Annual Report

Control and Representation of Articulated Objects: Insights from robots

PhD Program in Bioengineering Robotics
Curriculum: Cognitive Robotics, Interaction and
Rehabilitation Technologies

Student: Venkata Sai Yeshasvi Tirupachuri
Cycle: XXXI
.Tutors:

Dr. Daniele Pucci
Prof. Gabriel Baud-Bovy
Prof. Guilio Sandini

Year: 2
EMail : yeshasvi.tirupachuri@iit.it
Mobile: +39-327-974-3351

Contents

1 Research Activity

1.1 Scientific and Technological Objectives

To be effective in an ever changing complex environments that humans live, robots need to have certain cognitive capabilities on par with humans. Most importantly fruitful interactions with the complex dynamic environment is very crucial.

The research network PACE(Perception and Action in Complex Environments) deals with investigations into human motor control in highly complex environments. This is a fairly complex topic which can be explored at different fronts ranging from theoretical models to rehabilitation technologies. Coming from engineering and robotics background, I strongly believe that investigations in several areas of PACE research can result in endowing robots with humanlike cognitive capabilities for interaction with dynamic and complex environments.

The prime case of complex environments considered in my research topic is articulated objects, like doors, scissors etc. Because of their structural rigidity, most robots strain to manipulate kinematically-constrained articulated objects without causing large interaction forces. simply put these objects have different parts attached through joint(s) and hence the motion of the object parts is restricted according to the joint type(s). We humans with excellent perceptual and motor skills learn to recognize and manipulate them easily overtime. The same capability for robots is very much needed to be effective and productive in natural environments. Hence the title of my topic is "Control and Representation of Articulated Objects: Insights from robots".

The technical objectives can be summarized as

- A computational framework for articulation object model estimation

- Implementing cognitive models of sensorimotor signal processing in interaction

- Validating the framework with experiments using humanoid iCub robot

1.2 State of the art and Innovation

State-of-the-art research in articulation model estimation has been mainly focused on exploting visual perception. For effective manipulation of articulation objects adequate information of the shape, pose and the kinematic structure of the objects and its parts is crucial. Pose tracking generally assumes the shape of the objects or object parts and utilizes segmentation in the visual input. Shape reconstruction can be achieved using integration of multiple poses and then the kinematic structure of an articulated object can be achieved using the above information on the object parts.

The segmentation problem deals with extracting the region of interest in the visual input which contains the object to track.

- Single image based segmentation [19], [28]

- Motion based segmentation [5]

  Pose estimation or visual tracking can be achieved using

- Based on a known shape model of the object [34]

- Based on the sufficient texture extracting features [16] or optical flow [29], [25]

Shape reconstruction outputs a 3D model of an object by stitching together several object views into one coherent shape model using relative poses w.r.t the visual system. Object views are obtained using information from segmentation and pose estimation [15]. Visual SLAM is a popular approach combining tracking and shape reconstruction but for an entire scene assumed to be static. An integrated approach for the visual perception of articulated objects is recently proposed in [18] which combines tracking, shape reconstruction and estimation of kinematic structures.

Doors are the most likely experienced articulated objects in many robotic applications like rescue scenarios, elderly care, hospitality etc., and the earliest investigations started with a simple strategy to open an unknown door [22]. More recently, Advait Jain and Charles C. Kemp proposed the concept of equilibrium point control (EPC) [12] for the specific task of opening novel doors and drawers. In addition they implemented an articulation model estimation algorithm using the end-effector trajectory, assuming a stable grasp and planar motion. The algorithm returns an estimate of the location of the axis of rotation and the radius. Prismatic joint is estimated as a rotational joint with large radius.

Dov Katz and Oliver Brock [13] introduced the idea of interactive perception paradigm and highlights the need for extracting task specific perceptual information using the manipulation capabilities of a robot by interacting with the environment. Some of the limitations in [13] are removed in [14] and a robust algorithm is presented. Feature trajectories of moving objects are extracted using optical flow based tracking and a graph is built where vertices represent the tracked features. Highly connected sub-graphs are separated using min-cut algorithm and these sub-graphs represent the object parts. Then the articulation models are extracted from the information contained in the graph. Rotational joint is identified by rotating centers between two sub-graphs and prismatic joint by shifting movements of sub-graphs. They successfully demonstrated the use of interactive perception in extracting kinematic model of various tools to build a Denavit-Hartenberg (DH) parameter model and then use it to operate a tool. Furthermore, a symbolic learning based approach to manipulation is presented in [4] which uses relational representations of kinematic structures that are grounded using perceptual and interaction capabilities of the robot. They successfully demonstrated learning and generalization of manipulation knowledge to previously unseen objects.

Strum et al. [32] proposed a probabilistic learning framework using noisy 3D pose observations of object parts. They implemented predefined candidate joint models with parameters and also a non-parametric gaussian process model to which observed 3D pose trajectory data of object parts is fit to find kinematic structures of kinematic trees. Unlike the previous works this is a passive approach and they use pose observations from motion capture system, when the articulated object is manipulated by someone, and simulated data to validate their framework. Further, a stereo camera system is used to get dense depth images as input [31]. An approach to direct interaction with articulated objects to learn

articulation models and improve based on experience is presented in [30]. Furthermore, a unified framework with several extensions like dealing with kinematic loops and extended set of experiments is presented in [33].

