Characterization of Textile Grasping Experiments

Guillem Alenyà, Arnau Ramisa, Francesc Moreno-Noguer and Carme Torras

***Abstract*— Grasping highly deformable objects, like textiles, is an emerging area of research that involves both perception and manipulation abilities. As new techniques appear, it becomes essential to design strategies to compare them. However, this is not an easy task, since the large state-space of textile objects explodes when coupled with the variability of grippers, robotic hands and robot arms performing the manipulation task. This high variability makes it very difficult to design experiments to evaluate the performance of a system in a repeatable way and compare it to others. We propose a framework to allow the comparison of different grasping methods for textile objects.**

Instead of measuring each component separately, we therefore propose a methodology to explicitly measure the vision-manipulation correlation by taking into account the throughput of the actions. Perceptions of deformable objects should be grouped into different clusters, and the different grasping actions available should be tested for each perception type to obtain the action-perception success ratio. This characterization potentially allows to compare very different systems in terms of *specialized actions*, *perceptions* or *widely useful actions*, along with the cost of performing each action. We will also show that this categorization is useful in manipulation planning of deformable objects.

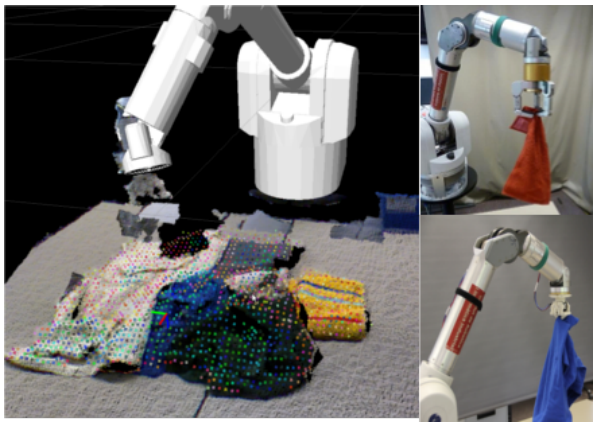


Fig. 1: A manipulator robot grasping several pieces of cloth using different gripper configurations and perceptions. (Left) Visualization of the robot, some selected interest points and the selected grasping point. (Right) Two grasping examples.

I. Introduction

Currently, one of the most exciting emerging topics in robotics research is general object manipulation in unprepared scenarios. During the latest years, many works have presented impressive advancements, like going to the fridge and bringing back a bottle of beer [1], folding a pile of towels [2], or even tasks involving multiple steps and coordination between various agents, like preparing pancakes [3].

As new methods to perform manipulation tasks appear, it becomes a critical scientific issue to find a solid methodology to compare them and assess their progress. However, the fact that manipulation tasks involve at the same time skills related to perception and action makes this systematic evaluation very hard. Furthermore, in a real system there is a high interdependency between actions and perceptions which prevents evaluating each one independently and, on top of that, initial conditions may be difficult to replicate between experiments, not to mention between different research groups with different hardware.

Initial conditions when only rigid objects are considered are rather easy to establish. Contrarily, except for some straightforward configurations, replicating initial conditions in experiments using non-rigid objects is very difficult, e.g. a scene including wrinkled textiles (see Fig. 1), as the pose of a textile is in general very difficult to fully characterize.

Textile manipulation experiments involve also uncertainty in the result of a grasping operation. Depending on the task, a failed grasp can be defined as grasping no objects or more than one (which is sometimes not obvious to observe by the robot).

In such scenario, decision-making frameworks that explicitly take into account uncertainty are naturally well suited. Among others we can cite two main families, Markov Decision Processes (MDPs) and its generalization to Partially Observable Markov Decision Process (POMDP) [4], and Bayesian Network Planning (BNP) and its extensions like PRADA [5]. All these methods have in common that planning is based on a finite list of actions along with their probabilities of success.

