

Insights for robotic cloth manipulation: A comprehensive analysis of a competition-winning system

*International Journal of Advanced
Robotic Systems*

March–April 2025: 1–10

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17298806251322582

journals.sagepub.com/home/arx



Victor-Louis De Gusseme , Remko Proesmans , Thomas Lips ,
Andreas Verleysen and Francis wyffels

Abstract

Robotic cloth manipulation will be important for assistive robots. To thoroughly evaluate progress in this field, a Cloth Manipulation and Perception Competition was organised at IROS 2022 and ICRA 2023. In this article, we present the system that won the folding track at IROS 2022 and the folding and unfolding tracks at ICRA 2023. By combining visual and tactile information with engineered motions, we built a system that can generalise to a range of patterned towels made from various materials, as required for the competition. We describe our system and its limitations, which we relate to future work with the goal of creating systems that can deal with any cloth, robot, or surface.

Keywords

Cloth manipulation, tactile sensing, dual-arm manipulation, household robotics, benchmarking

Date received: 29 July 2024; accepted: 21 January 2025

Topic: Robot Manipulation and Control

Topic Editor: Giuseppe Carbone

Associate Editor: Ankita Rajput

Introduction

Clothing is omnipresent in everyday human environments like households and healthcare facilities. Interacting with cloth is hence an important skill for assistive robots.^{1,2} However, due to high deformability, complex dynamics, self-occlusions, and a high-dimensional state space, robotic cloth manipulation, and perception remain challenging.³ In the subdomain of robotic laundering, various systems have been created over the years, focussing mainly on unfolding^{4–10} and folding.^{4,5,7} Significant progress has been made, but none of these systems approach human performance for arbitrary cloth pieces.

Comparing systems objectively is difficult. First of all, the specific cloth items used for evaluation have a significant impact on the performance due to the enormous diversity in possible shape, material properties, and appearance.¹¹ In addition, many aspects of the environment have an impact on

system performance, such as the roughness of the working surface.⁶ Furthermore, the procedure to obtain an initial cloth state can influence the outcome.¹¹ Indeed, some objective metrics have been widely adopted, like the coverage metric,^{6,12} but these do not perfectly correlate with the human notion of good folding or unfolding.

To better compare different systems and to advance progress through standardisation, Garcia-Camacho et al.¹¹ have defined an initial benchmark, consisting of a

AI and Robotics Lab (IDLab-AIRO), Ghent University—imec, Ghent, Belgium

Corresponding author:

Victor-Louis De Gusseme, AI and Robotics Lab (IDLab-AIRO), Ghent University—imec, Technologiepark-Zwijnaarde 126, B-9052 Ghent, Belgium.

Email: victorlouis.degusseme@ugent.be



standardised set of rectangular cloth items, a set of tasks and some suggested evaluation metrics. Based on this work, they have organised cloth manipulation competitions at IROS 2022 and ICRA 2023.¹³ Both competitions had three tracks: Perception, unfolding, and folding. Each participant chose which tracks to compete in and received the same standardised set of rectangular cloth items. The participants developed a solution using their own robot platform, and live demonstrations for the judges were held remotely.

In this article, we describe and formally evaluate our system that won the folding and unfolding track of this competition at ICRA 2023, as well as the folding track at IROS 2022 in an earlier stage. Our system is based on prior work on synthetic data generation for clothing,¹⁴ simulation-based fold optimisations,¹⁵ and the use of tactile sensing for unfolding.⁹ In particular, our work adds to the very limited set of fully integrated crumpled-to-folded cloth manipulation pipelines in literature.^{4,5} We believe this full integration gives us a more complete overview of possible failure modes, which we describe in detail. This article aims to share our insights on the challenges of robotic cloth manipulation and to relate these challenges to future research directions that can lead us to robust and generic robotic laundering.

In summary, our core contributions are as follows:

- **Novel end-to-end folding approach with tactile-based unfolding:** Tactile-sensing based edge tracing, used to simplify the unfolding stage, is seamlessly integrated into a complete end-to-end cloth folding pipeline.
- **Well-documented evaluation procedure on diverse towels:** The capabilities of the system are rigorously evaluated on a diverse set of towels, including larger sizes that pose significant challenges. This thorough testing emphasises the importance of evaluating cloth manipulation strategies on a wide range of items to ensure robustness against real-world variations.
- **Broadly applicable failure mode analysis:** The article offers an in-depth analysis of failures encountered during the end-to-end unfolding and folding process. These insights are valuable not only for refining our system but also for researchers using alternative strategies (e.g. learning-based approaches), as they highlight fundamental challenges inherent to cloth manipulation.

Related work

Unfolding

Cloth unfolding or smoothing is the task of bringing a piece of cloth from an arbitrary configuration into a flattened or canonical configuration.^{6,11} Initial work tackling this problem attempted to detect suitable grasp points on hanging cloth items using a dual-arm robot. The first grasp point can be found by detecting the lowest point

after lifting.^{4,5} The second grasp point was determined using geometric reasoning over folds and edges⁴ or by using machine learning techniques to predict manually annotated grasp locations.^{5,16}

More recently, tactile sensors have been used to complement these visual algorithms in an attempt to achieve better generalisation: Instead of directly finding the grasp point, which might be occluded or unreachable, these systems attempt to trace a cloth edge to reach an adjacent corner of the cloth item.^{9,10,17,18} We build on our prior work⁹ in this area for the unfolding strategy. Others have abandoned corner detection in hanging cloth and instead use learning-based approaches to unfold clothes based on visual information using planar drag action primitives,¹⁹ fling primitives⁶ or a combination thereof.^{7,8} These approaches show great potential, but so far, they have not shown the level of generalisation required to tackle the cloth manipulation competition. The metric that is typically used for unfolding is coverage, or the ratio between the flattened area of the cloth item and the actual area after unfolding.⁶

Folding

Cloth folding is usually interpreted as folding a flattened piece of cloth into a user-defined configuration. Most works use user-defined fold lines on a geometric template of the cloth to define this configuration.^{5,7,8,15,20–22} References^{4,5,7,8,15,20} use an open-loop control policy that estimates the cloth state once and then executes one or more folds. To do so, the user-defined fold lines are converted to a number of grasp poses and trajectories conditioned on a geometric template. Both the path and velocity profile of these trajectories are important¹⁵ and various fold trajectories have been proposed, such as g-folds,²⁰ circular folds²³ and bezier curves,^{15,24} which all leave one side of the cloth on the table. Reference⁴ uses a different technique and lifts the cloth entirely before folding.