Stefan et al. [27] introduced the exploration challenge for robots where the task is to perform explorative actions and learn the structure of the environment. One of their main contributions is probabilistic belief representation of articulation models including properties like friction and joint limits. They successfully demonstrated how the behaviour emerged from entropy-based exploration is more informative than explorative strategies based on heuristics. Hausman et at. [10] proposed a partcile filter based approach where the articulations models are represented as particles and implemented a probabilistic framework based on the concept of interactive perception to integrate visual observations with manipulation feedback from the robot. They also presented best action selection methods based on entropy and information gain which guides the robot to perform the most useful interactions with the object to reduce its uncertainty.

Martin-Martin Roberto and Oliver Brock [16] presented an online multi-level recursive estimation algorithm considering task-specific priors based on the concept of interactive perception. They used series of RGB-D image data as input to estimate articulation models including the joint configuration. Furthermore, they extended their approach [17] integrating information from vision, force-torque sensing and proprioception. In addition to kinematic model estimation they also generated a dynamic model of the articulated object.

The choice of a well established framework on human motor control to facilitate the basic software infrastructure for my research is a crucial step. Towards this end I did review of papers on the framework of Active Inference [8], [7] put forward by Prof. Karl Friston. Also I studied the framework of Passive Motion Paradigm(PMP) [20], [21] championed by our colleagues Dr. Vishwanathan Mohan and Prof. Pietro Morasso from the Italian Institute of Technology.

Visual information has been used as the most dominant sensory modality in articulation model estimation. But while manipulating an articulated object, another key information available is the interaction wrench. As a part of the Dynamic Touch and Interaction team and in collaboration with Dynamic Interaction Control team, I am exploring the possibilities of using interaction wrench along with visual perception to estimate the kinematic structure and dynamic properties of articulated objects. From the visual perception kinematic properties of the artiulated can be estimated but by also using interaction wrench information we believe one can estimate the dynamic properties too. This facilititates for better control during object manipulation. A bayesian framework is proposed in [23] for simultaneous state and dynamics estimation in presence of multiple redundant sensory information. But in the case of manipulation of articulated objects state information will estimated using visual inputs.

1.3 Methodology and Workplan

1.3.1 Modelling

We modelled two simple artitulated objects using Simulation Description Format (SDF) to be used in Gazebo simulation environment. The objects contain two links connected through a *prismatic* joint, P in one (Fig. 1a) and a *revolute* joint, R in the other (Fig. 1b). The reasoning behind this modelling is that most real life articulated objects can be represented in this simplified form. The cylindrical elements in black color are the handle links and the rectangular elements in yellow color are articulated object links. The handle links of the articulated objects are connected to object links through a fixed joint and one of the handles is anchored to the world in gazebo through a fixed joint and this also anchors the object link attached to that handle. The reasoning behind this choice is that many articulated objects have one of their links either grounded or are quite heavy to move compared to the freely movable link.
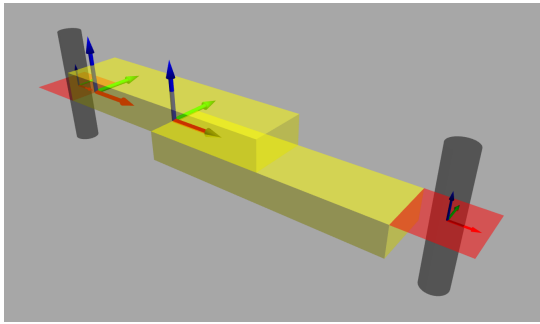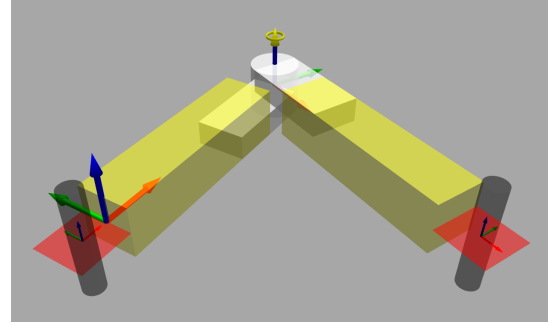

(a) Prismatic Model   (b) Revolute Model

Figure 1: Articulated Object Models

1.3.2 Experimental Scenario

The experimental scenario we want to test is as indicated in the figure 2, where the icub robot is holding the articulated object at the handles and performs explorative actions to estimate the articulation model present between the links of an articulated object.
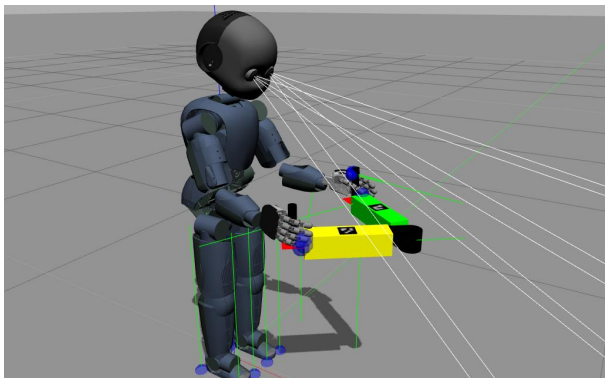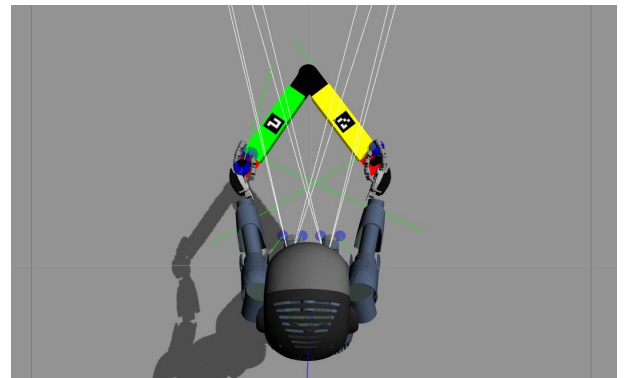
(a) Side View   (b) Top View

