2010 IEEE/RSJ International Conference on Intelligent Robots and Systems
Workshop on Grasp Planning and Task Learning by Imitation

Symbolic level generalization of in-hand manipulation tasks from human demonstrations using tactile data information

Ricardo Martins, Diego R. Faria and Jorge Dias

Abstract— **This work intends to contribute to the development of autonomous dexterous robotic hands by presenting an approach to describe the mechanisms underlying the human strategies during the execution of in-hand manipulation tasks. The work proposes a symbolic decription of the in-hand manipulation tasks. The in-hand manipulation tasks are demonstrated by a subject wearing an instrumented glove with a tactile sensing array on the palm and fingers region. The description of the manipulation movement consists on a sequence of hand contact states primitives with the object. The set of possible contact state primitives is defined previously to the demonstration.**

I. INTRODUCTION

Nowadays, different types of robotic platforms are starting to be massively introduced to new environments (domestic, healthcare, entertainment, education) where the robots have to deal with different challengs such as the ability to interact with persons and with objects on the environment. These new environments are dynamic, unpredictable and can not be completely known in advance, what has led to the development of mobile robotic platforms with multimodal systems: active vision systems, audition, multi-articulated arms and dexterous robotic hands. These multimodal modules provide a framework to develop artifical perception systems to autonomously deal with the dynamic of the environments, wide variety of objects and interact safely with humans.

One of the key elements of the performance of the robotic platforms is the ability to perform autonomous grasping, manipulation, explorationa and characterization of not completely known objects. To achieve these objectives there is a tendency on the field of robotic research to move the development of robotic hands from simple grippers towards human inspired articulated hands (mechanical stucture and number of degrees-of-freedom) and introducing on the robotic hand sensing devices such as tactile, temperature, force/torque sensors. The introduction of these new generation robotic hands places new challenges concerning the motion (fingers, palm, coordination fingers-fingers, fingers-palm) of the robotic hand, with such number of degrees-of-freedom, that will be executed to perform the intended task.

In order to provide robotic hands with human-like capabilities for handling objects and interact with the environment, one possible aaproach is to encode and learn those manipulation strategies from human demonstrations. The current approaches proposed in the literature, described on the next section, focus on developing algorithms that create generic representations of several classes of tasks. The generalization process consists on the extraction of the task essential components (constrains, regularities) across multiple demonstrations.

The neuroscience literature [1] proposes a decomposition of a typical human manipulation movement on different stages: reach, load, lift, hold, replace and unload. On in-hand manipulation tasks, the static hold phase is replaced by a segment designed by in-hand manipulation. On this segment, depending on the type of in-hand manipulation performed [2], it's possible to identify several movements that involve the internal consecutive regrasping and release of the object to perform its reorientation, fine positioning or a more complex interaction such as sequential rotation of the object. On this work we focus our attention on this segment.

II. RELATED WORK

Several approaches to solve the motion learning problem on different contexts and applications fields, such as human behavior modelling, learning by imitation are described on the literature. A simple approach was presented by Delson [3]. The authors simple make a statistical analysis of human demonstrations of a pick and place task and define the range of cartesian trajectories that can be performed to achieve that task. Tso [4] applies Hidden Markov Moddels to encode a trainning dataset built from a set of human demonstrations. Given a human demonstration as input, the system reproduces the trajectory of the trainning dataset with the highest likelihood. Calinon [5] proposes to extract continuous constrainsts from a set of demonstrations performed using different initial positions of the object. The cartesian trajectories of these demonstrations are projected using Principal Component Analysis and then the constraints are represented through Gaussian Mixture Models. To reproduce the task, the constraints are reprojected on the original data space and the generalized version of the cartesian trajectory is found by estimating the trajectory that satisfies all the constraints. The approches described previously propose the learning and encoding of movements at the trajectory level.