Estimating the cloth state prior to initiating a fold sequence is performed using segmentation and template matching^{5,7} or by localising a number of semantic keypoints.^{8,14} Various works aim to learn a closed-loop policy that omits the need for explicit state estimation or trajectory generation.^{21,22,25} Although such reactive policies have the potential to be more robust and precise, they tend to require a large number of interactions and have not shown sufficient generalisation to tackle the diverse set of cloth items used in this competition.

Measuring the performance of a folding system is complicated.²⁶ Coverage can be used but does not provide information about more fine-grained aspects that determine a good fold, such as wrinkles.¹¹ Determining metrics that better correlate with the human notion of a good fold is an open problem. Verleysen et al.²⁷ have made some progress in this direction, but the learned metric does not generalise to arbitrary embodiments.

Method

In this work, we describe a pipeline of robot actions that serves to first flatten crumpled towels and subsequently fold them twice. A typical run of the pipeline is shown in Figure 1. Our method starts the unfolding by tracing with infrared (IR) tactile sensors. We then interleave visual keypoint detection and several engineered motion primitives to complete the unfolding, align the towel, and fold it. The keypoint detection, unfolding, and folding are described in more detail in the following subsections.

Keypoint detection

During folding and unfolding, we use a deep learning-based keypoint detector to localise specific semantic locations on pieces of cloth directly. This marks a shift from traditional methods that often rely on template matching for state estimation of cloth on tables.²⁸ This method offers more versatility because this keypoint detector can be applied to cloth

in more complex configurations, such as when it is hanging or significantly deformed.

We formulate the keypoint detection problem as in the work of Lips et al.¹⁴: All keypoints are converted into Gaussian blobs on a target heatmap, and the model is trained to regress these heatmaps using a pixel-wise binary cross-entropy loss. The network has a U-net-inspired architecture with a pre-trained MaxViT²⁹ model as an encoder. This model was trained on a dataset of 10,000 synthetic RGB images. An improved version of the synthetic data pipeline used to generate these images was later published.³⁰ For towels, we take the four corners as the semantic locations of interest. The keypoint detector predicts spatial heatmaps from which the keypoints can be extracted, as shown in Figure 2.

Unfolding

The goal of the unfolding track of the competition is to flatten towels that are either crumpled or neatly folded



Figure 1. A typical run of the unfolding and folding pipeline shown as 16 key moments, grouped in four stages. Each row corresponds with a stage in the pipeline.

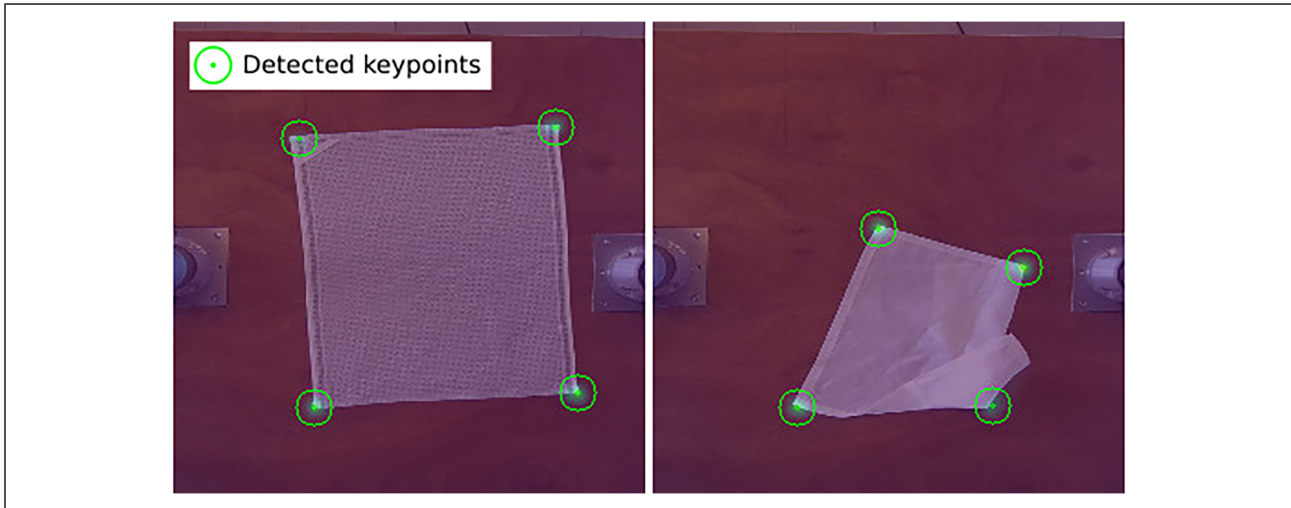


Figure 2. Two examples of images overlaid with the heatmap from the keypoint detector. The keypoint detector, trained on images of flat towels (left), also produces useful detections on deformed towels (right).

completely. The unfolding is considered successful if the surface area of the towel is maximised. The position and orientation of the flattened towel do not matter as long as it is laid entirely on the table.

