The methodology of training an underwater robot control system for operator actions

VA Kramar¹V V Alchakov¹

¹Sevastopol State University, 33, Universitetskayast., Sevastopol, 299053, Russia
E-mail:kramarv@mail.ru

Abstract. The article discusses the approach to building an underwater robot control system based on the “Programming by Demonstration” approach. The basic elements of the implementation of the approach are considered. The participants in the learning process are indicated. Input variables are defined, and a training model is selected. As a result, a training methodology for the underwater robot control system is formulated for the system to perform specified manipulations.

1. Introduction

The modern solution to the problem of constructing control of a multi-link manipulator with kinematic redundancy (after this referred to as a multi-link manipulator), for example, for the underwater robot (Fig. 1), is based on the use of the “Programming by Demonstration” (PbD) approach [1, 2, 3]. The approach based on teaching the robot how to perform various operations based on the results of the “teacher” action. The idea of the PbD approach borrowed from nature. For example, the basis of the child's learning is the repeatability of movements. A certain number of repetitions form the algorithm of the muscle fibres that provide the desired movement and translates this algorithm into the brain for the automatic execution of this movement in life [4].

The PbD approach provides a natural way to program robots, showing the desired task. It gives the manipulator with information on how to perform a specific task. In the learning process, an output vector \( y_i \) generated. This vector is the desired output value for each input vector \( x_i, i \in \{1, 2, ..., n\} \). The task of the robot is to form an I/O mapping \( \hat{y} = f(x) \), where \( \hat{y} \) is the predicted result that minimizes the error for a given training set.

Implementation of the PbD approach requires an answer to five central questions [5, 6, 7]: “Who provides the training?” “When to conduct the training?” “What to teach?” “The problem of conformity” and “How to evaluate training?” It is also required to answer the question on which of the two PbD levels the robot undergo training: “target task” or “trajectory of the manipulator” (high level or low level).

In this article, we consider the methodology of training a robot with multi-link manipulators to perform typical actions by manipulators in the framework of a given scenario.
2. Participants in the implementation of the PbD approach and teaching methods

Training of a multi-link manipulator with kinematic redundancy should be carried out by one “teacher” to preserve the identity of movements. Anthropomorphic characteristics of a “teacher” should be close to the main characteristics of a multi-link manipulator. In the case of a large amount of training, the use of two “teachers” with similar anthropomorphic characteristics is allowed. There are the questions “when and what” of teaching a robot? The answers to this question are closely related to the analysis of technological maps of various operations and the interpretation of the operation movements by the “teacher.” The issue of compliance and assessment lies in the responsibility of experts that evaluates the correctness of operations a “teacher.” Thus, when organizing the training of a multi-link manipulator, it is necessary to involve a “teacher,” which reproducing operations, and an expert, which evaluating the correctness of their implementation.

The implementation of various methods for training robots determined by the data received from the “teacher,” i.e., the type of data generation determines the type of training method. There are two approaches to training the movement of the manipulator by the operator. Joystick in learning allows controlling the position of the manipulator end-effector with the object of the operation. Moreover, there is no control over the behavior of all parts of the manipulator as a whole. Training using a copying suit corresponds to the manipulator. Training in the manipulator movement allows detailing the movement of all parts of the manipulator according to the movement of the “teacher’s” hand, performed according to the technological map of the operation. As a result, the whole dynamics of the natural movement of the human hand in six degrees of freedom will be transmitted to the multi-link manipulator. An example of such training for the SAR-400 robot is shown in Figure 2.

When applying the PbD approach, we are not talking about the optimality of the movement of the manipulator. Often, the trajectories of the manipulators are optimal in time or planned to comply with some joint spatial limitations, which can lead to paths other than human movements. In many cases, the manipulators should move according to the action of a person’s hand. It is known from clinical biomechanics that the natural movements of a person’s hand obey the law of energy-saving and, therefore, in the future, they can be considered optimal.
Thus the level of the robot training is the solution of the target task, based on the natural movement of the hand of the “teacher”.

Neurobiology data show that a sophisticated human movement, in particular, a hand movement consists of a set of primitive actions [8], which can underlie the complex movement of a manipulator. By combining primitives, any permissible can be performed, taking into account the kinematic restrictions of motion.