Figure 2: Experimental Scenario: iCub manipulating an articulated object

1.3.3   Implemetation Details

Visual Perception

I started working on the submodules of vision and force/torque in the perception module of the software architecture shown in figure 3. Obtaining robust 3D pose values of rigid bodies is still an open challenge in the field of computer vision. Vast amount of research has been carried on tracking rigid bodies either using markers, features or depth information. As visual perception is not the main goal of this work, I made use of aruco library [9] which is an opensource software for augumented reality applications based on OpenCV. Figure 4a shows an array of markers and figure 4b shows the markers attached on the two moving parts of a stapler with their detected 3D pose shows with an axes.
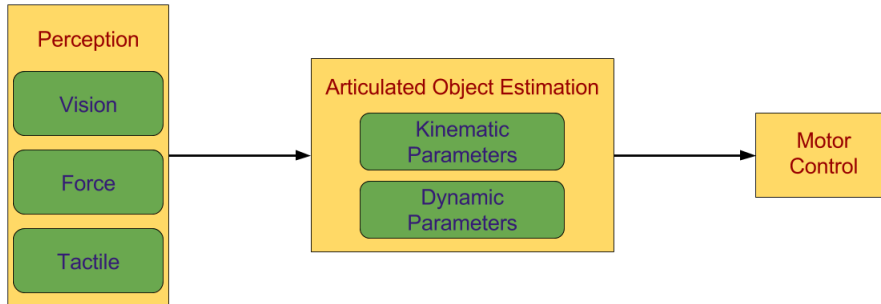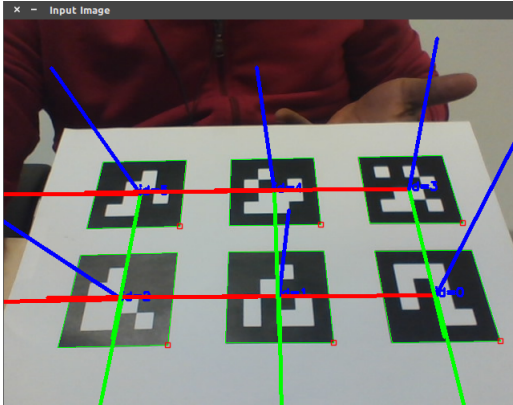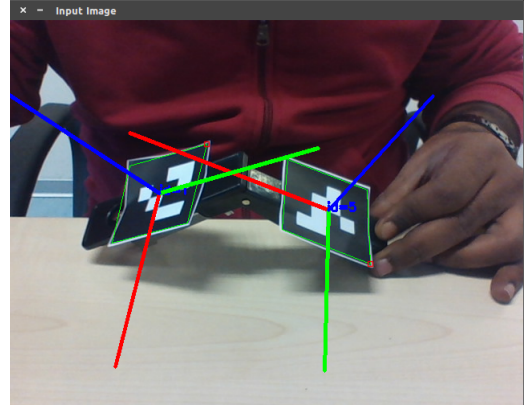

Figure 3: Software Architechture

| (a) Markers Board | (b) Markers on a stapler |

Figure 4: Augumented Reality Markers

Wrench Perception

Coming to force/torque perception, iCub robot is equipped with six proximal six-axis F/T sensors as shown in the figure 5. While using the iCub robot external wrenches can be estimated using the techniques develoed for whole-body control [24]. At this point the current icub model described using an URDF (Universal Robot Description Format) file in simulation environment did not include the hands of the robot. I created new URDF files [2] for the hands of the robot included along with collision elements and did contact tuning

[1] in gazebo simulation environment. Later two frames as shown in the figure 6 are defined at the center of the hands at which we can estimate the wrench acting on the robot.
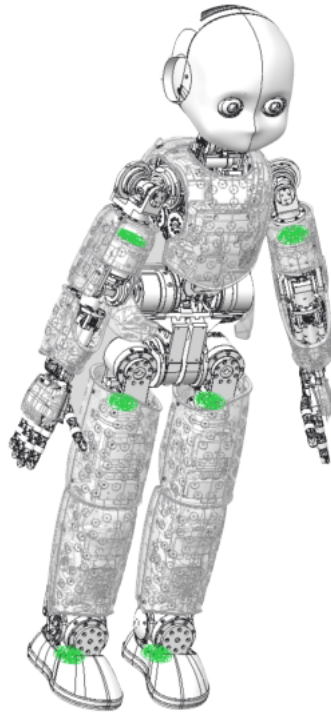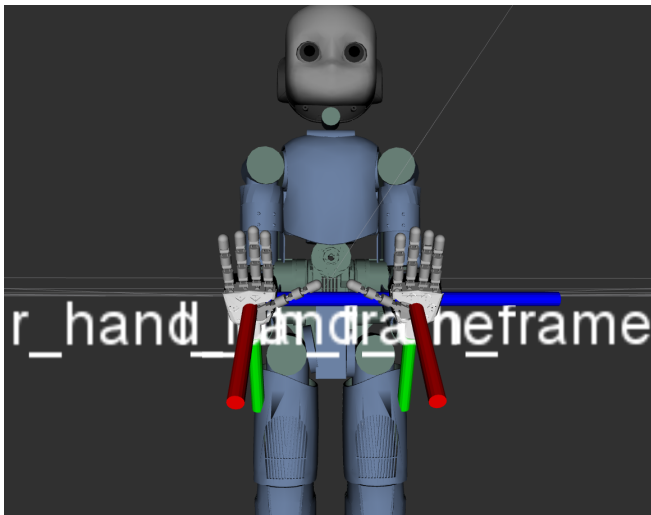