The unfolding methodology is based on our UnfoldIR pipeline,⁹ where edge tracing is performed using end-effector servoing based on tactile sensor feedback. This means the robot’s end-effector makes small, incremental movements, with the direction of each movement determined by the sensor readings. UnfoldIR lifts the crumpled cloth with a top-down grasp at the highest point of the cloth (Figure 1.1), and finds a first corner through lowest point detection (Figure 1.2), a common heuristic in prior work.^{4,5} Then, it relies on IR tactile fingertips to sense cloth contours between the fingers. The fingertips are shown in Figure 3, along with an exemplary sensor readout. One fingertip emits IR light, the other receives it. A grasped layer of cloth reduces the light captured by the receiver grid, and an adaptive threshold determines the edges of the dark cluster, localising the cloth in the sensor readout. By controlling the position of the dark cluster centre (blue circle in Figure 3), the robot arms can trace cloth edges. A corner function is fit to the edge of the dark cluster, and once its score is high enough, the system assumes that a second corner has been found between the tactile fingertips (Figure 1.4). If an edge tracing attempt fails, UnfoldIR can recover and retry without laying down the cloth.

In this work, we adapted UnfoldIR in several ways to handle the competition towels. Edge tracing is now done horizontally instead of vertically so that larger cloth can be unfolded. Secondly, the tactile fingertips were redesigned to be compliant so that they can be used for slide grasps as required for our folding pipeline. The fingertip designs have been open-sourced in Proesmians et al.³¹ Finally, in the interest of execution time, the system no

longer engages in recovery attempts until perfect edge tracing and corner detection have been achieved. Instead, one “best-effort” tracing sequence is executed until either a corner is detected or the cloth is about to slip away. In both cases, tracing ends and the cloth is laid down. Note that tracing can expose cloth edges and corners even when a fold is traced instead of a cloth edge, as in Figure 1.5. Subsequently, the two corners of the exposed edge are grasped in Figure 1.6, such that lifting the cloth unfolds it entirely, Figure 1.7–8.

Folding

The goal of the folding track of the competition is to fold towels in half twice. The first fold line has to be orthogonal to the long edges of the towel, and the second fold line has to be orthogonal to the first. The folding always starts from a towel that is completely flattened and spread out on the table, with a slight rotation or translation.

The core of our folding method is dual-arm circular arc trajectories. To perform these arcs within the constraints of our workspace, several auxiliary repositioning motions are required: Lifting and laying back down, dragging, and rotating the towel, see Figure 1.9–16. All motions during the folding stage are preceded by slide grasps, where we use the compliance of the fingertips to slide under the edges of the towel with minimal deformation. Another common property of all folding motions is that the distance between the grippers stays constant during the execution of each motion so that the tension in the cloth is maintained. All folding trajectories are position-based and executed with an open-loop controller.

To initiate folding, we detect the four corners of the towel. If the towel is misaligned or off-centre, it is repeatedly grasped by a short edge, lifted, and laid down (“Alignment” stage in Figure 1) to improve its positioning.

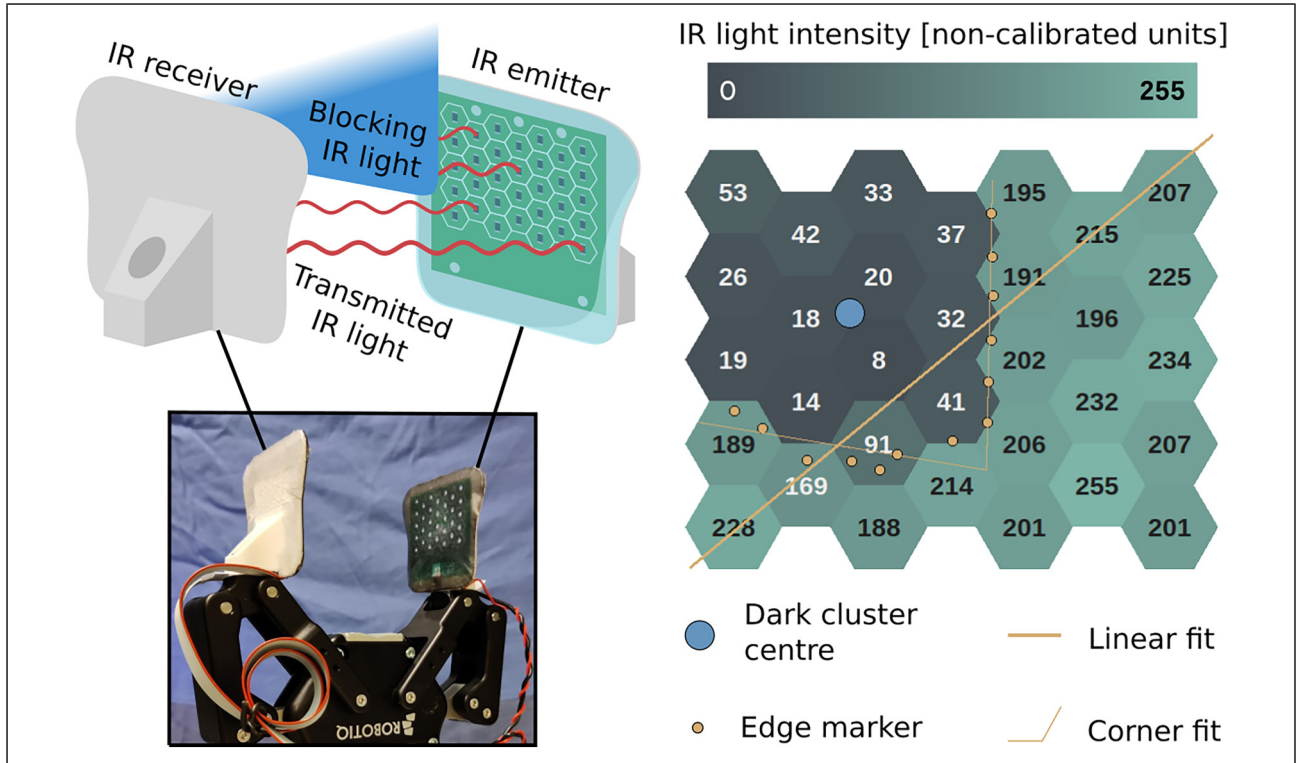


Figure 3. UnfoldIR⁹ tactile sensors. One fingertip emits infrared (IR) light, the other receives it. A grasped layer of cloth reduces the light captured by the receiver grid, as can be seen in the exemplary sensor readout (right).

