Robotic Grasping and Manipulation competitions
ICRA 2023, 29th May-2nd June

Cloth manipulation and perception
Track

Version 2.1

Latest update: 10 April, 2023

Editor: Irene Garcia-Camacho Institut de Robòtica i Informàtica Industrial

Organizers:
- Júlia Borras Institut de Robòtica i Informàtica Industrial, Spain
- Berk Calli Worcester Polytechnic Institute, MA, USA
- Adam Norton University of Massachusetts Lowell, MA, USA
- Guillem Alenyà Institut de Robòtica i Informàtica Industrial, Spain
Content

Introduction 3
  Background 3
  Overview 3
  Competition format and setup 5

Perception 6
  Task 1.1. Grasp point detection 6
    Task Description 6
    Input provided 6
    Expected output 8
    Scoring 8
    Procedures 9
    Rules 10

Manipulation 11
  Task 2.1. Unfolding 11
    Task description 11
    Input provided 11
    Expected output 12
    Scoring and ranking 12
    Procedures 13
    Rules 13
  Task 2.2. Folding 14
    Task description 14
    Input provided 14
    Expected output 14
    Scoring and ranking 15
    Procedures 15
    Rules 15

References 16
## Cloth perception and manipulation competition

### Document version

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>April 26, 2022</td>
<td>Initial version</td>
</tr>
<tr>
<td>1.1</td>
<td>July 04, 2022</td>
<td>Competition will be held virtually</td>
</tr>
<tr>
<td>1.2</td>
<td>August 28, 2022</td>
<td>More info in Competition format</td>
</tr>
<tr>
<td>1.3</td>
<td>September 06, 2022</td>
<td>Added procedure steps + evaluation details</td>
</tr>
<tr>
<td>2.0</td>
<td>January 30, 2023</td>
<td>Updated version for 2nd edition at ICRA 2023</td>
</tr>
<tr>
<td>2.1</td>
<td>April 10, 2023</td>
<td>Added penalization for false corners in perception task + minimum camera resolution for output</td>
</tr>
</tbody>
</table>
**Introduction**

**Background**

In the last decade, several competitions in robotic manipulation have been organised as a way to drive scientific progress in the field. They enable comparison of different approaches through a well-defined benchmark with equal test conditions. However, current competitions usually focus on rigid-object manipulation, leaving behind the challenges that suppose grasping deformable objects, especially highly-deformable ones as cloth-like objects. This competition focuses on the challenges of perception and manipulation of textile objects, searching to accelerate scientific progress in the domain of domestic service robots. It consists of a small set of tasks to benchmark in a common framework using the same set of objects and assessment methods.

![Assistive lab from the Institut de Robòtica i Informàtica Industrial](image)

*Fig. 1. Assistive lab from the Institut de Robòtica i Informàtica Industrial.*

**Overview**

This competition consists on tasks based on perception and a manipulation of cloth objects. For these tasks, objects from the Household Cloth Object Set [1] will be used. This object set presents a wide variety of household cloth objects selected to foster benchmarking in cloth manipulation.
In specific, the objects that will be used in the competition are the kitchen rags, the napkins, the towels and the pillowcases (not all of them). The kitchen rags are composed of 4 different rags of same size but different fabric materials, yarns and textures, what provides variability in the perception strategies and manipulation skills. The towels provide same fabric properties but different size. Whereas the napkins offer a different shape (squared) from the rest of the objects (rectangular). The pillowcases offer an additiona challenge due to its doble layer.


Competition format and setup

We do not restrict the characteristics of the table, but its **minimum size should be of 70x120cm** so the objects that will be used can fit completely inside the table. Any necessary considerations (e.g. specific color of the table) that teams require for their execution will have to be reported.

Any perception system can be used, but a **bonus scoring will be added for those perception systems which are part of a robotic system** (e.g. camera placed on the head of a mobile manipulator or included in a robotic arm that moves to sense environment). However, to properly benchmark the results of the tasks, it is required a specific format of the expected output (color or depth images + a csv file). In addition, teams will be asked to print the given **Aruco marker** in order to compute the pixel/centimeter ratio of each team to have a common unit (centimeters) with which to compare. This marker will be used to take an image of the final setup prior to executing the tasks (not used during the competition).

In similar way, any robotic system with grasping capabilities can be used in the manipulation tasks (2.1 and 2.2). This includes no restrictions on the gripper hardware, being able to use custom design grippers that ease the manipulation of textiles.

As this competition will be held virtually, some requirements on the setup are necessary to properly evaluate the tasks. **Each team will need 3 cameras**: one top down to have a zenithal view of the entire setup (table and objects), one from the side of the setup to view the robot execution and one hand-hold camera by a team participant zooming the process.
Perception

Task 1.1. Grasp point detection

Task Description

This task consists on locating suitable grasping points and their approach grasping vectors of a cloth object that can be either crumpled, folded or flat (see Fig. 4). These grasping points are the visible corners of a single layer of the cloth.