In this paper a method to characterize non-rigid manipulation approaches is proposed. It is based on identifying the perception-manipulation actions that a robotic system can perform and, for each couple, computing the success ratio. The advantages of the framework we propose are two-fold: we show that the aforementioned planning algorithms can be benchmarked with real and well established metrics, and different robotic systems can be compared by evaluating in a common framework the skills they have. The experimental section provides examples of both.

This work was supported by the Spanish Ministry of Science and Innovation under projects PAU+ DPI2011-27510, by the EU project INTELLACT 247947 FP7-269959 and by the Catalan Research Commission through SGR-00155.

The authors are with the Institut de Robotica i Informatica Industrial (CSIC-UPC), Llorens i Artigas 4-6, 08028 Barcelona, Spain {aramisa,galenya,fmoreno,torras}@iri.upc.edu

II. Previous work

Intrinsically linked to the idea of experiment replication there is the idea of system benchmarking. This is a core element in technological development, since it allows to objectively evaluate key system properties, and compare them with alternative solutions. It is also a requirement for most project funding grants to evaluate the developed methods with the standard benchmarks in the field.

Furthermore, providing benchmarks or benchmarking guidelines helps focus the research efforts on the solutions and concentrate on addressing the most pressing problems. Competitions can be organized around benchmarks to foster development. Examples of competitions in the field of object perception and manipulation for robotics are the "Solutions in Perception Challenge" [9] where participants must detect objects using a perception system mounted on a mobile robot, the "Robocup@Home", which involves perception and manipulation of rigid objects, or the "Sushi restaurant challenge" [1].

For purely perception tasks, such as image classification or human action recognition, creating benchmarks to compare different methods is usually straightforward, as it suffices to manually annotate the occurrences of the objects/events of interest in the images/videos. Examples are the Pascal Visual Object Challenge [6], ImageNet [7] and the Hollywood-2 dataset [8]. More robotics-oriented datasets are for example the "Solutions in Perception Challenge" dataset [9] and the RGB-D object dataset [10]. Again, evaluating the performance of a system can be done offline.

For tasks that couple perception and action, providing benchmarks that can be used offline is more difficult, and solutions usually resort to approximations or simulation. For rigid object grasping, it is commonplace to use simulators such as *GraspIt!* [11] to evaluate the quality of grasp points selected by an algorithm.

In [12] the authors construct a large database of sensor measurements, labeled as graspable or not after attempting about a hundred simulated grasps automatically using *GraspIt!*. Then, a new candidate grasp is compared to the constructed database, and its graspability is determined depending on the labels assigned to its nearest neighbors.

In [13], a dataset with thousands of synthetic objects and hundreds of thousands of simulated possible grasps with various robotic hands is evaluated with *GraspIt!*. Data in this dataset can be used to evaluate grasping algorithms by comparing the proposed grasps with the ones in the dataset.

Ciocarlie et al. [14] propose a dataset of 3D models of everyday objects graspable by a parallel jaw gripper. Acceptable grasp points for the objects have been annotated by testing them with a simulator. Additionally, the authors provide a dataset of robot sensor recordings collected while attempting 490 grasps of the objects, with ground truth information of object type and pose. This information is then used in the paper to empirically assess how well the quality metric given by *GraspIt!* works in practice, and it can be used in general to evaluate grasp planning and execution.

Ulbrich et al. [15] discuss the lack of standardized benchmarks for grasping related tasks in the robotics community and the fragmentation existent in grasping research in terms of environments, robots and software. As a solution they propose the OpenGRASP Benchmarking Suite for comparative evaluation of grasping algorithms (for rigid objects) in a simulated environment.

In the case of perception for manipulation of highly deformable objects such as textiles, simulation becomes much more complex, given the increase in degrees of freedom of the state space, and therefore there are no databases similar to those for rigid objects.

Consequently, state-of-the-art methods that deal with textile object manipulation [16], [2] evaluate their methods experimentally in a limited number of trials.