We have found that, once the towel is sufficiently centred and aligned, its behaviour is predictable enough so that no further feedback is required during folding. Hence, all subsequent motions are executed open-loop. For the first fold, we calculate a fold line and grasp poses based on the detected corners. Fold arcs are constructed by rotating the grasp positions around the fold line, and end 4 cm higher than the grasp positions. At the end of the fold arcs, we open the grippers only partially and move backwards slowly a few centimetres. This prevents the corners from folding over when falling out of the gripper and can even resolve small folds introduced during grasping. Afterwards, three motions (Figure 1.13–15) position the towel such that the second fold can be performed along the same direction as the first. The first of the three motions is a dual arm drag that recenters the towel. This is followed by two rotation motions, the first by 60° and the second by 30°. Between the rotations, we regrip the towel at different locations to avoid collisions and kinematic failures for all towel sizes in the competition. Finally, the second fold is performed, resulting in Figure 1.16.

Hardware

Our setup consists of two UR5e arms spaced 90 cm apart and equipped with Robotiq 2F-85 grippers. One gripper features tactile sensors, the other has custom fingers to

improve the grasp for tracing as explained in Proesmans et al.⁹ Two ZED 2i colour and depth (RGB-D) cameras are used. One is mounted 120 cm above the robots, providing a top-down view for keypoint detection. The second camera is mounted on the side to provide a lateral view. It is mounted vertically to maximise the vertical field of view. From this camera, the depth images are used during unfolding for highest and lowest point detection.

Evaluation

During the live demonstration for the judges of the competition, six unfolding and folding trials are executed. Three of those trials start from a crumpled state and three from a folded state. To more thoroughly characterise our system in this article, we ran 60 evaluation trials from randomised crumpled states with the same 12 towels used for the competition, shown in Figure 4. As opposed to the competition, we do not include any trials starting from a folded state because our unfolding system is not designed for this: A top-down grasp on a folded towel is likely to result in many layers of cloth being grasped, which breaks the lowest-point heuristic. Unfolding neatly folded cloth would require a layer singulation strategy.^{32,33}

In our experience, robotic unfolding strategies are sensitive to the initial state of the towel, so we perform five trials

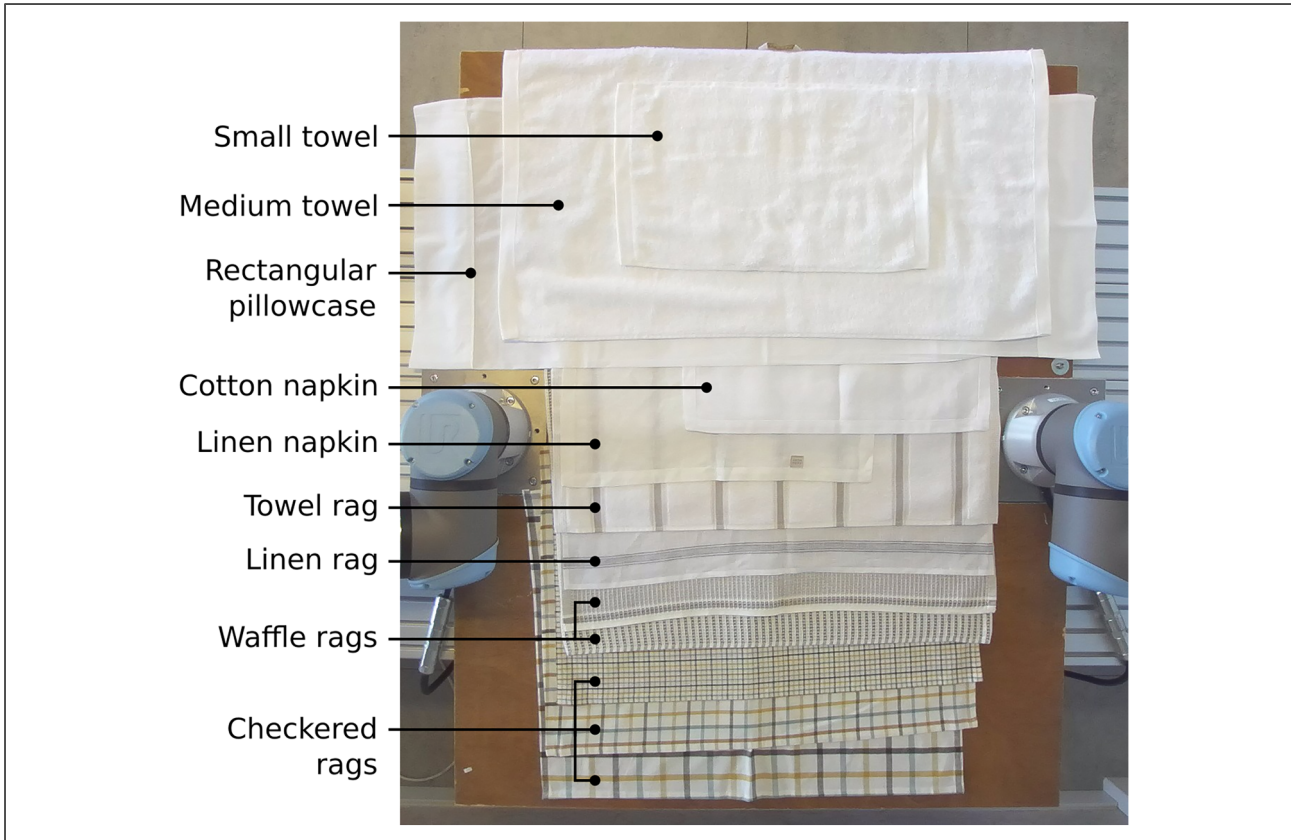


Figure 4. The 12 towels used for the evaluation from the Household Cloth Object Set.¹¹

per towel. We follow the methodology from the competition to initialise a trial: initial cloth states are classified using the visibility and location of the corners, for example, *one visible corner on the cloth and one on the table*.