Fig. 4. Example cloth configurations for Task 1.1. (Not real images used for grasp detection).

Input provided

The objects used will be part of the Household Cloth Object Set that will be sent to the participant teams.

For this task, a total of 10 trials will be performed. In each of these trials, the object and/or the configuration will be changed in order to provide variability for the perception strategies. In order to avoid overfitted solutions, any object/configuration combination of the following lists can be given. In order to provide a fair comparison, the same 10 object/configuration combinations will be determined the day of the competition for all the teams.

Objects:
- Rag towel
- Rag linen
- Rag waffle
- Rag chekered
- Napkin cotton
- Napkin linen
- Small towel
- Medium towel
- Rectangular pillowcase
Configurations:

- **Flat**: The 4 corners must be detected.

  ![Example of flat configuration.](image)
  
  *Fig. 5. Example of flat configuration.*

- **Crumpled**:
  - 1 or 2 visible corner inside the cloth.

  ![Example of corner in the cloth.](image)
  
  *Fig. 6. Example of 3 corner in the cloth.*

- 1 or 2 visible corner in the table.

  ![Example of corner in the table.](image)
  
  *Fig. 7. Example of 3 corner in the table.*

- 1 corner in the table and 1 inside the cloth.
Cloth perception and manipulation competition

- Folded:
  - 1 visible corner (notice that the correct corner to detect corresponds to the single layered one). To place it in this configuration follow the steps in Fig. 7. The folded cloth can be in any orientation.

![Fig. 7. Steps to fold the cloth.](image)

Expected output

Regardless the perception system or the methods used, to properly compare between teams, it is required to provide a color or depth image (in png or jpg format) of the perceived cloth with the selected grasping points clearly marked with red points and a csv file with their corresponding image coordinates. In addition, this image should have represented the approach grasping direction with a line and a csv file with the image coordinates of the vector’s origin and end (the end should correspond to the grasping point coordinates).

Scoring

- 1 point for each correct detected grasping point.
- 1 point for each correct approach grasping vector.
- -0.5 point for each false detection.

For each trial, teams will provide an image of the cloth in duplicate: one without markers that will be used to mark the ground truth and the other one with the resulting detected corners and grasping vector clearly marked. These two images will serve to compute the error of the corner and grasping vector detection of the team.

A detected grasping point (corner) is correct if the error is less than 1cm. A grasp approach direction is correct if it is inside the tolerance of ±45°.
Some examples:

![Examples of detected corners (in red) and grasp approach vectors (in blue) with tolerance (black).](image)

In addition, the execution time will also serve as evaluation parameter.

**Bonus**
- Hardware setup: Perception systems that are attached to a robotic system: 10 pts

**Total achievable points:**
Number visible corners (according to cloth config.) x 10 trials x 2 + bonus

**Procedures**

1. Fix a camera to take images of the entire cloth.
2. Take an image with this camera of the given Aruco template in the table.
3. Remove the Aruco template from the setup.
4. Place the given cloth in the indicated configuration.
5. Start the execution when indicated.
6. When the execution finishes, send the image of the final result.
Rules

1. During the competition, the **objects must be used in its original form** (without adding perception system markers or dying any part of the cloth). Anyway, the tags of the objects can be cut taking care of not damaging the sewing.
2. No manual or teleoperated intervention by human is allowed once the execution starts.
3. No restrictions on perception systems and sensors used.
4. Final result must be provided in the indicated format (color or depth image).
Manipulation

Task 2.1. Unfolding

Task description

This task consists on grasping the cloth and manipulating it until it is placed flat on the table. The cloth initial configuration can be either crumpled or folded and the final configuration of the task is flat.

This task can be performed in many different ways, having no limitations on the strategies used. That is to say, it can be performed for example with a bimanual robotic system by grasping directly the corners, performing edge tracing to the second corner or by pushing and sliding the cloth on the table until it is flat. However, the evaluation will take into account the quality of the final result through a zenital image of the cloth, as well as other parameters as the total execution time or the number of state transitions performed since the task starts.

![Initial and final cloth configurations for Task 2.1.](image)

Input provided

The objects used will be part of the Household Cloth Object Set that will be sent to the participant teams.

For this task, a total of 6 trials must be executed; 3 trials will be performed starting from the crumpled configuration of the cloth and 3 from the folded configuration.
Cloth perception and manipulation competition

Prior to start the execution, the team will place the indicated object in the given configuration. The jury then can ask the team to move a bit the cloth (move about 2cm the entire or some part of the object), so teams should not rely on harcoded positions of the object.

