

# Is Kinesthetic Teaching What Smart Factories Really Need?

Jessica Villalobos, Enrique Coronado, Alessandro Carfi<sup>1</sup>, Barbara Bruno<sup>2</sup>,  
and Fulvio Mastrogiovanni<sup>3</sup>

<sup>1</sup> University of Genoa, Via Opera Pia 13, 16145 Genoa, Italy  
`alessandro.carfi@dibris.unige.it`

<sup>2</sup> University of Genoa, Via Opera Pia 13, 16145 Genoa, Italy  
`barbara.bruno@unige.it`

<sup>3</sup> University of Genoa, Via Opera Pia 13, 16145 Genoa, Italy  
`fulvio.mastrogiovanni@unige.it`

**Abstract.** Programming by demonstration techniques have been investigated to facilitate and speed up the setup of new robot tasks. Kinesthetic teaching (KT), i.e., teaching by physically *guiding* a robot in the execution of a motion, has been adopted in industrial scenarios for its ease of use. In the work described here, we analyse and discuss limits and drawbacks of KT and suggest the adoption of a set of autonomous behaviors.

## 1 Introduction

The Industry 4.0 paradigm [3] requires techniques for fast and easy-to-attain robot task reconfigurations. *Programming by Demonstration* (PbD) addresses such requirements by making it possible to teach robots new tasks intuitively [4] [7] [1] [6] [5] [2]. PbD is characterized by two phases: *teaching*, in which one or multiple examples are shown by means of physical guidance, and *learning*, in which examples are generalized in order to obtain a resulting robot behavior. The execution of the task, then, is simply the autonomous repetition of the learned behavior. When adopted in industrial scenarios, PbD has the competitive advantage of not requiring any engineering knowledge for robot reconfiguration and task teaching. Therefore, robot operators are specialized laborers whose knowledge mainly derives from the experience gained when operating the robot manually. This lack of awareness of how robots work, their limitations, as well as the inherent differences between their and human motions, can lead to low quality robot motions, e.g., inefficient trajectories or increased execution time. Is it acceptable for the *quality* of robot motions to be highly sensitive to the operator's teaching skills?

In this discussion, when we refer to *quality*, we consider the time required for the teaching phase, in so far as it affects the duration of task execution. In particular:

- *pure playback* is the worst case, since execution time is equal to teaching time, and therefore any inefficiency in the teaching phase is replicated during execution;
- *way points playback* optimizes execution time but requires a longer teaching time (e.g., the operator must stop at each key way point), does not constrain motion between pairwise way points, and therefore their suboptimal selection can lead to inefficient motions;
- *generalization* over different examples reduces the influence of a single *bad* example on execution time, but increases the required teaching time.

## 2 Rationale and Hypotheses

PbD makes it possible to setup a robot for operators without engineering knowledge about how robots work, but with practical knowledge of industrial equipment. This can lead to (problem  $P_1$ ) sub-optimal solutions as far as a trajectory  $\tau$  is concerned, or to the need of inserting more complex and robust learning approaches to overcome sub-optimality, therefore resulting in a longer overall teaching time  $t_{|\tau|}$ . Since different operators have different working experience, (insight  $I_1$ ) an experienced operator is likely to obtain better results at teaching than a less experienced one. Furthermore, operators with different skills and experience may spend a different time for the same teaching goal, which could ( $P_2$ ) tempt less skilled operators to *speed up* the teaching procedure to the detriment of the final result. Indeed, ( $I_2$ ) the necessity for a fast reprogramming, lack of time, stress, inexperience or laziness are all factors that may shift focus to reduce teaching time, with possible drawbacks on the final result.

Starting from these considerations, we hypothesize that (hypothesis  $H_1$ ) the time required to teach a robot a given task, comprehensive of operator’s attempts and the required number of iterations, varies among operators.

Considering the robot’s end-effector trajectory  $\tau$  for a simple pick and place task, as executed at teaching time, ( $H_2$ ) the two spatial intervals related to *grasping* ( $\gamma$ ) and *releasing* ( $r$ ) are characterized by a higher density with respect to nearby points in the trajectory. As a consequence, it may be reasonable ( $I_4$ ) to reduce an operator’s influence during  $T_\gamma$  and  $T_r$  because of their criticality in the overall trajectory.

In order to validate what discussed above, we posit the following working hypotheses: ( $WH_1$ ) exposing operators to PbD training footage or to human-human interactions in which one human role-plays as a robot engaged in PbD, reduces the variance of teaching time; ( $WH_2$ ) the need to manually control the opening and closing of a robot’s gripper is a disturbing factor for the operator, which affects how grasping and releasing actions are taught; ( $WH_3$ ) the need to identify an appropriate grasping pose for the object to pick while teaching the task has a negative effect.

### 3 Methodology

We conducted a series of experiments involving 25 unpaid volunteers among students (and teachers) from a vocational education and training school in Italy, aged between 15 and 60. Volunteers have previous knowledge of industrial equipment, but not of collaborative robots, and may well be thought of as possible robot operators in *factories of the future*. The volunteers have been divided in two groups and asked to perform two activities, which we refer to as *human-human* interaction and *human-Baxter* interaction. The two groups perform the activities in different orders, i.e. with one group engaging in the human-human interaction activity before the human-Baxter interaction, and the other after.