Splitting complex arm movements on a primitive motion is caused by application the formula [9], which describes the relationship between the speed and accuracy of the movement of the manipulator

\[ MT = a + b \log_2 \left( \frac{2A}{W} \right), \]  

where \( MT \) is the duration of the movement, \( A \) is the amplitude of the movement equal to the distance to the target at the beginning of the movement, \( W \) is the width of the object, \( a \) and \( b \) are the coefficients. According to this formula, the required time of movement of the manipulator can be calculated.

To assess the speed of movement of the manipulator, we will use the formulas [10]

\[ \Delta \dot{x_d}(t) = \left( -\frac{60}{D^3} (x_t - \hat{x}(t)) - \frac{36}{D^2} \dot{x}_d(t) - \frac{9}{D} \ddot{x}_d(t) \right) \]

\[ \Delta \dot{x}_d(t) = \ddot{x}_d(t) \]

\[ \Delta x_d(t) = \dot{x}_d(t), \]  

where \( D = t_f - t \) is the remaining time to achieve the goal by the manipulator, \( t_f \) is the final time of the operation, \( \hat{x} \) is the assessment of the current position of the calculated point of the manipulator. In this case, preliminary coordination between the kinematic limitations of the manipulator, the manipulative capabilities of the human operator’s hands, and the control of the correctness of the operation by an expert.
3. Training Input and Learning model

The success of the PbD approach depends on the proper processing of training data obtained from the copy suit.

The robot training time is limited, and therefore, it is impossible to demonstrate all possible movements for various scenarios. Besides, the data received by the manipulators resulting from training on the actions of the “teacher”, due to the apparent reasons described above, have some error $\varepsilon$

$$y = f(x) + \varepsilon.$$  \hspace{1cm} (3)

Also, the data $f(x)$, most often, do not take into account external disturbances acting on multilink (anthropomorphic) manipulators.

The frequency of the transmitted data plays an important role. A reasonable compromise must be reached here. On the one hand, it is necessary to achieve acceptable accuracy of robot manipulation. On the other hand, it is necessary to ensure a small size of the knowledge base formed according to incoming data. Small changes in the position of the manipulator can radically change the movement of the manipulator generated as a result of training. Therefore we assume that the training of the manipulator and its movement, according to the results of the training, begins from the same position as the manipulators, which we call initial.

Given that the complex arm movements are broken down into simple, we will use the direct model of training \cite{11}. A direct learning model predicts the next state of a dynamic system, taking into account the current action and current state. The idea of a direct learning model is as follows: if we can observe the current state of the system $s_k$ and the action $a_k$ currently applied to the system, we can try to predict the next state $s_{k+1}$,

$$s_{k+1} = f_f(s_k, a_k)$$  \hspace{1cm} (4)

The function expresses the causally dependent physical properties of the system, and the processing of such mappings is carried out using standard regression methods.

4. The methodology of training the control system of the robot on the actions of the “teacher”

In the formation of the training methodology for the control system of the robot according to the actions of the “teacher”, we will adhere to the necessary requirements for the methodology as a specific procedure: realism; reproducibility; compliance with the goals and objectives of the planned action, validity; performance. The proposed training methodology can be extended to another robot with a similar structure.

The developed technique is based on building a connection between the sensors of the copying suit and the corresponding manipulator engines. The mapping of the sensor readings to the required rotation angles is non-trivial and non-identical since the space of the input state is large, and the unknown objective function is non-linear. In practice, not all copy suit sensors can be used. It is also necessary to note that the same sensors give different values for “teachers” with various anthropomorphic indicators.

In the training methodology, three phases are distinguished: the planning phase, the playback phase (direct learning) and the decision-making phase. Their ratio when learning a one-time movement is shown in Fig. 3.

Planning phase. The preparatory phase, which consists of planning the execution of the movement of the “teacher” on the flow charts of the operation and evaluating the duration of the movement and the speed of the manipulator.

Playback phase. The “teacher” controls the movement of the manipulator in real-time by moving his hand in a copy suit. Data from the sensor of the copying suit that is formed during the execution of the “teacher” is used for training the control system and test the selected algorithm for the operation. As a result, the manipulator performs the appropriate movements. It is assumed that the sensors of the
copy suit are calibrated.