Note that we are not including in the previous group, work in perception of textile objects as those of Miller et al. [17] or Ramisa et al. [18] since, although their domain of application is similar, the task addressed does not directly involve manipulation, but just perception, and therefore can be easily evaluated offline in a similar way as other perception algorithms. A somewhat mixed case is that of Willimon et al. [19], where the authors propose a method for interactive classification of clothes perceived when hanging from the manipulator, which can interactively rotate the cloth to help its classification. However, to evaluate this work rotations are simulated by making available to the classification module additional, pre-acquired, pictures of the object of interest at different angles.

III. Characterization of complex systems for textile manipulation

It is accepted that new approaches involving manipulation skills should contain exhaustive experimentation [20], and this becomes specially true when reliable simulation is difficult. To make an experiment replicable, there is a lot of data that should be recorded in the set-up, (e.g. camera intrinsic parameters, in-focus range, allowed precisions of robot controls, gripper close force thresholds, camera-base and tcp-hand calibration matrices, planner rules and rewards) and during the experiment (e.g. acquired images, actions to be performed at each execution step, list of positions sent to manipulator controller, success of action/failure detection, time-stamps)

This paper deals with *perceptions for manipulation*, i.e. perceptions that are designed to serve in manipulation operations. In such case, a manipulation action has to be necessarily performed to validate the perception. In the context with deformable clothes this raises the problem that the initial cloth configuration cannot be exactly replicated for each experiment.

On top of that, the processes that are taken into account are often non-deterministic due to the uncertainty in the perceptions and the actions carried out as a result of those

[1] http://mobilemanipulationchallenge.org/

perceptions. Consequently, trying to replicate one unique execution is unlikely to reproduce previous results.

Instead, we propose to statistically model the behaviour of the robot. The idea is to clearly identify the perception-action couples that constitute the skills of the robot, and find a behaviour table that includes the ratio of success when performing such perception-action.

In this paper the global task is the manipulation of textiles, and the skills that we will use to exemplify our approach consist on perception-actions that grasp a different number of objects located onto a flat surface.

To build such behaviour table we can use a brute force approach, performing several experiments for each skill, or use a guided or incremental method. There are a number of considerations that must be taken into account:

1) Initial conditions: the set-up used must be representative of the conditions at test time. Concerning the textiles, initial conditions like position, occlusions or type of wrinkles are very important. For example, if the experiment assumes that a human will manipulate the textiles prior to the robot operation, the probability of success of the skill has to be computed accordingly, using humans to set-up the initial conditions, or assuring that the robot after grasping will throw the textile onto the table with similar distributions as humans do.
2) Perceptions: the perception necessary for the associated action should be possible to compute for any scene. If this constraint cannot be satisfied, the procedure to handle a missing perception has to be clearly stated (e.g. repeat the experiment, use the most similar feature in the image).
3) Actions: in our framework the output of an experiment is defined as the number of textiles grasped depending on the number of textiles in the table. However, when actions are likely to fail for other reasons (e.g. out of working space, collision...) its influence on the success ratio for the outputs has to be specified (e.g. repeat the experiment, count as a failure).

Potentially this method for building a behavior table also enables the comparison between two different systems composed by distinct robots, actuators and perceptions. The comparison should be based on the skills the system is able to do and its performance when some perceptions are possible. The replication is then accomplished by building a system that mimics some of the required skills.

IV. Experiments

In order to give a proof of concept for the proposed method to compare cloth manipulation systems, the action-perception tables have been computed for two different robot systems. The first one uses 2 different perceptions based on *height* and a *wrinkledness* measure defined below (Fig 2), and 5 actions using a 3-finger hand in different configurations (Fig. 3). The second one uses 2 actions that are performed by a simple gripper (Fig. 4) and 5 different perceptions (Fig. 6) based on local descriptors that capture the distribution of the normals in a region around a central point, and are