**Human-Human Interaction:** At the beginning, a seven minutes video<sup>4</sup>, intended to demonstrate how KT works on Baxter, is shown to the group of volunteers. Then, volunteers divide in pairs, with one acting as the human *teacher* and the other as the robot *learner*, and are asked to apply the same methodology to teach a randomly-chosen task to each other. No verbal communication is allowed. Possible tasks include: screwing a jar lid, stacking six boxes to form a pyramid, ordering five bottles according to their weight, composing a square with four pens, picking and placing a box, folding a shirt, driving a screw in a piece of wood. The experiment is repeated twice swapping the roles of *teacher* and *learner*. During all the experiments, teacher and learner are in front of each other with a table in between. Once the teaching procedure ends, the learner is asked to repeat the task. Video were recorded during the experiments.

**Human-Baxter Interaction:** A Baxter dual-arm manipulator stands in front of a table on which two locations, namely *A* and *B*, are defined. In *A*, a 0.5 liter plastic bottle filled with water for about one tenth is located, whereas in *B* there is an open box with a  $30 \times 39$  cm base and a 12 cm height. The distance between *A* and *B* is 78 cm, while the distance between Baxter and the table is about 60 cm. Before the activity starts, an experimenter gives a practical demonstration on the use of KT to teach Baxter how to perform a pick and place task. Each experiment starts with the Baxter in the *untucked* pose, while we did not constrain the final pose. The experiment loosely follows this sequence:

- the volunteer stands in front of Baxter, on the other side of the table;
- the volunteer grasps the wrist of Baxter’s left arm to activate the zero gravity mode and starts the teaching procedure;
- the volunteer teaches Baxter how to relocate the bottle from *A* to *B* (i.e., how to put the bottle inside the open box) using KT;
- when appropriate, the volunteer pushes the buttons located on the robot’s wrist, to open and close its left gripper;
- the task is executed: if the robot does not succeed in performing it, the volunteer is asked to repeat the teaching procedure.

At all times, the trajectory  $\tau(t)$  is recorded expressed in joint space. It is noteworthy that no constraints on teaching time have been set for this experiment.

---

<sup>4</sup> <https://youtu.be/4FI7LwM3V38>

## 4 Discussion: What Next for Kinesthetic Teaching?

By analyzing density distributions in trajectories, we found that the *grasping* ( $T_\gamma$ ) and *releasing* ( $T_r$ ) phases group high density points. Furthermore, observations suggest that, in a typical pick and place task, the time spent in those phases is highly relevant, and amounts to about 38% of the total teaching time. These two findings prove that the process of teaching a robot how to grasp or release an object via KT is critical, and that any effort to improve it may lead to a relevant reduction of the required teaching (and possibly execution) time. The percentage of time spent during phases  $T_\gamma$  and  $T_r$  and the densities of points in these portions of the whole trajectory are characterized by a high variance among different volunteers. This is probably due to a difference in their skills or experience, and suggests the need to *reduce* the operator's *influence* on specific difficult phases of KT, by extending KT with a number of semi-autonomous robot behaviors.

To devise such behaviors we recur to human-human activities, and in particular to what happens when humans apply a simple form of KT to train each other. We observed that in the vast majority of cases, even without any verbal interaction the *learner* autonomously deduces that a given object (e.g., a box, a pen) must be picked, orients the hand in accordance with a suitable grasping pose, and autonomously closes the hand once deemed appropriate. A similar behaviour is observed during the release phase. These qualitative and quantitative observations lead us to identify two possible *disturbing factors* in the KT procedure: the need to control the opening and closing of the robot's gripper, as suggested by  $WH_2$ , and the need to select and reach the most appropriate grasping pose, as put forth by  $WH_3$ .

In order to lessen the consequences of these disturbing factors, we propose a research work plan that foresees the design and implementation of two semi-autonomous robot behaviors extending the basic KT paradigm: (i) the autonomous opening and closing of a robot's gripper when appropriate, which requires the robot to understand the operator's intentions and to reason about the configuration of the objects in its workspace, and (ii) the autonomous re-orientation of the gripper according to the identified grasping pose, which requires the robot to reason about object shapes and functions. The implementation of such abilities can allow a robot to actively participate in the teaching procedure, reducing the required teaching time and facilitating human-robot interaction via KT.

## 5 Acknowledgments

The authors would like to thank the teachers and students of the vocational education and training school "Centro Oratorio Votivo, Casa di Carità, Arti e Mestieri, Ovada" for their contribution to the drafting and execution of the experiments.

## References

1. Calinon, S., Guenter, F., Billard, A.: On learning, representing, and generalizing a task in a humanoid robot. *Neural Networks* **37**(2) (2007) 286–298
2. Dillmann, R., Kaiser, M., Ude, A.: Acquisition of elementary robot skills from human demonstration. In: *Proceedings of the International symposium on intelligent robotics systems (SIRS 1995)*, Pisa, Italy (1995) 185–192
3. H. Kagermann, W.W., Helbig, J.: Recommendations for implementing the strategic initiative industrie 4.0: Final report of the industrie 4.0 working group. *Produktion, Automatisierung und Logistik* (Frankfurt, 2013)
4. Halbert, D.C.: *Programming by example*. PhD thesis, University of California, Berkeley, USA (1984)
5. Inamura, T., Kojo, N., Inaba, M.: Situation recognition and behavior induction based on geometric symbol representation of multimodal sensorimotor patterns. In: *Proceeding of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, Beijing, China (October 2006) 5147–5152
6. Ito, M., Noda, K., Hoshino, Y., Tani, J.: Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks* **19**(3) (2006) 323–337
7. Kang, S.B., Ikeuchi, K.: A robot system that observes and replicates grasping tasks. In: *Proceedings of the 1995 IEEE International Conference on Computer Vision (ICCV 1995)*, Boston, USA (June 1995) 1093–1099