



2021 Intelligent Sensing Winter School

The CORSMAL challenge

7-10 December 2021

Speaker: Alessio Xompero



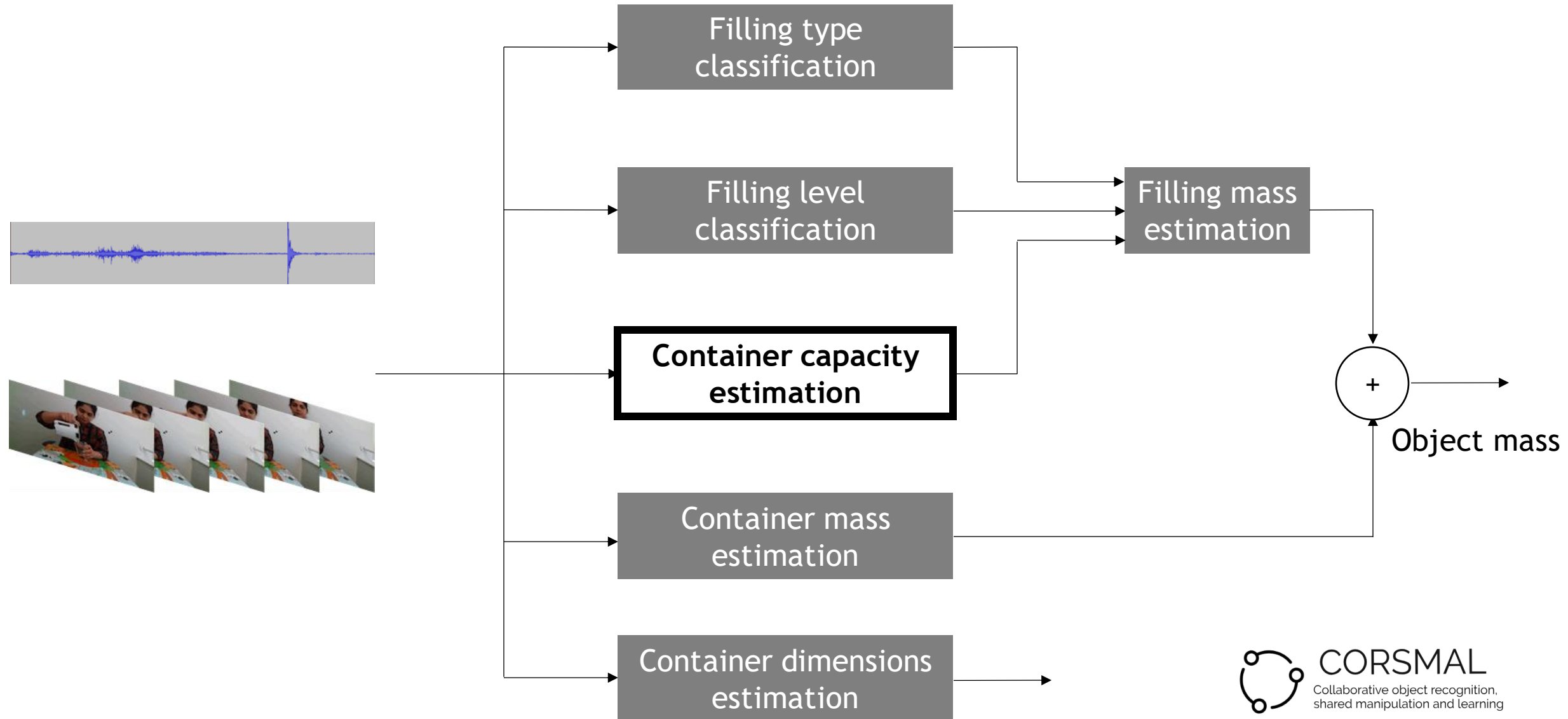
Introduction to the challenge



Can you estimate the properties of a container manipulated by a human?