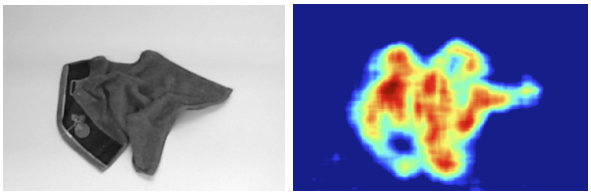
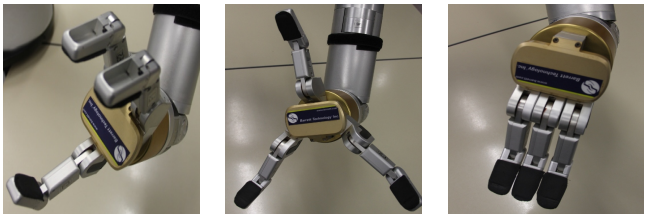


Fig. 2: Perception: wrinkledness detector delivers a heat map indicating most wrinkled parts that are good candidates as grasping points.



(a) straight (b) isometric (c) palm

Fig. 3: Different grasp configurations for System 1 that are to be combined with different grasping point selection methods.

compared to centroids previously learned using k-Means. In both experiments we use a Kinect camera as perception device. As a consequence, the vision system can use color information, as is common, but also can be combined with depth. In both systems depth is used to determine the final 3D position corresponding to the selected grasping point.

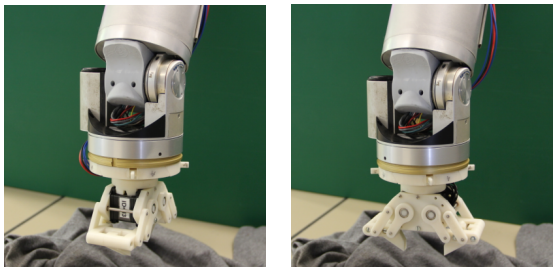
System 1: Perceptions for this approach use as grasp point selection that takes advantage of depth information:

1) The highest point of the cloth object, as typically done in related work [19], [2].
2) A wrinkledness measure [21] which is estimated for a point as the entropy of the distributions of the orientations of the normal vectors around each point in the input point cloud. The intuition is that a wrinkled area will have a mostly flat distribution, while that of a flat area will be very peaky, as all normals are parallel. In Figure 2 an example activation map of this measure can be seen.

Actions are performed with a 3-fingered hand located at different distances from the textile surface, being *deep* when the actual grasping is done 3cm below the selected grasping point and *shallow* when it is done 5cm over the selected point. The 3 fingers have been placed in the 3 different configurations shown in Fig. 3, and in 2 different pregrasping positions: completely open and half closed. The list of the 5 different grasping configurations we considered is:

1) deep/straight/pre-closed
2) deep/isometric/open
3) deep/palm
4) shallow/straight/pre-closed
5) shallow/isometric/pre-closed

System 2: Perceptions use a feature descriptor that captures the distribution of the orientations of the normals within an area around the pixel of interest. The descriptor used



(a) orientation 90 (b) orientation 0

Fig. 4: Different grasp configurations for System 2.

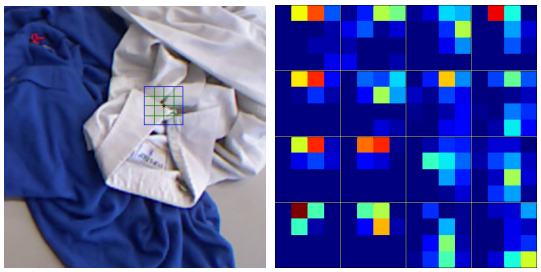


Fig. 5: Perception: The descriptor used for our second "proof of concept" system. In the first image the selected region, with the spatial subdivisions can be seen, and in the second image, there is a representation of the 2D angle histograms for each spatial subdivision (Best viewed in color).

