Efficient Object Search in Clutter via Image Torque

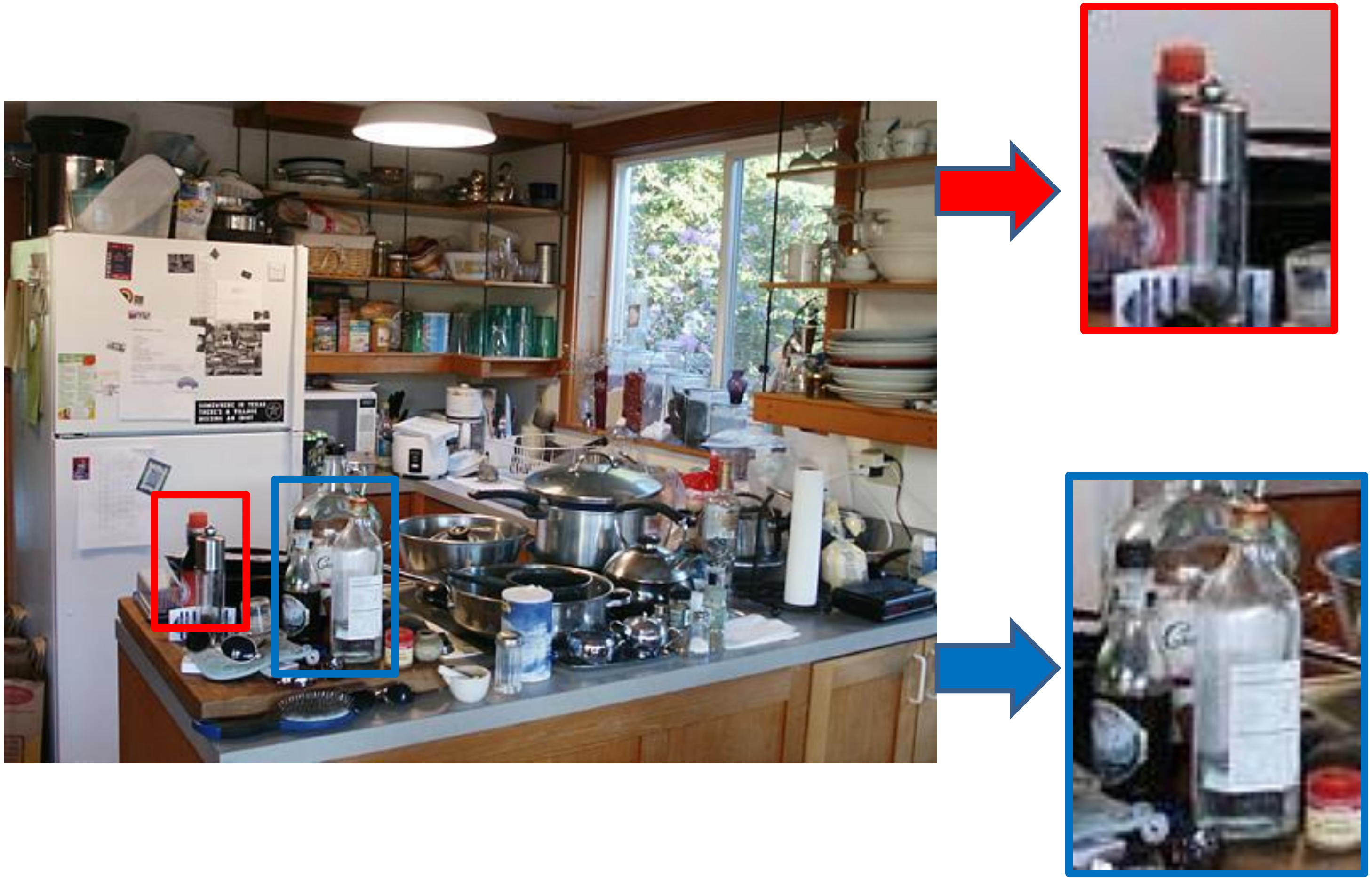
Ching L. Teo*, Austin Myers,
Cornelia Fermüller and Yiannis Aloimonos
University of Maryland

ICRA 2013

Sponsors:

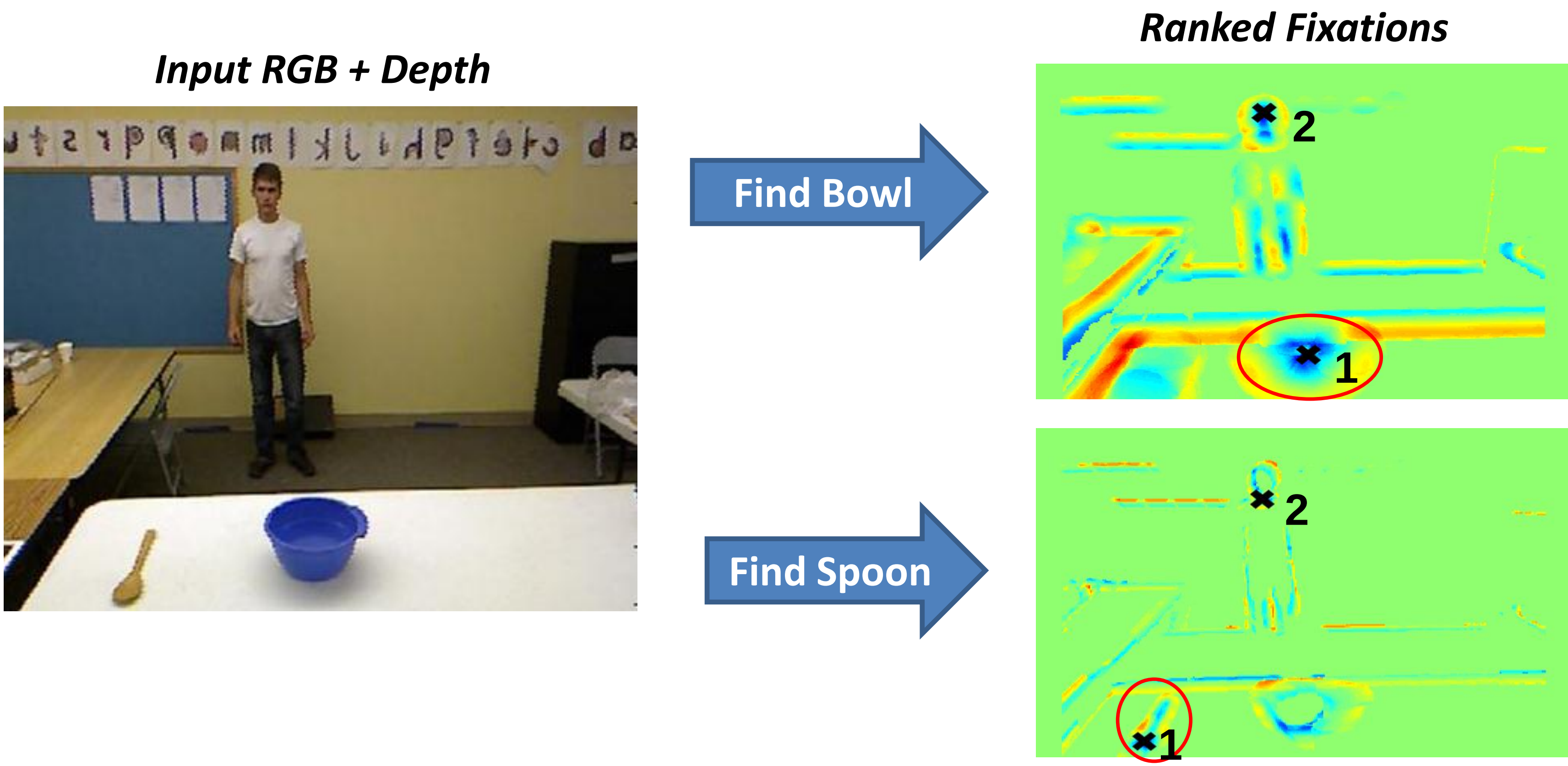
Where is the ‘bottle’ in the scene?

Where is the ‘bottle’ in the scene?



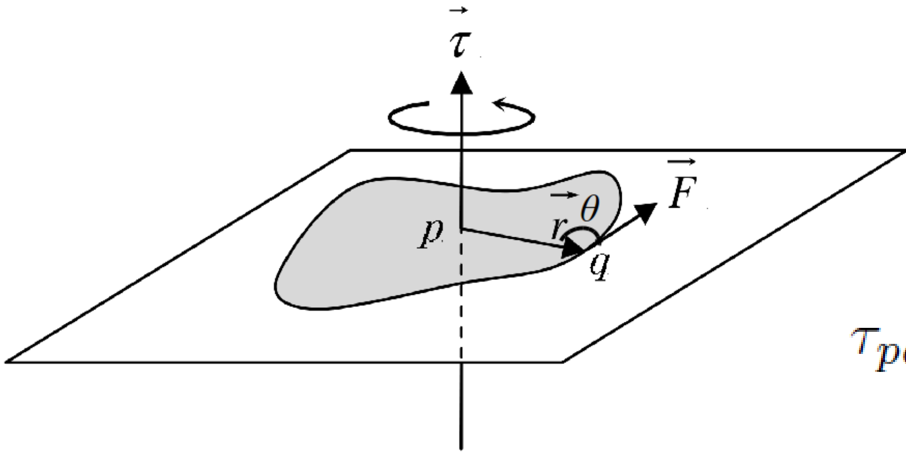
Our Contribution

- Propose a top-down mechanism for selecting *fixation points* that corresponds to the target via a *mid-level* contour grouping mechanism.