Figure 5: iCub robot: location of six proximal F/T Sensors

(a) Front View
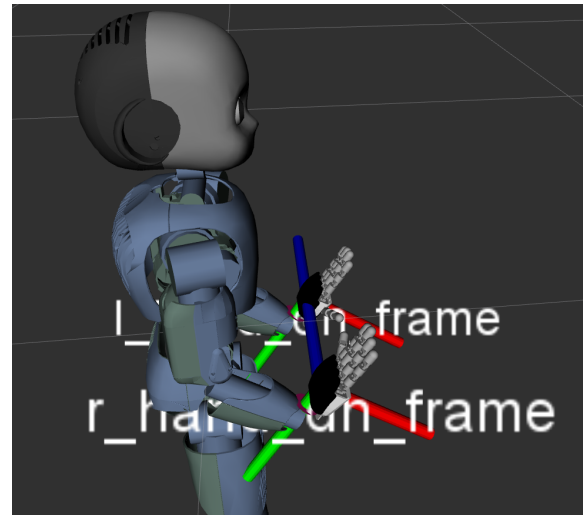
(b) Side View

Figure 6: iCub hands with new frames

A simulated 6 axis Force-Torque sensor plugin [11] is placed at the fixed joint connecting object link to its handle, to measure the external wrenches acting on the object links. In an ideal case where the iCub robot is having a stable power grasp of the object, the wrench measurements read from this Force-Torque sensor plugin should be similar to the measurements obtained at the hands of the robot provided that the direction of measurement is the same. But in practice the hands of the iCub robot are quite complex to achieve a stable power grasp.

Articulation Estimation Algorithm

Visual information has been used as the most dominant sensory modality in articulation model estimation. But while manipulating an articulated object, another key information available is the interaction forces and torques. We came up with an articulation estimation algorithm that uses momentum information and the interaction wrench information.

In this work we use spatial vectors [6], which are 6D vectors that are proven to be a powerful tool in analysing rigid-body dynamics. Unlike the standard notation of spatial vectors, in the spatial motion vectors we use the linear part first followed by the angular part and in the spatial force vectors we use forces first followed by moments. In a complex articulated system consisting of several links connected through joints of several degrees of freedom (e.g humanoid robots), the total momentum of the system is the sum of link momenta, provided that the link momenta are expressed in a common reference frame. For humanoid robots, it is commonly expressed after projecting each link momenta to the robot's Center of Mass (CoM, G) and hence it is called centroidal momentum [26].

$$^{G}h = \sum_{i=1}^{n} {}^{G}X_i \, h_i = \sum_{i=1}^{n} {}^{G}X_i \, M_i \, \mathrm{v}_i \tag{1}$$

Where,

- ^{G}h is the centroidal momentum of the system.

- n is the total number of links in the articulated system.

- $^{G}X_i$ is the spatial transformation from the body reference frame to the center of mass of the system.

For our modelling, the total momentum of the articulated object expressed with respect to the inertial frame of reference is given by,

$$h = X_1 \, M_1 \, \mathrm{v}_1 + X_2 \, M_2 \, \mathrm{v}_2 \tag{2}$$

In the case of a rigid body system containing a kinematic tree structure, the velocity of a body i, expressed in its body frame, is given by,

$$\mathrm{v}_i = {}^{i}X_{\lambda(i)} \, \mathrm{v}_{\lambda(i)} + \mathrm{v}_{Ji} \tag{3}$$

where,

- v_i is the twist of the i-th link expressed in its body frame.

- $\mathrm{v}_{\lambda(i)}$ is the twist of the parent link of i-th link, expressed in its body frame.

- $^{i}X_{\lambda(i)}$ is the spatial transformation from the parent link to the child link.

- v_{Ji} is the twist of the i-th joint, connecting link i to its parents, expressed in the body frame of link i.

The equation(3), in our modelling, simplifies to,

$$v_2 = {}^2X_1\, v_1 + v_J \tag{4}$$

The twist of a joint, expressed in the child link body frame, is given by,

$$v_J = S\dot{q} \tag{5}$$

In this work we considered only prismatic and revolute articulation models as they are present in most articulated objects. Assuming these two articulation model types we compute two joint twist quantities, using the relative transformation between the fixed link and movable link. In case of prismatic model hypothesis, the joint velocity $\dot{q}_p$ and the joint axis S_P are computed from the linear part of the relative transformation. In case of a revolute model hypothesis, the joint velocity $\dot{q}_R$ and the joint axis S_R are computed from the rotation matrix of the relative transformation.