Physical properties estimation

Task for the Winter School



15 containers
(5 cups, 5 drinking glasses, 5 food boxes)
3 filling types
3 filling levels



robot view

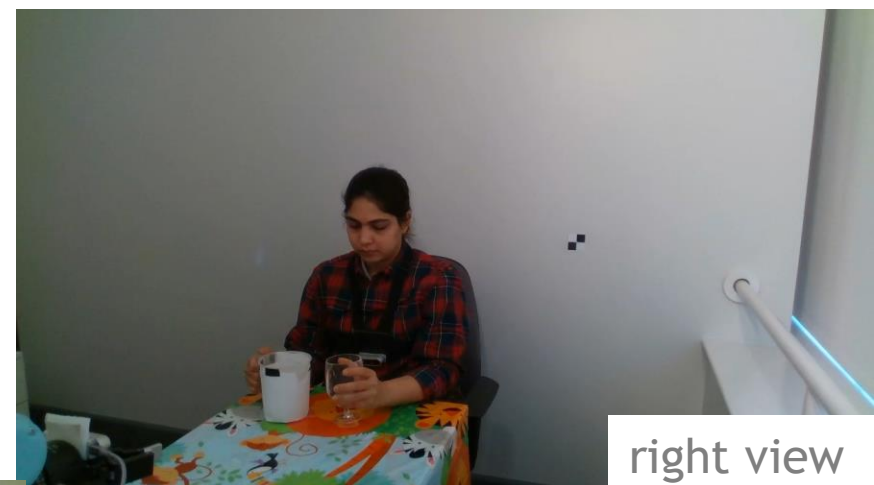
multi-view + multi-channel audio
4 views: 2 first-person + stereo pair
8 microphones: circular array