Other approches propose a symbolic learning and encoding of movements based on the supervised labelling and segmentation of the primitives during the learning stage.Kondo [6] proposes a method to describe in-hand manipulation demonstration movements by recognizing a sequence of contact state transitions between the human hand an the

This work is partially supported by the European project: HANDLE, FP7-231640. Diego R. Faria and Ricardo Martins are supported by Portuguese Foundation for Science and Technology (FCT). Ricardo Martins, Diego R. Faria and Jorge Dias are with Institute of Systems and Robotics, Department of Electrical Engineering and Computers, Faculty of Science and Technology, University of Coimbra Polo II, 3030-290 Coimbra, Portugal {rmartins, diego, jorge}@isr.uc.pt

manipulated object. The recognition algorithm is based on a Dynamic Programming approach by comparing the similarity of the contact state transition between an input sequence and template manipulation primitives. Bernardin [7] describes a technique to recognize continuous human grasping sequences using Hidden Markov Models. Twelve different grasp primitives are recnognized using combining data from hand palm tactile sensors and hand joints flexure levels from a data glove.

The work by Krugger [8] presents the automatic extraction of action primitives (without the necessity of pre-segmentation and manual labeling) and the corresponding grammar from continuous movements of several human demonstrations of grasping tasks. This approach considers that all the actions can be described by a set of elementary build blocks (action primitives) and there are a set of rules (grammar) that define how these actions primitives can be combined. The action primitives are represented by an extension of Hidden Markov Models, parametric Hidden Markov Models. The extraction of the motion primitives from the movements also take into account the changes in object state.

III. Approach overview

The approach proposed on this work follows the principle presented and described by some previous works referred previously that a human movement can be decomposed in a sequence of elementary primitives. Depending on the phase (reach, load, lift, hold, replace, unload, in-hand manipulation) of a typical manipulation movement, different types of signals (position/orientation of the fingers distal phalanges and wrist, joints flexure level, tactile sensing) dynamically change their role and importance on the control of object the object manipulation strategies of humans. On this particular context, the in-hand manipulation of objects, the contact signatures between the object and the different regions of the hand surface, as well as the configuration of the human hand joints flexure level are importante factors on the definition and characterization of those strategies. On this work, we focus our attention on the tactile signatures of some in-hand manipulation primitives. Through the temporal combination of those elementary primitives, it is possible to characterize the strategies used by humans to perform precise and complex movements during a in-hand manipulation task.

This work intend to implement a system to extract the correspondent generalized primitive sequence, based on tactile signatures, from human in-hand manipulation movements demonstration of a specific task. Figure 1 shows an overview of the global structure of the proposed system.

The human demonstrator performs an in-hand manipulaton task using an instrumented data glove with embeded hand joints flexure sensors and equiped with tactile sensors distributed on the hand palm and fingers surface region. During the execution of the task a sequence of the elementary primitives among the set of pre-defined primitives is detected. The primitives are individually detected. The coherence of the sequence of the primitives is verified by *task contraints*

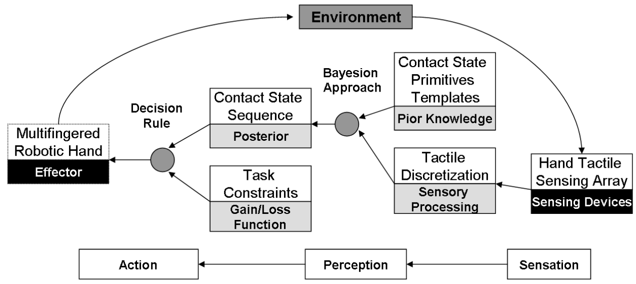

Fig. 1. Overview of the global structure of the proposed system

module. This module identifies the task class of the demonstration and guarantees that the output primitives sequence respects the general constraints of that class of tasks. This verification can be seen as a coherence verification between the detected raw sequence of primitives in order to avoid the introduction of the descontextualized primitives on that sequence due to unprecise movements performed during the demonstration or errors introduced during the primitives detection stage. The constraints referred previously are extracted from human demonstrations of different classes of tasks, performed by different subjects. The diversity of the demonstrations promotes the exploration of the variability of the strategies used by humans to perform the same task. The essential components of those strategies will emerge as the permanent elements. Then, it is possible to build the temporal and functional relations and rules between those elements in order to find a compact representation of those strategies.