Folding and unfolding performance of each trial during the competition is scored by the jury on a 100-point quality scale, with deductions for specific defects. We do not attempt to recreate this point-scoring system for our evaluation. Rather, we report the overall end-to-end success rate of our pipeline and the success rate of several key substeps in Table 1. Our focus is mainly on thoroughly describing the failures of our system and categorising them into a set of failure modes.

To allow comparison with previous work, we report the average coverage of the towels in Table 2. These coverages should, however, be interpreted with care: only trials that reached a specific point in the pipeline were included in the calculation, so we should expect an upward bias for later steps. Our average initial coverage of 28% is low and consistent with the coverages reported in FlingBot.⁶ This low initial coverage is a rough indication that the initial configurations were relatively challenging. The prime objective of our tracing is not solely increasing coverage but rather exposing a good edge to lift and achieve perfect unfolding. This is reflected in the large increase in

Table 1. Success rate of the most important substeps.

Step	Success rate
Edge tracing & unfolding (Figure 1.1–8)	23/60 (38%)
Edge tracing can start (Figure 1.3)	58/60 (97%)
Edge visible after tracing (Figure 1.4–5)	27/58 (47%)
Corner keypoints detected (Figure 1.5)	25/27 (93%)
Edge grasped for lift (Figure 1.6–8)	23/25 (92%)
Alignment (Figure 1.9–11)	18/23 (78%)
Folding (Figure 1.12–1.16)	17/18 (94%)
End-to-end	17/60 (28%)

Table 2. Coverage at the most relevant moments in the pipeline.

Step	Average coverage	Included trials
Initial (Figure 1.1)	28%	60
Tracing result (Figure 1.5)	62%	53
Lift result (Figure 1.8)	97%	23
First fold result (Figure 1.12)	51%	17
Second fold result (Figure 1.16)	27%	17

coverage from 62% after tracing to 97% after lifting. Finally, the folding decreases the coverage as expected, approximately halving the coverage with each fold.

Discussion

Trial outcomes

In 28% of runs the system performs flawlessly, producing twice-folded towels from a crumpled one. The average duration of the successful trials was 7 minutes and 17 seconds. The 43 failed end-to-end trials require further analysis.

In nearly all trials the lowest point grasp is successful and tracing can start. Only for the largest item, the 110 cm rectangular pillowcase, tracing was unable to start on two occasions due to the lowest point grasp not succeeding. Our system then relies on tracing to expose an edge of the towel (Figure 1.5), which it will subsequently grasp and lift to complete the unfolding (Figure 1.6–8). Exposing this edge was successful in about half of all runs. The 31 runs where this desired edge was not visible after tracing currently represent the main point of failure and can be divided into two categories of about equal size.

The first comprises 15 grasp-related failures, of which 4 examples are shown in Figure 5. These consist of 9 grip losses and 6 failed releases. The loss of grip occurs most during tracing and during the laying down after tracing. The failed releases happen mostly after the lowest point

grasp (i.e. the transition from Figures 1.2 to 1.3) and after laying down the traced towel (i.e. the transition from Figures 1.4 to 1.5). These failures are due to the cloth hooking around the edges of the fingertips. We attribute a considerable part of the grasp-related failures to degradation in the custom fingertips (e.g. wear of the 3D printed structure) and suboptimal finger shape. Adapting the fingertip design could significantly reduce these failures. In addition, adding tactile sensing to the second gripper could allow the system to recover from or even avoid grasp failures.

The second category consists of 16 “edge exposure” failures, two of which are shown in Figure 6. Tracing produces the best results when the tracing (see Figure 1.3) grasps only a single layer of fabric. When this is not the case, tracing still often produces results that can be unfolded entirely by the subsequent lift. However, in those 16 failed runs, the edge that we tried to grasp for lifting was not sufficiently visible. This suggests that the pretrace edge-exposure sequence from our UnfoldIR pipeline⁹ generalises suboptimally to larger cloth.

The remaining 10 failures are more diverse. Two failures can be attributed to the keypoint detector either missing or misplacing a corner keypoint. Two other failures were the failed grasping of an edge that was successfully exposed