left view

CORSMAL Containers Manipulation dataset

1,140 audio-visual recordings
drinking glass: 84
cup: 84
food box: 60



right view

12 people
3 scenarios
2 backgrounds
2 illuminations



body-cam view

Annotations

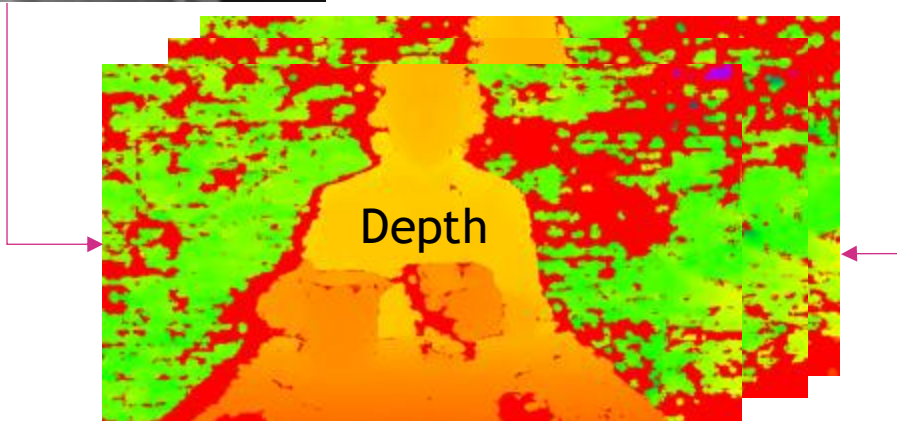
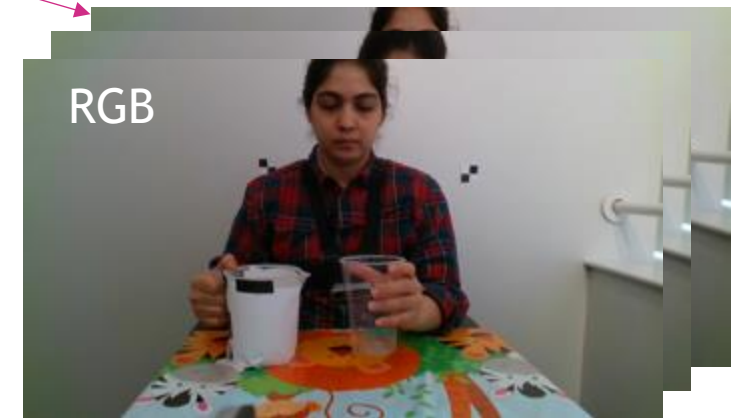
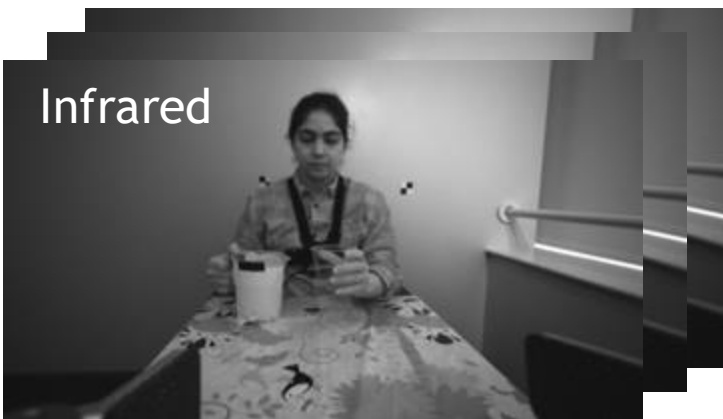
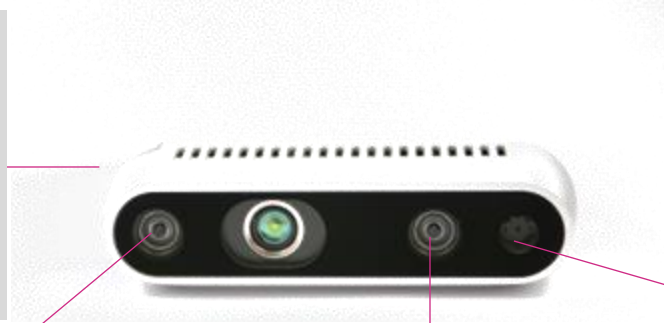
Filling type & level
Container volume
Container & filling masses