Now while computing the total momentum using equation(2), the information of articulation *model* is encapsulated, which can be represented in short as

$$h(model) \tag{6}$$

The total wrench acting on a system, expressed in the inertial frame A, is given by the summation of external wrenches and gravitational forces on each link.

$$W = \sum_{i=1}^{l} {}^AX_i\, f_i + \sum_{j=1}^{n} {}^AX_j\, m_j g \tag{7}$$

where,

- l is the number of force-torque sensors present in the articulated object.

- f_i is the total wrench measured by a force-torque sensor expressed in its local frame of reference.

- ${}^AX_i$ is the spatial transformation from the force-torque sensor local frame to the inertial frame, A.

- ${}^AX_j$ is the spatial transformation from the body referece frame B to the inertial frame, A.

In our modelling, the equation(7) is reduced to,

$$W = X_{ft_1}f_1 + X_{ft_2}f_2 + X_{com_1}m_1 g + X_{com_2}m_2 g \tag{8}$$

For any rigid body system, the sum of the external wrenches acting on it is equal to the rate of change of momentum.

$$W = \dot{h}(model) \tag{9}$$

Considering the two articulated model hypotheses, the model which solve the following optimization represents the true joint nature of articulation present between the two object parts.

$$model^* = \underset{model}{\mathrm{argmin}} \sum \|W - \dot{h}(model)\| \tag{10}$$

Currently our algorithm 1 is offline and the recorded data is passed as input. All the quantities denoted in the algorithm are computed with respect to the inertial frame of reference, A. The algorithm outputs two quantities,

- *Prismatic Hypothesis Error*, which is the value that corresponds to the mismatch between the actual prismatic joint motion and the prismatic articulation model hypothesis,

$$P_{hyp} = \sum_{data} \|W - \dot{h}(prismatic\ model)\| \tag{11}$$

- *Revolute Hypothesis Error* (*Rhyp*) which is the value that corresponds to the mismatch between the actual revolute joint motion and the revolute model hypothesis,

$$R_{hyp} = \sum_{data} \|W - \dot{h}(revolute\ model)\| \tag{12}$$

Algorithm 1 Calculate Joint Hypothesis - P_{hyp}, R_{hyp}

Require: : $Data : t, P_1, P_2, f_1, f_2$

1: **for** $i < size(Data, 1)$ **do**
2:     $dt \leftarrow diff(t)$
3:     $T_1 \leftarrow H(P_1)$
4:     $T_2 \leftarrow H(P_2)$
5:     ${}^1T_2 \leftarrow T_1^{-1}\, T_2$
6:     ${}^1P_2 \leftarrow linear({}^1T_2)$
7:     $q_P \leftarrow norm({}^1P_2)$
8:     $S_P \leftarrow \left[\frac{{}^1P_2}{d}, 0, 0, 0\right]^T$
9:     $\dot{q}_P \leftarrow \frac{diff(q_P)}{dt}$
10:    $v_J^P \leftarrow X(T_1)\, S_P\, \dot{q}_P$
11:    ${}^1R_2 \leftarrow rot({}^1T_2)$
12:    $axisAngleVector \leftarrow vrrotmat2vec({}^1R_2)$
13:    $q_R \leftarrow axisAngleVector(4)$
14:    $S_R \leftarrow \left[0, 0, 0, axisAngleVector(1:3)\right]^T$
15:    $\dot{q}_R \leftarrow \frac{diff(q_R)}{dt}$
16:    $v_J^R \leftarrow X(T_1)\, S_R\, \dot{q}_R$
17:    $p_1 \leftarrow linear(T_1)$
18:    $v_1 \leftarrow \frac{diff(p_1)}{dt}$
19:    $R_1 \leftarrow rot(T_1)$
20:    $\dot{R}_1 \leftarrow \frac{diff(R_1)}{dt}$
21:    $\omega_1 \leftarrow \dot{R}_1\, R_1^T$
22:    $v_1 \leftarrow [v_1, \omega_1]^T$
23:    $v_2^P \leftarrow v_1 + v_J^P$
24:    $v_2^R \leftarrow v_1 + v_J^R$
25:    $h^P \leftarrow [X_1^*\, M_1\, {}^1X]\, v_1 + [X_2^*\, M_2\, {}^2X]\, v_2^P$
26:    $h^R \leftarrow [X_1^*\, M_1\, {}^1X]\, v_1 + [X_2^*\, M_2\, {}^2X]\, v_2^R$
27:    $\dot{h}^P \leftarrow \frac{diff(h^P)}{dt}$
28:    $\dot{h}^R \leftarrow \frac{diff(h^R)}{dt}$
29:    $f_1 \leftarrow X_{left}^*\, f_1$
30:    $f_2 \leftarrow X_{right}^*\, f_2$
31:    $W \leftarrow f_1 + f_2 + X_{com_1}^*\, m_1 g + X_{com_2}^*\, m_2 g$
32:    $P \leftarrow W - \dot{h}^P$
33:    $R \leftarrow W - \dot{h}^R$
34: **end for**
35: $P_{hyp} \leftarrow norm(sum(P.^*P))$
36: $R_{hyp} \leftarrow norm(sum(R.^*R))$

1.4 Achieved results, dissemination, deviation from the original plan

The gantt chart in the figure 9 shows the proposed work plan during the project proposal and the actual work progress during the first year of PhD. The figure software architecture indicated in the figure 3 is realized by implementing various submodules to a functional

level to conduct our experiments in gazebo simulation environment.