is reminiscent of SIFT [22]: the normals of all the points inside the region of interest are quantized according to their orientation in spherical coordinates in a 2D angle histogram for each of the spatial subdivisions, and all the histograms are concatenated in a final descriptor. We used 16 spatial subdivisions and 16 bins in the angle histograms, therefore the dimension of our descriptors is 256. Figure 5 shows a representation of one feature descriptor. From a database of training descriptors, we ran k-Means with $k = 50$. Then we hand-picked one of the centroids, and we used it to rank the descriptors of a new perception. The descriptor closest to the selected centroid was selected as the grasping point. Note that we have chosen this simple method to select a grasping point for the purpose of constructing a proof of concept system, and we do not claim it to be better than any other given the perception information and features we are using.

The list of centroids considered (as seen in Figure 6), as well as their topographical interpretation, is:

1) *Planar:* Region with normals uniformly towards $-\pi$, and a pronounced elevation angle.
2) *Lower-right:* The upper-left part corresponds to a flat slope oriented towards $-\pi$ and an inclination of about 0.2π radians, while the lower-right is a wrinkled area.
3) *Upper:* The lower part corresponds to a flat slope, oriented towards π, and the upper part to a more wrinkled region, that in general has the opposite orientation.
4) *Upper-left:* Like *Lower-right* but swapping the slope and the wrinkled areas.
5) *Diagonal:* The top-left and the bottom-right parts correspond to flat slopes with an azimuth of approximately $-\pi$ and 0.75π, respectively, and a ridge crosses diagonally the descriptor area.

Id	Hand configuration	Objects grasped	Objects present 1	2	3	4
Perception: *height*						
1.1	deep straight pre-closed	0	0.10	0.10	0.10	0.10
		1	0.90	0.60	0.60	0.75
		2		0.30	0.20	0.10
		3			0.10	0.05
1.2	deep isometric open	0	0.10	0.10	0.10	0.15
		1	0.90	0.25	0.15	0.05
		2		0.65	0.70	0.60
		3			0.05	0.20
1.3	deep palm open	0	0.80	0.45	0.60	0.45
		1	0.20	0.45	0.25	0.40
		2		0.10	0.10	0.10
		3			0.05	0.05
1.4	shallow straight pre-closed	0	0.50	0.70	0.45	0.40
		1	0.50	0.20	0.40	0.45
		2		0.10	0.10	0.10
		3			0.05	0.05
1.5	shallow isometric pre-closed	0	0.40	0.55	0.45	0.45
		1	0.60	0.35	0.40	0.40
		2		0.10	0.10	0.10
		3			0.05	0.05
Perception: *wrinkle*						
1.6	deep straight pre-closed	0	0.20	0.20	0.30	0.30
		1	0.80	0.50	0.30	0.45
		2		0.30	0.30	0.20
		3			0.10	0.05
1.7	deep isometric open	0	0.10	0.05	0.20	0.30
		1	0.90	0.30	0.00	0.40
		2		0.65	0.60	0.10
		3			0.20	0.20
1.8	deep palm open	0	0.90	0.80	0.60	0.30
		1	0.10	0.15	0.30	0.45
		2		0.05	0.05	0.05
		3			0.05	0.05
1.9	shallow straight pre-closed	0	0.20	0.30	0.20	0.35
		1	0.80	0.60	0.40	0.40
		2		0.10	0.35	0.20
		3			0.05	0.05
1.10	shallow isometric pre-closed	0	0.20	0.20	0.20	0.50
		1	0.80	0.60	0.60	0.40
		2		0.20	0.20	0.10
		3			0.00	0.00