Sensor data

Inertial measurements
accelerometer, gyroscope

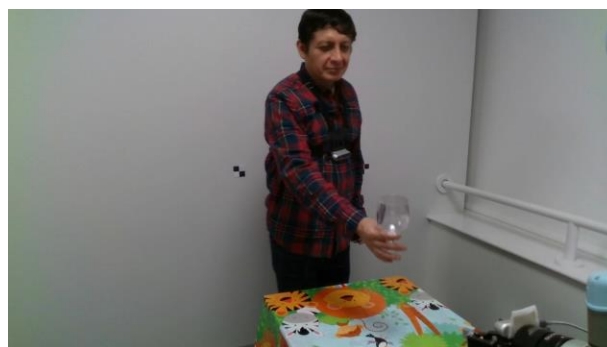
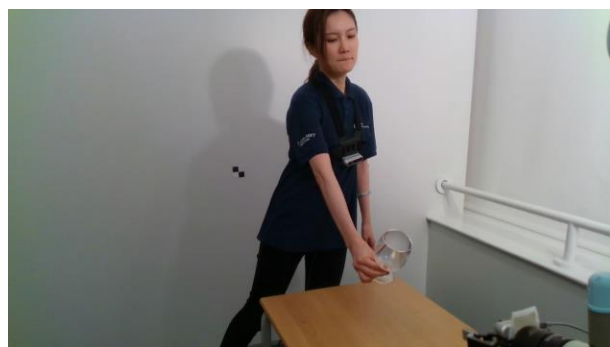
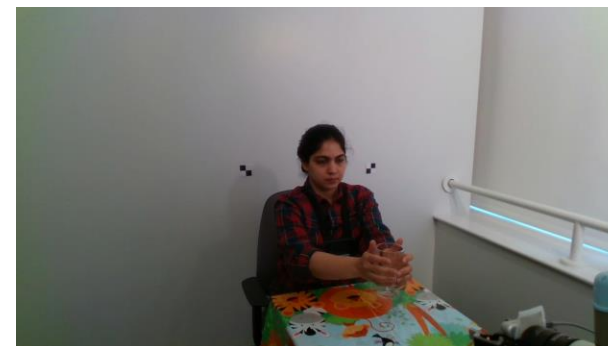
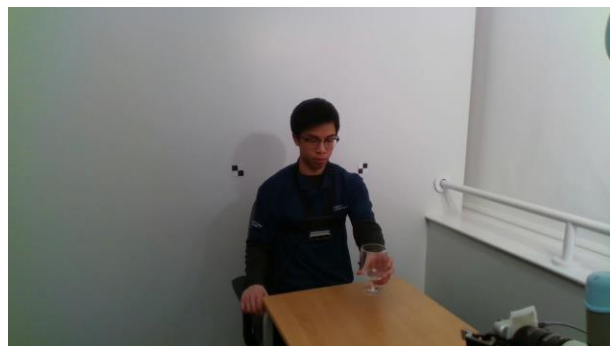
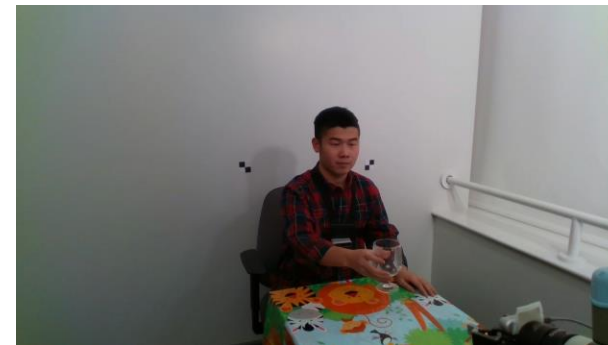
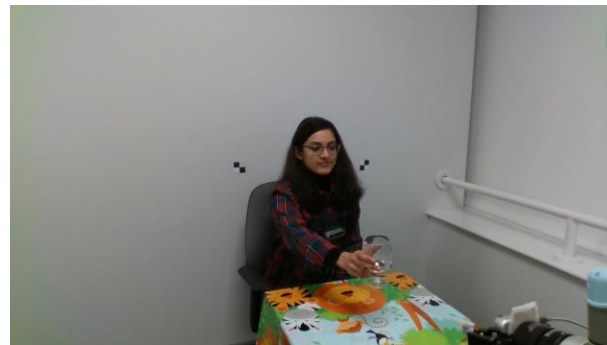
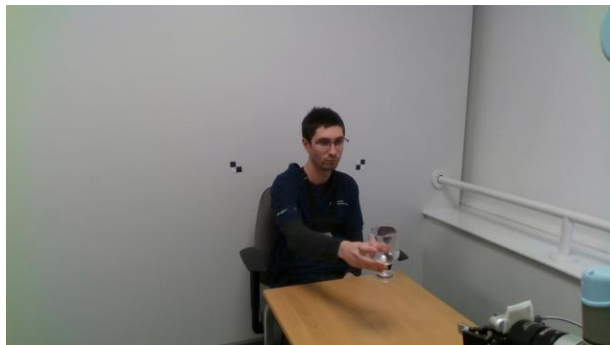
Calibration information
intrinsic, extrinsic parameters