Grouping contours via Image Torque[1]

- For every edge point *q* with center *p*:



$$\tau_{pq} = \|\vec{r}_{pq}\| \sin\theta_{pq}$$

- For a patch *P*:

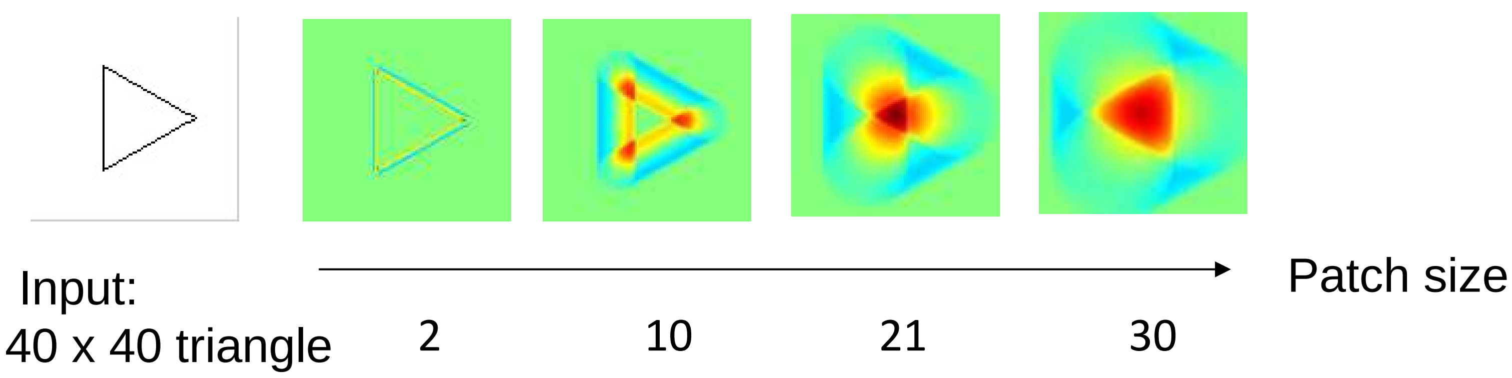
$$\tau_P = \frac{1}{2|P|} \sum_{q \in E(P)} \tau_{pq}$$

  – search over several scales for patch *P.*

[1] M. Nishigaki et al. The image torque operator: A new tool for mid-level vision. CVPR 2012

Properties of Image Torque

- τ_P has the following properties
  - Largest response for *closed contours* at a particular scale.
  - Ignores textured regions.
  - Maximum torque values occurs at region centroids
- Indicative of potential object locations.



Making τ_P Object Specific: τ_P^m

- By adjusting weights for each τ_{pq} :
  - Conforms to *shape* of target object within a predefined window *size:*

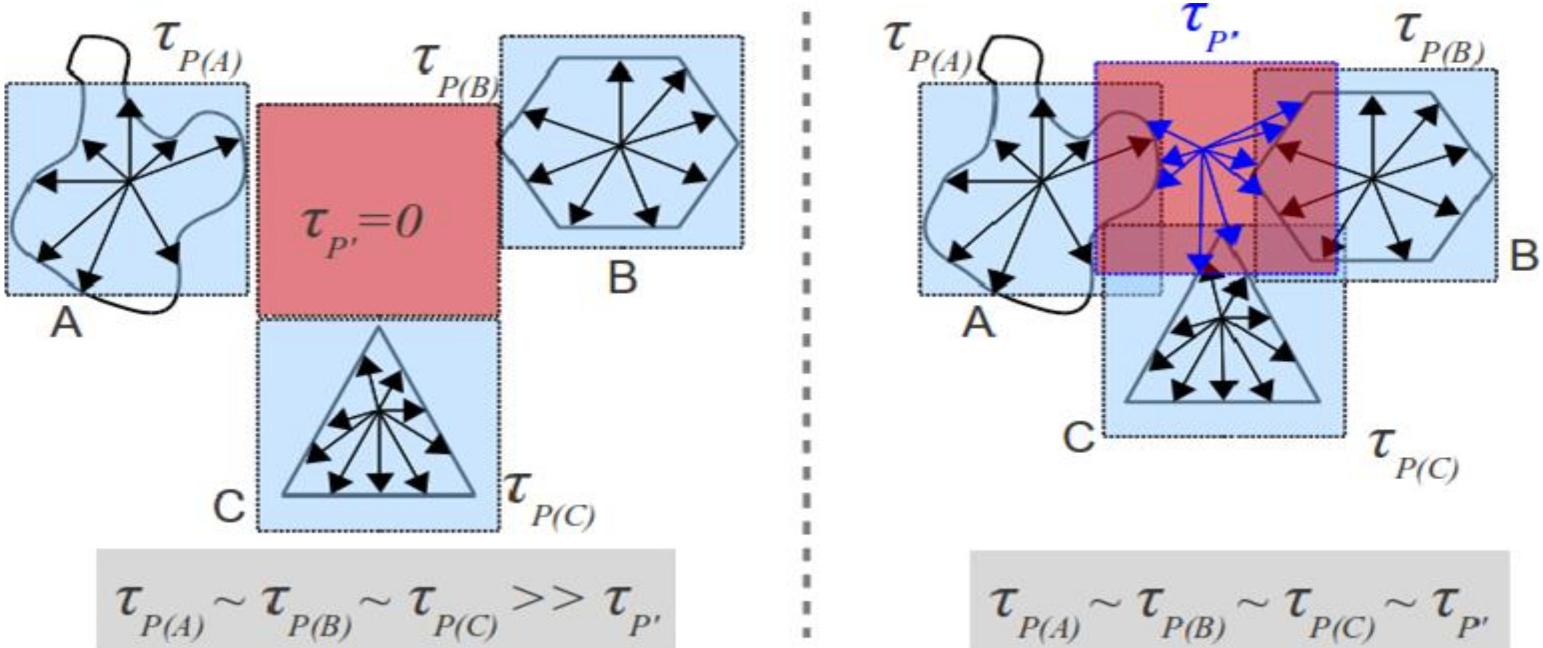
$$\tau_P^m = \frac{1}{2|P|} \sum_{q \in E(P)} m_{\mathcal{O}}(\tau_{pq})$$