Expected output

The task ends when the cloth lays flat completely inside of the table or when the maximum execution time is exceeded. Teams must provide zenithal color or depth images in png or jpg format of the entire flat cloth, from which the placement quality will be evaluated.

Scoring and ranking

- Unfolding quality points (0-100).
- -20pts if the cloth is not completely inside the table.
- -10pts for each bend (not wrinkle) in the cloth.

An unfolding is successfully performed when the cloth is completely flat inside the table without bends or wrinkles in the cloth. Having a zenithal image of the spread cloth, the quality of the unfolding is computed by measuring the area of the unfolded cloth through its contour and comparing it to the total area of the cloth when is perfectly spread. See examples of scoring in Fig. 11.

In addition, for each trial it will also be taken into account the execution time and the number of state transitions (steps performed by the robotic system during the execution) [3].

**Bonus**

- No necessary markers or modifications made to the cloth: 20 pts

*Total achievable points*: 100 pts x 6 trials + Bonus
Some examples of scoring:

Fig. 11. Some examples of unfoldings.

a) 99 (quality) pts.  

b) 97(quality)-10(bend)-10(bend)=77pts.  
c) 71 (quality) pts.

Procedures

1. Fix a camera so it can take zenithal images of the entire cloth.
2. Take an image with this camera of the given Aruco template in the table.
3. Remove the Aruco template from the setup.
4. Ensure that the cloth does not has wrinkles that affects the result of the task by ironing it (e.g. folding bends).
5. Place the given cloth in the indicated configuration.
6. Start the execution when indicated.
7. When the execution finishes, get a zenithal image of the final result.

Rules

1. No manual or teleoperated intervention by human is allowed once the execution starts.
2. Objects can be dyed or have markers to ease perception.
3. The tags of the objects can be cut taking care of not damaging the sewing.
4. No restrictions on robotic systems and sensors used.
5. There are no restrictions in the strategies used. Only the initial and final states of the cloth are defined.
6. Regardless the perception and robotic systems used, the result of the task must be presented in a zenithal color or depth image.
Task 2.2. Folding

Task description

This task consists on folding the cloth two times once is flat on the table. In order to correctly perform this task, the object must be folded following the folding protocol (see Fig. 7).

![Initial and Final Cloth Configurations for Task 2.2.](image)

This task can be a continuation of Task 2.1, being able to execute it in sequence right after the previous one, performing an end-to-end pipeline of the whole task of folding a cloth starting from a crumpled configuration.

Input provided

The objects used will be part of the Household Cloth Object Set that will be sent to the participant teams.

A total of 6 trials must be executed. This task starts always with the flat cloth laying on the table. Before starting the execution, the jury can ask the team to slightly move the cloth in order to avoid over fitted solutions.

Expected output

The expected result of the task is a cloth folded in halves two times following the folding sequence of the corresponding protocol in Fig. 7.

Zenithal color or depth images (in png or jpg format) of the final folded cloth will be required in order to evaluate the quality of the result.
Scoring and ranking

- Folding quality points (0-100).
- -20pts if the cloth is not completely inside the table.
- -10pt for each bend (not wrinkles) in the cloth.

A one fold manipulation is successfully performed when the corners of the original spread cloth are matching two by two and the area of the folded cloth is half of its initial area. Having the zenithal image of the folded cloth, we measure how well the corners match (folding quality) by measuring the ratio between the surface of the spread cloth before and after the fold through its contour. Because our task is restricted to folding in halfs, each fold needs to cut in half the area of the spread cloth on the table.

In addition, for each trial it will also be taken into account the execution time and the number of state transitions (steps performed by the robotic system during the execution) [3].

**Bonus**
- No necessary markers or modifications made to the cloth: 20 pts
- Executed in sequence with Task 2.1. without human intervention: 20 pts

Total achievable points: 100 points x 2 folds x 6 trials + 40 bonus

Procedures

1. Fix a camera so it can take zenithal images of the entire cloth.
2. Take an image with this camera of the given Aruco template in the table.
3. Remove the Aruco template from the setup.
4. Ensure that the cloth does not has wrinkles that affects the result of the task by ironing it (e.g folding bends).
5. Place the given cloth in the indicated configuration.
6. Start the execution when indicated.
7. When the execution finishes, get a zenithal image of the final result.

Rules

1. No manual or teleoperated intervention by human is allowed once the execution starts.
2. Objects can be dyed or have markers to ease perception.
3. The tags of the objects can be cut taking care of not damaging the sewing.
4. No restrictions on robotic systems and sensors used.
5. There are no restrictions in the strategies used. Only the initial and final states of the cloth are defined.
6. Regardless the perception and robotic systems used, the result of the task must be presented in a zenithal color or depth image.
References


