

Manipulative Tasks Identification by Learning and Generalizing Hand Motions

Diego R. Faria¹, Ricardo Martins¹, Jorge Lobo¹ and Jorge Dias¹,

¹ Institute of Systems and Robotics, DEEC, University of Coimbra, Portugal
{diego, rmartins, jlobo, jorge}@isr.uc.pt

Abstract. In this work, an approach for extracting features among multiple observations towards manipulation tasks recognition is proposed. The diversity of information such as hand motion, fingers flexure and object trajectory are important to represent a manipulation task. By using the relevant features we can generate a general form to represent a specific dataset of manipulation tasks. The hand motion generalization process is obtained and later, given a new observation, the task can be identified.

Keywords: Motion Patterns, Task Recognition, Task Generalization

1 Introduction

An important issue for modeling and recognition of human actions and behaviors are the motion patterns found during some activity. In different daily tasks the motion assumes an important key point to describe a specific action. The variety of human activity in everyday environment is very diverse; the same way that repeated performances of the same activity by the same subject can vary, similar activities performed by different individuals are also slightly different. The basic idea behind this is: if a particular motion pattern appears many times in long-term observation, this pattern must be meaningful to a user or to a task. In this work we are focusing on manipulative tasks at trajectory level to find similarities (significant patterns) given by multiple observations. The intention is to learn and to generalize a specific task by the hand movement including fingers motion, as well as object trajectory along the task for its recognition. This application is useful for task recognition in robot imitation learning. Although imitation learning is not the focus of this work, still the ability of learning and generalization are addressed.

2 Contribution to Sustainability

During the recent years, many research fields such as human-computer interface, medical rehabilitation, robotics, surveillance, and also sport performance analysis have focused some of their attention to the understanding and analysis of human behaviour

and human motions. Others examples can be seen in the field of entertainment such as games that use natural user interfaces where sensors grab the human motion to interact with the game. All these research fields are searching for new solutions to improve the standard quality levels of human living conditions, generalizing the access to high quality services. Motion pattern analysis is one of the key elements to the development of those services.

The contribution of this work is an approach to learn relevant features in manipulative tasks by finding similarities among hand motions where it is possible to generalize movements to be applied into different contexts. This kind of approach can be used in the future to endow robots by imitation learning, as well as for recognition of a specific action to interact in a human environment to assist people in different tasks, compensating the absence of specialized human resources. Robotics can be used in medicine to assist in surgeries, rehabilitation, and also for complex task which is dangerous for human beings. These applications can contribute to the technological innovation for sustainability.

3 Related Work

In [2], a programming by demonstration framework where relevant features of a given task are learned and then generalized for different contexts is presented. Human demonstrator teaches manipulative tasks for a humanoid robot. Through GMM/BMM the signals are encoded to provide a spatio-temporal correlation. The trajectories are then generalized by using GMR. In [3], an approach to find repeated motion patterns in long motion sequences is proposed. The authors state that if a point at a given instant of time, belongs to a set of repeated patterns, and then many similar shaped segments exist around that data point. They encode the characteristic point with partly locality sensitive hashing and find the repeated patterns using dynamic programming. In [4], a general approach to learn motor skills form human demonstrations is proposed. The authors have developed a library of movements by labeling each recorded movement according to task and context. By using Non-Linear differential equations they could learn and generalize the movements.

4 Proposed Approach

Inside the neuroscience field we can find a decomposition of a typical human manipulation movement on different stages [1]. Action phases are defined as manipulative activities involving series of primitives and events. In this work, the action phases are used to find motion patterns in each one. In Fig.1 is possible to identify the actions phases and the events that happens among them. In each phase, primitives may be detected to better describe an action. Those represented action phases are a high level segmentation of simple manipulative tasks. For the tasks that need to have in-hand manipulation (re-grasp or change the orientation of the object along the movement), the segmentation becomes more complex, the transport phase

can enclose or can be changed to in-hand manipulation movements. Fig.2 shows the steps of the proposed approach.