The action component module (hand fingers path and motion planning) of Figure 1 it is not discussed on this work. The average joint angles flexure level of the different configuration of the hand pose primitives is stored for future developments.

IV. Contact state template definition and detection

On this section the general model of the framework used to describe the different templates is defined, as well as, the set of pre-defined primitives used during the primitive detection stage. The methods implemented on that stage are also presented.

A. Primitives model definition

The general model of the framework used to describe the diferent templates of primitives is based on the output of the data acquisition devices used in the human demonstrations: the tactile sensing device - *Tekscan Grip System*[9] (Tekscan, Inc.; Boston, MA). The tactile sensing device consists of 360 sensing elements that are destributed along the hand palm and fingers surface. The output of each of the tactile sensing element is an 8-bit integer. As the tasks on these experiments are performed by the same subject on consecutive trials of a single session, the tactile sensing device was not calibrated. A method proposed by the tactile sensing device manufacturer called equilibration was applied to the device in

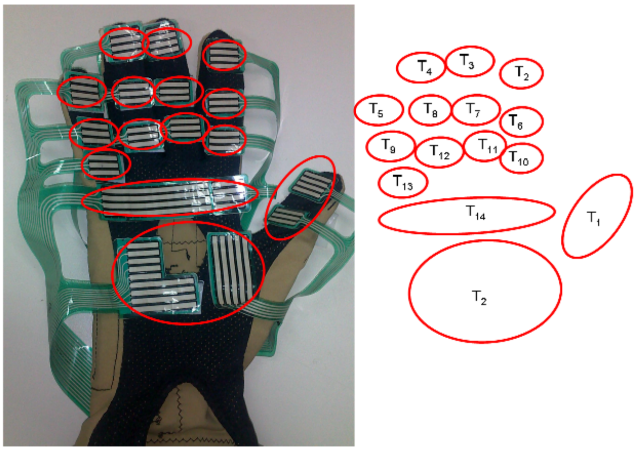

Fig. 2. Representation of the 15 tactile sensing regions defined from the 360 sensing elements of the tactile sensing device.

order to reduce/eliminate the variation between the individual sensing elements.

The output of the 360 tactile sensing elements are grouped on 15 regions as presented on the figure 2. Each region corresponds to different areas of the hand.

A variable T_i is assigned to each of this regions:

$$\mathbf{T} = \{T_1, T_2, ..., T_{15}\}$$

The domain for each of those variable can be defined as :

$$T_i \in \{NotActive, LowActive, HighActive\}$$

where the *NotActive*, *LowActive*,*HighActive* define the level of activation of that region during the in-hand manipulation task. The *NotActive* state of a variable T_i corresponds to an average output of the sensing elements corresponding to an output that is between 0 and 10. The *LowActive* corresponds to an output that is between 26-190 and to the *HighActive* state between 190-255.

The general model of the framework used to describe the different templates of primitives can be defined by the set of variables **T**.

B. Primitives set definition

The set of pre-defined templates comprises a total of seven templates. The contact state templates primitives are estimated from different seven grasp configurations. Six of those seven grasp configurations are demonstrated on the next figure. The remaining one correspond to a situation where there is no contact between the hand and the object (*Primitive7*).

The variable E designes a primitive. The domain definition of E is

$$E \in \{Primitive1, Primitive2, ..., Primitive7\}$$

In order to estimate the parameters of the templates parameters **T** of each of the 7 pre-defined defined primitives, several human demonstrations of the different static

Fig. 3. Representation of six of the pre-defined grasp configurations used to estimate the corresponding static pcontact state templates.

contact configurations of the human hand and the object. The templates are segmented on each demonstration and the probability distribution P(**T**/E) is built.