- with $$m_{\mathcal{O}}(\tau_{pq}) = \frac{\tau_{pq}}{d_{qs}}$$

- d_{qs} is the Euclidean distance between edge point q to the target shape.

Why does this work for clutter?

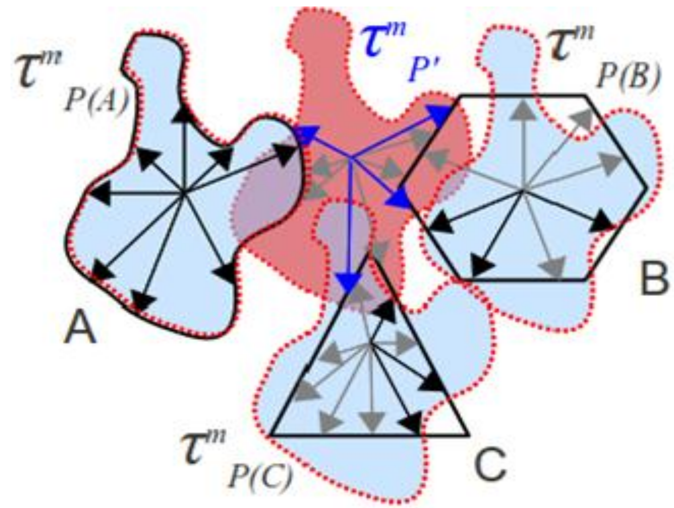
- τ_P gets confused by clutter:



- Left: no clutter – works well → $\tau_{P'}=0$
- Right: clutter – similar torque values with objects → $\tau_{P(A)} \approx \tau_{P(B)} \approx \tau_{P(C)} \approx \tau_{P'} \neq 0$

τ_P^m is robust to clutter

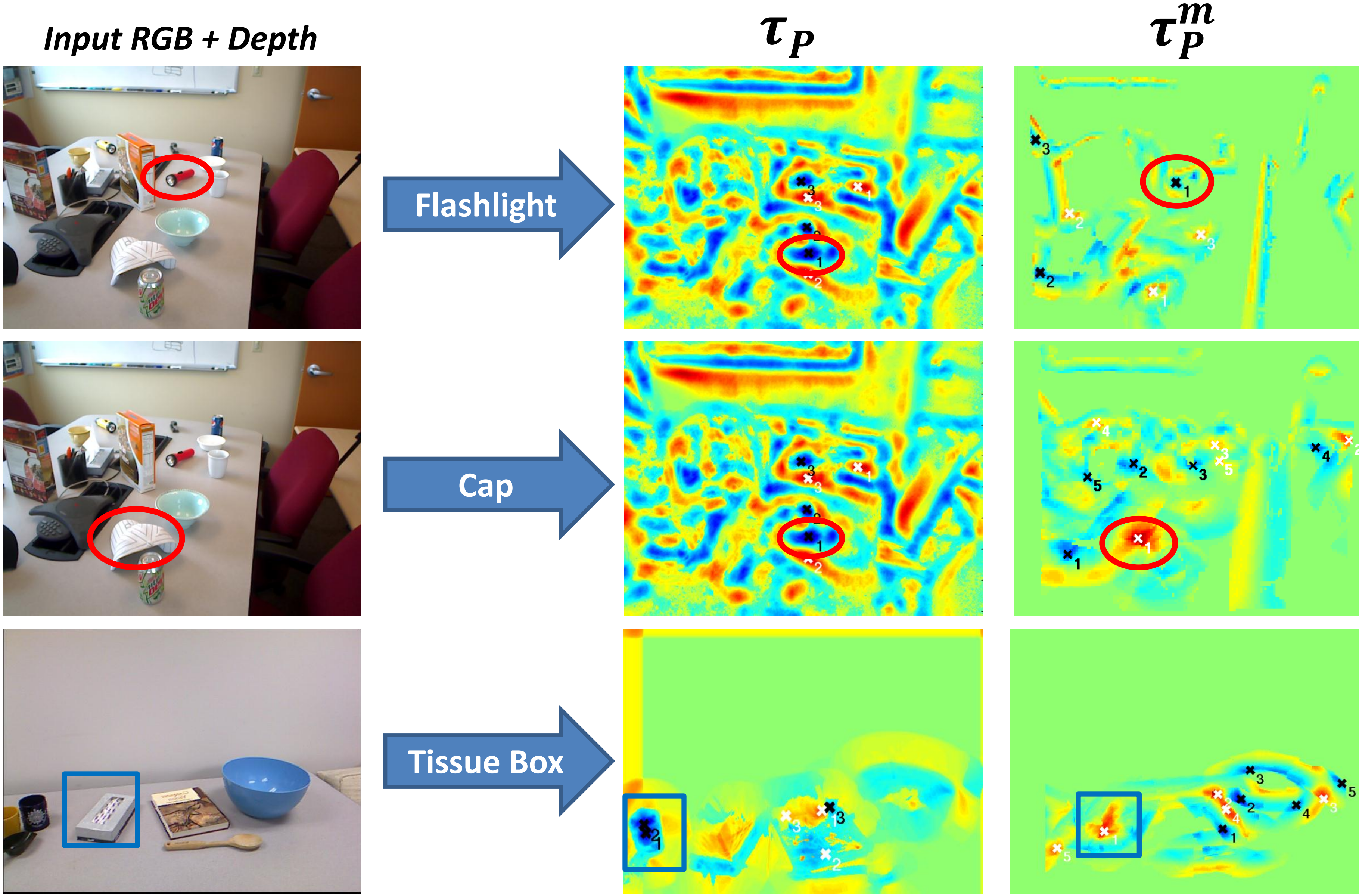
- E.g. Looking for shape A:



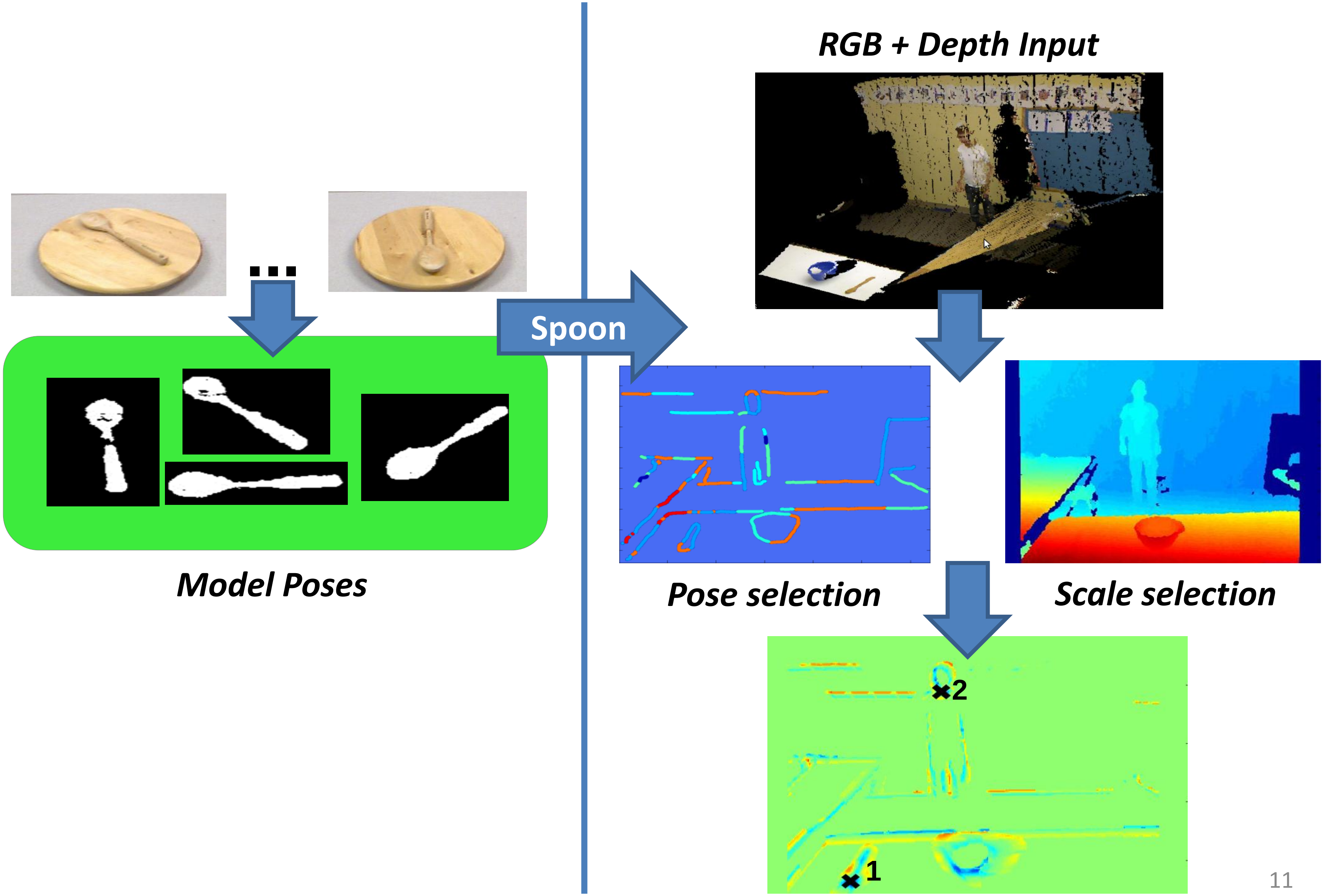
$$\tau^m_{P(A)} >> \tau^m_{P(B)} \sim \tau^m_{P(C)} \sim \tau^m_{P'}$$

