

# Integrating Object Affordances with Artificial Visual Attention

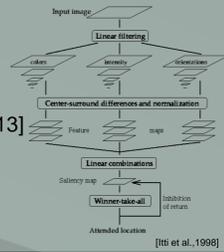
J. Tünnermann, C. Born, & B. Mertsching

# GET Lab

Prof. Dr.-Ing. Bärbel Mertsching  
 Universität Paderborn, Fakultät EIM-E  
<http://getwww.upb.de>

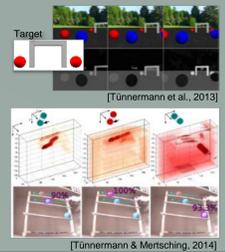
## Bottom-up attention (Saliency)

- 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?

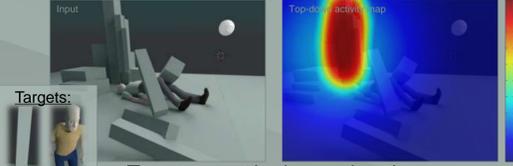
- Attention towards objects which are potential action targets.
- For a robot with an arm and gripper: estimate grasp affordances.

## Bottom-up attention (Saliency) – contd.



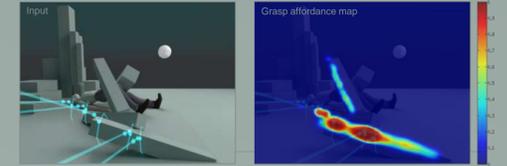
- Task relevant objects are often not salient.
- Salient objects are often not task relevant.

## Top-down attention – contd.

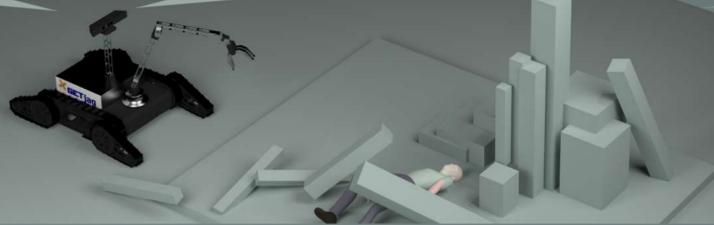


- Targets must be known in advance.
- Their appearance may vary widely.

## Affordance-based attention? – contd.

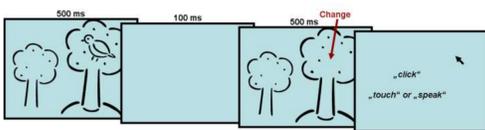


- Not overly general, not overly specific.
- In combination with bottom-up and top-down attention?



©Tünnermann, B. Mertsching, J. Tünnermann, Human-robot model: blender2009.com / brightspace

## The Change-Blindness Paradigm and Attention

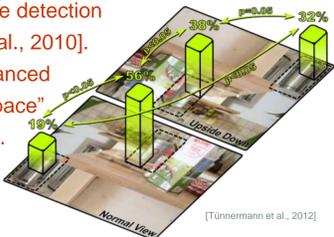


Participants often fail to see the change

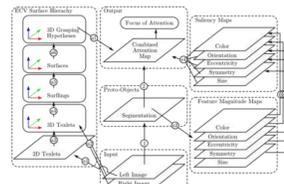
- The disruption removes the signal that guides attention [Rensink et al., 1997].

Attention-related manipulations that selectively reduce CB

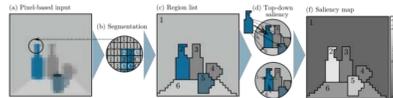
- “Central interest” changes vs. “marginal interest” changes [Rensink et al., 1997].
- “Gaze cues attenuate change blindness in the flicker paradigm” [Langton & O'Donnell, 2003].
- “The effects scene inversion on change blindness” [Shore & Klein, 2000].
- “Enhancing implicit change detection through action” [Tseng et al., 2010].
- “Change detection is enhanced for objects in the action space” [Tünnermann et al., 2012].