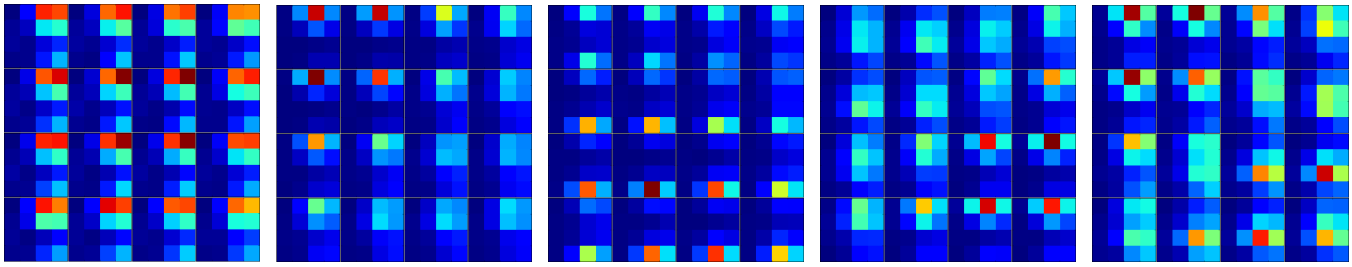
TABLE I: Computed probability distribution of success in grasping pieces of cloth in presence of different number of clothes for *System 1*.

The grasping actions are executed using a simple custom-made gripper (see Fig. 4). In contrast to the 3-fingered hand, in this case there is not much flexibility to choose different grasping strategies. As the perception descriptors are not invariant to rotations, it makes sense to have different orientations of the gripper. In this experiment, the distance is constant at the *deep* position. The list of grasping configurations is:

1) deep/orientation 0
2) deep/orientation 90

A. Comparing the two robotic systems

We have executed several grasping experiments in each of the two systems. Each one represents a solution to the same problem of grasping pieces of cloth, but they are very different in nature. The first one uses rather simple perceptions (height and wrinkle) combined with complex



(a) planar (b) lower-right wrinkle (c) upper wrinkle (d) upper-left wrinkle (e) diagonal wrinkle

Fig. 6: Perception: representation of the five centroids used in the experiments.

Id	Perception	Objects grasped	Objects present 1	2	3
Action: *orientation 90*					
2.1	planar	0	0.60	0.15	0.05
		1	0.40	0.70	0.95
		2		0.15	0.00
		3			0.00
2.2	lower-right	0	0.60	0.10	0.00
		1	0.40	0.90	0.90
		2		0.00	0.10
		3			0.00
2.3	upper	0	0.70	0.10	0.20
		1	0.30	0.70	0.70
		2		0.20	0.10
		3			0.00
2.4	upper-left	0	0.30	0.10	0.10
		1	0.70	0.80	0.80
		2		0.10	0.10
		3			0.00
2.5	diagonal	0	0.10	0.05	0.25
		1	0.90	0.90	0.70
		2		0.05	0.05
		3			0.00
Action: *orientation 0*					
2.6	planar	0	0.70	0.05	0.05
		1	0.30	0.80	0.85
		2		0.15	0.10
		3			0.00
2.7	lower-right	0	0.20	0.40	0.10
		1	0.80	0.50	0.90
		2		0.10	0.00
		3			0.00
2.8	upper	0	0.60	0.20	0.00
		1	0.40	0.70	0.90
		2		0.10	0.10
		3			0.00
2.9	upper-left	0	0.40	0.30	0.10
		1	0.60	0.70	0.80
		2		0.00	0.10
		3			0.00
2.10	diagonal	0	0.20	0.10	0.25
		1	0.80	0.90	0.75
		2		0.00	0.00
		3			0.00

TABLE II: Computed probability distribution of success in grasping pieces of cloth in presence of different number of cloth for *System 2*

grasping strategies (different finger positions and grasping distances). Conversely, the second one uses a more complex perception algorithm (with region descriptors) and a simple grasping strategy (oriented gripper).

The interest here is to build a model by measuring the success of each combination of perception-action when a different number of textiles is present on the table. The exact number of pieces is known when building the models. The output is defined as the number of textiles grasped, because the task to solve with each system can be very different e.g. to grasp only one piece of cloth at each action, or grasp as many pieces as possible.

