, the overall label for the action segment would probably be correct for most cases.

TABLE II
CLASSIFICATION RESULTS FOR INDIVIDUAL ACTIONS AND SUBJECTS. THE MICRO F_1-SCORE CORRESPONDS TO THE ACCURACY.

	F_1-score [%]		
	Micro	Macro	Weighted
Total	84.66	57.92	82.72
Total subject 1	83.60	56.44	81.34
Total subject 2	85.65	59.10	84.11
Wipe	90.08	48.15	88.92
Stir	91.88	48.21	90.36
Roll	78.21	35.46	69.72
Peel	85.97	46.90	84.61
Cut	78.24	35.15	71.97

Figure 5 shows the confusion matrix for all motions. Even though we used manually labeled data as ground truth data in this work, it is important to keep in mind that this labeling is far from perfect as the segmentation points can be imprecise. For example, the motion onsets of the hands might be detected too early or too late. This explains why quite a few motions are automatically detected as *loosely-coupled* but manually labeled as *unimanual* motions or vice versa. As can be seen in Figure 5, confusion also frequently occurs within *tightly-coupled* actions. On the one hand, this is due to the fact that the conditions of the rule-based classification are not always met exactly. For example, the condition of constant offset between the hands is not always met during rolling due to small changes of the hand pose during the execution of this action. Further, the automatic segmentation generates different key points compared to the manual segmentation, that might also be more correct than the manually labeled key points. For example, at the end of the *wiping* action there is often a segment that is labeled as *tightly_asym_left*. This is due to the fact that no more wiping is performed on the plate, but the plate is pulled away from under the sponge in a transitional movement to put it down.

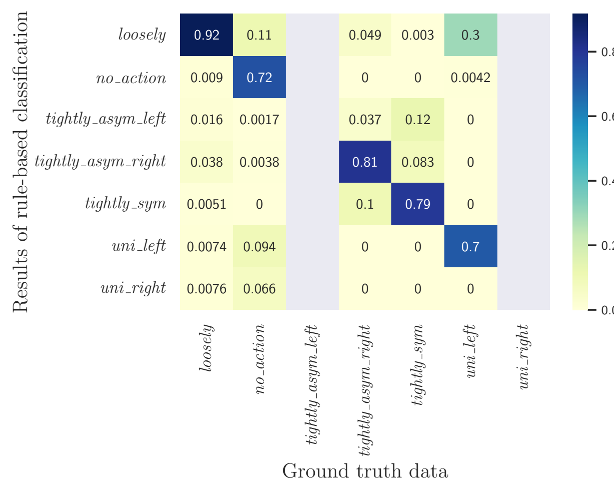

Fig. 5. Accumulative classification correctness over all actions and subjects depicted as normalized confusion matrix. Grey columns correspond to labels which were not contained in the ground truth data.

VI. CONCLUSION AND FUTURE WORK

This work introduces a taxonomy for bimanual manipulation inspired by prior work in robotics, neuroscience, and rehabilitation science. The taxonomy differentiates different coordination patterns in bimanual manipulation tasks based on the key aspects of coordination and interaction between the hands, role of the hands in the task as well as the symmetry in the task execution. The conceptually proposed taxonomy was analyzed using a motion capture dataset of human bimanual manipulation in daily tasks to investigate whether and which of the bimanual categories proposed in the taxonomy appear in human demonstrations. We represent bimanual tasks as contact graphs that describe contact relations between hands and objects in the scene and define motion features needed for the classification of bimanual tasks. Based on this, a rule-based classification is used to determine bimanual categories in human demonstrations of daily cooking activities. Although our method is limited by the accuracy of the manually labeled ground truth data, we showed that we can detect categories with an accuracy of $84.66\,\%$. The taxonomy can contribute to a better understanding of bimanual coordination in humans and the transfer of human strategies in executing bimanual tasks to robots. For example, the taxonomy can be used to describe bimanual manipulation tasks as a sequence of bimanual categories which implicitly encode relevant spatial and temporal constraints that are needed for the execution. This representation would allow the selection of appropriate controllers for different phases of a bimanual task as well as the switching between different control strategies, e. g. switching between a leader-follower control scheme for tightly coupled actions and independent controllers for uncoordinated actions. Furthermore, the taxonomy can be used to improve action and intention recognition in human-robot interaction tasks as well as to improve bimanual interactions with prostheses as shown in [29].

References

[1] S. Swinnen and J. Gooijers, "Bimanual Coordination," in *Brain Mapping*, 2015, pp. 475–482.

[2] S. Kantak, S. Jax, and G. Wittenberg, "Bimanual coordination: A missing piece of arm rehabilitation after stroke," *Restorative neurology and neuroscience*, vol. 35, no. 4, pp. 347–364, 2017.

[3] N. Kamakura, M. Matsuo, H. Ishii, F. Mitsuboshi, and Y. Miura, "Patterns of static prehension in normal hands," *American Journal of Occupational Therapy*, vol. 34, no. 7, pp. 437–445, 1980.

[4] M. R. Cutkosky *et al.*, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 269–279, 1989.

[5] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragić, "The grasp taxonomy of human grasp types," *IEEE Transactions on human-machine systems*, vol. 46, no. 1, pp. 66–77, 2015.

[6] V. Arapi, C. Della Santina, G. Averta, A. Bicchi, and M. Bianchi, "Understanding human manipulation with the environment: A novel taxonomy for video labelling," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6537–6544, 2021.

[7] I. M. Bullock, R. R. Ma, and A. M. Dollar, "A hand-centric classification of human and robot dexterous manipulation," *IEEE transactions on Haptics*, vol. 6, no. 2, pp. 129–144, 2012.