![Diagram of the training methodology](attachment:training_methodology_diagram.png)

**Figure 3.** The diagram of the training methodology

The data from the sensors are then processed and used as input data for creating a knowledge base for training the control system, with which the output angles of the joints of the manipulator are calculated. Data is discarded if it exceeds the boundary limits, critical values of the joints of the manipulator, or if the difference between each new calculated angle and the previous exceeds a predetermined threshold.

Decision-making phase is a phase when the expert decides (this is a subjective decision) whether the resulting movement of the manipulator was satisfactory.

The last two phases repeated a specified number of times. The number of repetitions is determined based on specific considerations. For example, based on the hypothesis of minimum dispersion [12], according to which the variance of the position of a given point of the manipulator (capture) should be minimized during the tests.

Finally, based on a generalization of the previous, we formulate a methodology for teaching the control system of the robot.

**Planning phase:**
- Select the type of operation to be performed.
- Dividing the movement into primitives based on the calculated duration of the movement of the manipulator.
- The calculation of the speed of movement of the manipulator for each of the primitive movements.

**Playback phase and decision-making phase:**
- Setting the TCP interface between the copy suit (joystick) and the control system.
- Setting the interface between the control system and the robot manipulators.
- Transferring manipulators to the starting position.
- With the help of a copying suit (joystick), the “teacher” perform movement under the supervision of an expert over the movement correct execution.

**Assessment of movement by an expert:**
- In the case of a successful movement, the data is saved to the knowledge base.
- Analysis of the number of completed movements.
- Rationing knowledge base data.
- Performing test movements by manipulators according to the results of training.
– If the movements are satisfactory, then the preparation of the knowledge base data for the construction of control laws.

Final phase:
– Construction of the laws of manipulator control.
– Performing test actions with manipulators on the simulator.
– Performing test actions with manipulators.

Since the knowledge of the robot is not static, it is necessary to maintain regular updates to the knowledge base.

5. Conclusion
The article developed a methodology for teaching the robot control system on the actions of the "teacher". The article substantiates the application of the “Programming by Demonstration” approach as the main one for teaching the robot control system according to the actions of the “teacher”. The advantages of using this approach are indicated in comparison with traditional ones based on the construction of mathematical models of the system in the form of systems of differential and finite-difference equations, as well as based on modelling the kinematics of the system, for example, the Denavit-Hartenberg coordinate transformation method. The article describes the choice of participants in the learning process of the robot and also defines the method of training - training using a copy suit. The requirements for the training input data are defined. Based on the previous, the methodology of teaching the robot control system on the actions of the "teacher" is finally formulated.

6. Acknowledgements
Studies were supported by the Ministry of Science and Higher Education of the Russian Federation (agreement № 14.578.21.0264, project RFMEFI57818X0264).

References
[1]. Billard A, Calinon S, Dillmann R, Schaal S 2008 Springer-Verlag Berlin Heidelberg 1591
[2]. SkoglundA, 2009 Orebro University 486
[3]. AlchakovV, Kramar V, LarionenkoA, 2020IOP Conf. Series: Materials Science and Engineering709 022092
[4]. Stanley J, Elliott M. 2018Springer 689
[5]. Dautenhahnand K Nehaniv C 2002 Cambridge, MA: MIT Press
[6]. Münch S., Kreuziger J., Kaiser M., Dillmann R. 1994 Proceedings of the 24th International Symposium on Industrial Robots
[7]. Zuyuan Zhu, Huosheng Hu 2018Robotics, 7(17), 1-25
[8]. BizziE, Mussa-IvaldiF A, and Giszter S1991 Science253(5017) :287–291
[9]. Paul M. Pitts 1954 Journal of Experimental Psychology47(6) 381-391
[10].Shadmehr Reza, Wise Steven P. 2005National Institute of Mental Health in Bethesda Maryland, MIT Press 575
[11].Siciliano B, Khatib O 2016Springer-Verlag Berlin Heidelberg - 2227
[12].Harris CM Wolpert DM 1998Nature20;394(6695)780-4