PERSPECTIVE
https://doi.org/10.1038/s42256-019-0078-4

nature machine intelligence

Robotic manipulation and the role of the task in the metric of success

V. Ortenzi[1]*, M. Controzzi[2,3], F. Cini[2,3], J. Leitner[4], M. Bianchi[5,6], M. A. Roa[7] and P. Corke[4]

Humans perform object manipulation in order to execute a specific task. Seldom is such action started with no goal in mind. In contrast, traditional robotic grasping (first stage for object manipulation) seems to focus purely on getting hold of the object—neglecting the goal of the manipulation. Most metrics used in robotic grasping do not account for the final task in their judgement of quality and success. In this Perspective we suggest a change of view. Since the overall goal of a manipulation task shapes the actions of humans and their grasps, we advocate that the task itself should shape the metric of success. To this end, we propose a new metric centred on the task. Finally, we call for action to support the conversation and discussion on such an important topic for the community.

Many manufacturing companies have modernized plants and factories, adopting high-tech solutions that involve robotic manipulators achieving high-throughput execution of repetitive jobs. Their presence in industrial settings is eased by the structured nature of such environments. However, fully useful robots are still a chimaera in more unstructured environments such as domestic settings. Industrial factories are laid out so that objects are placed in positions that are known a priori, and robots perform actions that can be defined with utmost certainty. In contrast, homes are an example of unstructured environments, as they are usually dynamic and unpredictable.

Two solutions have been recently introduced to account for uncertainty in object position and to improve adaptability of robotic hands with respect to the environment[1]: compliance in the robotic structure (for example, variable stiffness actuators) and soft grippers. However, object manipulation involves a high degree of unpredictability, both when creating products in industry and, increasingly, when helping and assisting humans. In this view, there is a push in industry (for example, the current trend of automation named industry 4.0 (ref. [2]) to allow robots and humans to collaborate in smart factories. In these shared environments, robots are expected to seamlessly interact with their surroundings and other concurrent agents—humans and potentially other robots—for successful cooperation and collaboration. To this end, robots need to be able to effectively manipulate objects not only in structured environments, but also in unstructured environments.

Dexterous object manipulation is the basis of human interaction with the environment. We interact with objects to accomplish a task we have in mind[3]. We know our goal, we perceive the environment and we devise an action plan that we can modify on the go while acting. In this frame of mind, the first action we commit to is to grasp the object we want to use. Therefore, grasping and manipulation are conceived by humans as task-oriented or purposive actions. These actions are performed to accomplish a task and they are heavily defined and parameterized by it. At the level of human motor control, similarities in the main building blocks exist across different tasks, such as element of geometrical basis, or synergies. However, a task specialization needs the enrolment of additional elements of such a geometrical basis for action execution[4,5]. Grasp choice, hand shaping and hand position on the object are parameters heavily influenced by the task to perform[6–12].

Intuitively, we choose to place and shape our hands on the object to fulfil the requirements of the task. Different tasks might induce different hand shaping and hand placement even on the same object. When signing a document, we most probably grasp a pen leaving the tip free from obstruction while the pen itself is held stably, as in Fig. 1. Such a grasp allows the pen to be moved and to leave the intended traces of ink on the paper[13]. Should we want to pass the pen to another person to sign the document, we would grasp the pen in a different way—most likely leaving the handle free for the second person to grasp the pen and write immediately after the handover[14]. Research in human cognitive and neurological models has long attributed the 'how to grasp' with the 'why to grasp', stemming from a separate process[15]. More recently, physiological data suggest that, instead of computing all decisions in an abstract form, decisions are made in the same regions of the brain that guide movement execution[16,17]. Neuroscientific data point to the sensorimotor regions of the frontal and parietal cerebral cortex as well as parts of the basal ganglia.

Robotic manipulation—especially grasping—has advanced admirably in recent years, but the focus has mainly been on the success of a grasp that has no further goal than to grasp an object, correlating to what the sensorimotor control of the human grasping process would be. It is completely oblivious of the inherent purposive character of the object grasping and manipulation. Although some contributions have investigated robotic grasping and manipulation as being task-oriented, there seems to be a big divide between the analysis of the human action of object grasping and manipulation, and what we have reproduced on robotic platforms so far. In other words, higher-level decision making based on the task is not yet an input to the grasping process.

In this Perspective, we argue that an approach more focused on the task might further the success of robotic grasping and manipulation. We also suggest that a metric used to judge the quality and

[1]Extreme Robotics Laboratory, University of Birmingham, Birmingham, UK. [2]The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy. [3]Department of Excellence in Robotics and AI, Scuola Superiore Sant'Anna, Pisa, Italy. [4]ARC Centre of Excellence for Robotic Vision, Queensland University of Technology, Brisbane, Queensland, Australia. [5]Centro di Ricerca "E. Piaggio", Università di Pisa, Pisa, Italy. [6]Department of Information Engineering, Università di Pisa, Pisa, Italy. [7]DLR—German Aerospace Center, Institute of Robotics and Mechatronics, Wessling, Germany. *e-mail: v.ortenzi@bham.ac.uk

