Benchmarking cloth manipulation using action graphs: an example in placing flat

Irene Garcia-Camacho, Júlia Borràs and Guillem Alenyà

Abstract— **Benchmarking robotic manipulation is complex due to the difficulty in reproducing and comparing results across different embodiments and scenarios. Cloth manipulation presents additional challenges due to the complex object configuration space. Traditional cloth manipulation papers do not have well defined metrics to evaluate the success of a task or the quality of the result, and are tailored to each evaluation. In this paper we propose to evaluate cloth manipulation segmenting a task into steps that can be evaluated independently, and to study how their success measures influence in the next segment and relate to task as a whole. In particular, we study a popular task such as placing a cloth flat on a table. We propose a benchmark with simple but continuous evaluation metrics that explore the influence of grasp location into the quality of the task. Our results show that grasp location doesn't need to be precise on corners, that quality measures focused on evaluating different cloth parts can enlighten issues to solve and that success definition of a segment has to consider its influence on the ability to perform successfully the next segment of action.**

I. Introduction

A manipulation task is usually composed of several steps or sub-tasks. When benchmarking manipulation, there is a discussion on whether to evaluate the performance of a task as a whole or as each of the components required to accomplish it [1, 2]. The vast majority of the literature usually propose binary success/fail metrics for evaluating the final result of the task, without any analytical evaluation of their phases [3, 4].

In our early research [5] we already considered the need of evaluating a manipulation task not only as a whole but also as a sequence of its steps. With this aim, we divided both tasks in phases and evaluated them individually. Two limitations where found: first, the division in phases was ad-hoc and difficult to generalise. In this paper we propose the use of an action graph [6] to identify the phases. Second, the proposed metrics concerning the outcome of the intermediate phases were binary and this is quite uninformative. Here, we make an effort to provide a continuous quality measure.

As any complex system, cloth manipulation tasks can be accomplished with a great variety of strategies that may involve variate grasp sequences and require very different skills to accomplish the tasks. In this paper, we claim that it is important to be able to evaluate both. To do so, it becomes very important to have a clear model of what a task is and what it involves, to be able to model different strategies following a common framework that enables comparison of both the parts and the whole.

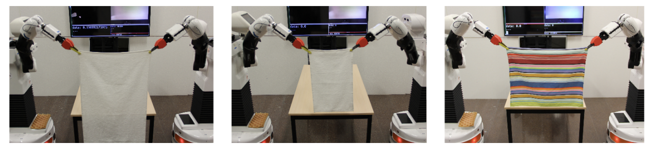

Fig. 1: Placing flat setup. From left to right, robots grasping objects A (large towel), B (small towel) and C (kitchen cloth), respectively.

In one of our previous works we defined a scene state that segments each manipulation task in clear steps [6]. With the idea of studying how the quality measure of one step influences the next one, in this paper we study the task of placing a cloth flat on a table, a sub-task often required for many tasks such as folding, making a bed or putting a tablecloth. We assume that the robot already has grasped one corner and reaches the second corner performing an edge tracing over the edge. Based in our experience [7], we now that autonomously detecting when to stop the tracing motion to prevent the cloth to slip out of the grasp is not trivial, specially if we need to reach exactly the corner. However, we have identified from our data obtained from human observation in [6] that reaching the very corner may not be necessary. In this paper we present a study on how relevant it is to reach the corner of the cloth to be able to perform the task of placing flat successfully. We will see how the quality measures of both consecutive steps are not independent. This is common in almost all manipulation tasks.

The main contribution of the paper is the proposal of measures to evaluate cloth manipulation tasks and a study on how the influence of quality of the manipulation steps can affect the quality of the successive steps in a task.

II. The Benchmark

A. Setup description

Any bimanual robotic system can be used for performing the task and applying the benchmark. In our case we used two TIAGo robots placed facing each other one at each side of the table so that both robots can spread larger cloths without the need of using the mobile platforms.

The objects used are shown in Fig. 1 and their reference can be consulted in the related website[1]. It includes different rectangular cloths (two of them already used in [5] for

This work receives funding from the Spanish State Research Agency through the BURG project (CHIST-ERA - PCIN2019-103447) and the María de Maeztu Seal of Excellence to IRI (MDM-2016-0656).

Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain
{igarcia, jborras, galenya}@iri.upc.edu

[1] http://www.iri.upc.edu/groups/perception/#PlacingFlat

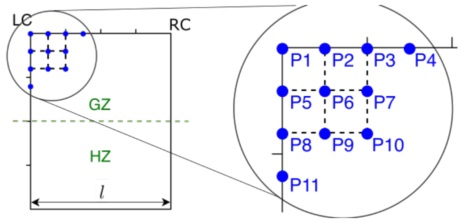

Fig. 2: Grasping points labelling. LC and RC stands for left and right corner and GZ, HZ are for grasping zone and hanging zone.

benchmarking) and a table, which is necessary as support the object. Instead of specifying a specific table model we specify an interval of sizes. The limits are necessary to constraint the workspace and normalize the difficulty of the task.

B. Initial grasping points

The benchmark we are presenting focuses on quantifying the quality of the spread cloth according to the grasp from which it starts.

Following notation in Fig. 2, let l be the length of the cloth edge between the two robotic arms. Then, we defined the grasping area of the fabric obtained between the left corner of the fabric (LC) and the 25% of l in each of the adjacent edges of LC. This area is then divided in a grid of 9 points, which vertices correspond to the grasping points. We also considered adding two additional points at a distance of 37.5% of l. Therefore, each of the grasping positions of the cloth is labelled as the distance to the adjacent edges of the goal corner as p=(x,y), where x and y can be $0\%l$, $12.5\%l$, $25\%l$ or $37.5\%l$.

C. Evaluation metrics

This benchmark aims to offer metrics to evaluate the result of spreading a cloth on a table based on simple tools, providing a vision algorithm which using simple hardware can autonomously give a measure of the quality of the placement according to a predefined template, giving also insight of where the error is in the cloth. The template used to compare the baseline results is obtained as a successful placement executed by a human which has then flattened the cloth.

The algorithm provides a percentage of the error of the baseline result performing shape matching with the template using zenith images of the entire spread cloth. The shape matching is performed by segmentation of the object from the background, finding the contour of the cloth and applying a rotation and translation to fit it with the template contour. Fig. 3 shows four examples of baseline results with their input image and the output of applying the vision algorithm. The advantage of this vision algorithm is that it does not need depth information or special lightning conditions as other works [8, 9] to detect the deformations that appear in the placed cloth produced by wrinkles or bends. Therefore, it only needs a basic RGB camera for taking the images as it is based on image difference. The code is publicly available in the website[1].

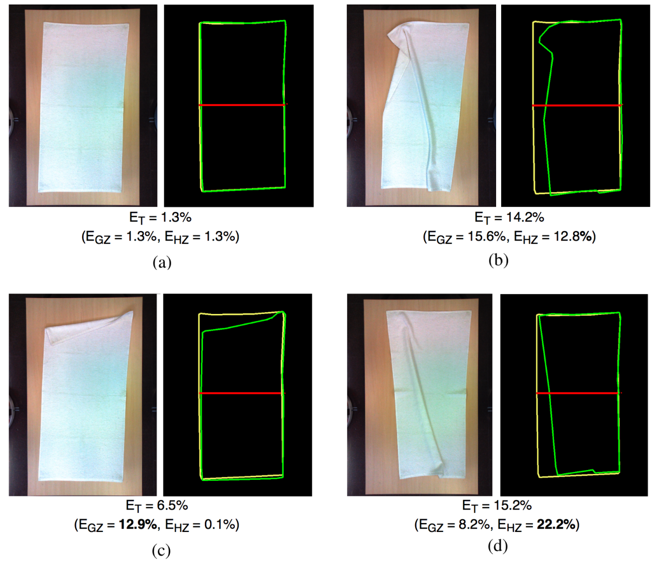

Fig. 3: RGB input images and shape matching (template contour and baseline result contour in yellow and green, respectively).