C. Primitives detection on raw data input

In order to proceed to the dectection of the pre-defined primitives, the raw data input produced during the in-hand manipulation demonstration is integrated during equal time intervals. The integrated data during each time slot $\mathbf{T}_t$ is classified according with the following expression and it is assigned a template label E_t to that period of time. The template with maximum likelihood is the template assigned to that timeslot.

$$P_t(E/\mathbf{T}) = \frac{P_t(\mathbf{T}/E)P(E)}{P_t(\mathbf{T})}$$

$P_t(\mathbf{T}/E)$ is achieved from the primitives demonstration trainning session. $P(E)$ is the probability of a template (1/7) and $P_t(\mathbf{T})$ is the probability of a model measurement.

The previous expression can be rewritten as follows.

$$P_t(E = primitive_i/\mathbf{T} = (t_1, ..., t_{15})) =$$

$$\frac{P_t(\mathbf{T} = (t_1, ..., t_{15})/E = primitive_i)P(E = primitive_i)}{\sum_{j=1}^{7} P_t(\mathbf{T} = (t_1, ..., t_{15})/E = primitive_j)P(E = primitive_j)}$$

The output of the primitives detection stage is a raw temporal sequence of the templates corresponding to the pre-defined primitives.

V. Experimental results

A. Experimental setup

During the human demonstrations of the in-hand manipulation tasks, the subject wears on the right hand an instrumented glove (*Cyberglove II*) with a tactile sensing array (*Tekscan Grip System*) attached to the palm and fingers surface region. The objects that are placed on the top of a table are manipulated only with one hand (right hand). The subject is seated during the in-hand manipulation tasks demonstrations. The data from the tactile sensing array is sampled at 500 Hz. The connfiguration of the tactile sensing array, as well as the typical configuration of the experimental area during the task demonstration are shown on figure 4.

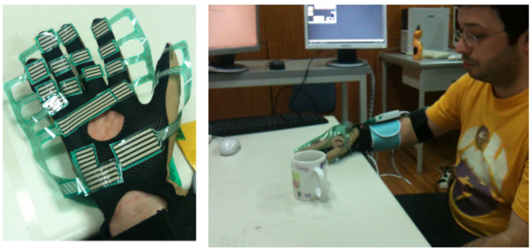

Fig. 4. On the left: Configuration of the tactile sensing array on the hand surface. On the Right: Configuration of the experimental area.

TABLE I

Template1 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	H	H	H	H	L
T_2	H	H	H	H	L
T_3	N	N	N	N	N
T_4	N	N	N	N	N
T_5	N	N	N	N	N
T_6	N	N	N	N	N
T_7	N	N	N	N	N
T_8	N	N	N	N	N
T_9	N	N	N	N	N
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	N	N	N	N	N

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

B. Primitives set trainning

In order to estimate the parameters of $\mathbf{T}$ for each of the seven pre-defined contact state template primitives, five demonstrations of each grasp configuration presented previously were performed by a subject. The trainning result for each of the primitives templates is shown on Tables I, II, III, IV, V, VI and VII respectively. The conditional probability density distribution functions of $\mathbf{T}$ for each contact state template, extracted from the trainning results, is shown on tables VIII, IX, X, XI, XII, XIII, and XIV respectively.

C. Primitives detection on raw data input

In order to test the primitives detection approach of the pre-defined contact state templates, two tasks were defined. On both tasks the manipulated object is a mug and the starting configuration (position and relative orientation to the subject) is the same. *Task I* consists on the reorientation of the mug in order to positionate the grasp of the mug in a configuration suitable to be grasped by handle by the subject right hand. *Task II* consists on grasping the mug without reorientation and elevate it. Each of the tasks was performed 2 times. The demonstrated tasks are segmented on blocks of 250 tactile sensing samples (500 miliseconds).The detection of the primitives is made using the average value of each of the $\mathbf{T}$ tactile inputs. The results for the primitives detection on raw data inputs for *Task I* are shown on Tables XV and XVI.