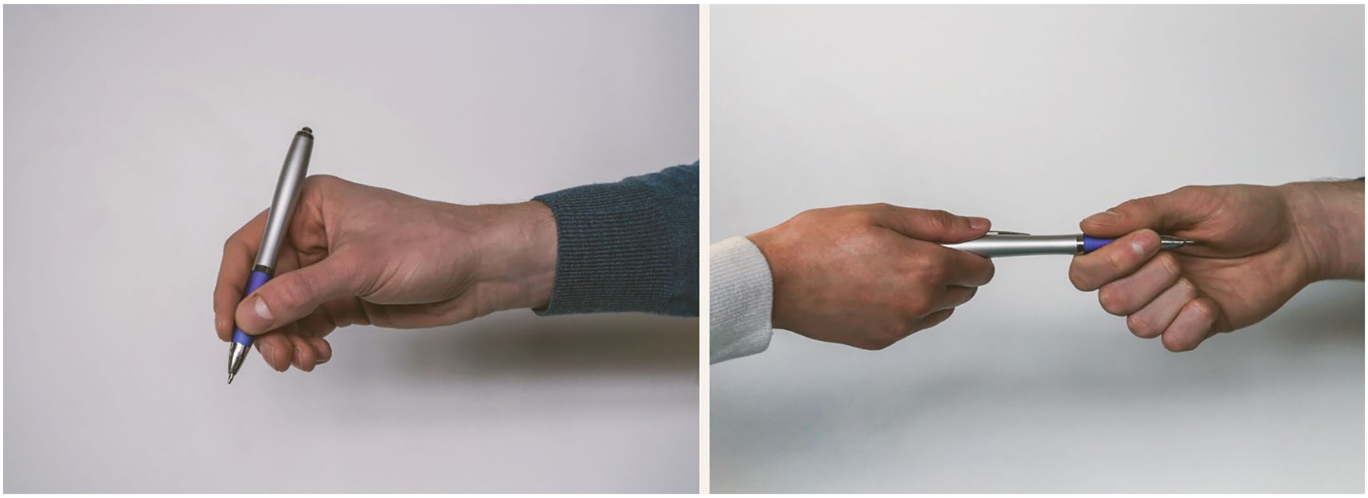

Fig. 1 | How a grasp changes because of a task. Left: grasping a pen for writing. Right: handing over the pen from a person in the right side to a person in the left side.

success of a grasp should include the grasp and manipulation task alongside more common quality indexes such as mean picks per hour (MPPH) and stability.

The need for object grasping and manipulation

Manipulation is arguably one of the skills that have allowed humans to be successful[18,19]. Human ability at manipulation has enabled important innovations such as constructing tools, building homes for shelter, fencing open fields to subordinate animals and mastering agriculture. It has also enabled human artistic expressions such as the *Pietà* by Michelangelo and the *Mona Lisa* by Leonardo, which are important for the soul of the human species.

A result of the human ability to invent and construct things is technology. Technology has pervaded the human world and it is now an integral part of it. So much so that at times, technology is asked to augment human capabilities, or at least to assist us while performing our tasks. Numerous are the examples of how technology is relied on blindly, such as how we constantly rely on our phones to remind us the schedule for the day.

Robotics is a technology that is expected to help us. However, numerous sci-fi movies and books have fuelled people's imaginations, creating expectations that are unrealistic. The result is a considerable gap between what is expected from robotics and what robots are able to deliver at the moment. In the contexts of households and hospitals, robots must face challenges brought by the environment, including but not limited to unpredictable contacts and interactions, partial knowledge and a difficult interpretation of the available perceptions. For example, a robot nurse might be asked to support a patient with impaired movements while standing up from the bed, or might be required to fetch specific medicine from a specific cabinet. The first scenario requires the robot to understand the intention of the human patient, and offer the required supporting force to help them stand up. In the second scenario, the robot has to locate the cabinet, navigate through a dynamic environment (avoiding collisions with obstacles and humans), open the cabinet, grasp the right medicine and finally give the medicine—without dropping it—to the patient, possibly alongside a glass of water.

In parallel, robots are increasingly expected to interact with human workers even in factories, for tasks such as passing tools, using tools, and collaboratively moving heavy weights or objects that are excessively large to be handled by a single operator. These scenarios showcase the necessity of natural and efficient human–robot interaction in factories. Versatile, adaptive and robust object manipulation is a pivotal skill that is heavily requested of robots if they are expected to work alongside and assist us.

Object manipulation involves a number of fine-grained subactions that happen contemporaneously and/or in sequence. Perception and control are deeply entwined in a synthesis of feedforward, feedback and learning[3,20,21]. Also, when a joint action is needed (such as a handover), the requirement for coordination and understanding of the partner adds an additional layer of complexity to the whole process[22].

While being a very intricate and non-trivial process, a successful manipulation often starts with an accomplished grasp of the object to be manipulated. We are aware that object manipulation can also be carried out without grasping. Such is the case of the manipulation of some deformable materials—for example, making pizza dough. Manipulation by pushing an object also does not involve grasping. However, for the scope of this Perspective, we consider only object manipulation tasks that require grasping.

Purposive action for humans

Humans do not grasp objects for the sake of it. Grasping is hardly a nonsensical action that we perform to waste time. Developmental psychologists might object, as when toddlers, we play with objects to understand the world around us and to understand our motor plans[23]. However, during adulthood, human grasping has the clear purpose of preparing the object to be used for the intended manipulation task that motivated the whole macro-action. In these terms, grasping represents the first fine-grained sub-action of the object manipulation task. And as such, grasping is the significant first step to a successful manipulation. We intend object manipulation as the macro-action of interacting and modifying the environment through action, especially with hands; in such context, grasping is one of the possible sub-actions, and consists of gripping the object.

By the mid-twentieth century, studies in human physiology established that grasping is a purposive action[24]. In particular, Napier affirms[24] that "it is the nature of the intended activity that finally influences the pattern of the grip." He asserts that the choice of hand shaping depends ultimately on the manipulation task to perform. Further studies focused on the impressive number of characteristics and features that make a selected grasp successful for the intended task. Even if a general consensus is usually absent, most of the published literature agrees in categorizing grasps into three main classes: precision grasps, where the emphasis is put on manipulability and dexterity (for example, the grasp used on a little ball); power grasps, where the emphasis is put on stability (for example, the grasp used on a hammer); and grasps that present both characters to the same degree and thus are classified as intermediate grasps (for example, the grasp used on a pen when writing).

