

CORSMAL demo: Localisation and shape estimation of containers

<http://corsmal.eecs.qmul.ac.uk>

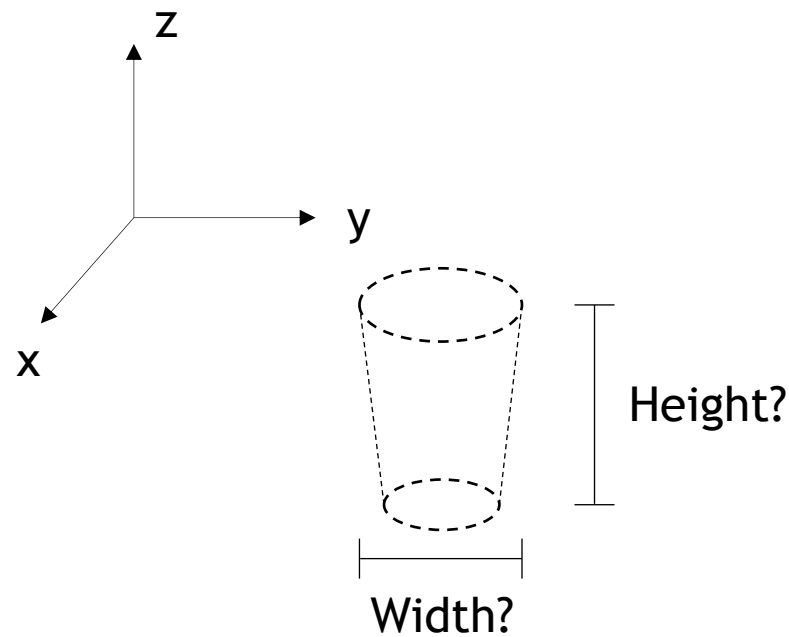
The problem



View 1



View 2

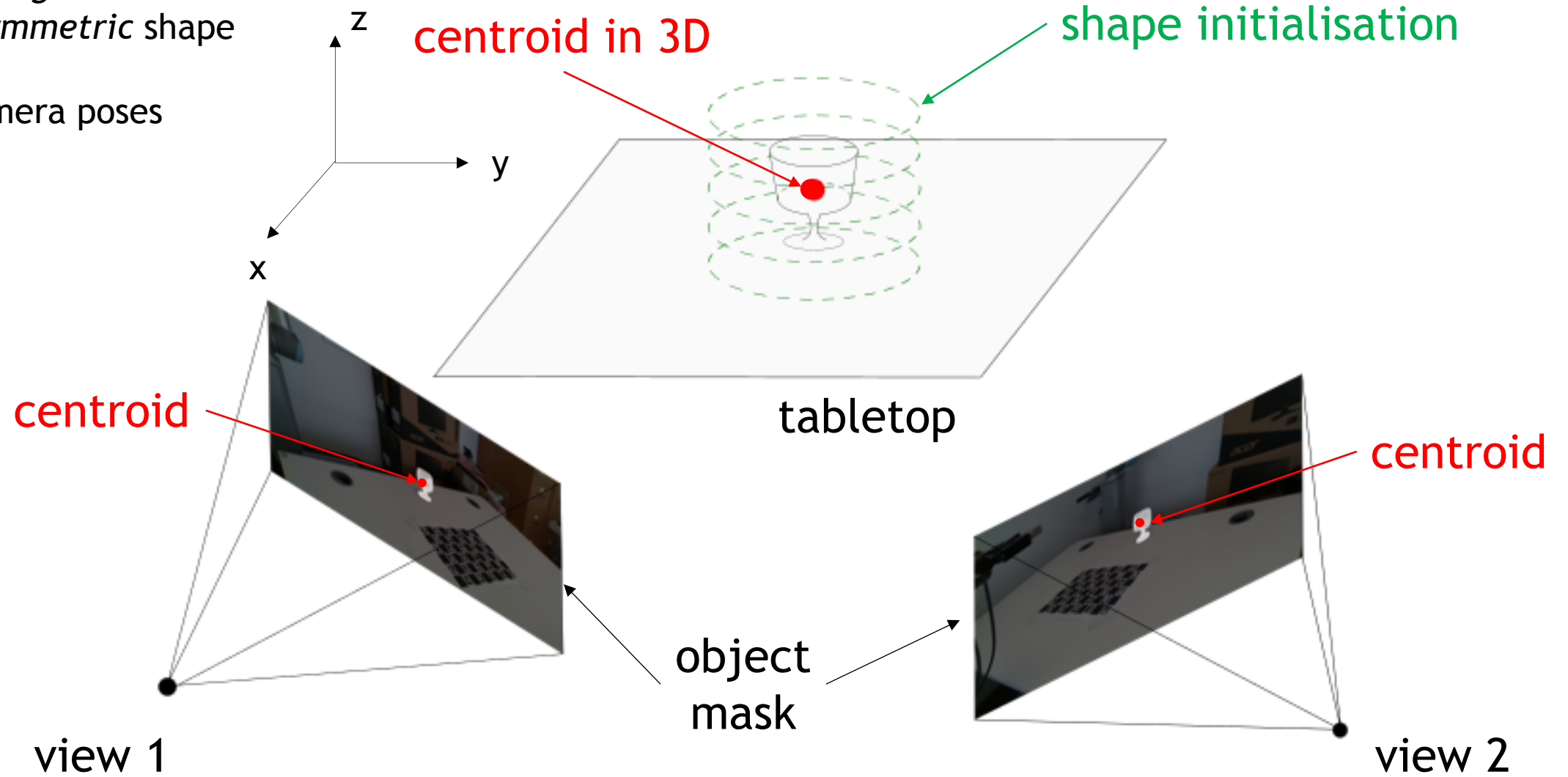


How to estimate the dimensions in 3D of unknown objects?

LoDE: Localisation and object Dimension Estimator

Objects *upright* and with
circular symmetric shape

Known camera poses



Iterative multi-view 3D-2D shape fitting

For each circumference l :

(Re)-sample N
3D points, $\mathbf{Q}_{n,l}^{(i)}$

Image point
Camera pose and calibration parameters

$$\mathbf{u}_{n,l}^c = \pi \left(\mathbf{Q}_{n,l}^{(i)}, \mathbf{C}^c, \boldsymbol{\theta}^c \right)$$

Projective geometry

Object masks

$$\sum_{n=1}^N \mathbf{m}^1(\mathbf{u}_{n,l}^1) + \mathbf{m}^2(\mathbf{u}_{n,l}^2) \stackrel{?}{=} 2N$$