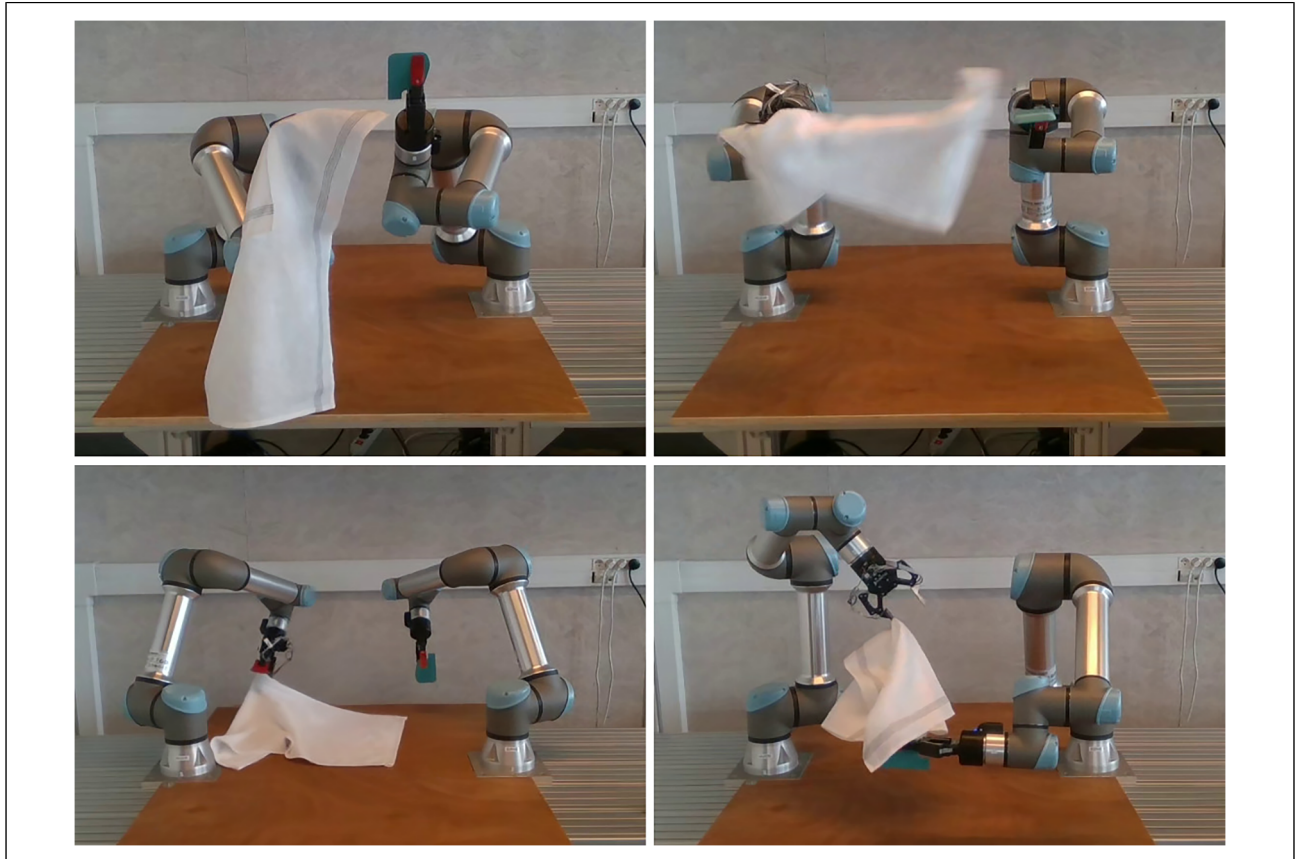


Figure 5. Four examples of grasp-related failures. Top left: Grip lost during tracing. Top right: Grip lost when laying down the result of tracing. Bottom left: Cloth sticks to gripper after laying down the result of tracing. Bottom right: Cloth does not release from the gripper during pretrace.