It appears that there are at least five determinant factors behind the choice of a grasp[6,25–28]. Object features (factor 1) such as its shape, size and embedded function all have a clear influence as they introduce kinematic and dynamic constraints on the grasp. For instance, objects with sides wider than 5 cm and a mass over 250 g usually demand a power grasp. Contrarily, small and light objects are easier to manipulate but, counterintuitively, they are said to not force a direct preference for precision grasps[29].

Similarly, a task (factor 2) introduces constraints such as force and mobility to act on the object and possibly the environment. For example, if a person uses a ladle directly, it is likely that the ladle will be grasped by its long handle. However, if the goal is to pass the ladle to another person and this receiver is expected to use the ladle directly after grasping it, chances are that the ladle will be carefully offered to the receiver with the handle as unencumbered as possible[14,15,30].

Factor 3 is represented by the human hand. Although incredibly robust and capable, the human hand has kinematic and dynamic constraints that cannot be easily overcome. Its length and size limit the size of the objects that can be grasped and manipulated. We find it difficult to get hold of particularly bulky items and often resort to the use of both hands—when possible—and/or tools to assist us.

Another factor that influences and primes the choice of a grasp is prior experience (factor 4). The habits of the person approaching the grasp appear to inherently shape the choice, due possibly to the things learned during previous grasping attempts. Moreover, experience helps to build knowledge about the dynamic properties of objects, which provides the grasper with an understanding of how objects will react to a manipulation attempt[11]. Such experience can be seen in the case of passing a plate to another person. If the passer has prior experience in waiting tables, the use of a non-prehensile grasp is more likely than if the passer has never waited tables.

Factor 5 is chance. Happenstance seems to affect primarily the approach to the object and then the dynamic part of the grasp. In fact, the static part of the grasp—the final position of the hand with respect to the object—can be achieved comfortably by readjusting the object in the hand. Details such as the initial position of the object and the environmental constraints are unknowns to be considered in each situation (for example, a pen can be on a table, in a pen holder or in a person's hand: each of these situations offers the object in a different initial position and with different environmental constraints). Environmental constraints could be a factor on their own, as they provide limitations of motion, support and reaction forces that could be used to facilitate the manipulation task[1,31,32].

Finally, two additional factors are safety and social convention, which seem to be most important when there are multiple actors involved. For example, if a person uses a knife directly, it is highly likely that the knife will be grasped by its hilt. However, if the goal is to hand the knife over to another person, chances are that the knife will be carefully grasped and offered to the receiver with the hilt completely free.

These insights invited many researchers to delve deeper into the action of grasping and how this action prepares for manipulation. The high number of parameters involved has directed the focus of the investigation into the analysis of subsets of them. By and large, there is not much consensus on the characterization of a grasp. As already mentioned, most of the research seems to agree that each grasp has two fundamental defining characteristics: the stability given to the object and the manipulability with which it is possible to move the grasped object[33–35]. The stability of the grasp is a measure of its robustness and is defined as the action of the hand on an object that prevents its motion relative to the hand, possibly in the face of disturbing forces acting on the object itself. If the stability is preponderant, then the grasp is classified as a power grasp. The manipulability is defined as the capability of changing the position and orientation of the manipulandum from a given reference configuration to a different configuration. It implies contacts and low forces at the fingertips; thus, if this characteristic is the most influential, the grasp is classified as a precision grasp. This division is very general and simplistic in its stark classification; thus, it suffers from many drawbacks. First, there are intermediate classes of grasps for which the belonging to either of the general classes would be an imposed stretch. As a result, intermediate grasps appear in the literature as between power and precision grasps[6].

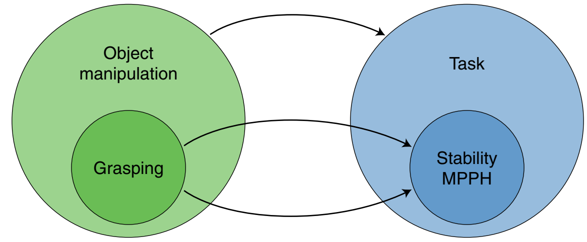

Fig. 2 | Manipulation and grasping. Left: action level. Right: requirements (and measures) of success. Object manipulation has an overall task that dictates the ultimate success or failure of the action. Grasping is one of the sub-actions of manipulation, and metrics to judge grasping must include stability and velocity (MPPH).

Second, such broad classes do not help when classifying grasps, as they leave broad space to interpretation and arbitrary choices during the classification. Indeed, most of the grasp taxonomies include finer-grained grasp types. For instance, the taxonomy in ref. [36] contains over 30 different types of grasp.

However, an all-inclusive taxonomy is impossible to obtain and, as such, much is left to the arbitrary choice of the classifying arbiter when associating the label to a grasp. To strive to achieve a complete taxonomy might be considered an ill-posed problem as it is virtually impossible to count the innumerable variations of grasps normally used by humans. Furthermore, a more marked focus on the purpose of the grasp would push towards the over-imposing duty of analysing an infinite number of potential tasks that demand objects to be grasped. The good news is that some underlying principles seem to exist.

The success of a grasp

Defining a successful grasp is a tricky problem—a well-accepted opinion in the robotics community. What makes a grasp, or for that matter, any performed action, successful?

Humans continuously adapt while grasping. In contrast, robots do not yet possess this reactiveness, and generally the 'most right' grasp has to be calculated a priori. In these terms, a grasp is generally classified as successful if the object is stably in the robotic end-effector for some arbitrary time. Yet this is usually not what is considered to be the end of the action and therefore a successful outcome of human grasping.