Regarding *System 1*, Table I presents the result of the different perception-action couples. As can be observed, actions 1.1, 1.2, an 1.7 are specialized in grasping one object when one object is present, but actions 1.2 and 1.7 fail to grasp only one object when two are present. At the same time they are the best actions to grasp two objects when two or more objects are present. It could seem that using action 1.1 to grasp one object is the best choice, but observe that this action has a high probability of grasping more than one object. Depending on how failures are defined in the task, i.e. if grasping two objects is penalized, another action is preferable, like 1.3, 1.4, 1.5 or better 1.8 that can fail at grasping one object but have lower probability of grasping several. Note that, as perception is noisy and occlusions easily appear, the exact number of pieces on the table can be unknown when performing real experiments.

System 2 is described in Table II. It shows that action 2.5 is very specialized at taking only one object, and that grasping two objects with the proposed grasps is difficult but possible mainly with actions 2.1, 2.6 and 2.7. Observe that grasping three or more objects is improbable with these perception-action couples.

For an easy comparison we are considering that all perception-action couples are always possible, e.g. the highest point, or the most wrinkled point can be always computed. In a general system this does not have to be true, as some perceptions, e.g. a diagonal wrinkle, are dependent on the scene and may not appear. Here, for a given scene, we computed all the descriptors and used the one with the minimum distance to the selected centroid to execute one action, but this criterion can be different and has to be clearly described.

B. Application to decision making

As has been shown in the previous section, perception and actions in textile manipulation scenarios are intrinsically uncertain in both perception and manipulation outcomes, and additionally both are interdependent. Probabilistic planning is a useful tool that naturally suits this context as uncertainty is taken explicitly into account.

3 objects	Goal reached	Total actions	Variability in actions	Actions to first removal
POMDP partial	90%	12.1	36%	3.8
MDP partial	80%	7.8	32%	2.9
POMDP occluded	70%	15.3	41%	5.0
MDP occluded	20%	6.0	32%	3.0

TABLE III: Comparison of two different methods to solve the same task. POMDP behaves better as uncertainty is better modelled. Partially reproduced from [23].

Monsó et. al [23] used the characterization of System 1 to train two different decision-making algorithms to solve the task of removing one by one an unknown number of textiles from a pile. They used some of the presented actions to move pieces of cloth from one pile to another with the objective of isolating one piece. When uncertainty about having one piece isolated was low enough, the textile was moved to a basket. Two different approaches are compared, one using an MDP algorithm and the other using a POMDP. Perception for planning was the number of objects in each pile. The main difference is that MDP considers perceptions to be certain while POMDP does not (it considers uncertainty also in perceptions).

The experiments involved 3 pieces of cloth in two different configurations: *partially occluded* when all were visible, and *occluded* when there was a complete occlusion between cloth. Table III shows the results of both approaches. Here the comparison is fair as both decision making algorithms used perception-action skills with the same success ratio.

V. Conclusions

Accurately replicating textile manipulation experiments requires storing a huge number of parameters and may even be impossible due to the difficulty of establishing exact initial conditions. Simulation-based benchmarking, widely used in solid object grasping research, is of limited applicability given the difficulty of reliably simulating something with as many degrees of freedom as a textile object.

In this work we have introduced a method to characterize different textile grasping systems at the level of perception-action couples and skills. The method is based on constructing tables linking perceptions and actions through success probability.

Elements of the mentioned tables can be used to reason about the practical possibilities of the grasping system. This potentially allows comparison between different robotic systems. This method abstracts the underlying hardware and propose the idea that experiments are replicable in the measure that perception-action tables can be reproduced. The major drawback is that computing such tables can be time consuming and that initial conditions and perception assumptions need be clearly identified.

Furthermore, this ability to reason about the practical possibilities of the system enables the training of decision-making algorithms that take advantage of explicitly modelling the underlying uncertainty on perceptions and actions.

Acknowledgements

The authors would like to thank J.L. Rivero for his help in programming and performing the experiments, and P. Monsó for fruitful discussions about MDP and POMDP applications.

References

[1] Willow-Garage, "Beer Me Robot." `http://www.willowgarage.com/blog/2010/07/06/beer-me-robot`, 2010.

[2] J. Maitin-Shepard, M. Cusumano-towner, J. Lei, and P. Abbeel, "Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 2308–2315, 2010.

[3] M. Beetz, U. Klank, and I. Kresse, "Robotic roommates making pancakes-look into perception-manipulation loop," in *ICRA Workshop on Mobile Manipulation: Integrating Perception and Manipulation*, pp. 9–13, 2011.

[4] S.Thrun, W.Burgard, D.Fox, *Probabilistic Robotics*. MIT Press, 2005.

[5] T. Lang, M. Toussaint, "Planning with noisy probabilistic relational rules," *Journal of Artificial Intel. Research*, vol.39, no.1, pp.1–49,2010.

[6] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results." http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html.

[7] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *Proc. IEEE Int. Conf. on Computer Vision & Pattern Recognition*, pp. 710–719, June 2009.

[8] M. Marszałek, I. Laptev, and C. Schmid, "Actions in context," in *Proc. IEEE Int. Conf. on Computer Vision & Pattern Recognition*, 2009.

[9] "Solutions in perception challenge." `http://solutionsinperception.org/index.html`.

[10] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multi-view rgb-d object dataset," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 1817–1824, IEEE, 2011.

[11] A. T. Miller and P. K. Allen, "GraspIt! A Versatile Simulator for Robotic Grasping," *IEEE Robotics and Automation Magazine*, vol. 11, no. December, pp. 110–122, 2004.

[12] L. Zhang and M. Ciocarlie, "Grasp Evaluation With Graspable Feature Matching," in *RSS Workshop on Mobile Manipulation: Learning to Manipulate*, 2011.

[13] C. Goldfeder, M. Ciocarlie, and H. Dang, "The Columbia grasp database," *and Automation, 2009.*, pp. 1710–1716, May 2009.

[14] M. Ciocarlie, C. Pantofaru, K. Hsiao, G. Bradski, P. Brook, and E. Dreyfuss, "Three Data Sets for Mobile Manipulation in Human Environments," *IEEE Robotics and Automation Magazine, Special Issue: Towards a WWW for Robots*, vol. 18, no. June, p. 44, 2011.

[15] S. Ulbrich, D. Kappler, T. Asfour, N. Vahrenkamp, A. Bierbaum, M. Przybylski, and R. Dillmann, "The OpenGRASP benchmarking suite: An environment for the comparative analysis of grasping and dexterous manipulation," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 1761–1767, 2011.

[16] M. Cusumano-towner, A. Singh, S. Miller, J. F. O'Brien, and P. Abbeel, "Bringing Clothing into Desired Configurations with Limited Perception," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 3893–3900, 2011.

[17] S. Miller, M. Fritz, T. Darrell, and P. Abbeel, "Parametrized Shape Models for Clothing," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 4861–4868, 2011.

[18] A. Ramisa, G. Alenya, F. Moreno-noguer, and C. Torras, "Using Depth and Appearance Features for Informed Robot Grasping of Highly Wrinkled Clothes," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2012.

[19] B. Willimon, S. Birchfield, and I. Walker, "Classification of Clothing using Interactive Perception," in *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 1862–1868, 2011.

[20] F. Bonsignorio, J. Hallam, and A. P. del Pobil, "Good Experimental Methodology guidelines v0.9,", European Robotics Network, 2008.

[21] A. Ramisa, G. Alenya, F. Moreno-noguer, and C. Torras, "Determining Where to Grasp Cloth Using Depth Information," in *Int. Conf. of the ACIA, Frontiers in Artificial Intel. and Applications*, pp.197–205, 2011.

[22] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[23] P. Monsó, G. Alenyà, and C. Torras, "POMDP Approach to Robotized Clothes Separation," in *submitted*, 2012.