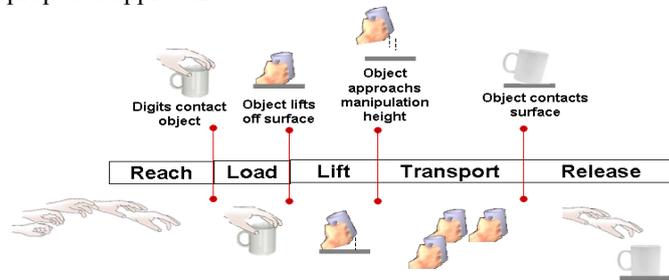


Fig. 1. Defined action phases for the manipulative tasks presented in this work.

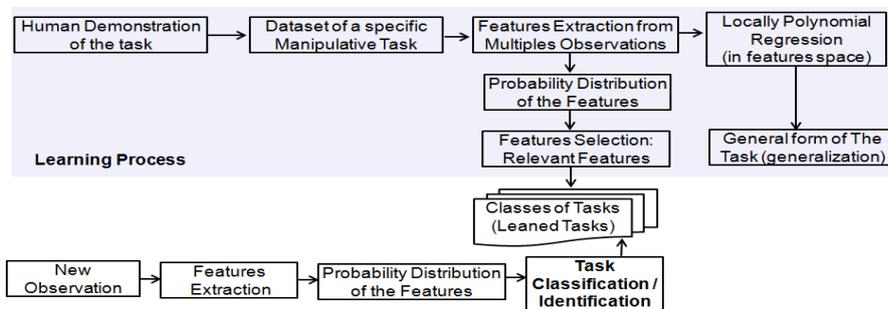
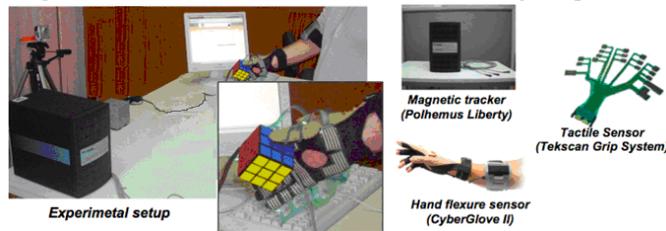


Fig.2. Steps of the proposed approach.

4.1 Experimental Setup and Data Acquisition

The data acquisition is performed using: Polhemus magnetic motion tracking system [5]; TekScan grip [6] a tactile sensor for force feedback and CyberGlove II [7] for fingers flexure measurement. Each magnetic sensor has 6DoF (3D position and Euler angles). The magnetic sensors were attached to the fingertips to track the hand and fingers movements and also to track the object pose. The setup (Fig.3) for the experiments



is composed of a wooden table, and the experiments were executed by a subject seated in front of the table.

Fig. 3. Experimental Setup: Devices for the experiments.

To facilitate the detection of each action phase by analyzing the sensors data, the sensors synchronization was needed. A distributed data acquisition was adopted, where a trigger defines the start and end of the acquisition. Thus, by looking to the

data it is possible to identify some events, thus allowing detecting the beginning and end of an action phase in a movement. Some assumptions as described in previous work [8] are adopted, where some heuristic rules need to be satisfied, for example, reaching: it happens while there is hand motion, the object is static, no force pressure from tactile sensing, and variance on the fingers flexure; load phase: hand motion, force feedback, no movement of the object.

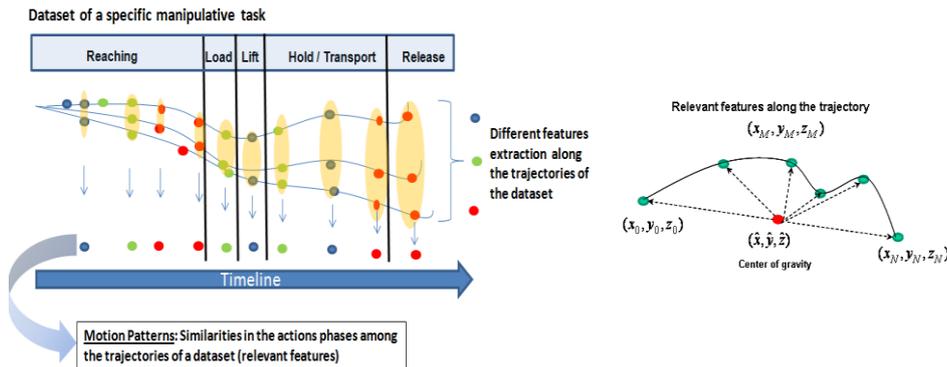


Fig.4. – (a) Motion Patterns: Similarities detection in the action phases of the trajectories of a dataset of a manipulative task; (b) Distance of the learned features from the center of gravity.