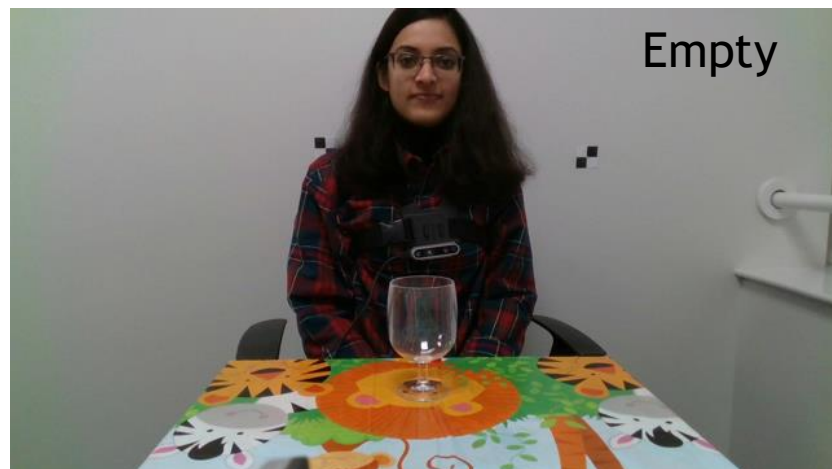
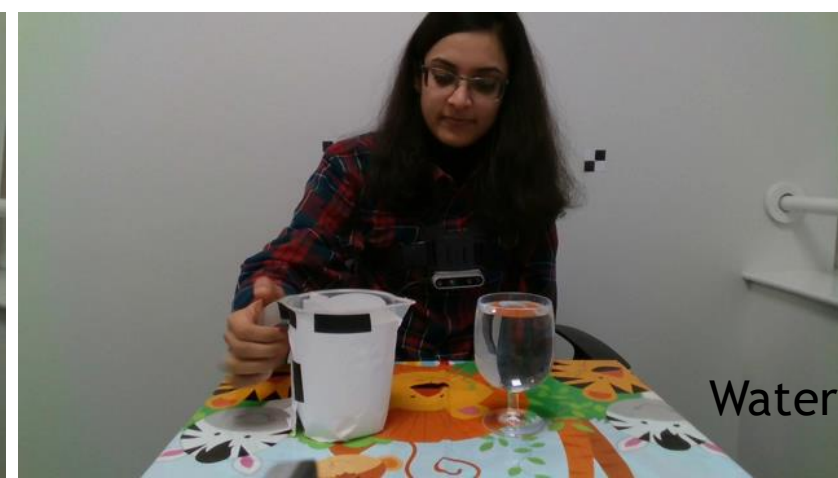
RGB and depth images are **spatially aligned**

(1280x720 pixels, 30 Hz)