Having px_t as the total number of pixels of the successful placement template and px_b as the total number of pixels of the baseline placement, the proposed percentage error of the spreading is:

$$E_T = 1 - \frac{px_b}{px_t}. \tag{1}$$

After closely examining the images of the placement results of all the trials, we observe a difference on the types of deformations that appear on each zone of the cloth, being more present the bends on the grasping zone (GZ) and wrinkles on the hanging zone (HZ). For instance, Fig. 3c presents a bent corner in the grasping zone, what concentrates the error on that zone. Instead, in Fig. 3d we observe a wrinkle that occupies the whole hanging zone causing the error in that zone to be much higher than in the grasping one. To capture these differences, we propose a second metric, which provides the percentage error differentiating between two zones of the cloth: grasping zone (GZ) and hanging zone (HZ). Being px_{tgz} and px_{thz} as the number of pixels of the placed template of the GZ and the HZ, respectively, and px_{bgz} and px_{bhz} as the number of pixels of the baseline placement result in each of the zones, the proposed percentage error of each zone is

$$E_{gz} = 1 - \frac{px_{bgz}}{px_{tgz}} \text{ and } E_{hz} = 1 - \frac{px_{bhz}}{px_{thz}}. \tag{2}$$

Example evaluations for our baseline are shown in Fig. 3.

III. Baseline system evaluation

To provide a baseline to evaluate the benchmarking we designed a bi-manual trajectory to extend the cloth on top of the table. It describes a linear path from the starting pose (hanging cloth) to the goal pose (on top of the table). The trajectory has different final goal poses according to the size of the fabric in order to always place the cloth in the center of the table. The trajectory was defined to place

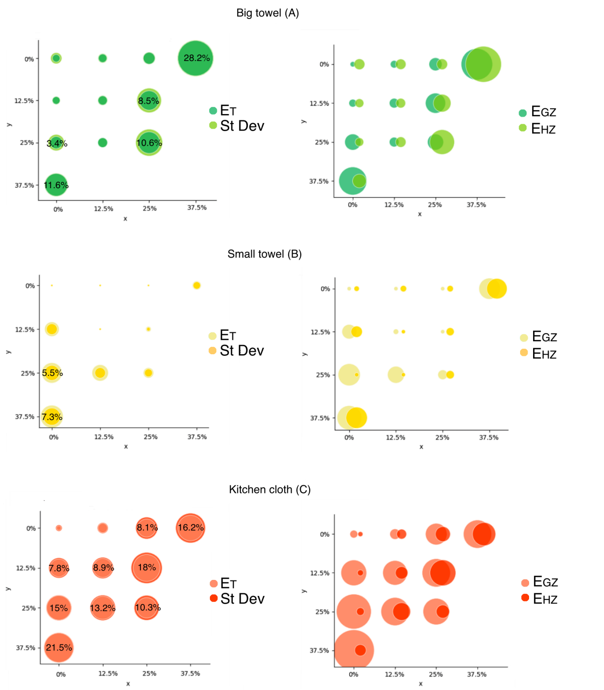

Fig. 4: Left column: Mean E_T error and standard deviation. Right column: Mean E_{gz} and E_{gz}.

successfully the objects when grasping the exact corner point $P = (0\%l, 0\%l)$. When grasping the other points, lateral and vertical offsets were applied to maintain an equivalent path.

Three trials were executed for each object and each initial grasping position, what resulted in 99 repetitions. Fig. 4 represents the obtained results for the three objects using the proposed metrics E_T, E_{gz} and E_{hz}, represented in a grid of disks distributed according to the grasping point location with respect to the corner, following Fig. 2. On the left column, each disk represents the mean error (E_T) of the three trials in solid color, and the standard deviation in a lighter color. As we expected, it is clear how as the grasping point moves away from the corner, the error on the placement increases. Nevertheless, for objects A and B this error is very small in most of the points and does not correspond to any particular wrinkle or bent corner, having only a bigger error when the cloth is grasped the furthest to the corner. With respect to object C, we observe that the error is bigger as soon as we are not in the exact corner. This is due to the characteristics of the fabric, since it is less rigid than the towels and therefore it tends to generate wrinkles and bendings more easily. We also see that the standard deviation on the trials error is generally very small. We also observe that, generally, the error is slightly minor when the grasping distance to the corner moves in the horizontal direction compared to the vertical one.