4.2 Motion Patterns: Features Extraction and Similarities among Trajectories

An example of features selection is presented in Fig.4 (a). Given a dataset of hand trajectories of a task, we want to find the similarities among all trajectories, repeated patterns that are relevant features to generate a generalized one. The idea is to detect features in each action phase of all trajectories of a dataset, and then the probability distribution of these features is computed. The algorithm selects the type of feature with higher occurrence in all trajectories by looking for each feature coordinates. At the end, the features with high similarities among the trajectories are kept and by applying an interpolation among the features positions we can have a general trajectory. The classes of features that are used to describe a trajectory are curvatures and hand orientation that vary during the task performance. In previous work [9], a probabilistic approach for hand trajectory classification where curvatures and hand orientation where detected in 3D space was developed. Here we are following the same idea for feature extraction.

4.3 Patterns from Different Sensors Modality

Other type of features is also taken into account: the fingers distances (thumb to index finger; index to middle finger and so on). The distances variance happens alongside the hand trajectory, examples include hand aperture, grasping, etc., helping also to differentiate each phase of a task. At each point of the finger trajectories a mean

distance of the sum of squared Euclidean distances of the fingers is computed (1). Inside of each action phase there are many 3D points so that it is possible to compute N distances. An alternative to represent each action phase is to compute the average of all computed mean distances (1).

$$D_{avg} = \frac{1}{N} \sum_{i=1}^N D_i \quad (1)$$

where

$$D = \sum_{k=1}^N (d_k)^2 \quad (2)$$

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} . \quad (3)$$

The signals of the cyber-glove represent degrees of flexure (e.g. 0-255). An easy way to use this information is by defining some grasp types such as cylindrical grasping, spherical grasping, in different levels of flexure and also defining some type of gestures, i.e. extension and flexion of the hand in different degrees of flexure (e.g. low, medium, high). After a learning stage by analyzing many observations of the same gesture it is possible to know the degree of flexure for each finger for each grasp type. In each task, it is necessary to identify the types of the defined grasping/gesture and then computing the probability distribution $P(Grasp | Observation)$ of each one alongside the action phases of the task.

4.4 Task Representation by General Form of the Trajectories

The representation of a dataset of a specific task at trajectory level is given by the general form of the data which is achieved after selecting the relevant features and then applying a regression on the data to generalize it. The spatio-temporal information is used to apply a polynomial regression to fit the data to have a smoothed trajectory of the manipulative task. The polynomial regression was chosen due to the curvilinear response during the fit and it can be adjusted because it is a special case of multiple linear regressions model. We are adopting the quadratic form of the model, a polynomial regression of second order. Although polynomial regression fits a nonlinear model to the data, as a statistical estimation problem, it is linear, in the sense that the regression function is linear in the unknown parameters that are estimated from the data. The general model of second order polynomial regression is given by:

$$Y_i = \beta_0 + \beta_1 x_i + \beta_{11} x_i^2 + \varepsilon_i, \quad (4)$$

where $x_i = X_i - \bar{X}$ and ε is an unobserved random error with mean zero conditioned on a scalar variable; ε can be computed as error of least square fitting; β minimizes the least square error.

In our case, due to type of trajectories used in this work, to correctly fit the curves, the regression needed to be done locally, at subdivisions of the trajectory. Example of regression in sub-regions of the trajectories is shown in Fig.5.

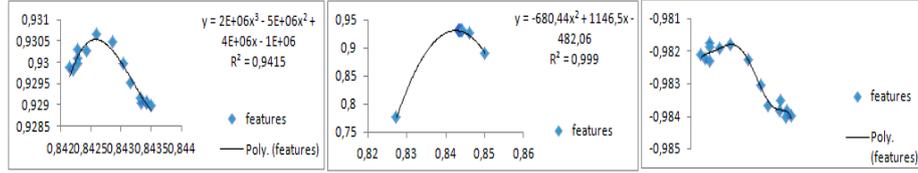


Fig.5. – Regression applied on sub-regions of an action phase of a manipulative task. 2D view: left and middle images: x, y view; right image: x, z .

