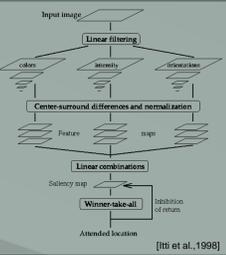


Saliency and Affordance in Artificial Visual Attention

J. Tünnermann & B. Mertsching

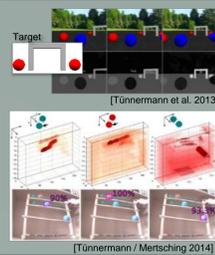
Bottom-up attention (Saliency)

- Classic and region-based models
 - [Itti et al. 1998]
 - [Aziz / Mertsching 2008]
 - [Tünnermann / Mertsching 2013]
- Concepts
 - Local contrasts.
 - Feature Integration.
 - FOA selection.



Top-down attention

- Classic and region-based models
 - [Navalpakkam / Itti 2006]
 - [Aziz / Mertsching 2008]
 - [Tünnermann et al. 2013]
- Concepts
 - Task specific feature weighting.
 - Template-based visual search.



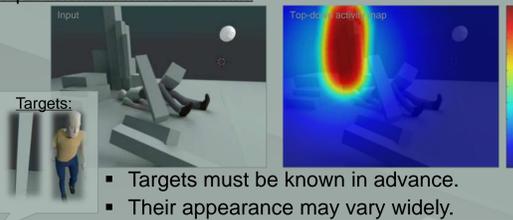
Affordance-based attention?

- Attention towards objects which are potential action targets.
- For a robot with an arm and gripper: estimate grasp affordances.

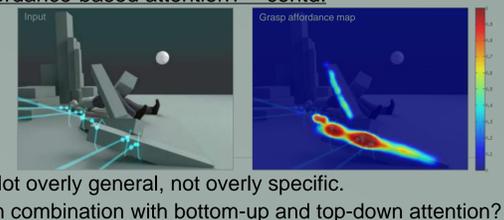
Bottom-up attention (Saliency) – contd.



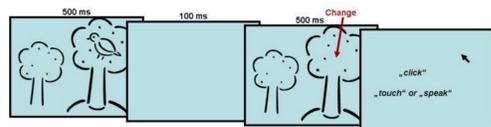
Top-down attention – contd.



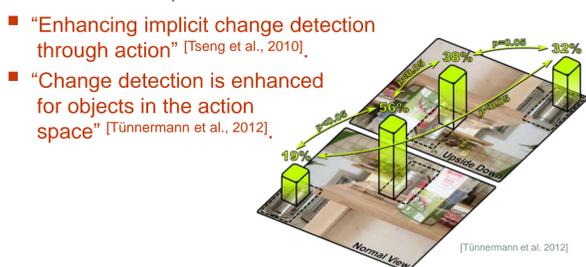
Affordance-based attention? – contd.



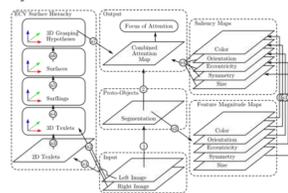
The Change-Blindness Paradigm and Attention



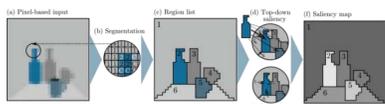
- Participants often fail to see the change
 - The disruption removes the signal that guides attention [Rensink et al. 1997].
- Attention-related manipulations that selectively reduce CB
 - “Central interest” changes vs. “marginal interest” changes [Rensink et al. 1997].
 - “Gaze cues attenuate change blindness in the flicker paradigm” [O'Donnell / Langton, 2003].
 - “The effects scene inversion on change blindness” [Shore / Klein, 2000].



A Combined Model for Artificial Saliency and Affordance



The region-based attention model [A/M 2008], [T/M 2014].



Graspability representation [T/K/M/M in rev.]

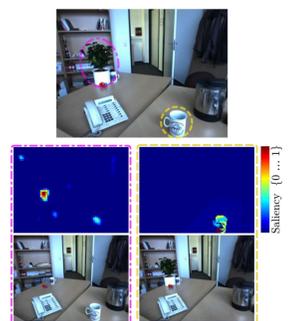
- Early Cognitive Vision.
 - Sparse 3D representation.
 - Developed by Norbert Krüger's group in Odense, Denmark.
- Affordance estimation:
- Grasps are projected into 2D.
 - The relative density of grasp points is calculated for regions.