[8] J. Borràs and T. Asfour, "A whole-body pose taxonomy for loco-manipulation tasks," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2015, pp. 1578–1585.

[9] F. Krebs, A. Meixner, I. Patzer, and T. Asfour, "The kit bimanual manipulation dataset," in *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, 2021, pp. 499–506.

[10] J. Barral, B. Debû, and C. Rival, "Developmental changes in unimanual and bimanual aiming movements," *Developmental Neuropsychology*, vol. 29, no. 3, pp. 415–429, 2006.

[11] Y.-C. Hung and W. Zeng, "Accuracy constraints improve symmetric bimanual coordination for children with and without unilateral cerebral palsy," *Developmental neurorehabilitation*, vol. 23, no. 3, pp. 176–184, 2020.

[12] T. M. Roebuck-Spencer, S. N. Mattson, S. D. Marion, W. S. Brown, and E. P. Riley, "Bimanual coordination in alcohol-exposed children: Role of the corpus callosum," *Journal of the International Neuropsychological Society*, vol. 10, no. 4, p. 536, 2004.

[13] S. P. Swinnen and N. Wenderoth, "Two hands, one brain: Cognitive neuroscience of bimanual skill," *Trends in cognitive sciences*, vol. 8, no. 1, pp. 18–25, 2004.

[14] S. S. Obhi, "Bimanual coordination: An unbalanced field of research," *Motor Control*, vol. 8, no. 2, pp. 111–120, 2004.

[15] Y. C. Nakamura, C. A. O'Sullivan, and N. S. Pollard, "Effect of object and task properties on bimanual transport," *Journal of motor behavior*, vol. 51, no. 3, pp. 245–258, 2019.

[16] Y. Guiard, "Asymmetric Division of Labor in Human Skilled Bimanual Action," *Journal of Motor Behavior*, vol. 19, no. 4, pp. 486–517, 1987.

[17] M. Kimmerle, C. L. Ferre, K. A. Kotwica, and G. F. Michel, "Development of role-differentiated bimanual manipulation during the infant's first year," *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, vol. 52, no. 2, pp. 168–180, 2010.

[18] R. L. Sainburg, "Evidence for a dynamic-dominance hypothesis of handedness," *Experimental brain research*, vol. 142, no. 2, pp. 241–258, 2002.

[19] C. Shirota, J. Jansa, J. Diaz, S. Balasubramanian, S. Mazzoleni, N. A. Borghese, and A. Melendez-Calderon, "On the assessment of coordination between upper extremities: Towards a common language between rehabilitation engineers, clinicians and neuroscientists," *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, pp. 1–14, 2016.

[20] A. Billard and D. Kragić, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, June 2019.

[21] C. Smith, Y. Karayiannidis, L. Nalpantidis, X. Gratal, P. Qi, D. V. Dimarogonas, and D. Kragić, "Dual arm manipulation—a survey," *Robotics and Autonomous Systems*, vol. 60, no. 10, pp. 1340–1353, 2012.

[22] L. Pais Ureche and A. Billard, "Constraints extraction from asymmetrical bimanual tasks and their use in coordinated behavior," *Robotics and Autonomous Systems*, vol. 103, pp. 222–235, 2018.

[23] A. L. Pais and A. Billard, "Encoding bi-manual coordination patterns from human demonstrations," in *ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 264–265.

[24] M. Laghi, M. Maimeri, M. Marchand, C. Leparoux, M. Catalano, A. Ajoudani, and A. Bicchi, "Shared-autonomy control for intuitive bimanual tele-manipulation," in *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, 2018, pp. 1–9.

[25] D. Paulius, Y. Huang, J. Meloncon, and Y. Sun, "Manipulation motion taxonomy and coding for robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2019, pp. 5596–5601.

[26] K. Yao, D. Sternad, and A. Billard, "Hand pose selection in a bimanual fine-manipulation task," *Journal of Neurophysiology*, vol. 126, no. 1, pp. 195–212, 2021.

[27] R. Zöllner, T. Asfour, and R. Dillmann, "Programming by demonstration: Dual-arm manipulation tasks for humanoid robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2004, pp. 479–484.

[28] D. Surdilović, Y. Yakut, T.-M. Nguyen, X. B. Pham, A. Vick, and R. Martin-Martin, "Compliance control with dual-arm humanoid robots: Design, planning and programming," in *IEEE-RAS Int. Conf. on Humanoid Robots (Humanoids)*, 2010.

[29] R. Volkmar, S. Dosen, J. Gonzalez-Vargas, M. Baum, and M. Markovic, "Improving bimanual interaction with a prosthesis using semi-autonomous control," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, p. 140, 2019.

[30] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, "Shared control–based bimanual robot manipulation," *Science Robotics*, vol. 4, 2019.

[31] H. A. Park and C. G. Lee, "Dual-arm coordinated-motion task specification and performance evaluation," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2016, pp. 929–936.

[32] J. R. Boehm, N. P. Fey, and A. M. Fey, "Online recognition of bimanual coordination provides important context for movement data in bimanual teleoperated robots," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2021, pp. 6248–6255.

[33] C. Mandery, Ö. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "Unifying Representations and Large-Scale Whole-Body Motion Databases for Studying Human Motion," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 796–809, 2016.

[34] R. Kartmann, F. Paus, M. Grotz, and T. Asfour, "Extraction of physically plausible support relations to predict and validate manipulation action effects," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 4, pp. 3991–3998, 2018.

[35] M. Wächter and T. Asfour, "Hierarchical segmentation of manipulation actions based on object relations and motion characteristics," in *International Conference on Advanced Robotics (ICAR)*, 2015, pp. 549–556.