4.5 Task Identification

A new task can be recognized by matching a prototype (learned task) to a new observation or via classification where there are multi-classes. In the case of the hand trajectory, some properties of the learned features (translation invariance) as shown in Fig.4 (b) can also be used in the matching. The learned features are used for the matching between a prototype (generalized information) and a new observation.

Another alternative is applying continuous classification based on multiplicative updates of beliefs by Bayesian technique (5) taking into consideration the learned observations (relevant features of the general form of signals). First, it is identified if the task to be classified has the same action phases of the learned tasks to classify it.

$$P(G_k | c_k, i) = \frac{P(c_k | G, i) P(c_obj_k | G, i) P(o_k | G, i) P(h_k | G, i) P(G)}{\sum_j P(c_k | G, i) P(c_obj_k | G, i) P(o_k | G, i) P(h_k | G, i) P(G)} \quad (5)$$

To understand the general classification model, some definitions are done: g is a known task goal from all possible G ; c is a certain value of feature C (Curvature types) found in the hand trajectories; C_obj : curvatures found in the object trajectories; H : the grasping type learned from the data-glove signals; o : a certain value of feature O (hand orientation types); i is a given index from all possible action phases A . For more details about a methodology for some features extraction and their probability distribution see [9]. The probability $P(c | g, i)$ that a feature C has certain value c can be defined by learning the probability distribution $P(C | G, A)$; $P(o | g, i)$ of feature O learning $P(O | G, A)$; $P(h | g, i)$ learning $P(H | G, A)$ and $P(c_obj | g, i)$ of feature C_obj learning $P(C_obj | G, A)$. These likelihoods are obtained by histogram techniques.

5 Results

The trajectories that were used to evaluate the proposed approach are pick-up and place and pick-up and lift. Fig.6 (a) shows the raw data of the dataset correspondent to the task pick-up and place (hand trajectories); (b) shows the detected action phases using the sensors information. Fig.7 (a) shows an example of the 3D positions of the features extracted (curvatures: trajectory directions) from all observations before finding similarities; (b) relevant features selection by analyzing the probability distribution of the features to know which type of feature is more relevant, later by

computing the least square among all features points of the trajectories dataset we can estimate the coordinates of them; (c) example of interpolation of the features points as a function of arc length along a space curve.

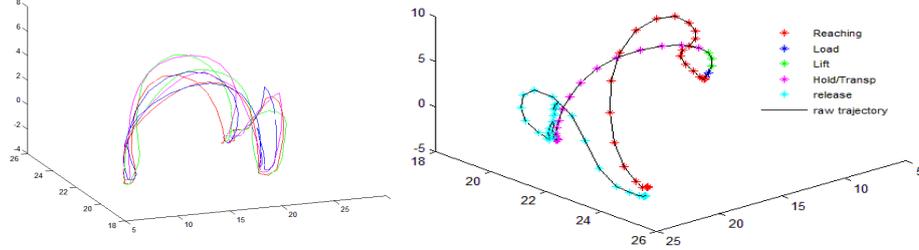


Fig.6. – Left: Raw data(in inches): trajectories dataset (object displacement); Right: Trajectory segmentation by action phase.

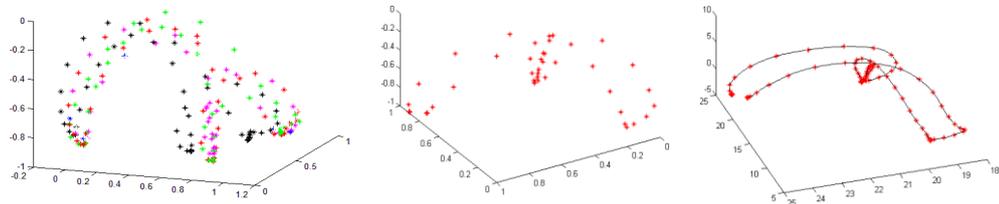


Fig.7. – (a) Extracted Features; (b) Relevant Features (similarities among all trajectories); (c) Generalized Trajectory

