A Formalization for Hand-Object Interaction

Mohamad Alameh, Alessandro Carfì, and Fulvio Mastrogiovanni

***Abstract*— Humans interact with objects using their hands, and to be able to operate in the same environment robots should be able to do the same. To achieve this goal, the study of human hand-object interaction is fundamental. In this abstract we will propose a formalization of the hand-object interaction to help autonomous systems to characterize this interaction.**

I. Introduction

Humans interact with objects using different parts of their body but of course the hand is one of the most used. Humans practice hand-object interaction starting early ages, and their skills are unmatched. The hand-object interaction has been subject of different studies whether in human's motor control, or in robotics applications [1], [2]. Hand-object interaction have been addressed in the literature from different point of views, whether in the definition of taxonomies and classification, or in the description of different interaction phases, e.g., Reach, Grasp and Manipulate [3], [4].

In this work we aim to summarize the state of the art to provide a comprehensible, and machine processable formalization of the hand-object interaction in order to provide a standard definition that can be used and extended in related research.

II. Hand-Object interaction

The hand-object interaction is often described as a three phases process "Reach-Grasp-Manipulate" [3], [4]. However, this description is too simple to fully describe the hand-object interaction and all its declinations. For this reason, we introduce in our formalization a state machine with three object states: Off-hand, Reached and In-Hand, see Figure 1.

In the off-hand state the object is apart from the hand and no physical interaction can occur. However, while in this state the hand can move, conveying to possible observer the intention to interact with a certain object. Once the hand arrives in the proximity of the object, the hand-object interaction is in the reached state. In this state, the hand is close enough to physically interact with the object, but the object weight is not supported by the hand. Finally, the in-hand state is reached once the weight of the object is completely supported by the human. In this state the human is free to alter the object status. Notice that the states are defined with respect to the object, but the transitions are caused by the active subject of the interaction i.e. the human hand. To carefully describe the hand-object interaction, it is necessary to describe the transition between each state. In particular, as we can see in Figure 1, four transitions have been introduced: reach, depart, load, and unload. These four transitions can be considered as couples of opposites: reach-depart and load-unload. In fact, during the reach transition

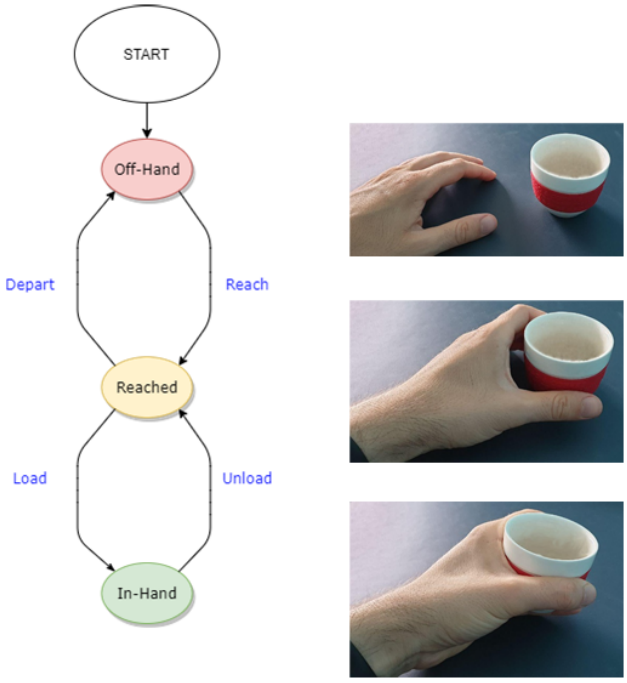

Fig. 1. Hand-object interaction state machine representation from object perspective

the hand approaches the object to reduce the distance with it, and to obtain a posture that allows the hand to grasp the object, while during the depart transition the hand leaves a posture close to the object, suggesting the end of the hand-object interaction. Both these two transitions generally involve the motion of other body parts, such as the arm or even a full body motion in case the object is far from the initial position of the hand. Instead, the load and unload transitions imply mainly a motion of the hand and the fingers. Once the object is in the reached state, the hand can grasp the object and later load its weight. On the contrary, the unload is performed when the object is in-hand and it is released, the unload can even be accidental. Notice that while defining the transitions between the three states, we introduced the "grasp" term. For hand-object interaction the grasp is a fundamental concept which will be described in the next subsection.

A. Grasp

"A grasp is every static hand posture with which an object can be held securely with one hand, irrespective of the hand orientation" [5]. Different efforts to classify all the possible grasping poses have been done in the literature. One of the

first grasping hierarchical classification in the literature, is the Cutkosky's grasp taxonomy [6] that counterposes the dexterity associated with a grasping pose to its power. The Cutkosky's grasp taxonomy highlights an intuitive concept, the grasp is highly influenced by the object physical characteristics, i.e. shape, weight, texture, and stiffness, and by the considered task, i.e. do I need more power or dexterity? Therefore, to handle an object, the hand should have a pose, relative to the object, that is coherent with the object's physical characteristics and the ongoing task. This pose, usually referred as grasping pose, is defined by the relative pose of the hand with respect to the object and by the fingers' configuration. Consequently, the action of grasping can be defined as applying a force to an object, while the hand is in a grasping pose. As we have seen, grasping an object is fundamental to load it in-hand. However, the grasping does not imply the loading. In fact, an object could be grasped in the reached state without implying an immediate transition to the in-hand state, e.g. grasping the knob of a stereo. Grasping an object not only can anticipate a loading but a manipulation as well.