Most of the robotic research so far has focused attention on stability and the role of forces during a grasp[37]. The concepts of form-closure and force-closure[38] are most helpful while describing the static characteristics of a grasp. Form closure defines stability in geometric terms regardless of the external forces (if these forces do

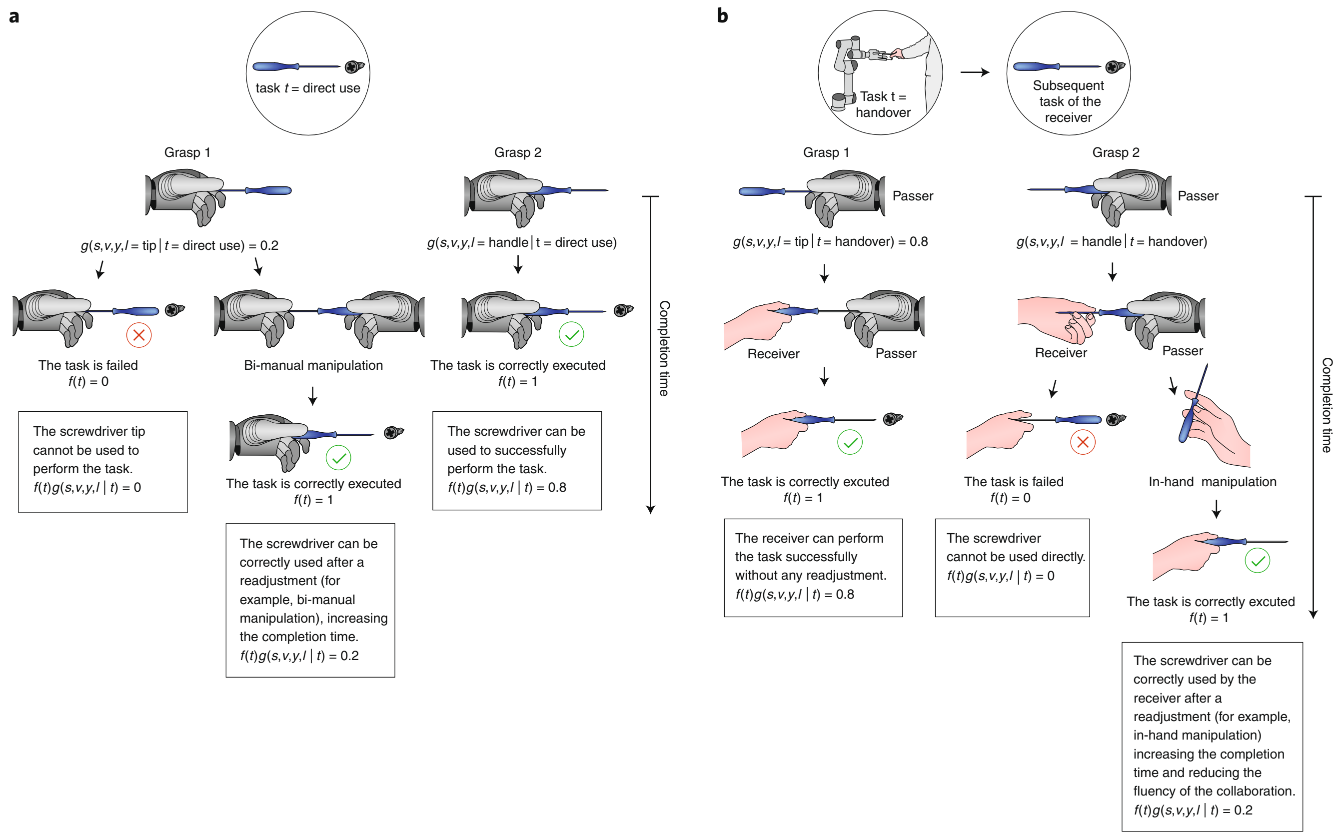

Fig. 3 | The case of a screwdriver. In such a context, grasps are changed when using a screwdriver directly (A) with respect to handing it over to another person for this other person to use it[14] (B). For a grasp with force s, velocity v and grasp type y, the function g can be expressed as follows: if the grasp location is on the handle (grasp 2 in A and B), then $g(s, v, y, l = \text{handle} \mid t = \text{direct use}) = 0.8$, while $g(s, v, y, l = \text{handle} \mid t = \text{handover}) = 0.2$. If the grasp is on the tip or close to the tip (grasp 1 in A and B), then $g(s, v, y, l = \text{tip} \mid t = \text{direct use}) = 0.2$, while $g(s, v, y, l = \text{tip} \mid t = \text{handover}) = 0.8$. In this example, velocity of grasping is not a key factor and is therefore not accounted for in the final value. When the aim is to pass it to another person, the screwdriver is not usually grasped on the handle[14]. Thus, a high-quality metric value (0.8) is given to a grasp on the tip, while a low value (0.2) is given to a grasp on the handle. Such a grasp would imply a potential difficulty in receiving the object, compelling the receiver to compensate by performing unnatural movements during the reaching movement and/or readjusting the object in their hands prior the task, increasing the completion time. These values could, for example, be devised using a statistical evaluation of grasp occurrence in the same human action and revised exploiting a learning algorithm[60].

not cause a mechanical collapse of the grasping device). Force closure defines stability based on the equilibrium of the forces at play, and the ability to resist and hold on to the object in the presence of external perturbations as well. So, if the direction of a force changes, then stability can be affected.

Taxonomies in refs. [6,36] have had deep influence over the robotics community. Moreover, other taxonomies such as in refs. [39,40] offer further insights into pressure patterns on fingers and the palm associated with different classes of grasps and force pattern distributions. Furthermore, a recent taxonomy of grasps based on biomedical data[41] relates taxonomies designed with qualitative descriptions of the grasps to a quantitative hierarchy.

Considering the definition of stability mentioned above and what has been established in studies on human grasping and manipulation, manipulation actions require a varying degree of stability and manipulability. Thus, stability appears to be a good indicator of success. However, it seems reductive to account only for stability while judging a grasp. In fact, a stable grasp can still be a 'wrong' grasp if it hinders the fulfilment of the ultimate goal of the action. In the example of the pen, a power grasp placed close to the writing tip is stable and would be classified successful in terms of stability. However, such a grasp would not allow the grasper to perform the writing task, so a re-adjustment would be needed. Hence, stability per se is necessary, but not sufficient to reach any conclusion about the success of a grasp from a task perspective.