We encountered a couple of problems to conduct intented experiments. Firstly, the 3D pose values of the object links using markers is inconsistent, particularly the orientation part. This is due to the noisy visual observations. Secondly, as mentioned before the hands of the iCub robot are quite complex to maintain a stable power grasp. We then decided to make some assumptions that can enable us to perform experiments to validate our momentum based articulation estimation algorithm approach.

As visual perception is not the main goal of this work, we acquire the pose values directly from the simulation environment using a plugin. Instead of using the robot to manipulate the articulated objects, we applied external forces in simulation to move the object in the direction of allowed motion by the joint. The joints have a damping value of 0:1 and static friction value of 0:1. Link inertial properties i.e., mass, inertia and center of mass are assumed to be known apriori. The range of motion for the prismatic joint is 0:14 meters and for the revolute joint, 95 degrees. We recorded data with five different forces applied at the handle attached to the freely movable link. While the object moves, we record the simulation time, object links 3D pose values and the external wrench values acting on the object links.

For the data of prismatic joint motion under the influence of different forces, the Fig.(7) shows the value of P_{hyp} is less than R_{hyp}. This indicates that for this joint motion, prismatic articulation model is a close match to the actual joint type present in the articulated object.
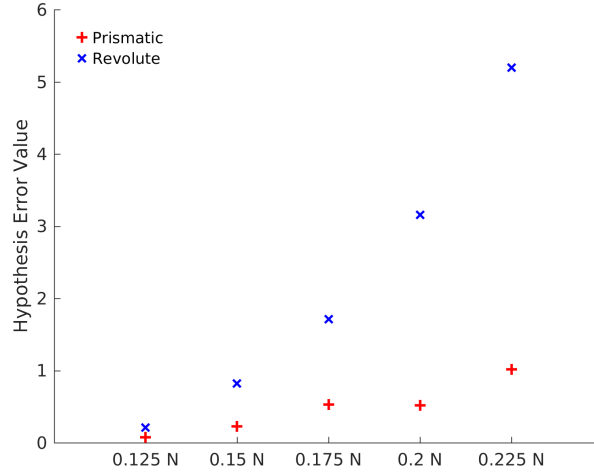

Figure 7: Prismatic Motion Hypothesis Error Value

For the data of revolute joint motion under the influence of different forces, the Fig.(8) shows the value of R_{hyp} is less than P_{hyp}. This indicates that for this joint motion, revolute articulation model is a close match to the actual joint type present in the articulated object.
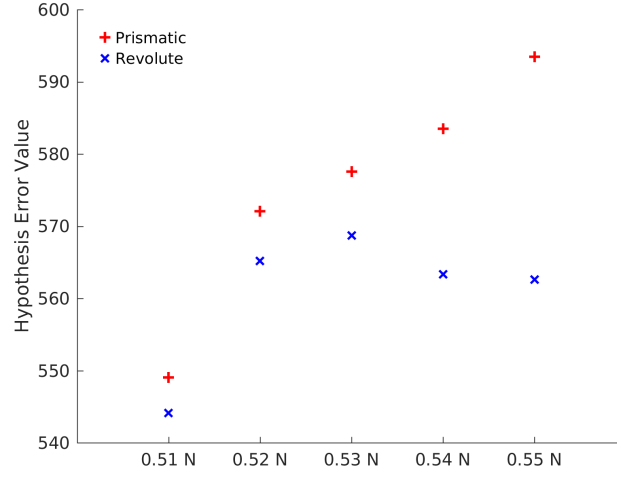
Figure 8: Revolute Motion Hypothesis Error Value

Even though we made several assumptions about the availability of rigid body pose and inertial parameters of the object models, Gazebo is a fixed time-step simulator, of 1 *ms*, which adds considerable numerical errors for a stiff mechanical system considered in our experiments. In such a noisy environment, our work is a proof of concept proving that we can distinguish between articulation models using the kinematic evolution and interaction wrench information available while manipulating an articulated object.