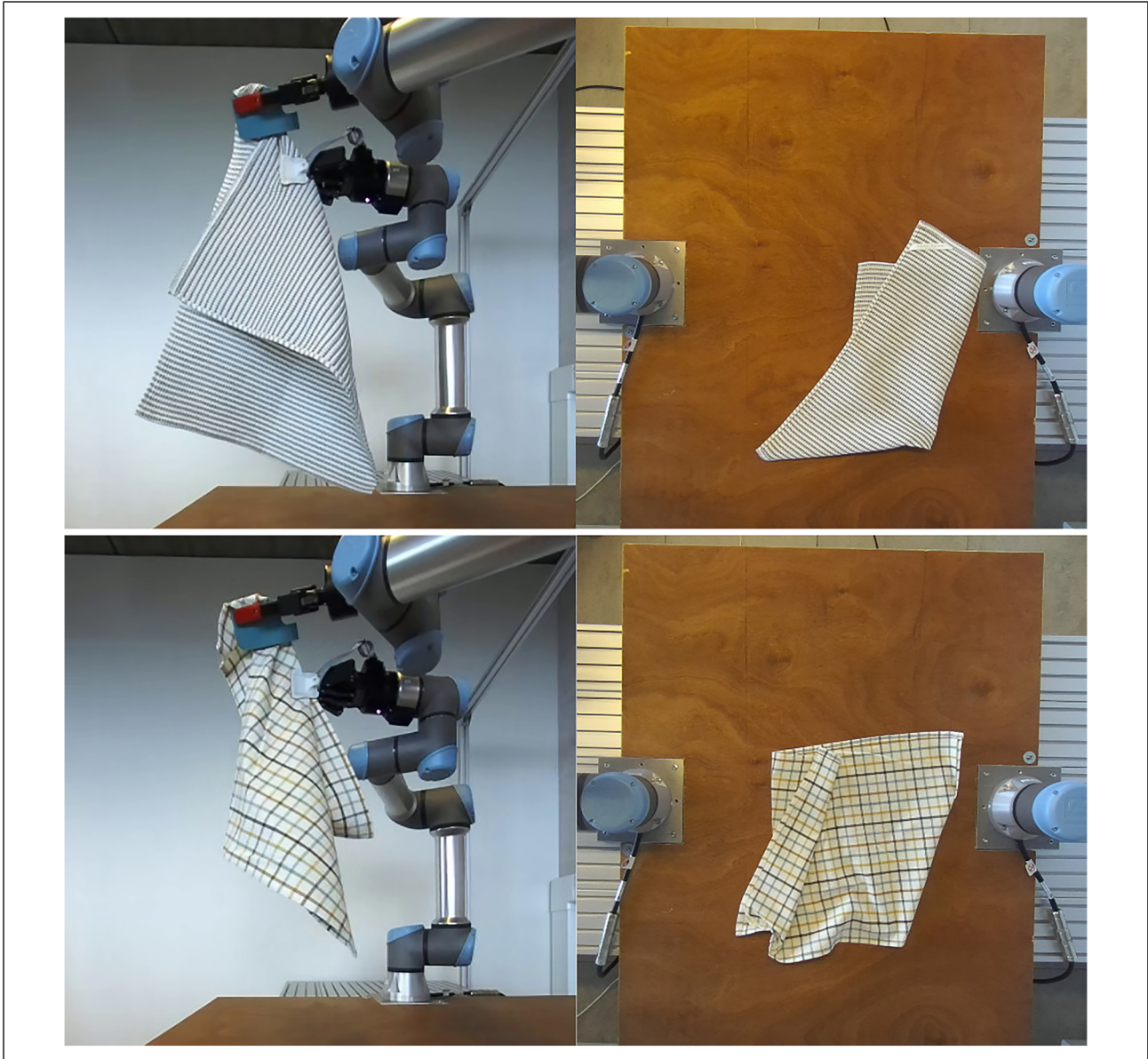


Figure 6. Two examples of edge exposure failures. On the left the grasp before tracing is shown and on the right the corresponding tracing result.

and detected on the table after the first lift. Two more were dual-arm collisions when grasping the short edge of a small towel, which is avoidable through collision checking. Another three errors are due to the lifting motions for alignment deforming the towels in unintended ways, causing subsequent steps to fail. One final failure happened during the folding, where the cloth stuck to one of the grippers after the first fold.

While our system already works well for many initial states, its end-to-end success rate is low. Previous end-to-end pipelines^{4,5} have reported higher success rates, but these cannot be directly compared for two reasons. First, the range of materials and towel sizes that these works tested on is not reported in detail. Second, the success rate for

unfolding and folding is ambiguous, there is no clear definition of which defects count as failures.²⁶ It would be interesting to update their methods with modern tools such as deep neural networks, and also to evaluate them on the standardised set of towels from the competition, enabling a one-to-one comparison. Avigal et al. recently proposed a learned approach for cloth unfolding and folding.⁷ They defined a fixed set of motion primitives, from which the optimal one is chosen given a certain cloth state. The flexibility and speed of this method is promising, but their generalisation still has to be examined further, as they evaluated only one towel, and three items in total.

This work has thus found that a hybrid tactile-vision approach to unfolding is promising, as demonstrated by

the 17 perfectly executed trials. However, our analysis also revealed a key area for improvement: A more robust grasp selection strategy is crucial for initiating the tracing process. In 31 trials, tracing was initiated but failed to lead to a graspable edge, preventing the completion of the unfolding process. We identify the development of more sophisticated grasp selection techniques, incorporating both visual and tactile information, as a critical direction for future work.

Research opportunities

We have found that a small set of motion primitives is expressive enough to complete the cloth manipulation tasks. The use of primitives limits the number of required decisions throughout the pipeline. However, executing a sequence of action primitives can lead to compounding errors. It is therefore crucial that each primitive is sufficiently robust to variability. To achieve this, we advocate for a pervasive integration of both visual and tactile feedback. Recent work has already made some progress in this regard.^{10,18} For our pipeline, improving the pretrace sequence with visual edge detection similar to Sunil et al.,¹⁰ could greatly increase the quality of the grasp for the tracing to start.

Even more important than incorporating feedback in the individual primitives is deciding when and how to use them properly. In our current pipeline, the sequence of actions is mostly predetermined. Though we spent significant effort in engineering this state machine, it still lacks flexibility to, for example, deal with failed motion primitives or to omit unnecessary steps, such as detecting when tracing should be executed a second time, or not at all. Learning when and how to use the motion primitives to complete the task could greatly improve the end-to-end performance of the system and reduce the engineering effort. This could be achieved by extending approaches based on spatial action maps such as Ha and Song⁶ and Avigal et al.⁷ with new primitives, like tracing.

An alternative would be to discard the motion primitives entirely and learn the task end-to-end based on demonstrations or (delayed) feedback. Such systems could reduce the engineering effort and would be more versatile, but they currently require prohibitive amounts of data to generalise to the diversity that characterises cloth manipulation.

Conclusion

This article introduced a novel approach to end-to-end cloth unfolding. Our key innovation is the successful integration of advanced tactile sensing throughout the entire unfolding pipeline. Our victories in the Cloth Manipulation Competition at IROS 2022 and ICRA 2023 demonstrated the effectiveness of this approach. To provide an even more extensive assessment of our method, we developed a detailed evaluation methodology that demonstrated the potential of our approach, as evidenced by the 17 perfectly executed unfolding trials.

While our system demonstrated strong performance, the evaluation also highlighted areas for future work. Specifically, we identify that 31 failures are due to the tracing not producing a graspable edge. To resolve this, we recommend investigating new methods for grasp selection that leverage both visual and tactile information to identify optimal grasp points for initiating the tracing process. Beyond grasp selection, future research should also prioritise evaluation on standardised cloth sets, participation in robotics competitions, and exploration of cloth manipulation in realistic home environments. Open communication about limitations will be essential to drive progress in this field.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Research Foundation Flanders (FWO) under Grant Numbers 1SD4421N, 1S15923N, 1S56022N and the euROBIn Project (EU grant number 101070596).