- Using τ_P^m – edges that are non-conforming are given less weights → $\tau^m_{P(A)} >> \tau^m_{P(B)} \approx \tau^m_{P(C)} \approx \tau^m_{P'} \neq 0$

Example Results

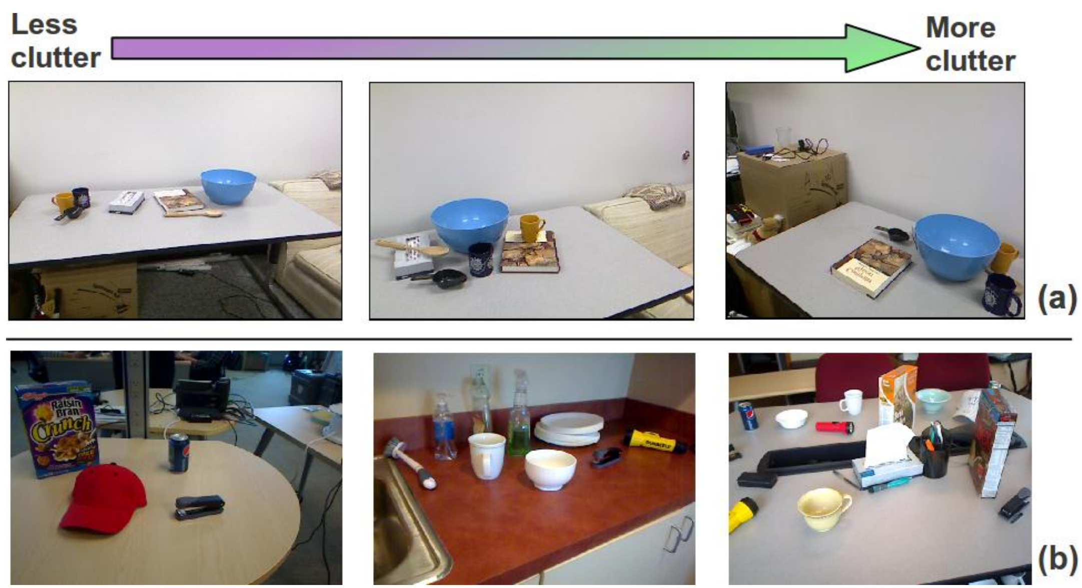


Implementation Details



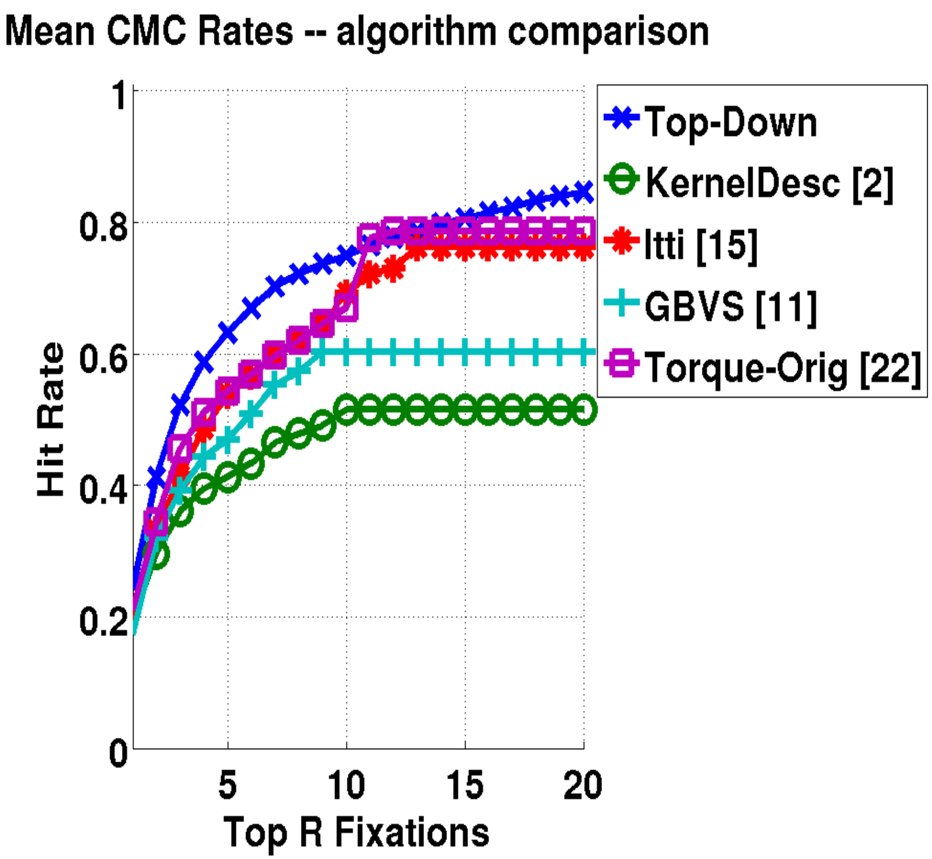