Estimated width
and height of
circumference l

$$(w_l, h_l)$$

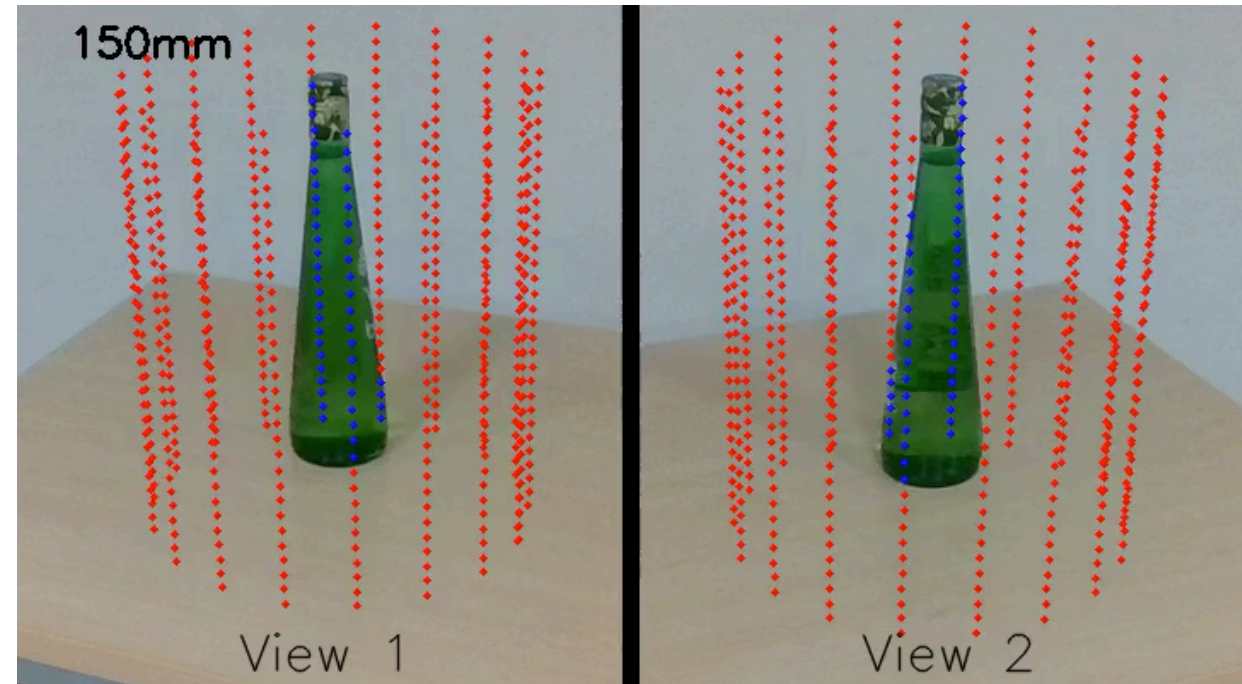
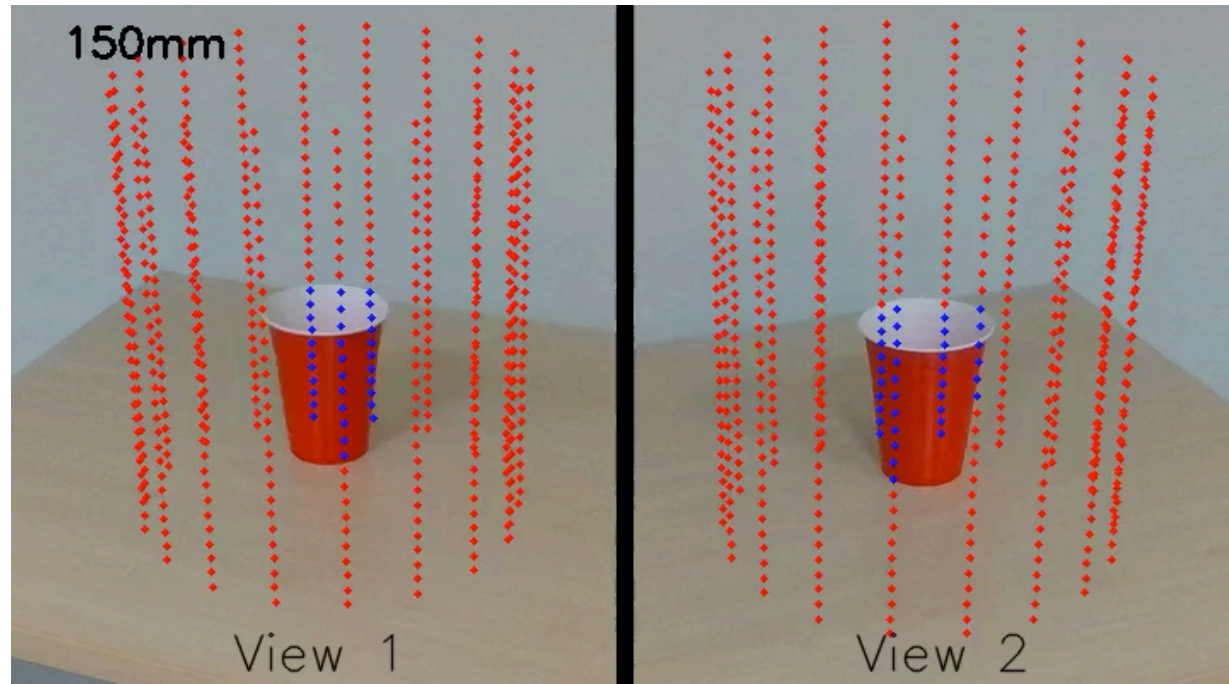
$$r_l^{(i)} \rightarrow r_l^{(i+1)}$$