TABLE II

Template2 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	N	N	N	N	N
T_2	L	L	L	L	L
T_3	H	H	H	H	H
T_4	H	H	H	H	H
T_5	L	L	L	N	N
T_6	L	L	N	N	N
T_7	H	H	L	L	L
T_8	H	H	L	L	L
T_9	L	L	N	N	N
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	H	H	H	H	H

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE III

Template3 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	H	H	H	H	H
T_2	H	H	H	H	H
T_3	H	H	H	H	H
T_4	H	H	H	H	H
T_5	L	L	L	N	N
T_6	H	H	H	H	H
T_7	H	H	H	H	H
T_8	L	L	L	L	L
T_9	L	L	N	N	N
T_{10}	H	H	H	H	H
T_{11}	L	L	L	L	L
T_{12}	L	L	L	L	L
T_{13}	N	N	N	N	N
T_{14}	L	L	L	L	L
T_{15}	N	N	N	N	N

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE IV

Template4 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	H	H	H	H	H
T_2	N	N	N	N	N
T_3	H	H	H	H	H
T_4	N	N	N	N	N
T_5	N	N	N	N	N
T_6	N	N	N	N	N
T_7	L	L	N	N	N
T_8	N	N	N	N	N
T_9	N	N	N	N	N
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	N	N	N	N	N

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE V

Template5 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	H	H	H	H	H
T_2	H	H	H	H	H
T_3	H	H	H	L	L
T_4	H	L	L	N	N
T_5	H	L	L	N	N
T_6	N	N	N	N	N
T_7	N	N	N	N	N
T_8	N	N	N	N	N
T_9	N	N	N	N	N
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	N	N	N	N	N

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE VI

Template6 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	L	L	L	L	L
T_2	L	L	L	L	L
T_3	L	L	L	L	L
T_4	L	L	L	L	L
T_5	L	L	N	N	N
T_6	H	H	H	H	H
T_7	H	H	H	H	H
T_8	H	H	H	H	H
T_9	L	L	L	L	L
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	H	H	H	H	H

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE VII

Template7 TRAINNING RESULT

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
T_1	N	N	N	N	N
T_2	N	N	N	N	N
T_3	N	N	N	N	N
T_4	N	N	N	N	N
T_5	N	N	N	N	N
T_6	N	N	N	N	N
T_7	N	N	N	N	N
T_8	N	N	N	N	N
T_9	N	N	N	N	N
T_{10}	N	N	N	N	N
T_{11}	N	N	N	N	N
T_{12}	N	N	N	N	N
T_{13}	N	N	N	N	N
T_{14}	N	N	N	N	N
T_{15}	N	N	N	N	N

Legend: N-*NotActive*; L-*LowActive*; H-*HighActive*;

TABLE VIII

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive1*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive1})$
(H,H,N,N,N,N,N, N,N,N,N,N,N,N,N)	4/5
(L,L,N,N,N,N,N, N,N,N,N,N,N,N,N)	1/5
other	0

TABLE IX

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive2*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive2})$
(N,L,H,H,L,L,H, H,L,N,N,N,N,N,H)	2/5
(N,L,H,H,L,N,L, L,N,N,N,N,N,N,H)	1/5
(N,L,H,H,N,N,L, L,N,N,N,N,N,N,H)	2/5
other	0

TABLE X

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive3*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive3})$
(H,H,H,H,L,H,H, L,L,H,L,L,N,L,N)	3/5
(H,H,H,H,N,H,H, L,N,H,L,L,N,L,N)	2/5
other	0

TABLE XI

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive4*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive4})$
(H,N,H,N,N,N,L, N,N,N,N,N,N,N,N)	2/5
(H,N,H,N,N,N,N, N,N,N,N,N,N,N,N)	3/5
other	0

TABLE XII

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive5*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive5})$
(H,H,H,H,H,N,N, N,N,N,N,N,N,N,N)	1/5
(H,H,H,L,L,N,N, N,N,N,N,N,N,N,N)	2/5
(H,H,L,N,N,N,N, N,N,N,N,N,N,N,N)	2/5
other	0