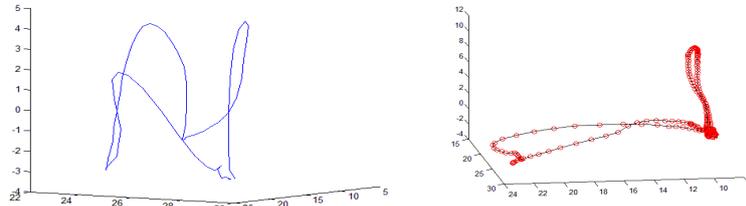


Fig.8. – (a) New observation: trajectory to be classified (pick-up and place); (b) Trajectory of dataset pick-up and lift.

TABLE I
CLASSIFICATION RESULT

Action Phases	Pick-up and place %	Pick-up and lift %
Reaching	45.00	55.00
Load	48.10	51.90
Lift	59.32	40.68
Transport	69.83	30.17
Release	78.00	22.00

The 2nd and 3rd columns show the % of the new observation belonging to pick-up and place or pick-up and lift task in each instant (part of the task). We have detected the relevant features in each phase using their probabilities to classify the new observation. The new trajectory (Fig. 8-a) is classified as pick-up and place correctly. We can see that the new observation is similar to the learned tasks during the reaching until lift phase, after that, the new observation start to differ from pick-up and lift task and it is correctly classified as pick-up and place with 78%.

The action phases for both dataset happen in different period. Given a new trajectory we want to recognize the task class. The classification variables are updated in each action phase and at the end the variables keep the final result of the classification. Tab.1 shows the result of the classification of a new observation of pick-up and place. For that, just the learned curvature features in the classification model (5) were used; in future work we intend to implement the complete model to reach better results.

6 Conclusion and Future Work

We have proposed an approach to represent and recognize a manipulative task by multiple observations performed by a human demonstrator. By finding the relevant features of a dataset, we can obtain a general form to represent a task, allowing also its identification. The preliminary results motivate us to follow this proposed methodology since we could obtain valid classifications. In future work the proposed approach will be extended by using multimodal data to test and evaluate different manipulative tasks.

Acknowledgments. This work is partially supported by the European project HANDLE within the 7^o framework FP7. Diego Faria and Ricardo Martins are supported by the Portuguese Foundation for Science and Technology (FCT).

References

1. R. S. Johansson and J. R. Flanagan, "Coding and use of tactile signals from the fingertips in object manipulation tasks," *Nat Rev Neurosci*, vol. 10, pp. 345–359, 2009.
2. S. Calinon, F. Guenter A. Billard, "On Learning, Representing and Generalizing a Task in a Humanoid Robot", *IEEE Transactions on Systems, Man and Cybernetics, Part B*, Special issue on robot learning by observation, 37:2, 286-298.
3. K. Ogawara, Y. Tanabe, R. Kurazume, and T. Hasegawa, "Detecting repeated motion patterns via dynamic programming using motion density", in *Proc. 2009 IEEE Int. Conf. on Robotics and Automation (ICRA'09)*, pp. 1743–1749.
4. P. Pastor, H. Hoffmann, T. Asfour, S. Schaal, "Learning and generalization of motor skills by learning from demonstration", *Int. Conference on Robotics and Automation ICRA'09*.
5. Polhemus Liberty 240/8 Motion Tracking System: http://www.polhemus.com/?page=Motion_Liberty.
6. TekScan Grip Sensor: <http://www.tekscan.com/medical/system-grip.html>.
7. CyberGlove II: <http://www.cyberglovesystems.com>.
8. D. R. Faria, R. Martins, J. Dias. "Learning Motion Patterns from Multiple Observations along the Actions Phases of Manipulative Tasks". *Workshop on Grasping Planning and Task Learning by Imitation: IEEE/RSJ IROS'2010*, - Taipei, Taiwan – Oct. 2010.
9. D. R. Faria and J. Dias. "3D Hand Trajectory Segmentation by Curvatures and Hand Orientation for Classification through a Probabilistic Approach", In *Proceedings of The IEEE/RSJ Int.Conf. on Intelligent Robots and Systems, IROS'09*, St. Louis, MO, USA.