B. Manipulate

According to the linguistic definition, to manipulate something means to control, use, or change something with skill [7]. Formally speaking, we define the manipulation as "*the action that changes the state of an object*." Where the state of an object comprehends its pose in the space with respect to a global reference frame and its internal degrees of freedom (DoF), if any. Although object manipulation can be performed with whichever body part, our concern is about hand manipulation. To manipulate an object, the hand should be able to exercise a force on the object. Considering the hand-object interaction formalization previously introduced, two object states allow the manipulation i.e., "reached" and "in-hand". In the **reached state** the hand can manipulate the object by mean of a simple contact, i.e. pushing an object, or can grasp the object for a more complex manipulation, i.e. rotating a glass while keeping it on the table. Instead, in the **in-hand state**, the human is in full control of the object and he can alter substantially its pose. A further distinction should be done on the body parts that characterize the motion. If it is true that we are considering manipulations in which the hand is in direct contact with the object, this does not imply that the manipulation is driven by the hand, e.g. pulling a lever is a manipulation driven by the arm. Therefore, it is necessary to introduce another keyword to describe hand driven manipulations, within-hand.

A **within-hand** manipulation is performed exclusively using the fingers and therefore it implies a change of the object pose with respect to the hand reference frame[8] or the change of an internal DoF of the object. In the literature, the within-hand manipulation concept sometimes is referred as in-hand manipulation[9]. However, even if the within-hand manipulation often occurs while the object is in-hand, this is not a prerequisite and we can have within-hand manipulation in the reached state. Therefore, we keep the in-hand and within-hand concepts disjoint. It is noteworthy that a manipulation can be a pure within-hand manipulation or, more often, the result of a within-hand manipulation coordinated with a motion of another body part. According to the literature related to human hand skills [9] [10], within-hand manipulation is the ability to hold and move an object within one hand; The person may be shifting objects on the surface of the fingers or among the fingers, rotating objects between the fingers, moving objects from palm to fingers (palm-to-finger translation), or moving objects from fingers to palm (finger-to-palm translation) [9]. A person may perform these movements while no other object is in the hand or while simultaneously stabilizing another object or objects in the hand [10].

According to this definition, within-hand manipulation skills are divided into three major categories: translation (palm-to-finger and finger-to-palm), shift, and rotation (simple and complex)[9].

- Translation: is the ability to move objects from the fingertips to the palm or from the palm to the fingertips, such as moving coins from the palm to the fingertips to place in a vending machine, or picking up pennies and moving them from the fingertips to the palm.
- Shift: is the ability to move an object in a linear manner with the fingertips, with respect to its original position, such as re-positioning the pencil in the fingers for writing.
- Rotation: is the ability to turn an object around in the pads of the fingers and thumb (simple rotation), or turning an object from end to end (complex rotation) such as flipping a pencil from writing end to eraser.

Within-hand manipulation is a sequence of these three manipulation categories, that can exist in any order. Notice that, these three manipulation categories have been introduced in the medical literature to assess the manipulation skills of patients. These categories give a high-level description of the manipulation and are not intended to be used as primitives.

Finally, the influence that the gravity plays on the manipulation should be included in our formalization. To the best of our knowledge, in the literature the contribution of the gravity has been mentioned [5] but usually ignored. To take into account the gravity and the possible effects that can have on the manipulation, our formalization includes a description of the palm's pose, which can be: i) horizontal with the palm facing down, ii) palm vertical with thumb pointing up, since the other vertical palm configuration with the thumb pointing down is not a common manipulation pose, and iii) horizontal with the palm facing up.

III. Conclusion

In this work, we have presented a formalization of the hand-object interaction describing some relevant components such as the grasping and the manipulation. This formalization is going to be useful to decompose the hand-object interaction problem in sub-problems easier to be solved by a robotic system. Furthermore, this formalization could be used to characterize human hand-object interaction demonstrations.

REFERENCES

[1] A. Shumway-Cook and M. H. Woollacott, Motor Control: Translating Research Into Clinical Practice. Lippincott Williams & Wilkins, 2007, google-Books-ID: BJcL3enz3xMC.

[2] F. J. Valero-Cuevas and M. Santello, "On neuromechanical approaches for the study of biological and robotic grasp and manipulation," Journal of NeuroEngineering and Rehabilitation, vol. 14, no. 1, p. 101, Oct. 2017. [Online]. Available: https://doi.org/10.1186/s12984-017-0305-3

[3] S. B. Kang and K. Ikeuchi, "Toward automatic robot instruction from perception-temporal segmentation of tasks from human hand motion," IEEE Transactions on Robotics and Automation, vol. 11, no. 5, pp. 670–681, Oct. 1995, conference Name: IEEE Transactions on Robotics and Automation.

[4] K. Stollenwerk, A. Vögele, B. Krüger, A. Hinkenjann, and R. Klein, "Automatic Temporal Segmentation of Articulated Hand Motion," 2016.

[5] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The GRASP Taxonomy of Human Grasp Types," IEEE Transactions on Human-Machine Systems, vol. 46, no. 1, pp. 66–77, Feb. 2016.

[6] M. Cutkosky, "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, Jun. 1989.

[7] "Oxford Learner's Dictionaries." [Online]. Available: https://www.oxfordlearnersdictionaries.com/

[8] 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, Apr. 2013, conference Name: IEEE Transactions on Haptics.

[9] C. E. Exner, "Development of hand functions," Occupational therapy for children, 1989.

[10] Exner, Charlotte E, "The zone of proximal development in in-hand manipulation skills of nondysfunctional 3-and 4-year-old children," American Journal of Occupational Therapy, vol. 44, no. 10, pp. 884–891, 1990.