Subjects



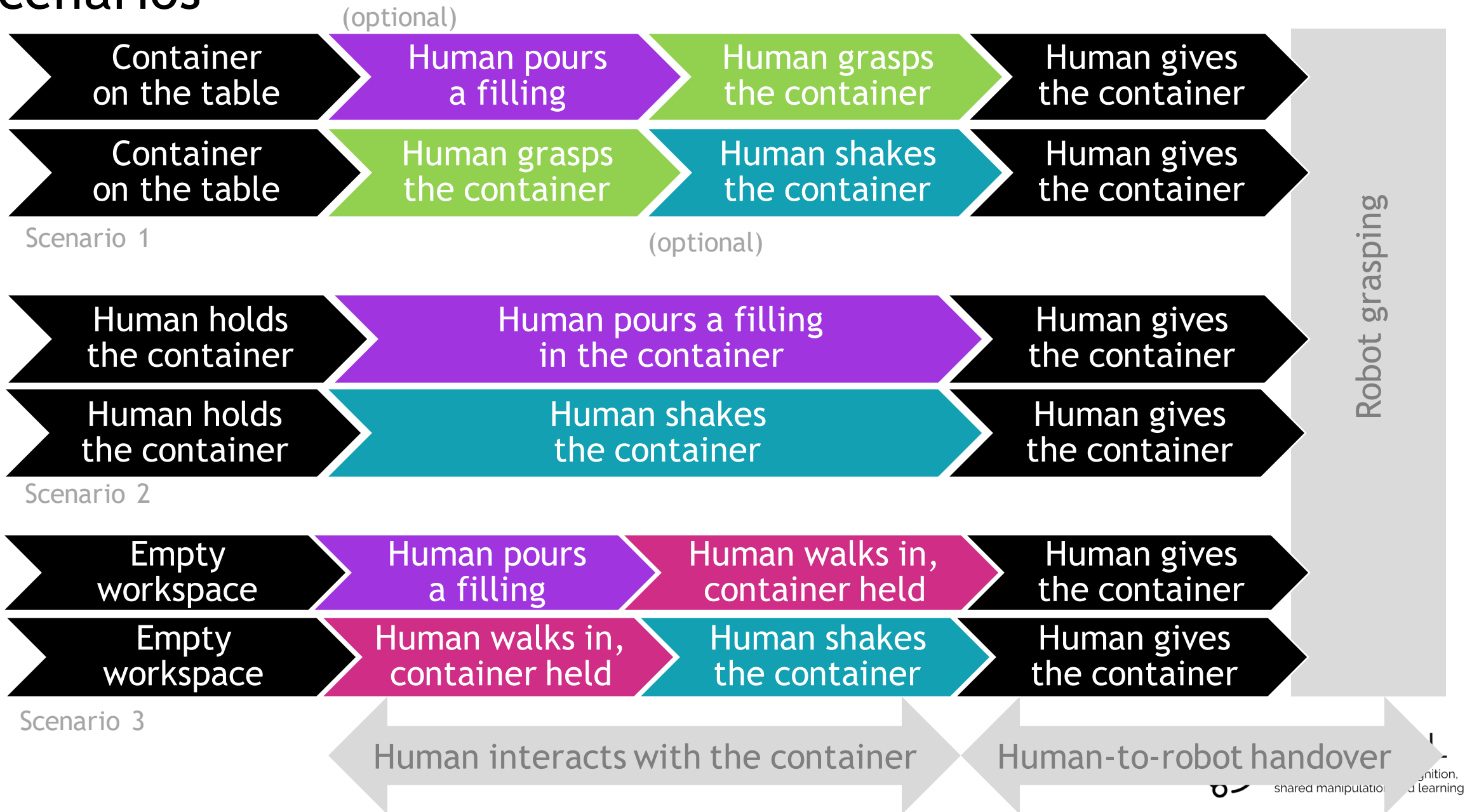
Fillings



Backgrounds and illuminations



Scenarios



Scenario 1

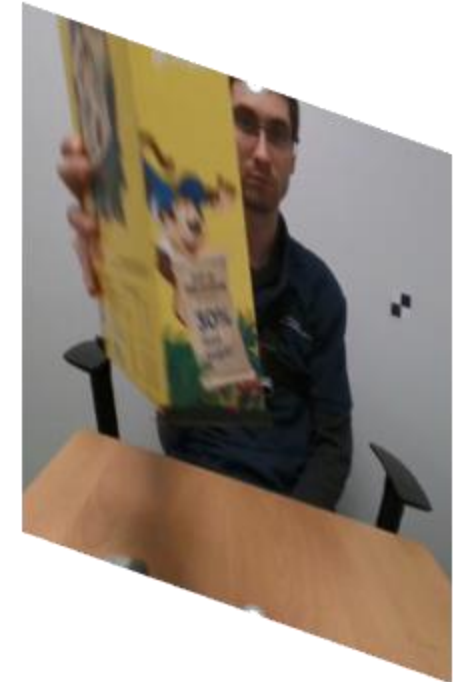
Container on the table

Human grasps the container

Human shakes the container

Human gives the container

Robot grasping



also

Container on the table

Human pours a filling

Human grasps the container

Human gives the container

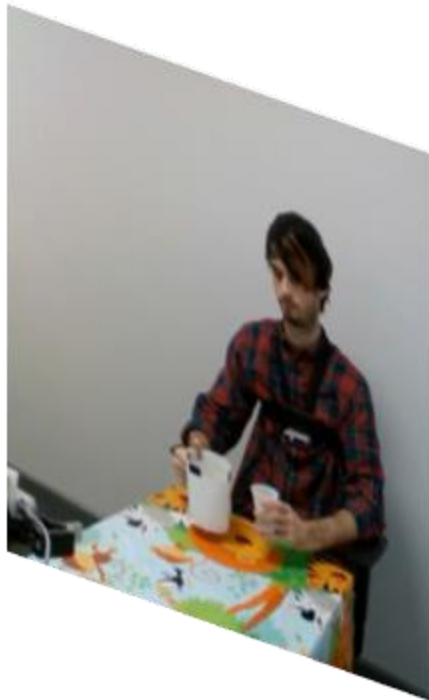
Robot grasping

RSML
Creative Object Recognition
and manipulation and learning



Scenario 2

Human holds the container



Human pours a filling in the container



Human gives the container



Robot grasping

also

Human holds the container

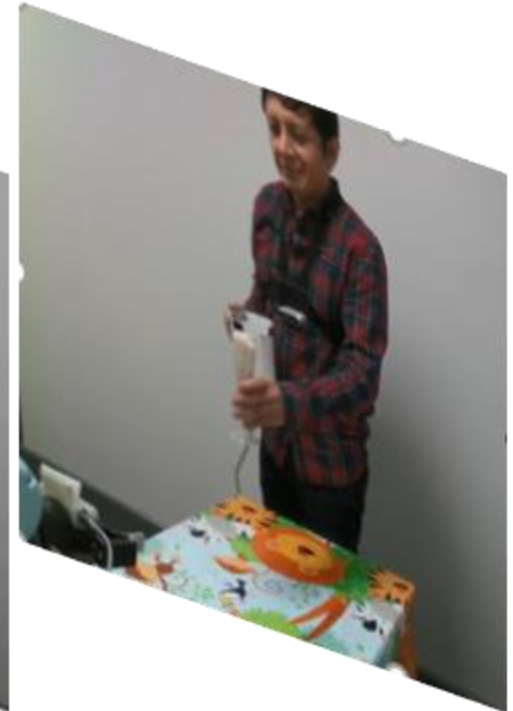
Human shakes the container

Human gives the container

Robot grasping
RSMAL
Creative Object recognition
and manipulation and learning

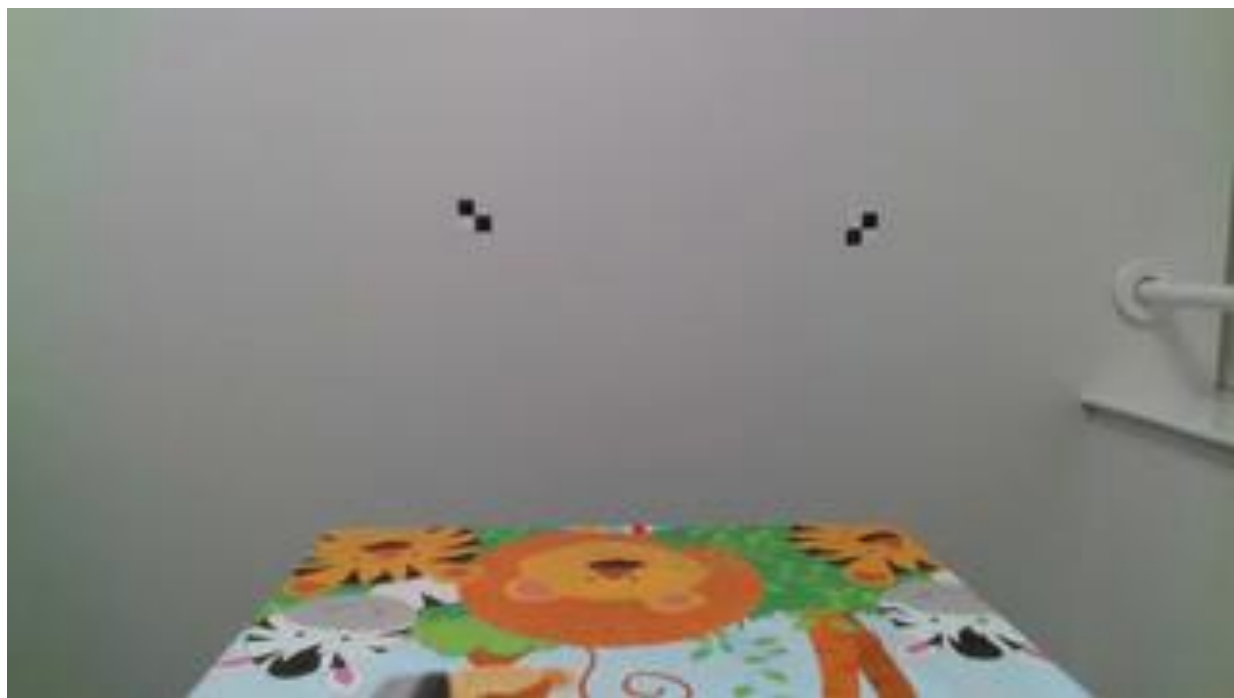


Scenario 3



also





Summary

Scenario 1



Scenario 2



Scenario 3



Increasing challenges (occlusions, entering field of view, human dynamics)

Multiple actions

holding container, grasping container, walking in, pouring content into container, shaking container, giving container, standing, sitting