Moreover, great effort has been put on the realization of grasp actions that are as fast as possible. MPPH is sometimes used as another measure of quality and success[42]. MPPH is defined as the product of the mean grasp rate and the mean grasp reliability (success rate). This metric is particularly useful in logistics. In terms of production, it is of paramount importance that a robot is able to pick objects quickly because the overall output of a factory—and ultimately the revenue—depends on it. Also, the task is generally simple, so performance speed is sufficient. Although speed is heavily requested in many industrial applications, speed alone does not seem to be a fair and just measure of success for a grasp generally. Another example from the logistics domain depicts a very different picture: sometimes workers not only have to pick the object, but also read a barcode. This implies orienting the object in a way that the reader recognizes such a barcode. In this case, speed cannot be the only measure of quality for the grasp. In our opinion, a metric based on picking speed suffers from the same drawbacks of the stability-only metric. It does not account for the task—the whole story behind a grasping attempt. A workaround can be designed in MPPH, accounting for the manipulation task while defining the grasp reliability. But that is equivalent to state that the purpose of

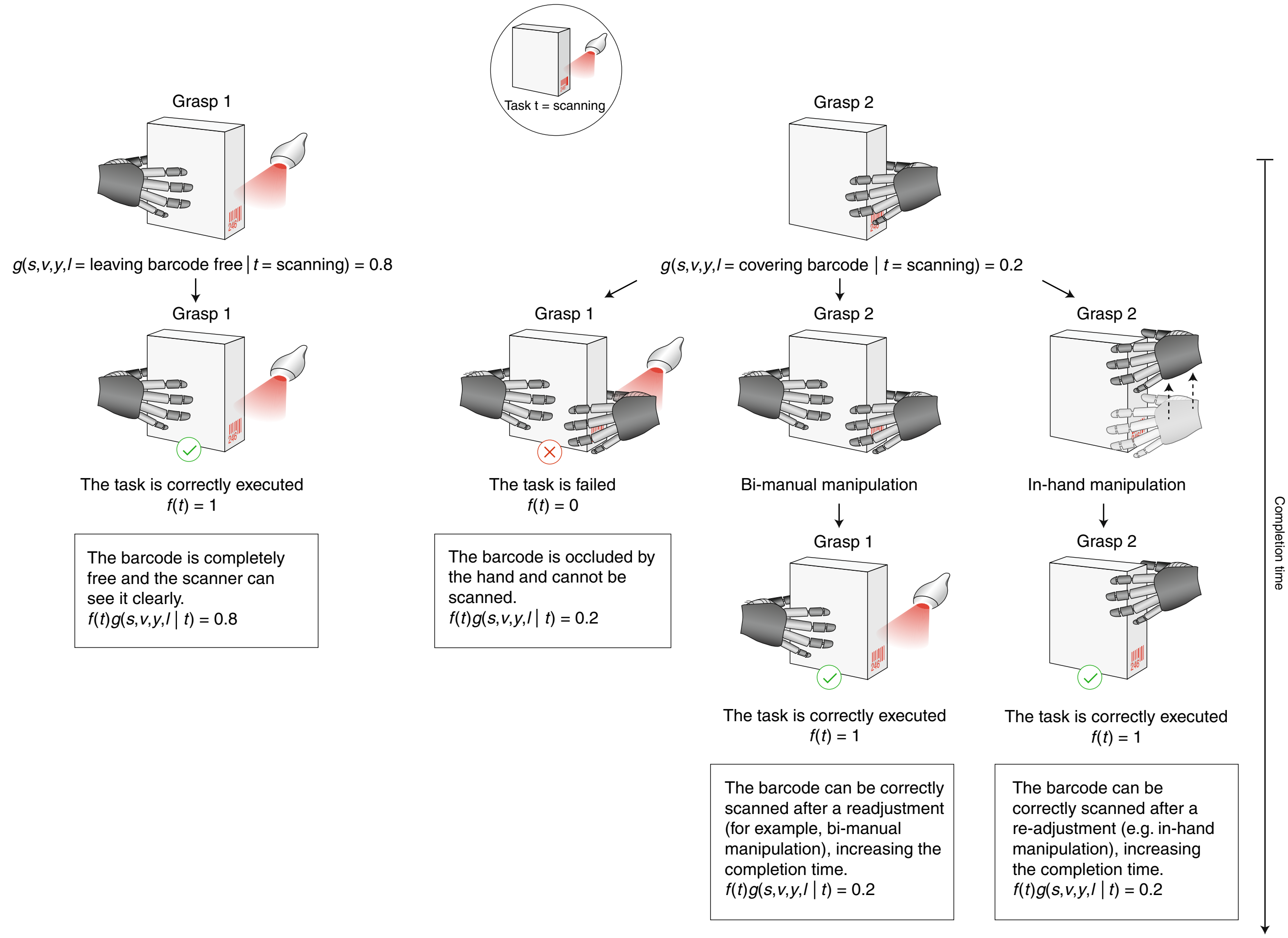

Fig. 4 | The case of a robotic arm having the task *t* to scan a box of cereals. Consider two grasps with the same velocity v, same grasp type y, but with different grasping forces $s_2 > s_1$ and location l_2 = covering barcode and l_1 = leaving barcode free. Then, if the readjustment of the grasp is not possible, $\rho = f(t)g(s_2, v, y, l_2$ = covering barcode | t = scanning) = 0 as the task failed ($f(t) = 0$), $g(s_1, v, y, l_1$ = leaving barcode free | t = scanning) = 0.8. This example shows how such a metric would account for stability and velocity with respect to the task to perform, discounting stability when the grasp is wrongly placed.

a grasp has to be accounted for when expressing a measure of its success, which is effectively our thesis.