Reduce radius

Convergence condition

i : iteration index
 n : point circumference index
 c : camera view index

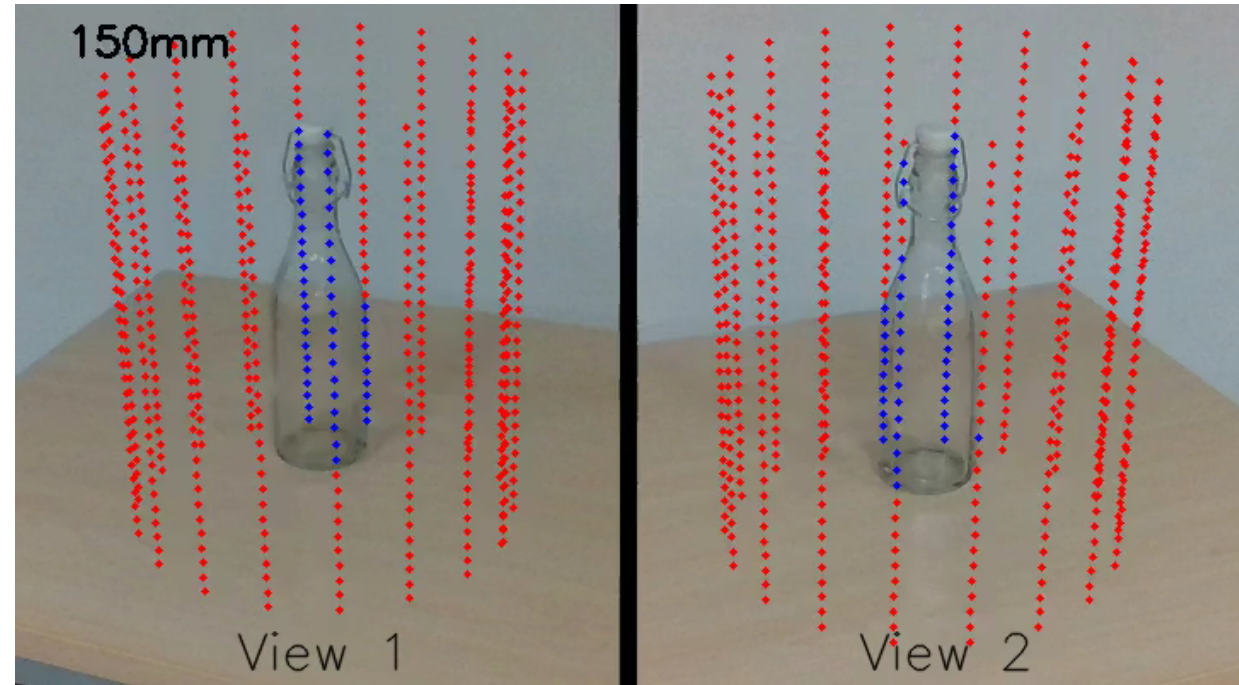
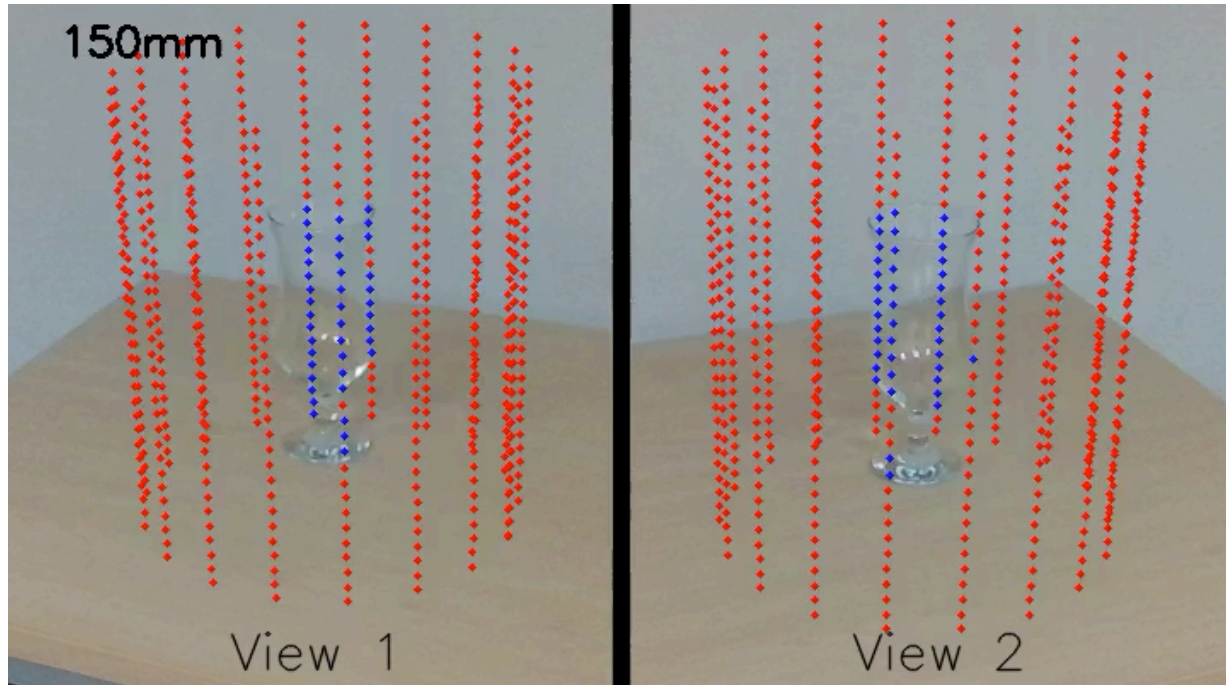
Iterative multi-view 3D-2D shape fitting



- 500 circumferences (1mm separation)
- 20 points per circumference
- Radius schedule: {150.0, 149.5, ..., 1.5, 1.0} mm

- projected points lying outside the segmentation masks
- not converged points lying inside the segmentation masks
- converged points lying inside the segmentation masks

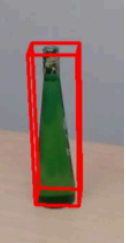



Iterative multi-view 3D-2D shape fitting



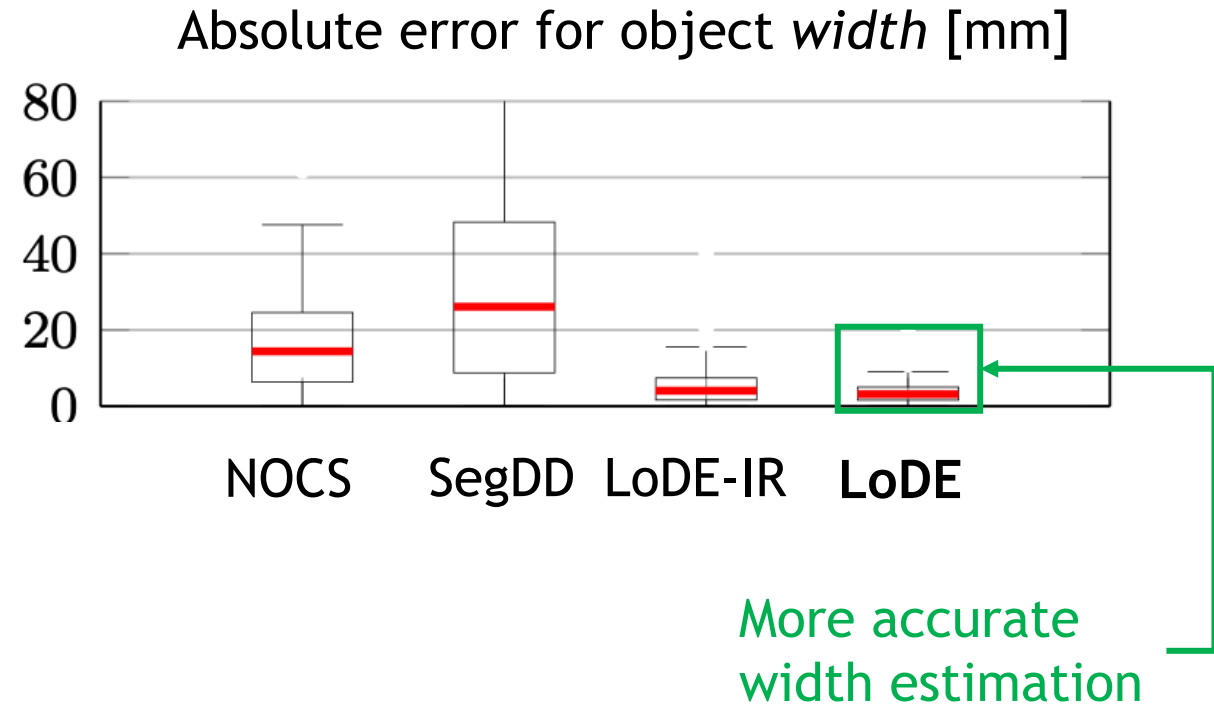
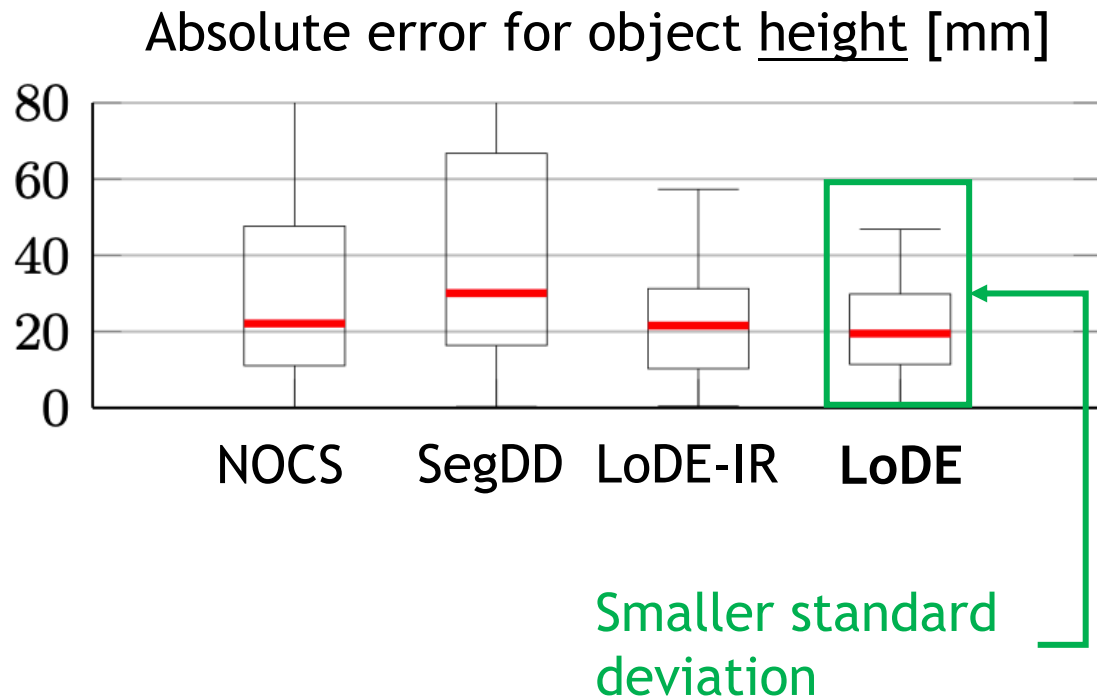
- 500 circumferences (1mm separation)
- 20 points per circumference
- Radius schedule: {150.0, 149.5, ..., 1.5, 1.0} mm

- projected points lying outside the segmentation masks
- not converged points lying inside the segmentation masks
- converged points lying inside the segmentation masks

Experimental setup

	Method	# of views	RGB	Depth	Infrared (IR)
	NOCS [Wang2019CVPR]	Single	✓	✓	
	SegDD: Segmentation on RGB-D images + back-projection	Single	✓	✓	
	LoDE-IR	Two (narrow baseline)			✓
	LoDE	Two (wide baseline)	✓		

Object dimensions results



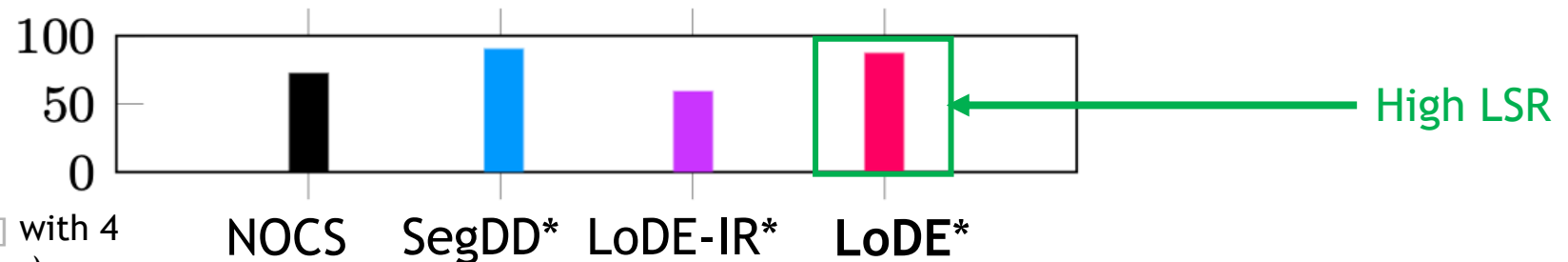
CORSMAL Container dataset (207 configurations)
Results provided for localised object

Object localisation results via instance segmentation

Some transparent objects not found in either of the two views
Inaccurate masks (truncated or not fitting the real object boundaries)

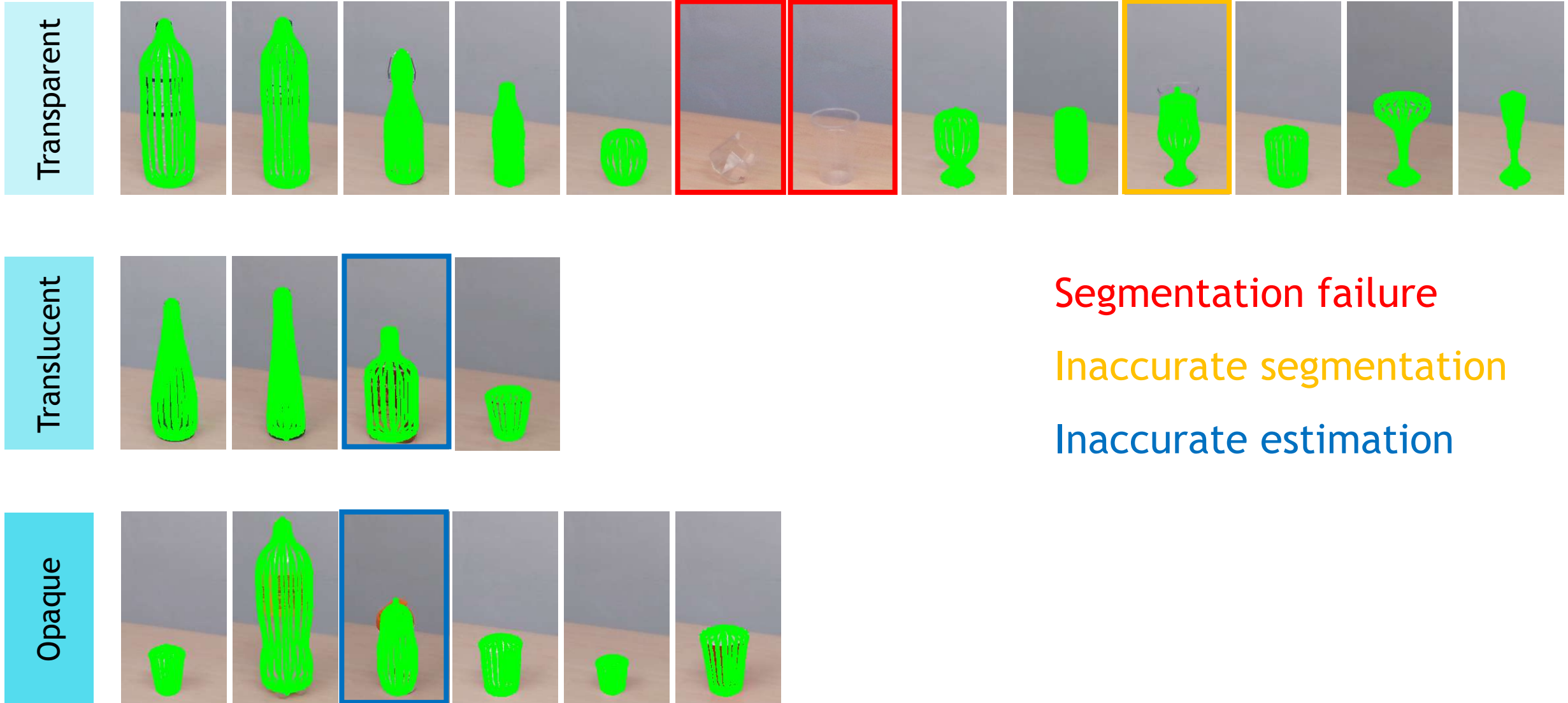


Localisation success ratio (LSR): $\frac{\# \text{localisations}}{\# \text{total configurations}} [\%]$



* Re-trained Mask R-CNN [He2017ICCV] with 4 classes (cups, wine glass, bottle, vase)

Shape fitting at convergence



Segmentation failure

Inaccurate segmentation

Inaccurate estimation

Qualitative comparison

NOCS



SegDD



LoDE-IR



LoDE



Conclusions

LoDE: localisation and object dimensions estimation (in 3D)

- Iterative multi-view 3D-2D shape fitting algorithm
- Upright objects with circular symmetric shape
- No depth, no prior 3D object models, no markers

Future work:

- Occlusions
- General object shapes
- Different object poses



<https://corsmal.eecs.qmul.ac.uk/LoDE.html>

Xompero, A., Sanchez-Matilla, R., Modas, A., Frossard, P., Cavallaro, A.

Multi-view shape estimation of transparent containers

IEEE International Conference on Acoustic, Speech and Signal Processing, May 2020



CORSMAL

Collaborative object recognition,
shared manipulation and learning