Comparing Model Predictions with Human Performance in the Change-Blindness Task

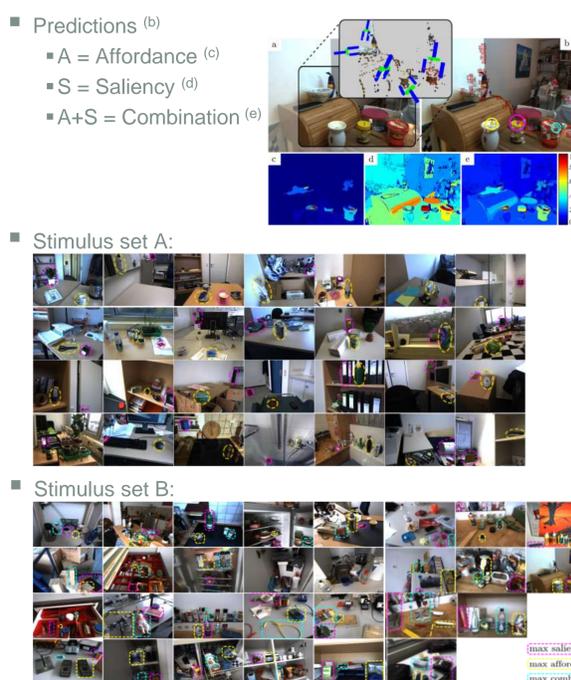
- Experiment 1:** 28 images (set A) / 40 participants.
 - Itti et al. predictions vs. the proposed affordance pred.
- Experiment 2:** 28 images (set A) / 40 participants.
 - Control for experiment 1, images shown upside down.
- Experiment 3:** 28 images (set A) / 10 participants.
 - Pilot to test the change-blindness flicker paradigm.
- Experiment 4:** 29 images (set B) / 40 participants.
 - Region-based saliency vs. affordance vs. combination.

Stimulus creation procedure:

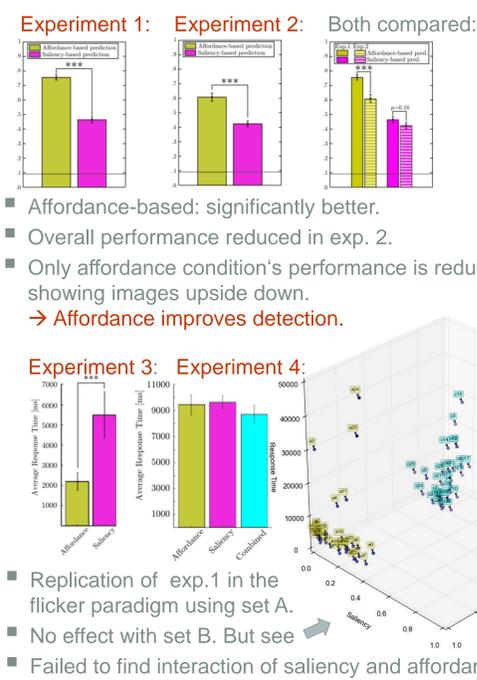
- Capture original image.
- Run model on orig. image. Remove predicted object.
- Capture changed image.
- Locally blend in the changes to create the altered versions.
- (6) Repeat (2) to (5) for all models to be compared.



The Stimulus Material



Results



Discussion & Outlook

- Affordances are important in the guidance of attention
 - Support from experiments 1, 2, 3 by outranking saliency.
 - Support from the literature:
 - Affordances influence attention in cueing experiments [Roberts / Humphreys 2011], [Garrido-Vasquez / Schubö, 2014].
 - EEG and brain imaging studies [Handy et al. 2003].
- Affordance and saliency may not be combinable in all situations (experiment 3).
 - Some scenes too difficult? →
 - Combination strategies?
 - Prefer affordances if available?
- Further Goals
 - Deep integration with feedback between attention, the scene representation, and affordance.
 - Real time implementation in a robot.
 - Practical evaluation.
 - Integration with action execution.