We are not trying in any way to discount the impressive advances introduced by the classical ways to address robotic grasping. Instead, we would like to propose a more general practice to assess the success of a grasp. It seems that robots (and roboticists) are currently stuck on making low-level picking robust. While this is important, it often leads to the question of 'what next?' We strongly believe that this is an important part to include from the start in the metric to qualify a grasp. Such a metric should account for classical characteristics such as stability and speed, and concurrently of the ultimate purpose of such tasks.

The robotic community has recognized many of the aspects of human manipulation as fundamental to reproduce similar skills in robots. From the five factors mentioned in the previous section, robotics researchers are actively pursuing research in most of them. In particular, the developments in computer vision and machine learning over the past decade have created robotic capabilities along factor 1 that are getting close to human levels—that is, object detection, localization and even defining an object's 'graspability' (or grasp quality). The advent of deep learning has progressed the state of what a computer system can learn and apply to a new task, possibly also imitating human behaviour[43]. This is a step towards an experience-based grasping system for robotic systems (factor 4). Recently, reactive grasping solutions have become more prominent and have delivered better systems[44–46] by being able to deal with changing placements (reacting to the changes in the environment, possibly due to factor 5).

Task-oriented grasping is not a novel concept in the robotic community, with papers investigating grasping and manipulation strategies heavily based on the task to be performed. The problem of purposive robotic grasping can be found in refs. [47–51]. A concept that often recurs is object affordances[15,30,52–55].

Ref. [15] says, "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill." This concept is borrowed from the field of ecological psychology, and it has been exploited in robotics to further constrain and guide the choice of a grasp in order to maximize the success of the ultimate manipulation task. This seems to be a sensible choice as tasks' and objects' constraints are two of the factors influencing the human grasping behaviour.

Another way to include the purpose of the task into a grasping strategy can be using learning—and deep learning—when extracting information from visual feedback. Learning from imitation is also a successful paradigm to repeat actions shown by a human. However, the scaling of such an approach to more generalized actions is still critical[56].

Furthermore, other work has investigated how to account for and exploit the environmental constraints[57,58]. For example, when we face the task to move a book lying on a table (flat surface on flat surface) to the most appropriate spot on a bookshelf, we consider using the back wall at the end of the table to rotate the book, or alternatively, to slide the book to the edge of the table. In the first case, we would be exploiting the reaction force of the wall to hinge the rotation of the book; in the second scenario, we would use the table as a supportive surface for the sliding motion and then use its ending (the edge) to grasp the book more comfortably.

A new metric needs to be on the higher level of manipulation

The task of a grasp should be considered in metrics of quality and success. While the current metrics—stability and MPPH—account for fundamental characteristics, those characteristics alone are necessary but not sufficient to determine the success of the whole action. Introducing the purpose of the action would potentially allow a fairer and just assessment of the success of a grasp—and of the manipulation task as a whole.

We propose a change of perspective (Fig. 2). Instead of basing these metrics at the level of the grasp (sub-action), we suggest to raise the level to the manipulation task, thus considering primarily the completion of the ultimate goal of the whole action. Such a metric would still account for the classical metrics used to judge a grasp (stability and MPPH). However, the requirements in terms of stability and speed are dependent on the purpose of the grasp, which is dictated by the manipulation task. Thus, we advance a new metric ρ:

$$\rho = f(t) \times g(s,\ v,\ y,\ l|t)$$

where t is the task, s is the grasping force, v is a measure of picking velocity, y is the grasp type and l is the grasp location. $g(s, v, y, l|t)$ is a function that depends only on the robot inputs (grasping force, speed, grasp type and location) given a task t. $g(s, v, y, l|t)$ returns a value between 0 and 1, judging how good the grasp is based on the outcome measures of task t (each manipulation task can be carried out in multiple ways). The proposed metric can act as MPPH when picking speed is of tantamount importance, while it is flexible enough to allow for tasks whose success or quality of performance is not defined (solely) by picking speed. In other words, the function g is shaped in accordance with the outcome measures of the task t. This entails that if speed is important, then speed will shape the metric. Speed could potentially have a negligible role in other tasks, and this would be reflected in the metric seamlessly. And $f(t)$ is a Boolean function that returns 0 if the task failed and 1 if succeeded, evaluated on the single attempt at a specific manipulation task.

In our opinion, this metric would be beneficial not only to judge the quality of a grasp (a posteriori), but also to synthesize grasps directly targeting the overall manipulation task (for example, it can be used to predict the quality of the grasp and to choose the most appropriate grasp for the task). Additionally, this metric could be used in a learning framework[59] to shape the reward function in a reinforcement learning paradigm.

Figures 3 and 4 show two examples of how this metric works.

Conclusion

We presented a novel metric to judge the success of a manipulation task. Generally, stability and MPPH judge the quality and success of a grasp. In our opinion, this is a limitation as grasping is a purposive action, and as such, it should be evaluated in the wider scenario of the manipulation task. Hence, we propose a metric that embeds these features of grasping, which are necessary to the overall success of the manipulation, but is shaped by the task of the manipulation.

Our formulation of metric leaves the definition of the function g open. The quality measure depends on the outcome measures of the task; thus, it cannot be univocally defined and it must be considered case by case, similarly to what happens to the work-around in the MPPH metric when considering the task in the mean grasp reliability.

We acknowledge the difficulty in considering the task and we also acknowledge that the community recognizes the role of the task in selecting a grasp. However, we believe that putting more effort into this challenge would greatly benefit robotic manipulation. For this reason, we invite the reader to a thorough conversation on the topic. We are open to dialogue and invite readers to contact us to discuss further.