TABLE XIII

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive6*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive6})$
(L,L,L,L,L,H,H, H,L,N,N,N,N,N,H)	2/5
(L,L,L,L,N,H,H, H,L,N,N,N,N,N,H)	3/5
other	0

TABLE XIV

CONDITIONAL PROBABILITY DENSITY DISTRIBUTION - *Primitive7*

$\mathbf{t}_k$	$P(\mathbf{T} = \mathbf{t}_k / E = \textit{primitive7})$
(N,N,N,N,N,N,N, N,N,N,N,N,N,N,N)	1
other	0

TABLE XV
PRIMITIVES DETECTION ON RAW DATA INPUT - *Task I-Trial 01*

Timeslot(ms)	Trial 01	Estimation
0-500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
500-1000	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
1000-1500	(H,H,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive1*
1500-2000	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
2000-2500	(H,H,H,L,L,N,N,N,N,N,N,N,N,N,N)	*primitive5*
2500-3000	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
3000-3500	(H,H,L,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive5*
3500-4000	(H,H,H,L,L,N,N,N,N,N,N,N,N,N,N)	*primitive5*
4000-4500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
4500-5000	(H,H,L,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive5*
5000-5250	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*

TABLE XVI
PRIMITIVES DETECTION ON RAW DATA INPUT - *Task I-Trial 02*

Timeslot(ms)	Trial 02	Estimation
0-500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
500-1000	(H,H,H,H,H,N,N,N,N,N,N,N,N,N,N)	*primitive5*
1000-1500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
1500-2000	(H,H,H,H,H,N,N,N,N,N,N,N,N,N,N)	*primitive5*
2000-2500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
2500-3000	(H,H,L,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive5*
3000-3500	(N,H,N,N,N,N,N,N,N,N,N,N,N,N,N)	*not recognized*
3500-4000	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*

The results for the primitives detection on raw data inputs for *Task II* are shown on Tables XVII and XVIII.

The estimated primitive corresponding to the input data of each segment is made by calculating $P_t(E = \mathit{primitive}_i/\mathbf{T} = (t_1, ..., t_{15}))$ for each primitive of the set of pre-defined primitives given the input tactile data $\mathbf{T}$ and selecting the primitive that maximizes the previous expression.

The application of the proposed primitives detection approach to the acquired data during the demonstrations of task, decomposes the input data on a sequence of segments. Typically the first segments correspond to the template of *Primitive7*, where the is no contact between the hand and the object. This period corresponds to the movement of the hand towards the object that is going to be manipulated.

Task I manipulation movements were segmented in a repetitive sequence of grasping and release of the object, in order to reorientate the mug placed on the top of the table to be grasped correctly. This sequence of grasp-realease allows, the subject performing the experiment, to repositionate the hand on the object, adapting the grasp configuration to the new pose of the object, to maximize the effectiveness of the subsequent hand actuation on the object. The fingers involved on the reorientation of the mug are predominantly the thumb, index and middle fingers. The ring and little fingers have a less intensive participation on those movements, although the assigned primitive is the same.

TABLE XVII
PRIMITIVES DETECTION ON RAW DATA INPUT - *Task II-Trial 01*

Timeslot(ms)	Trial 01	Estimation
0-500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
500-1000	(H,H,L,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive5*
1000-1500	(H,H,H,H,N,H,H,L,N,H,L,L,N,L,N)	*primitive3*
1500-2000	(H,H,H,H,L,H,H,L,L,H,L,L,N,L,N)	*primitive3*
2000-2500	(H,H,H,H,L,H,H,L,L,H,L,L,N,L,N)	*primitive3*

TABLE XVIII
PRIMITIVES DETECTION ON RAW DATA INPUT - *Task II-Trial 02*

Timeslot(ms)	Trial 02	Estimation
0-500	(N,N,N,N,N,N,N,N,N,N,N,N,N,N,N)	*primitive7*
500-1000	(H,H,H,H,N,H,H,L,N,H,L,L,N,L,N)	*primitive3*
1000-1500	(H,H,H,H,L,H,H,L,H,L,L,N,L,N)	*primitive3*
1500-2000	(H,H,H,H,L,H,H,L,L,H,L,L,N,L,N)	*primitive3*

The reorientation of the mug is performed faster on the second trial, requiring less grasp-release cycles. This can probably be related with the experience acquired by the human demonstrator or with the magnitude of the movements applied to the mug. The magnitude of the tactile inputs and number of fingers mobilized is higher during the initial grasp-release cycles.