Figure 9: Gantt Chart PhD Year II
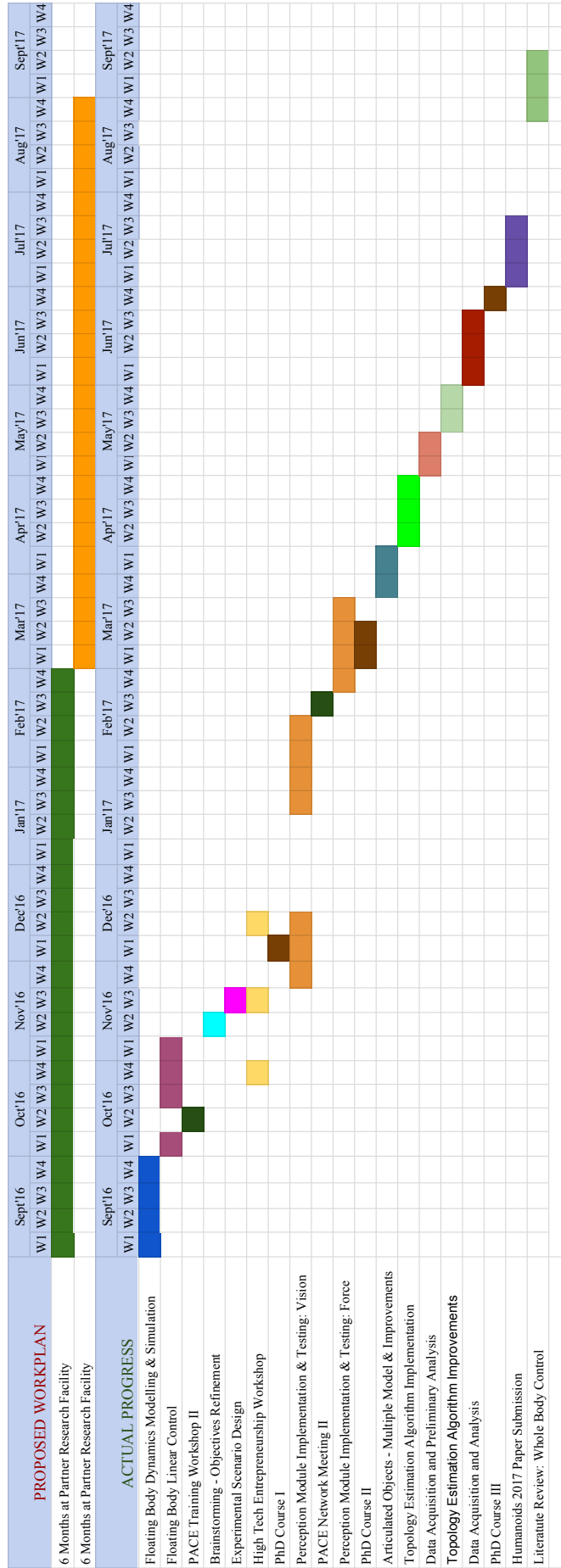
1.5 References

[1] icub hands contact tuning. `https://github.com/robotology/icub-gazebo/pull/36`, (accessed September 11, 2017).

[2] icub hands urdf files. `https://github.com/Yeshasvitvs/icub_hands_urdf/`, (accessed September 11, 2017).

[3] Leveraging machine learning alternatively. `http://itn-pace.eu/leveraging-machine-learning-alternatively-by-venkata-tirupachuri/`, (accessed September 12, 2017).

[4] Dov Katz Yuri Pyuro Oliver Brock. Learning to manipulate articulated objects in unstructured environments using a grounded relational representation. *Robotics: Science and Systems IV*, page 254, 2009.

[5] Shao-Yi Chien, Shyh-Yih Ma, and Liang-Gee Chen. Efficient moving object segmentation algorithm using background registration technique. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(7):577–586, 2002.

[6] Roy Featherstone. *Rigid body dynamics algorithms*. Springer, 2014.

[7] Karl Friston, Jérémie Mattout, and James Kilner. Action understanding and active inference. *Biological cybernetics*, 104(1-2):137–160, 2011.

[8] Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. Active inference and epistemic value. *Cognitive neuroscience*, 6(4):187–214, 2015.

[9] S. Garrido-Jurado, R. Mu noz Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280 – 2292, 2014.

[10] Karol Hausman, Scott Niekum, Sarah Osentoski, and Gaurav S Sukhatme. Active articulation model estimation through interactive perception. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3305–3312. IEEE, 2015.

[11] Enrico Mingo Hoffman, Silvio Traversaro, Alessio Rocchi, Mirko Ferrati, Alessandro Settimi, Francesco Romano, Lorenzo Natale, Antonio Bicchi, Francesco Nori, and Nikos G Tsagarakis. Yarp based plugins for gazebo simulator. In *International Workshop on Modelling and Simulation for Autonomous Systems*, pages 333–346. Springer, 2014.

[12] Advait Jain and Charles C Kemp. Pulling open novel doors and drawers with equilibrium point control. In *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*, pages 498–505. IEEE, 2009.

[13] Dov Katz and Oliver Brock. Interactive perception: Closing the gap between action and perception. In *ICRA 2007 Workshop: From features to actions-Unifying perspectives in computational and robot vision*, 2007.

[14] Dov Katz and Oliver Brock. Manipulating articulated objects with interactive perception. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 272–277. IEEE, 2008.

[15] Michael Krainin, Peter Henry, Xiaofeng Ren, and Dieter Fox. Manipulator and object tracking for in-hand 3d object modeling. *The International Journal of Robotics Research*, 30(11):1311–1327, 2011.

[16] Roberto Martin Martin and Oliver Brock. Online interactive perception of articulated objects with multi-level recursive estimation based on task-specific priors. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2494–2501. IEEE, 2014.

[17] Roberto Martın-Martın and Oliver Brock. Building kinematic and dynamic models of articulated objects with multi-modal interactive perception. 2017.

[18] Roberto Martın-Martın, Sebastian Höfer, and Oliver Brock. An integrated approach to visual perception of articulated objects.

[19] Ajay Mishra, Yiannis Aloimonos, and Cornelia Fermuller. Active segmentation for robotics. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3133–3139. IEEE, 2009.

[20] Vishwanathan Mohan and Pietro Morasso. Passive motion paradigm: an alternative to optimal control. *Frontiers in neurorobotics*, 5:4, 2011.