Received: 11 May 2019; Accepted: 9 July 2019;
Published online: 9 August 2019

References

1. Santina, C. D. et al. Learning from humans how to grasp: a data-driven architecture for autonomous grasping with anthropomorphic soft hands. *IEEE Robot. Autom. Lett.* **4**, 1533–1540 (2019).
2. Østergaard, E. H. *White Paper: The Role of Cobots in Industry 4.0* (Universal Robots, 2018).
3. Flanagan, J. R., Bowman, M. C. & Johansson, R. S. Control strategies in object manipulation tasks. *Curr. Opin. Neurobiol.* **16**, 650–659 (2006).
4. Santello, M., Flanders, M. & Soechting, J. F. Postural hand synergies for tool use. *J. Neurosci.* **18**, 10105–10115 (1998).
5. Santello, M. et al. Hand synergies: integration of robotics and neuroscience for understanding the control of biological and artificial hands. *Phys. Life Rev.* **17**, 1–23 (2016).
6. Cutkosky, M. R. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Trans. Robot. Autom.* **5**, 269–279 (1989).
7. Ansuini, C., Santello, M., Massaccesi, S. & Castiello, U. Effects of end-goal on hand shaping. *J. Neurophysiol.* **95**, 2456–2465 (2006).
8. Ansuini, C., Giosa, L., Turella, L., Altoè, G. & Castiello, U. An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* **185**, 111–119 (2008).
9. Cohen, R. G. & Rosenbaum, D. A. Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Exp. Brain Res.* **157**, 486–495 (2004).
10. Friedman, J. & Flash, T. Task-dependent selection of grasp kinematics and stiffness in human object manipulation. *Cortex* **43**, 444–460 (2007).
11. Lukos, J., Ansuini, C. & Santello, M. Choice of contact points during multidigit grasping: effect of predictability of object center of mass location. *J. Neurosci.* **27**, 3894–3903 (2007).
12. Crajé, C., Lukos, J. R., Ansuini, C., Gordon, A. M. & Santello, M. The effects of task and content on digit placement on a bottle. *Expe. Brain Res.* **212**, 119–124 (2011).
13. Feder, K. P. & Majnemer, A. Handwriting development, competency, and intervention. *Dev. Med. Child Neurol.* **49**, 312–317 (2007).
14. Cini, F., Ortenzi, V., Corke, P. & Controzzi, M. On the choice of grasp type and location when handing over an object. *Sci. Robot.* **4**, eaau9757 (2019).
15. Gibson, J. J. *The Ecological Approach to Visual Perception* (Houghton Mifflin, 1979).
16. Johnson-Frey, S. H. The neural bases of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78 (2004).
17. Cisek, P. & Thura, D. in *Reach-to-Grasp Behavior: Brain, Behavior, and Modelling Across the Life Span* (eds Corbetta, D. & Santello, M.) Ch. 5 (CRC, 2018).
18. Gibson, K. Tools, language and intelligence: evolutionary implications. *Man* **26**, 255–264 (1991).
19. Iriki, A. & Taoka, M. Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philos. Trans. R. Soc. Lond. B* **367**, 10–23 (2012).
20. Johansson, R. & Cole, K. J. Sensory-motor coordination during grasping and manipulative actions. *Curr. Opin. Neurobiol.* **2**, 815–823 (1993).
21. Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. Principles of sensorimotor learning. *Nat. Rev. Neurosci.* **12**, 739–751 (2011).
22. Sacheli, L. M., Arcangeli, E. & Paulesu, E. Evidence for a dyadic motor plan in joint action. *Sci. Rep.* **8**, 5027 (2018).
23. Corbetta, D. & Snapp-Childs, W. Seeing and touching: the role of sensory-motor experience on the development of infant reaching. *Infant Behav. Dev.* **32**, 44–58 (2009).
24. Napier, J. R. The prehensile movements of the human hand. *Bone Joint J.* **38-B**, 902–913 (1956).
25. Landsmeer, J. M. F. Power grip and precision handling. *Ann. Rheum. Dis.* **21**, 164–70 (1962).

26. Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F. & Miura, Y. Patterns of static prehension in normal hands. *Am. J. Occup. Ther.* **34**, 437–445 (1980).
27. Iberall, T. The nature of human prehension: three dextrous hands in one. In *Proc. 1987 IEEE International Conference on Robotics and Automation* 396–401 (IEEE, 1987).
28. Iberall, T. Human prehension and dexterous robot hands. *Int. J. Robot. Res.* **16**, 285–299 (1997).
29. Feix, T., Bullock, I. & Dollar, A. M. Analysis of human grasping behavior: object characteristics and grasp type. *IEEE Trans. Haptics* **7**, 311–323 (2014).
30. Osiurak, F., Rossetti, Y. & Badets, A. What is an affordance? 40 years later. *Neurosci. Biobehav. Rev.* **77**, 403–417 (2017).
31. Eppner, C., Deimel, R., Álvarez Ruiz, J., Maertens, M. & Brock, O. Exploitation of environmental constraints in human and robotic grasping. *Int. J. Robot. Res.* **34**, 1021–1038 (2015).
32. Nakamura, Y. C., Troniak, D. M., Rodriguez, A., Mason, M. T. & Pollard, N. S. The complexities of grasping in the wild. *In Proc. 2017 IEEE-RAS 17th International Conference on Humanoid Robotics* 233–240 (IEEE, 2017).
33. Bicchi, A. Hands for dexterous manipulation and robust grasping: a difficult road toward simplicity. *IEEE Trans. Robot. Autom.* **16**, 652–662 (2000).
34. Controzzi, M., Cipriani, C. & Carrozza, M. C. in *The Human Hand as an Inspiration for Robot Hand Development* (eds Balasubramanian, R. & Santos, V. J.) 219–246 (Springer, 2014).
35. Roa, M. A. & Suárez, R. Grasp quality measures: review and performance. *Auton. Robots* **38**, 65–88 (2015).
36. Feix, T., Romero, J., Schmiedmayer, H., Dollar, A. M. & Kragic, D. The grasp taxonomy of human grasp types. *IEEE Trans. Hum. Mach. Sys.* **46**, 66–77 (2016).
37. Falco, J., Van Wyk, K., Liu, S. & Carpin, S. Grasping the performance: facilitating replicable performance measures via benchmarking and standardized methodologies. *IEEE Robot. Autom. Mag.* **22**, 125–136 (2015).
38. Bicchi, A. & Kumar, V. Robotic grasping and contact: a review. In *Proc. 2000 IEEE International Conference on Robotics and Automation* 348–353 (IEEE, 2000).
39. Gonzalez, F., Gosselin, F. & Bachta, W. Analysis of hand contact areas and interaction capabilities during manipulation and exploration. *IEEE Trans. Haptics* **7**, 415–429 (2014).
40. Abbasi, B., Noohi, E., Parastegari, S. & Zefran, M. Grasp taxonomy based on force distribution. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication* 1098–1103 (IEEE, 2016).
41. Stival, F. et al. A quantitative taxonomy of human hand grasps. *J. Neuroeng. Rehabil.* **16**, 28 (2019).
42. Mahler, J. et al. Guest editorial open discussion of robot grasping benchmarks, protocols, and metrics. *IEEE Trans. Autom. Sci. Eng.* **15**, 1440–1442 (2018).
43. Goldberg, K. Robots and the return to collaborative intelligence. *Nat. Mach. Intell.* **1**, 2–4 (2019).
44. Leitner, J., Frank, M., Forster, A. & Schmidhuber, J. Reactive reaching and grasping on a humanoid: towards closing the action-perception loop on the iCub. In *2014 11th International Conference on Informatics in Control, Automation and Robotics* 102–109 (IEEE, 2014).
45. Redmon, J. & Angelova, A. Real-time grasp detection using convolutional neural networks. In *Proc. IEEE International Conference on Robotics and Automation* 1316–1322 (IEEE, 2015).
46. Morrison, D., Corke, P. I. & Leitner, J. Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach. *Robotics: Science and Systems* https://doi.org/10.15607/RSS.2018.XIV.021 (2018).
47. Hjelm, M., Ek, C. H., Detry, R. & Kragic, D. Learning human priors for task-constrained grasping. In *Proc. 10th International Conference on Computer Vision Systems* (eds Nalpantidis, L., Krueger, V., Eklundh, J.-O. & Gasteratos, A.) 207–217 (Springer, 2015).
48. Song, H. O., Fritz, M., Goehring, D. & Darrell, T. Learning to detect visual grasp affordance. *IEEE Trans. Autom. Sci. Eng.* **13**, 798–809 (2016).
49. Nguyen, A., Kanoulas, D., Caldwell, D. G. & Tsagarakis, N. G. Detecting object affordances with convolutional neural networks. In *Proc. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems* 2765–2770 (IEEE, 2016).
50. Detry, R., Papon, J. & Matthies, L. Task-oriented grasping with semantic and geometric scene understanding. In *Proc. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* 3266–3273 (IEEE, 2017).
51. Kokic, M., Stork, J. A., Haustein, J. A. & Kragic, D. Affordance detection for task-specific grasping using deep learning. In *Proc. 2017 IEEE-RAS 17th International Conference on Humanoid Robotics* 91–98 (IEEE, 2017).
52. Chemero, A. An outline of a theory of affordances. *Ecol. Psychol.* **15**, 181–195 (2003).
53. Montesano, L., Lopes, M., Bernardino, A. & Santos-Victor, J. Learning object affordances: from sensory-motor coordination to imitation. *IEEE Trans. Robot.* **24**, 15–26 (2008).
54. Osiurak, F., Jarry, C. & Gall, D. L. Grasping the affordances, understanding the reasoning: toward a dialectical theory of human tool use. *Psychol. Rev.* **117**, 517–540 (2010).
55. Hsu, J. Machines on mission possible. *Nat. Mach. Intell.* **1**, 124–127 (2019).
56. Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P. & Schaal, S. Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput.* **25**, 328–373 (2013).
57. Kappler, D., Chang, L. Y., Pollard, N. S., Asfour, T. & Dillmann, R. Templates for pre-grasp sliding interactions. *Robot. Auton. Sys.* **60**, 411–423 (2012).
58. Puhlmann, S., Heinemann, F., Brock, O. & Maertens, M. A compact representation of human single-object grasping. In *Proc. of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems* 1954–1959 (IEEE, 2016).
59. Cheng, G., Ramirez-Amaro, K., Beetz, M. & Kuniyoshi, Y. Purposive learning: Robot reasoning about the meanings of human activities. *Sci. Robot.* **4**, eaav1530 (2019).
60. Kupcsik, A., Hsu, D. & Lee, W. S. in *Robotics Research* Vol. 1 (eds Bicchi, A. & Burgard, W.) 161–176 (Springer, 2018).

Acknowledgements

V.O. is supported by the UK National Centre for Nuclear Robotics initiative, funded by EPSRC EP/R02572X/1. M.B. is partially supported by the EU H2020 project 'SOFTPRO: Synergy-based Open-source Foundations and Technologies for Prosthetics and RehabilitatiOn' (no. 688857), and by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence). M.A.R.'s work has been partially funded by the European Commission's Eighth Framework Program as part of the project Soft Manipulation (grant number H2020-ICT-645599). J.L. and P.C. are supported by the Australian Research Council Centre of Excellence for Robotic Vision (project no. CE140100016). F.C. and M.C. are supported by the European Research Council (project acronym MYKI; project number 679820). The authors want to thank K. Goldberg for his invaluable comments and constructive criticism.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to V.O.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.