During the primitive extraction from the *Task I* - trial 02 input signals, one of the segments was not classified. This was caused by the not very large extension of the primitives templates trainning datasets. This can be improved by increasing the trainning datasets extensions, as well as, introducing in the algorithm some similarity measures between input data signals that are going to be classified.

The second manipulation task, *Task II*, was decomposed on a serie of primitives that involve the participation of high extensions of the fingers surface. The task doesn't require the execution of grasp-release sequences of movements. The intermediate primitive, *Primitive5*, is not detected at the beginning of trial 02, as it is on trial 01. This can be caused by the speed of execution of the movement and by the segmentation of the movement on segments of equal pre-defined extension, as well as, the average of input signals on that period.

VI. CONCLUSIONS AND FUTURE WORK

This work has presented the global structure and methodology of the proposed approach to describe (symbolic description) the mechanisms underlying the human strategies used to execute in-hand manipulation movements. The results presented and discussed previously, represent the outcomes and analysis of this initial approach to devellop a broader system, already presented on figure 2, showing some effectiveness of this type of approache to achieve those goals. The results of the primitives detection stage can be improved by defining a larger set of pre-defined primitives in order to have a better symbolic description resolution of the performed task. Other possible future development is the implementation of unsupervised learning methods to estimate the parameters of the models of the primitives. The action module (fingers path and motion planning) which controls the interaction of the robotic dexterous hand with the object can be develloped and tested in a virtual environment simulator, using as input the contact state primitives sequences, in order to evaluate the effectiveness of the proposed approach.

Other component of the system that is going to be improved is the task primitive sequence representaction constraints extraction. This development will allow the automatic construction of a grammar for the task symbolic representation framework. This grammar defines, for a set of tasks, the set of rules that define the temporal and functional relations between the different primitives. These rules constitute the canonical representation of the relations between primitives, allowing the generalized encoding and consequent synthesis of in-hand manipulation movements.

VII. Ackowledgements

This work is partially supported by the European project HANDLE (23-16-40). Ricardo Martins and Diego R. Faria 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] C. E. Exner, *Development of hand skills in the child*. J. Case-Smith & C. Pehoski, 1992, ch. In-hand manipulation skills, pp. 35–45.
[3] N. Delson and H. West, "Robot programming by human demonstration: adaptation and inconsistency in constrained motion," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, vol. 1, 22-28 1996, pp. 30 –36 vol.1.
[4] S. Tso and K. Liu, "Hidden markov model for intelligent extraction of robot trajectory command from demonstrated trajectories," in *Industrial Technology, 1996. (ICIT '96), Proceedings of The IEEE International Conference on*, 2-6 1996, pp. 294 –298.
[5] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 37, no. 2, pp. 286 –298, April 2007.
[6] M. Kondo, J. Ueda, and T. Ogasawara, "Recognition of in-hand manipulation using contact state transition for multifingered robot hand control," *Robot. Auton. Syst.*, vol. 56, no. 1, pp. 66–81, 2008.
[7] K. Bernardin, K. Ogawara, K. Ikeuchi, and R. Dillmann, "A sensor fusion approach for recognizing continuous human grasping sequences using hidden markov models," *Robotics, IEEE Transactions on*, vol. 21, no. 1, pp. 47 – 57, feb. 2005.
[8] V. Kruger, D. Herzog, S. Baby, A. Ude, and D. Kragic, "Learning actions from observations," *Robotics Automation Magazine, IEEE*, vol. 17, no. 2, pp. 30 –43, june 2010.
[9] http://www.tekscan.com/industrial/grip-system.html.