## A Combined Model for Artificial Saliency and Affordance



The region-based model [A&M, 2008], [T&M, 2014]:

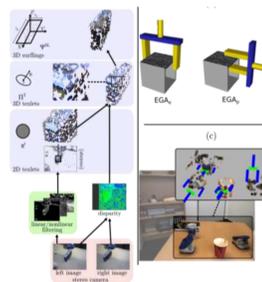


Integrating graspability-based affordances: [Tünnermann et al., in review]:

- 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.

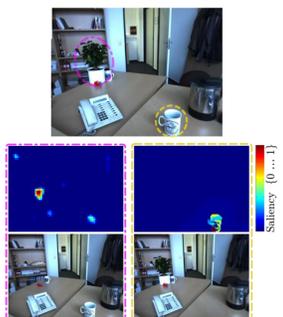
- Pilot to test the change-blindness flicker paradigm.

Experiment 4: 29 images (set B) / 40 participants.

- Region-based saliency vs. affordance vs. combination.

Stimulus creation procedure:

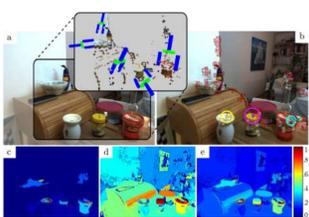
- Capture original image.
- Run model on orig. image. Remove predicted object.
- Capture changed image.
- Locally blend in the changes to create the altered versions.
- (6) Repeat (2) to (5) for all models to be compared.



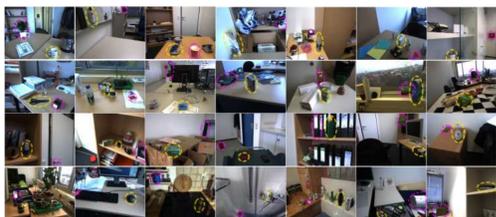
## The Stimulus Material

Predictions (b)

- A = Affordance (c)
- S = Saliency (d)
- A+S = Combination (e)



Stimulus set A:

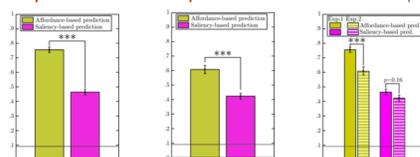


Stimulus set B:



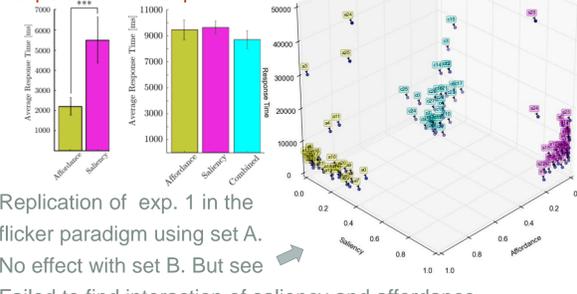
## Results

Experiment 1: Experiment 2: Both compared:



- Affordance-based: significantly better.
- Overall performance reduced in experiment 2.
- Only affordance condition's performance is reduced by showing images upside down.
- Affordance improves detection.

Experiment 3: Experiment 4:



- Replication of exp. 1 in the flicker paradigm using set A.
- No effect with set B. But see
- Failed to find interaction of saliency and affordance.

## Discussion & Outlook

Affordances are important in the guidance of attention

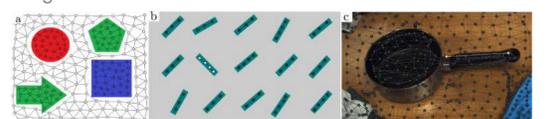
- Support from experiments 1, 2, 3 by outranking saliency.
- Support from the literature:
  - Affordances influence attention in cueing experiments [Roberts et al., 2011], [Garrido-Vasquez et al., 2014].
  - EEG and brain imaging studies [Handy et al., 2003].

Affordance and saliency may not be combinable in all situations (experiment 3)

- Some scenes too difficult? →
- Combination strategies?



A deeper integration of attention and affordances based on Growing Neural Gas?



- Proto-objects based on Growing Neural Gas (a).
- Saliency and further attentional mechanisms (b).
- Applying monocular affordance estimation, as proposed by [Song et al, 2011], on these network structures (c).