ORCID iDs

Victor-Louis De Gusseme  <https://orcid.org/0000-0003-2046-573X>

Remko Proesmans  <https://orcid.org/0000-0002-5925-625X>

Thomas Lips  <https://orcid.org/0000-0001-9530-5349>

Andreas Verleysen  <https://orcid.org/0000-0001-5071-5619>

Francis wyffels  <https://orcid.org/0000-0002-5491-8349>

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. Jiménez P and Torras C. Perception of cloth in assistive robotic manipulation tasks. *Nat Comput* 2020; 19: 409–431.
2. Zhu J, Cherubini A, Dune C, et al. Challenges and outlook in robotic manipulation of deformable objects. *IEEE Robot Autom Mag* 2022; 29: 67–77.
3. Yin H, Varava A and Kragic D. Modeling, learning, perception, and control methods for deformable object manipulation. *Sci Robot* 2021; 6: eabd8803.
4. Maitin-Shepard J, Cusumano-Towner M, Lei J, et al. Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. In: 2010 IEEE International Conference on Robotics and Automation, pp.2308–2315. IEEE.
5. Doumanoglou A, Stria J, Peleka G, et al. Folding clothes autonomously: a complete pipeline. *IEEE Trans Robot* 2016; 32: 1461–1478.

6. Ha H and Song S. Flingbot: the unreasonable effectiveness of dynamic manipulation for cloth unfolding. In: Conference on Robot Learning, pp.24–33. PMLR.
7. Avigal Y, Berscheid L, Asfour T, et al. Speedfolding: learning efficient bimanual folding of garments. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.1–8. IEEE.
8. Canberk A, Chi C, Ha H, et al. Cloth funnels: Canonicalized-alignment for multi-purpose garment manipulation. *arXiv preprint arXiv:221009347*, 2022.
9. Proesmans R, Verleysen A and wyffels F. Unfoldir: Tactile robotic unfolding of cloth. *IEEE Robot Autom Lett* 2023; 8: 4426–4432.
10. Sunil N, Wang S, She Y, et al. Visuotactile affordances for cloth manipulation with local control. In: *Conference on Robot Learning*, pp.1596–1606. PMLR.
11. Garcia-Camacho I, Borrás J, Calli B, et al. Household cloth object set: fostering benchmarking in deformable object manipulation. *IEEE Robot Autom Lett* 2022; 7: 5866–5873.
12. Hoque R, Shivakumar K, Aeron S, et al. Learning to fold real garments with one arm: A case study in cloud-based robotics research. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.251–257. DOI: 10.1109/IROS47612.2022.9981253.
13. García-Camacho I, Borrás J, Calli B, et al. Cloth manipulation and perception competition. In: 2nd Workshop on Representing and Manipulating Deformable Objects, 2022, ICRA.
14. Lips T, De Gusseme VL, et al. Learning keypoints from synthetic data for robotic cloth folding. *arXiv preprint arXiv:220506714*, 2022.
15. De Gusseme VL and wyffels F. Effective cloth folding trajectories in simulation with only two parameters. *Front Neurorobot* 2022; 16: 989702.
16. Doumanoglou A, Kargakos A, Kim TK, et al. Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning. In: 2014 IEEE International Conference on Robotics and Automation (ICRA), pp.987–993. DOI: 10.1109/ICRA.2014.6906974.
17. Sahari KSM, Seki H, Kamiya Y, et al. Edge tracing manipulation of clothes based on different gripper types. *J Comput Sci* 2010; 6: 872–879.
18. Zhang F and Demiris Y. Visual-tactile learning of garment unfolding for robot-assisted dressing. *IEEE Robot Autom Lett* 2023; 8: 5512–5519.
19. Seita D, Ganapathi A, Hoque R, et al. Deep imitation learning of sequential fabric smoothing from an algorithmic supervisor. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.9651–9658. IEEE.
20. Miller S, Van Den Berg J, Fritz M, et al. A geometric approach to robotic laundry folding. *Int J Rob Res* 2012; 31: 249–267.
21. Matas J, James S and Davison AJ. Sim-to-real reinforcement learning for deformable object manipulation. In: Conference on Robot Learning, pp.734–743. PMLR.
22. Lin X, Wang Y, Olkin J, et al. Softgym: benchmarking deep reinforcement learning for deformable object manipulation. In: Conference on Robot Learning, pp.432–448. PMLR.
23. Petrík V, Smutný V, Krsek P, et al. Single arm robotic garment folding path generation. *Adv Robot* 2017; 31: 1325–1337.
24. Li Y, Yue Y, Xu D, et al. Folding deformable objects using predictive simulation and trajectory optimization. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.6000–6006. IEEE.
25. Jangir R, Alenya G and Torras C. Dynamic cloth manipulation with deep reinforcement learning. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp.4630–4636. IEEE.
26. Garcia-Camacho I, Lippi M, Welle MC, et al. Benchmarking bimanual cloth manipulation. *IEEE Robot Autom Lett* 2020; 5: 1111–1118.
27. Verleysen A, Biondina M and wyffels F. Learning self-supervised task progression metrics: a case of cloth folding. *Appl Intell* 2023; 53: 1725–1743.
28. Miller S, Fritz M, Darrell T, et al. Parametrized shape models for clothing. In: 2011 IEEE international conference on robotics and automation, pp.4861–4868. DOI: 10.1109/ICRA.2011.5980453.
29. Tu Z, Talebi H, Zhang H, et al. Maxvit: Multi-axis vision transformer. In: *European Conference on Computer Vision*, pp.459–479. Springer.
30. Lips T, De Gusseme VL and Wyffels F. Learning keypoints for robotic cloth manipulation using synthetic data. *IEEE Robot Autom Lett* 2024; 9: 6528–6535.
31. Proesmans R, Verleysen A and wyffels F. Augmenting off-the-shelf grippers with tactile sensing. *arXiv preprint arXiv:230605902*, 2023.
32. Chen LY, Shi B, Lin R, et al. Bagging by learning to singulate layers using interactive perception. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.3176–3183. DOI: 10.1109/IROS55552.2023.10341634.
33. Tirumala S, Weng T, Seita D, et al. Learning to singulate layers of cloth using tactile feedback. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.7773–7780. DOI: 10.1109/IROS47612.2022.9981341.