Containers and data splits for Winter School challenge

Training set
(684 audio-visual recordings)



Test set

84 audio-visual recordings/glass

84 audio-visual recordings/cup

60 audio-visual recordings/box

Evaluation

Performance score assessing the effectiveness of the solution (50%)

Average Capacity Score \longrightarrow $ACS = \frac{1}{J} \sum_{j=1}^J \mathbf{1}_{\left\{ \frac{|\gamma_j - \hat{\gamma}_j|}{\hat{\gamma}_j} \leq \epsilon \right\}}$

Number of configurations

Indicator function (0, if capacity not estimated for configuration j)

γ_j : estimated capacity value
 $\hat{\gamma}_j$: annotated capacity value

Evaluation toolkit: <https://github.com/CORSMAL/CORSMALChallengeEvalToolkit/tree/intelsensing-winter-school-21>

Evaluation by a panel of judges for

- innovation & creativity of the solution (20%)
- quality of the presentation (20%)
- distribution of workload across team members (10%)

What you need to submit

1. PDF of your **presentation** - see next slide (max 10 MB)

Format: TeamName.pdf

2. 3 csv files with the **results** on the test set of each training/test split

Format: TeamName_splitN.csv

3. Link to an unlisted video (e.g., YouTube) of max 5 mins as a screen-capture of the **running code**

Submission deadline: Fri 10 Dec, 11.30am GMT

- Submit to: a.xompero@qmul.ac.uk

Presentation

- **Template:** PPT within the CORSMALChallengeEvalKit (see [link](#))
- Maximum **duration:** 5 mins (after which you will be muted)
- Focus on describing **your solution**, and:
 - Report the modalities (RGB images, audio, etc.) employed
 - Properly reference existing tools, methods, papers, etc., employed
 - Report the hardware used for the experiments
 - Describe the roles of each team member in the project
 - Report the advantage of the multi-modal solution over the used uni-modal baseline

No animations as you will submit a PDF file

Schedule and communication (GMT times)

Day 1

- Challenge begins (3pm GMT)
- Teams definition/assignment (20-30 mins)
- To communicate the **team name** and **corresponding person** by 5 pm

Day 2-3: Zoom slot for Q&A, 11am-12pm, with Alessio

Day 4

- 11:30am: **submission deadline**
- 1pm: presentation of the results by the teams

Important communications will be given by email:

a.xompero@qmul.ac.uk

Let the challenge begin!
Good luck!