[21] Vishwanathan Mohan, Pietro Morasso, Giulio Sandini, and Stathis Kasderidis. Inference through embodied simulation in cognitive robots. *Cognitive Computation*, 5(3):355–382, 2013.

[22] Günter Niemeyer and J-JE Slotine. A simple strategy for opening an unknown door. In *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, volume 2, pages 1448–1453. IEEE, 1997.

[23] Francesco Nori, Naveen Kuppuswamy, and Silvio Traversaro. Simultaneous state and dynamics estimation in articulated structures. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 3380–3386. IEEE, 2015.

[24] Francesco Nori, Silvio Traversaro, Jorhabib Eljaik, Francesco Romano, Andrea Del Prete, and Daniele Pucci. icub whole-body control through force regulation on rigid non-coplanar contacts. *Frontiers in Robotics and AI*, 2:6, 2015.

[25] Peter Ochs, Jitendra Malik, and Thomas Brox. Segmentation of moving objects by long term video analysis. *IEEE transactions on pattern analysis and machine intelligence*, 36(6):1187–1200, 2014.

[26] David E Orin, Ambarish Goswami, and Sung-Hee Lee. Centroidal dynamics of a humanoid robot. *Autonomous Robots*, 35(2-3):161–176, 2013.

[27] Stefan Otte, Johannes Kulick, Marc Toussaint, and Oliver Brock. Entropy-based strategies for physical exploration of the environment's degrees of freedom. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 615–622. IEEE, 2014.

[28] Jeremie Papon, Alexey Abramov, Markus Schoeler, and Florentin Worgotter. Voxel cloud connectivity segmentation-supervoxels for point clouds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2027–2034, 2013.

[29] Jörg Stückler and Sven Behnke. Efficient dense 3d rigid-body motion segmentation in rgb-d video. In *BMVC*, 2013.

[30] Jürgen Sturm, Advait Jain, Cyrill Stachniss, Charles C Kemp, and Wolfram Burgard. Operating articulated objects based on experience. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 2739–2744. IEEE, 2010.

[31] Jürgen Sturm, Kurt Konolige, Cyrill Stachniss, and Wolfram Burgard. 3d pose estimation, tracking and model learning of articulated objects from dense depth video using projected texture stereo. In *RGB-D: Advanced Reasoning with Depth Cameras Workshop, RSS*, 2010.

[32] Jürgen Sturm, Vijay Pradeep, Cyrill Stachniss, Christian Plagemann, Kurt Konolige, and Wolfram Burgard. Learning kinematic models for articulated objects. In *IJCAI*, pages 1851–1856, 2009.

[33] Jürgen Sturm, Cyrill Stachniss, and Wolfram Burgard. A probabilistic framework for learning kinematic models of articulated objects. *Journal of Artificial Intelligence Research*, 41:477–526, 2011.

[34] Manuel Wüthrich, Peter Pastor, Mrinal Kalakrishnan, Jeannette Bohg, and Stefan Schaal. Probabilistic object tracking using a range camera. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3195–3202. IEEE, 2013.

2 Traning Related to the PhD Programme

1. Artificial Cognitive Systems

   - Type: PhD Course
   - Organizer: IIT
   - Duration: 10 Hours
   - Credits: 3
   - Verification: Pass

2. An Introduction to Spatial (6D) Vectors and Their Use in Robot Dynamics

   - Type: PhD Course
   - Organizer: IIT
   - Duration: 15 Hours
   - Credits: 5
   - Verification: Pass

3. Machine Learning: A Computational Intelligence Approach

   - Type: PhD Course
   - Organizer: UNIGE
   - Duration: 18 Hours
   - Credits: 6
   - Verification: EXAM TO BE HELD ON 18-SEP-2017

3 Other Activities

- PACE second Thematic Workshop(TW2) on Active Inference at l'Université de Lyon, 12 - 14 October 2016, Lyon, France

- Online E-Learning module on Kinematic data (arm + eye): Concepts and analysis by Prof. Joan Lopez Moliner from Universitat de Barcelona (UB), 24 November 2016

- Workshop on High Tech Entrepreneurship organized by IIT and University of Genoa, 28 October - 15 December 2016, Genoa, Italy

- PACE Second Network Meeting(NM2) at Universitat de Barcelona (UB), 20 - 22 February 2017, Barcelona, Spain

- Science communication project [3] of PACE in partnership with the International School for Advanced Studies - SISSA, Trieste, Italy

- Online E-Learning module on Introduction to robotics and haptic device control by Prof. Gabriel Baud-Bovy from IIT, Italy, 4 May 2017

- Online E-Learning module on Haptic devices by Prof. Gabriel Baud-Bovy from IIT, Italy, 24 May 2017

- Online E-Learning module on Human-robot social interactions by Dr. Alessandra Sciutti from IIT, Italy, 15 June 2017

- Online E-Learning module on Robot assisted rehabilitation by Dr. Jacopo Zenzeri from IIT, Italy, 20 June 2017

- Online E-Learning module on Control theory dy Dr. Daniele Pucci from IIT, 29 June 2017

- Online E-Learning module on An example of and advanced robot: iCub by Prof. Giulio Sandini and Dr. Francesco Rea from IIT, Italy, 25 July 2017

Signatures,
 PhD Student   Tutor(s)