On the right column in Fig. 4, the couple of disks in each position represent the mean E_{gz} and E_{hz}. We can observe that for the towels, the error is similar at both zones, while for the kitchen cloth there is a clear difference in both zones, having a higher percentage of error on the side where the robots are grasping the cloth (GZ). This occurs because flexible textiles tend to wrinkle more with contacts, while for the more rigid ones the errors on the grasping points can also translete to the hanging zone. In future works, we will analyse if these results can be improved maintaining tension between the grasped points during the placement [10] or using a trajectory that drags the cloth against the edge.

IV. Conclusions

With this work, we can conclude that it is important to evaluate the quality of the skills performed in each of the steps of a manipulation task since they will affect the final result of the whole task. In addition, using continuous measures of error instead of binary ones allows giving a margin on the definition of success but also can help anticipating whether a task can be performed successfully with a measure of uncertainty, helping on the decision process of possible strategies for error recovery. In the task presented in this paper, we identified a clear dependency between the initial grasping position and the placing result but also that we have margin to perform placing tasks of cloths without needing to arrive to the exact corner, what is complex and requires complex vision algorithms and precision in grasping.

We also show how metrics that evaluate the quality of the spread cloth differentiating by zones can help to identify where the errors are occurring with more frequency and therefore give an idea of how to improve the strategy implementation. It also would be useful to combine this type of metrics with more complex ones that provide the type of deformation found, giving information for instance, whether there is a wrinkle or a bend, since both errors require different strategies for removing them.

References

[1] G. Grunwald, C. Borst, and J. Zöllner, "Benchmarking dexterous dual-arm/hand robotic manipulation," in *IROS Workshop on Perf. Eval. and Benchmarking for Intelligent Robots and Syst.*, 2008.

[2] R. Dillmann, "Ka 1.10 benchmarks for robotics research," *Project funded by the European Communityunder the "Information Society Technologies"Programme (1998-2002)*, 2004.

[3] H. Yuba, S. Arnold, and K. Yamazaki, "Unfolding of a rectangular cloth from unarranged starting shapes by a dual-armed robot with a mechanism for managing recognition error and uncertainty," *Adv. Robot.*, vol. 31, no. 10, pp. 544–556, 2017.

[4] A. Doumanoglou, A. Kargakos, T. Kim, and S. Malassiotis, "Autonom. active recognition and unfolding of clothes using random decision forests and prob. planning," in *IEEE Int. Conf. Rob. Aut.*, 2014.

[5] I. Garcia-Camacho, M. Lippi, M. C. Welle, H. Yin, R. Antonova, A. Varava, J. Borras, C. Torras, A. Marino, G. Alenyà, and D. Kragic, "Benchmarking bimanual cloth manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1111–1118, 2020.

[6] J. Borràs, G. Alenyà, and C. Torras, "Encoding cloth manipulations using a graph of states and transitions," *arXiv:2009.14681*, 2020.

[7] I. Geer, M. Maceira, J. Borràs, C. Torras, and G. Alenyà, "Interaction identification through tactile sensing during cloth manipulation using a 3-axis touch sensor," in *IROS Workshop on Robotic Manipulation of Deformable Objects*, pp. 1–6, 2020.

[8] L. Sun, G. Aragon-Camarasa, S. Rogers, and J. P. Siebert, "Accurate garment surface analysis using an active stereo robot head with application to dual-arm flattening," in *IEEE Int. Conf. on Robotics and Automation*, pp. 185–192, 2015.

[9] Y. Li, X. Hu, D. Xu, Y. Yue, E. Grinspun, and P. Allen, "Multi-sensor surface analysis for robotic ironing," *IEEE Int. Conf. Rob. Aut.*, 2016.

[10] A. Lee, H. Lu, A. Gupta, S. Levine, and P. Abbeel, "Learning force-based manipulation of deformable objects from multiple demonstrations," in *IEEE Int. Conf. on Rob. and Autom.*, pp. 177–184, 2015.