Experiments

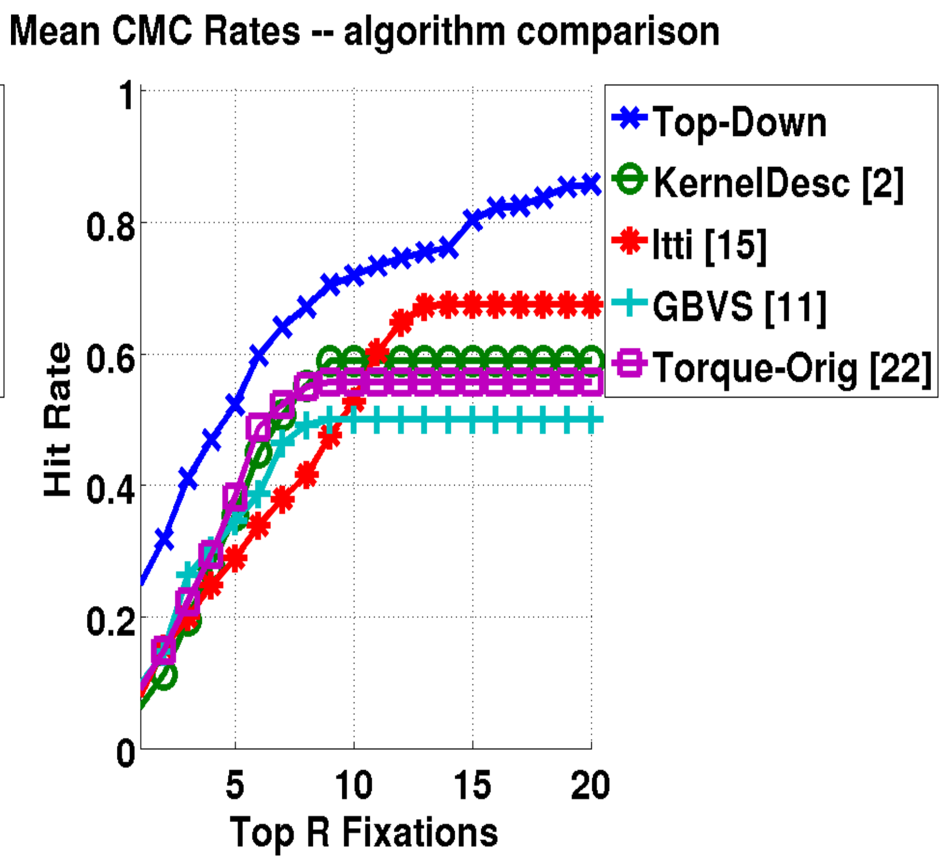
- (a) Cluttered data collected at UMD – 3 sequences, 7 objects
- (b) U Washington Dataset of common objects – 8 sequences, 6 objects



- Evaluate by counting the rank of correct fixations from total fixations returned.

Results



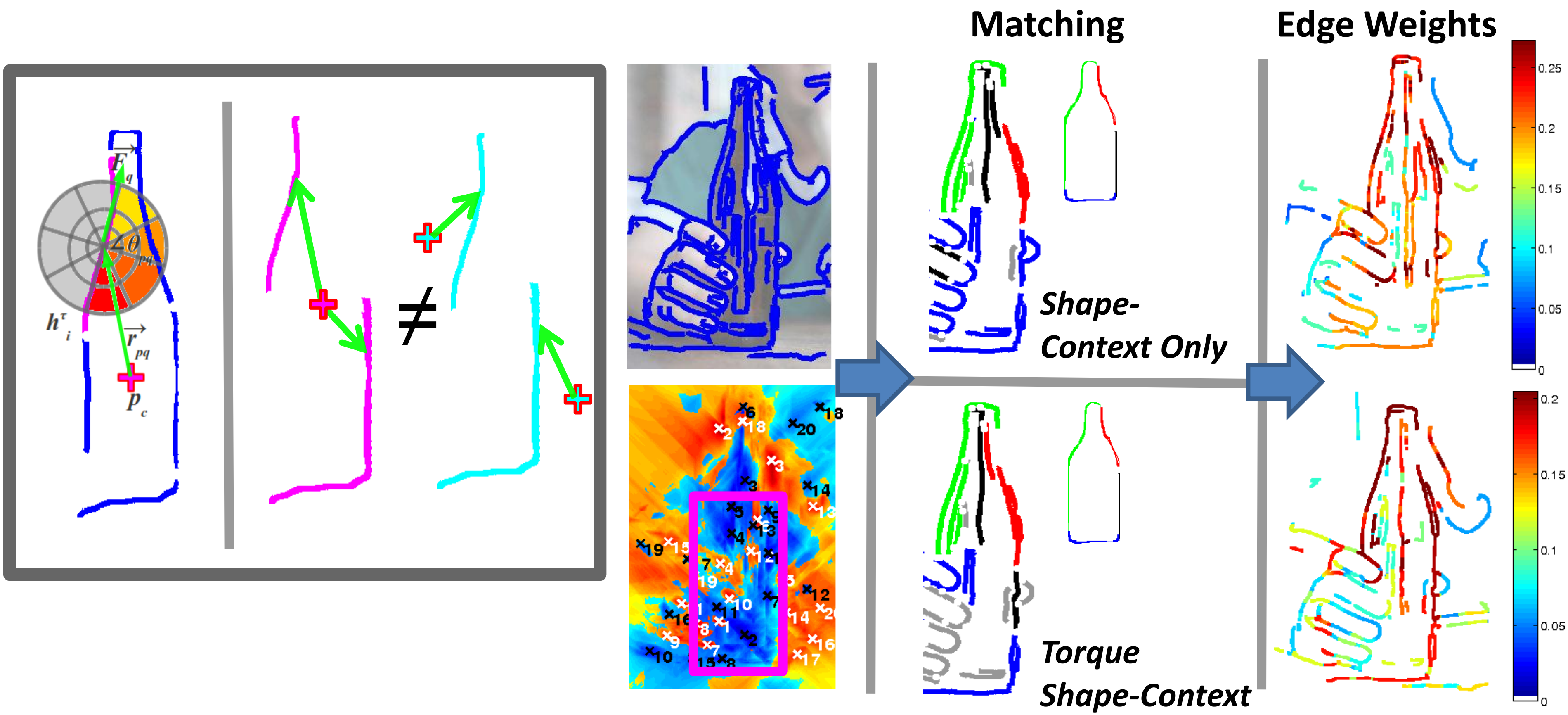


(L) CMC (Cumulative Match Curves) for Clutter sequences,
(R) CMC for U Washington sequences

On-going improvements

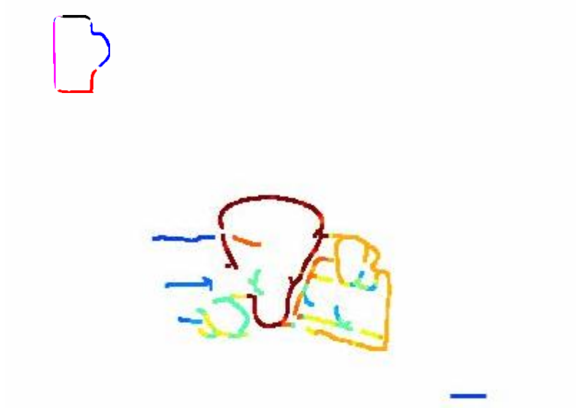
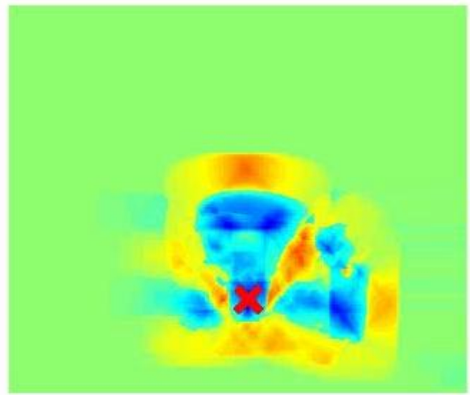
- Searching over full shape model is not fast in practice.
  - **Learn** salient contours from annotated examples
  - **Match** partial contour fragments to models
  - **Reweigh** torque based on matching scores
- Learned contours examples (ETHZ-Shapes):

Matching: Torque Shape-Context

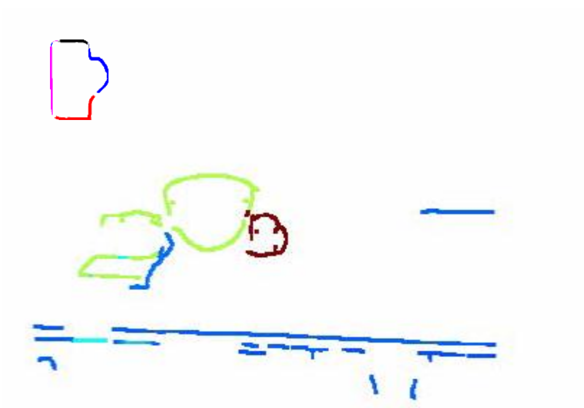
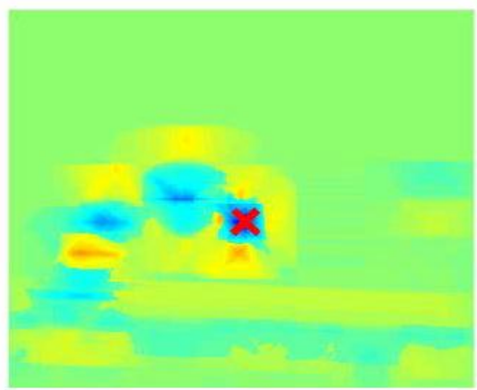


Example results

Edge Weights τ_P^m **RGB Input**


Seq 2


Seq 3

Outlook and Future Work

- Modify the image torque in other ways:
  - Generic set of *shape primitives* common among a set of objects.
- Linking torque to higher-level descriptions: e.g. *attributes*.
- Extend approach with object segmentation for recognition.

Thanks! Questions?

- Contacts:
  - Ching L. Teo cteo@cs.umd.edu

- Acknowledgements: