GET Lab
Prof. Dr.-Ing. Bärbel Mertsching
University of Paderborn
http://getwww.upb.de

Saliency and Affordance in Artificial Visual Attention

Jan Tünnermann & Bärbel Mertsching

July 13, 2014

Motivation & Introduction

- Like humans, autonomous robots need to perceive and act in dynamic environments.

→ Mimic human attention mechanisms to filter relevant from irrelevant information at early stages.

Motivation & Introduction

- A testbed for autonomous mobile robots:

RoboCup Rescue

Autonomous GETbot

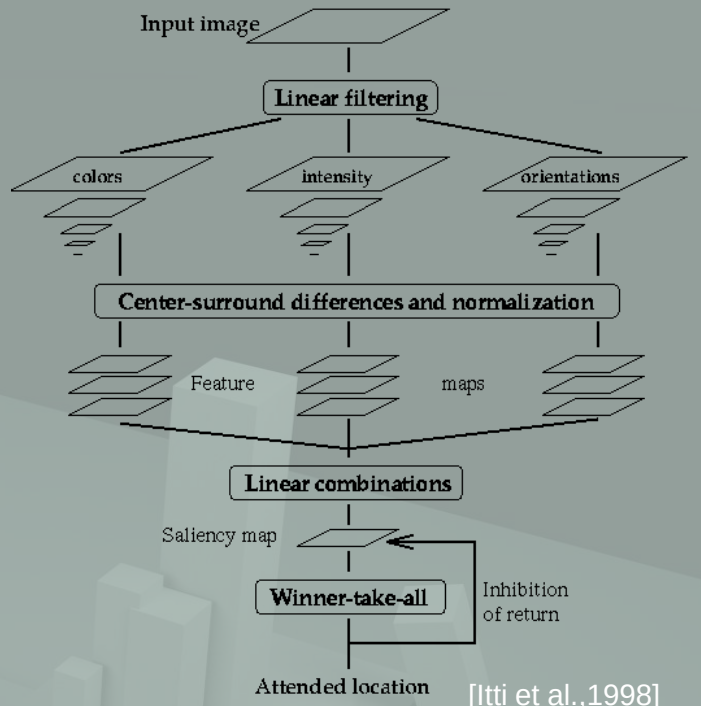
Highly mobile GETjag

[Source: Warwick Mobile Robotics]

Motivation & Introduction

Bottom-up attention (Saliency):

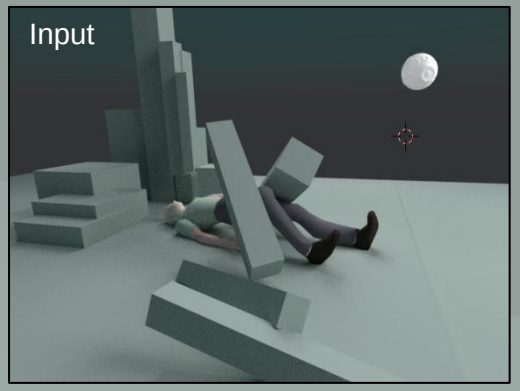
- Classic and region-based models:
  - [Itti et al. 1998]
  - [Aziz / Mertsching 2008]
  - [Tünnermann / Mertsching 2013]
- Concepts:
  - Local contrasts.
  - Feature Integration.
  - FOA selection.

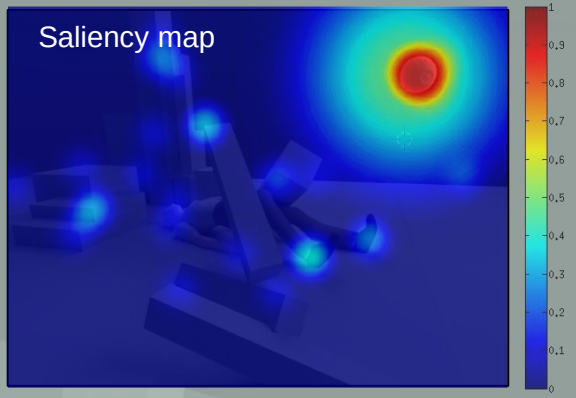


[Itti et al.,1998]

Motivation & Introduction

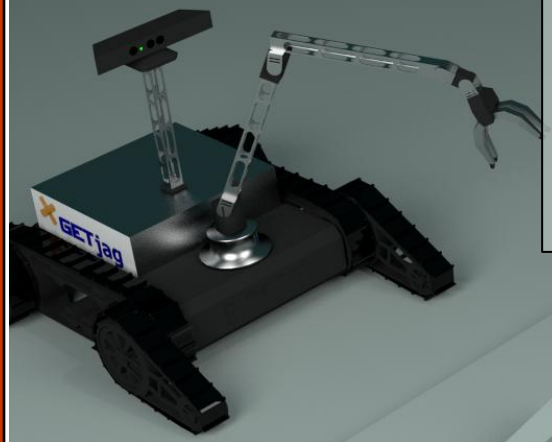
Bottom-up attention (Saliency):





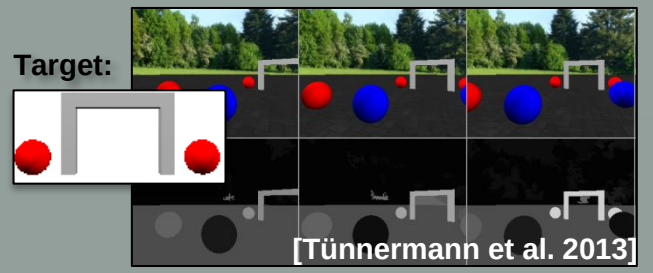
- Task relevant objects are often not salient.
- Salient objects are often not task relevant.

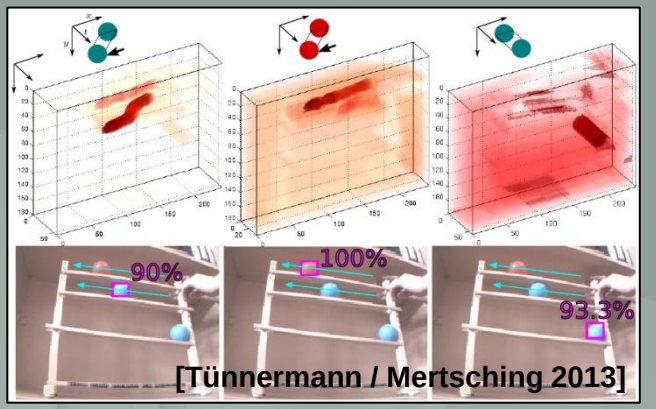
Motivation & Introduction



Top-down attention:

- Classic and region-based models:
  - [Navalpakkam / Itti 1998]
  - [Aziz / Mertsching 2008]
  - [Tünnermann et al. 2013]
- Concepts:
  - Task specifc feature weighting.
  - Template-based visual search.

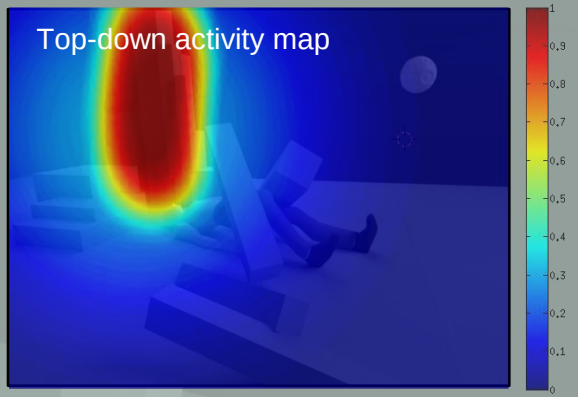




Motivation & Introduction

Targets:

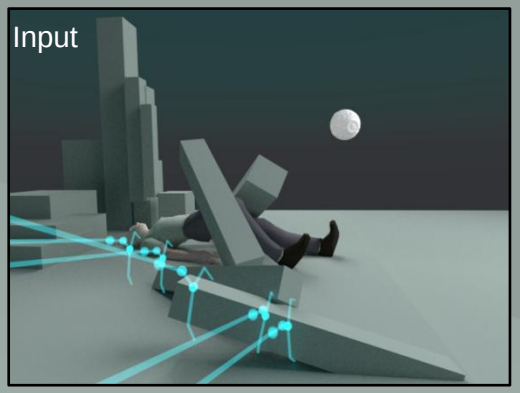
Top-down attention:

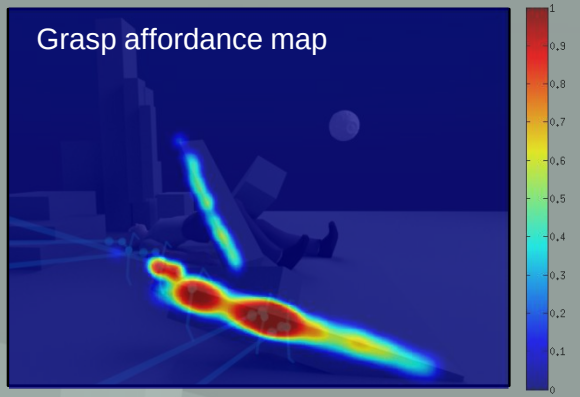


- Targets must be known in advance.
- The appearance of targets may differ in various situations widely.

Motivation & Introduction

Affordance-based attention?





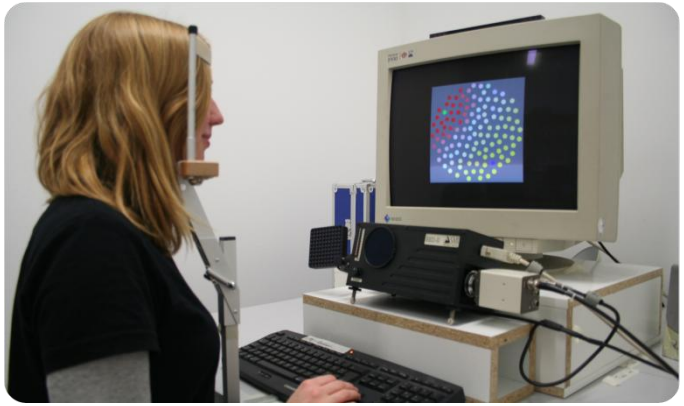
- Attention towards potential action targets.
- Not overly general, not overly specific.
- In combination with saliency?

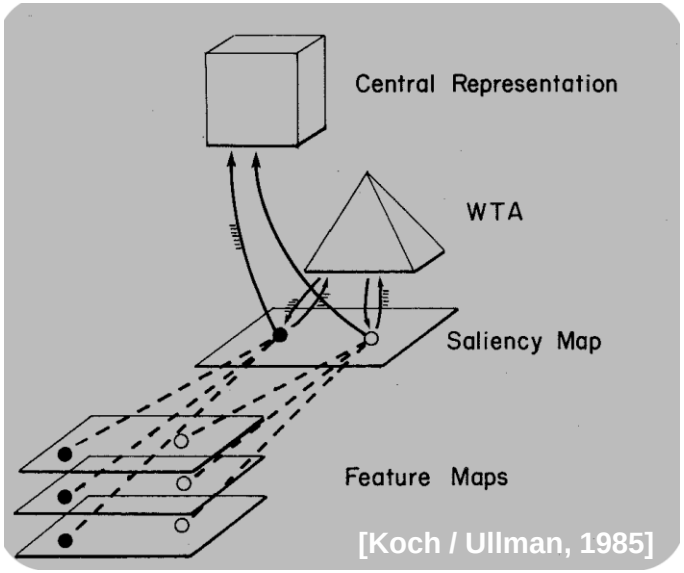
Outline

- Introduction
- Measuring Human Attention in Natural Scenes
- Affordance-based Attention
- Combining Affordance and Saliency
- Discussion and Outlook

Measuring Human Attention in Natural Scenes

- Common practice
  - Eye-tracking experiments.
  - Psychophyiscal experiments (reactions times, detection accuracy, etc.).
- Difficulties in eye-tracking with regard to natural scene perception
  - It only reveals overt attention shifts.
  - It is highly influenced by hard-to-control trop-down influences [Yarbus 1967].
  - Center bias, etc.
- Difficulties with most psychophysical tasks
  - Use of highly artificial stimuli, very unlike natural scenes.
- A promising option
  - The change-blindness paradigm.





Measuring Human Attention in Natural Scenes

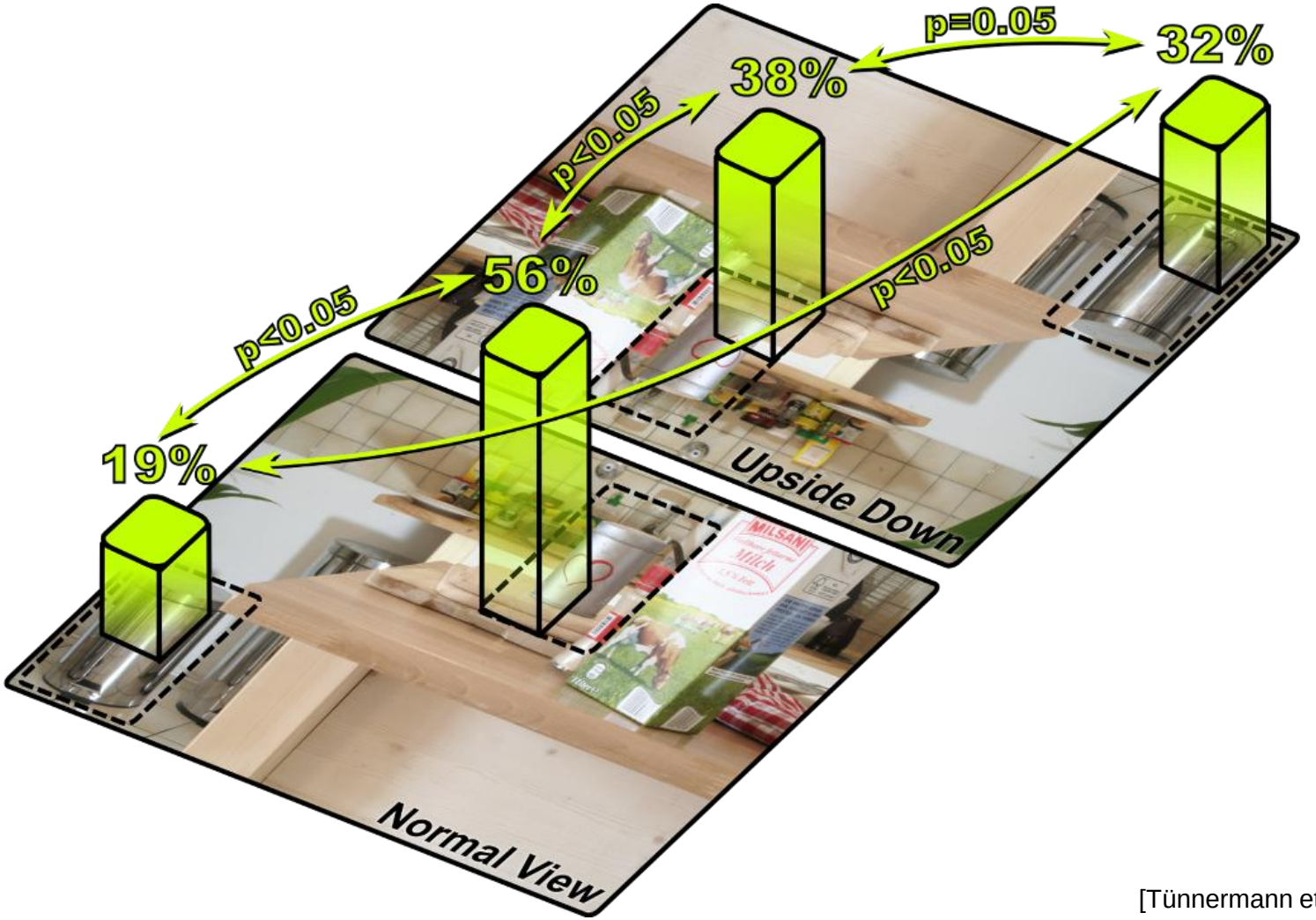
The Change-Blindness Effect: How is attention involved?

- 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].
- "Enhancing implicit change detection through action" [Tseng et al., 2010].
- "Change detection is enhanced for objects in the action space" [Tünnermann et al., 2012].

Measuring Human Attention in Natural Scenes



[Tünnermann et al. 2012]

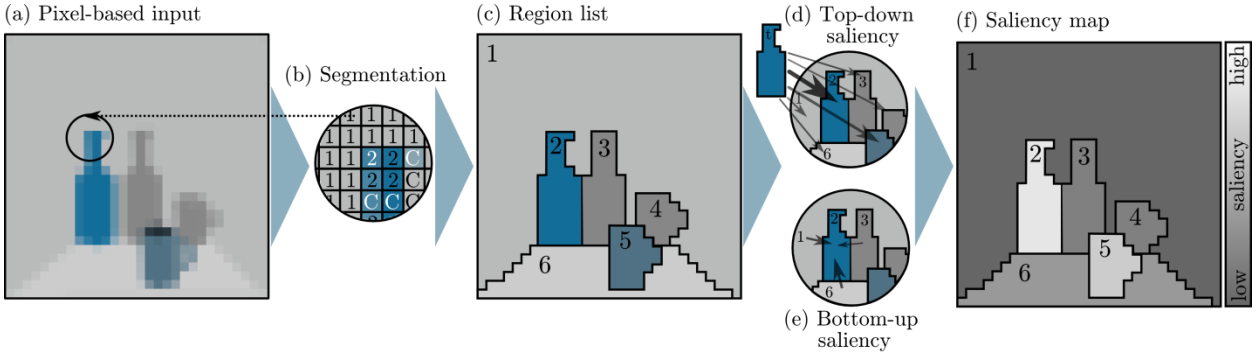
Outline

- Introduction
- Measuring Human Attention in Natural Scenes
- Affordance-based Attention
- Combining Affordance and Saliency
- Discussion and Outlook

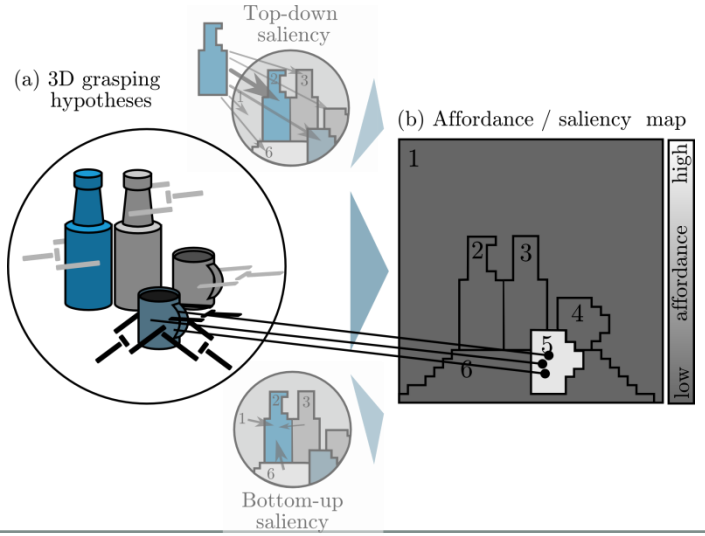
Affordance-based Attention

General approach

- Region-based saliency [GET Lab]:



- Include a graspability estimate:



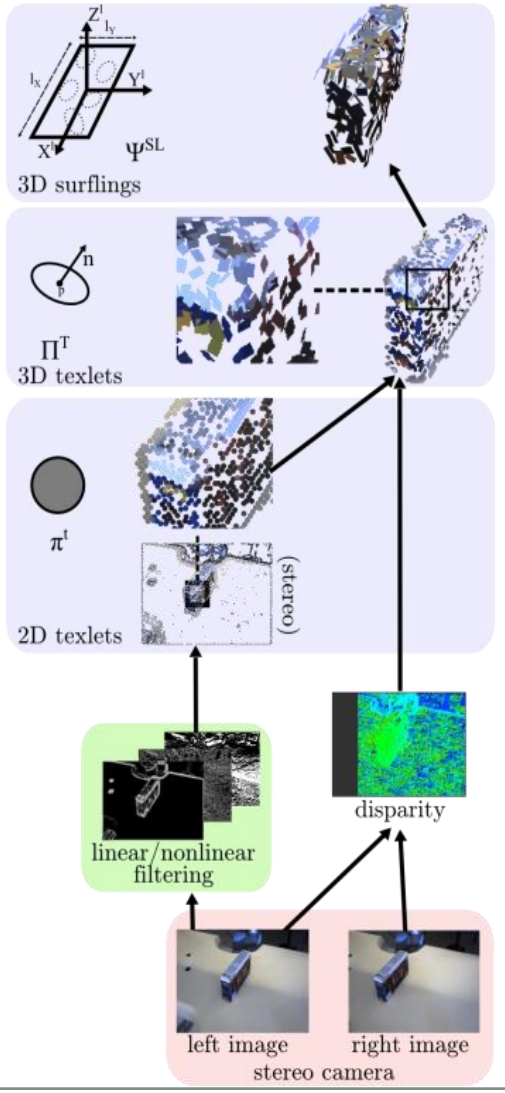
Affordance-based Attention

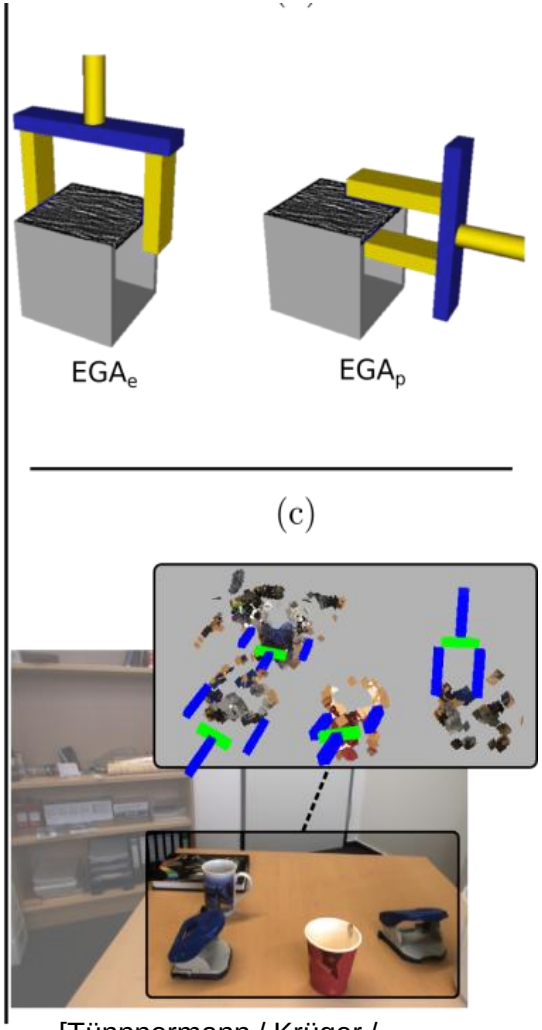
Graspability representation:

- ECV (Early Cognitive Vision).
- A sparse 3D scene representation.
- Developed by Norbert Krüger's group in Odense, Denmark.

Affordance estimate:

- Grasp points are projected into 2D.
- The relative density of grasp points is calculated for every region.



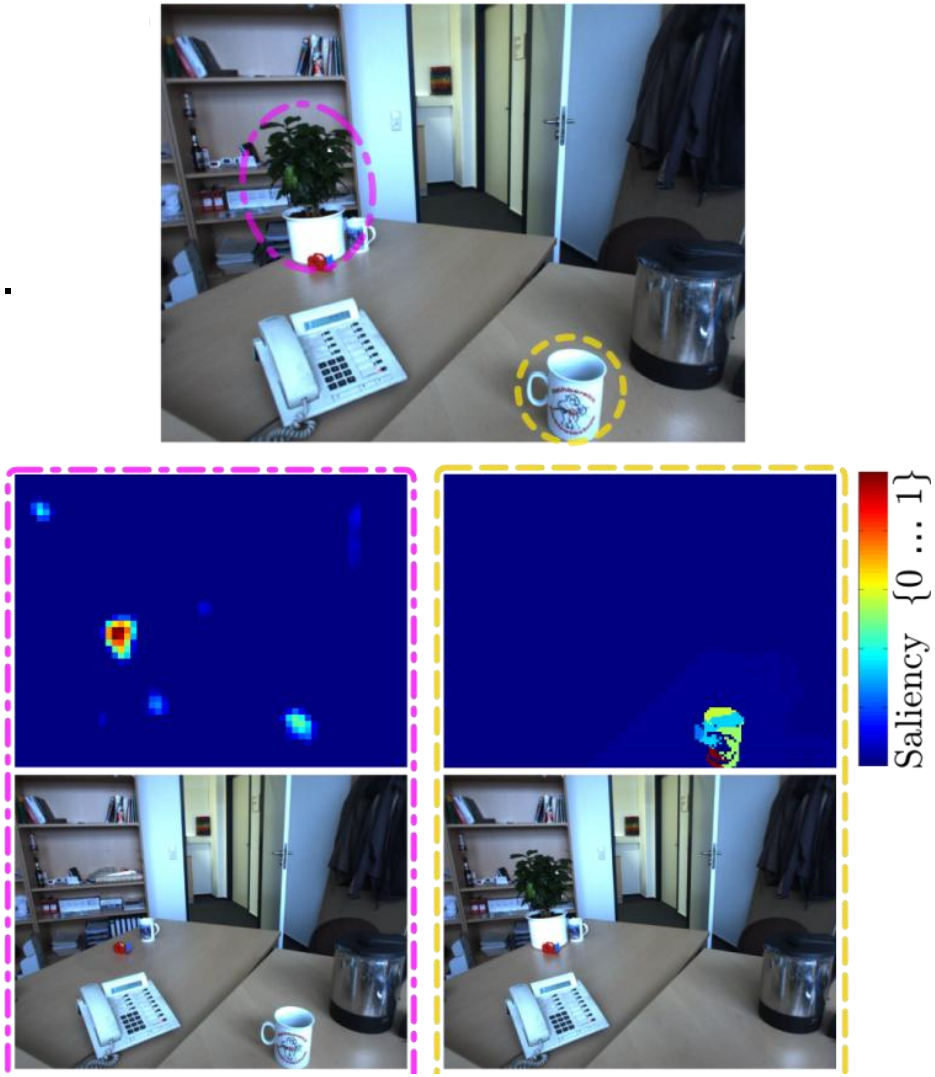


[Tünnermann / Krüger / Mertsching / Mustafa; in review]

Affordance-based Attention

Design of Experiment 1 & 2

- A pure region-based affordance model is compared to the popular bottom-up saliency model by Itti et al. in a change detection task.
- Removals are based on each model's prediction.
- The single shot change-blindness paradigm was used.
- 40 participants in experiment 1 and 40 participants in experiment 2 (a control with images shown upside down).



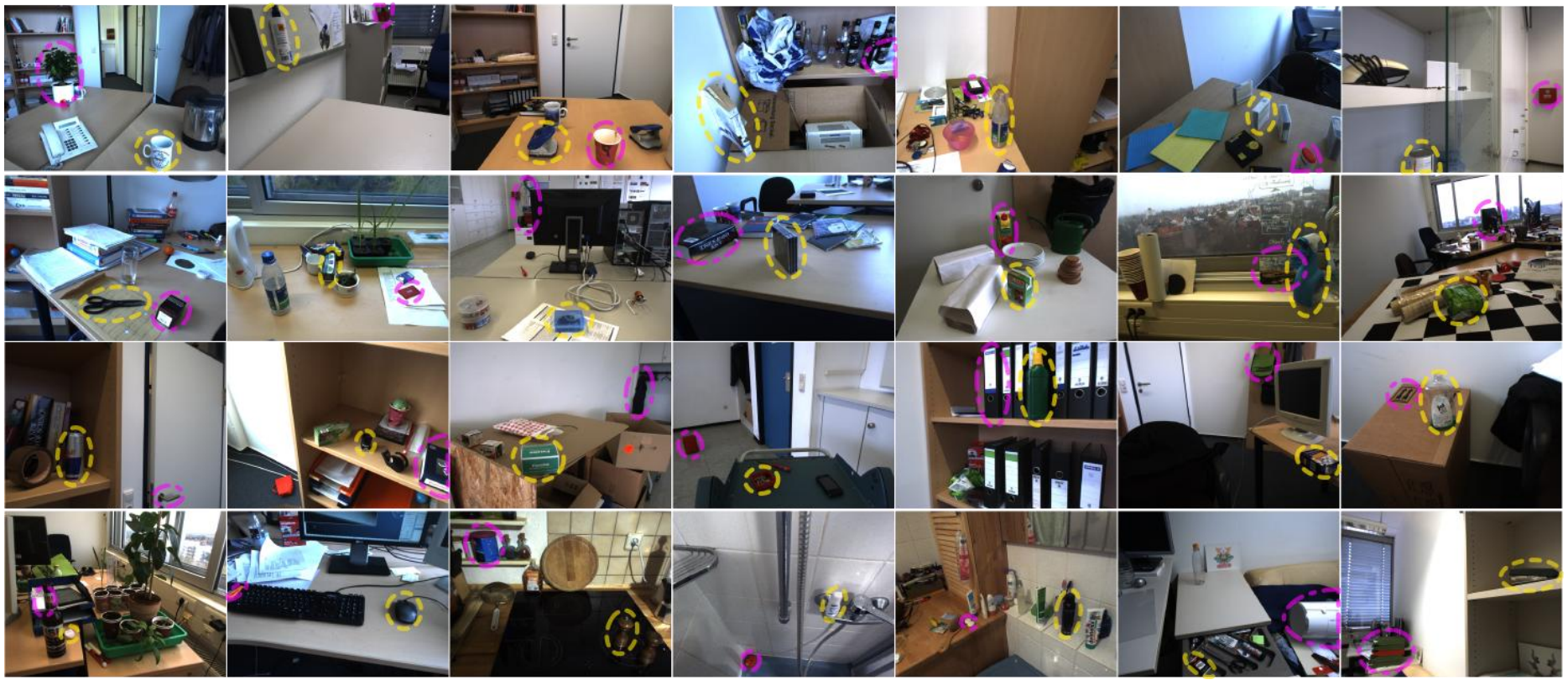
Fachbereich

Fachbereich

Fachbereich

Affordance-based Attention

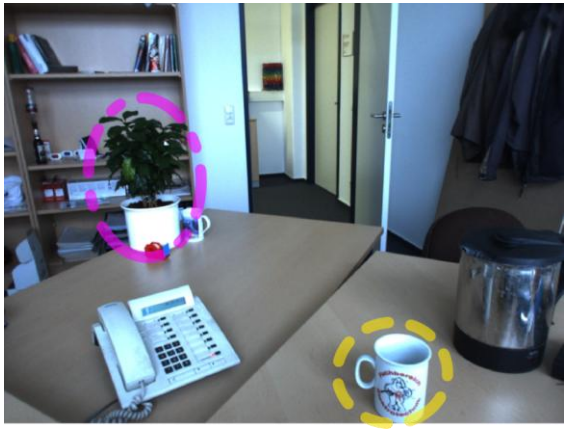
Stimulus Material of Experiment 1 & 2

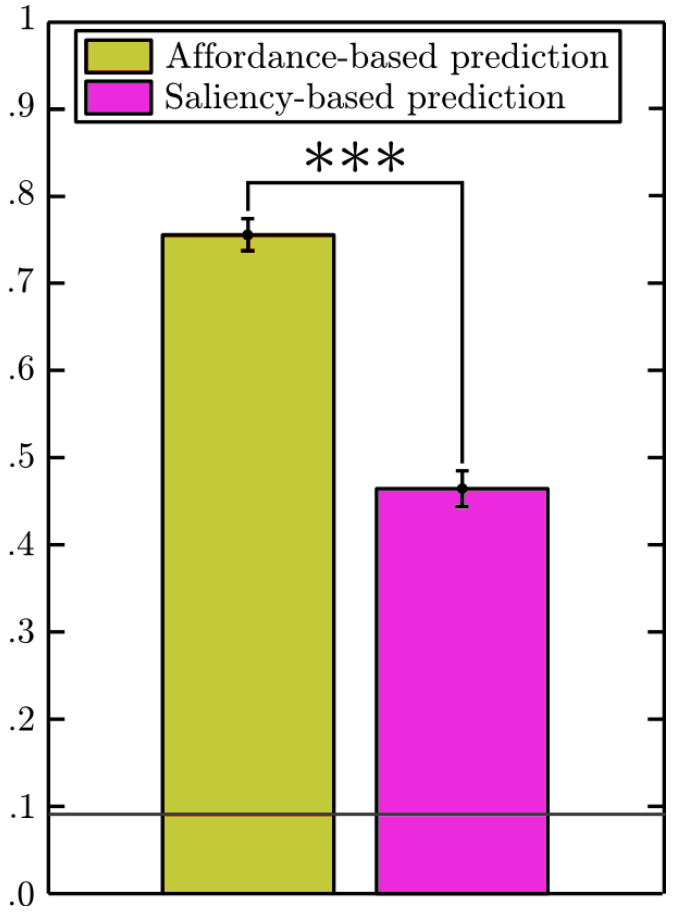


Affordance-based Attention

Results of Experiment 1

- Affordance-based predictions have a significantly better detection performance.

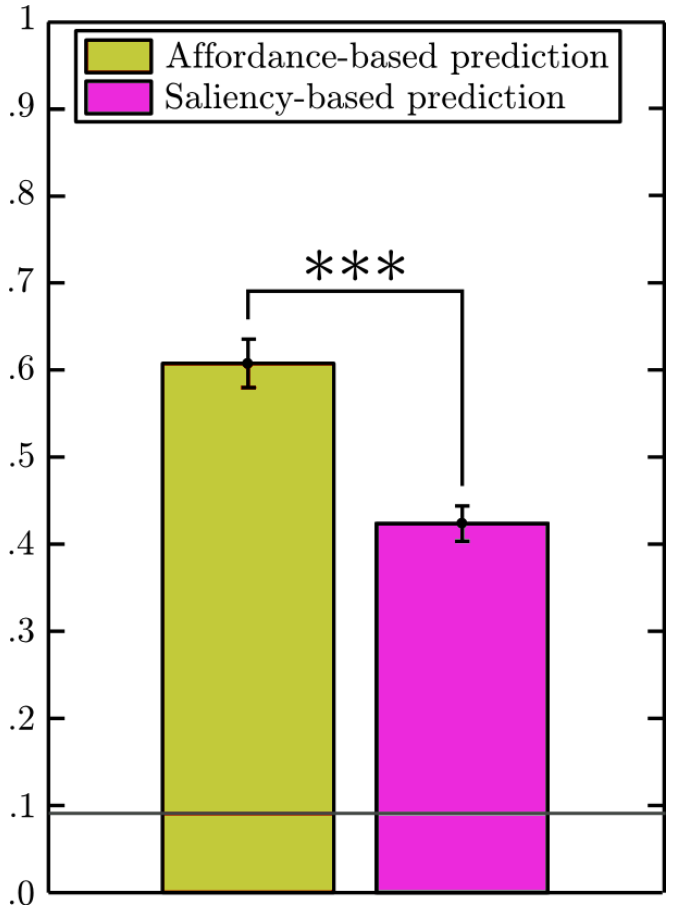




Affordance-based Attention

Results of Experiment 2

- Affordance-based predictions have a significantly better detection performance.
- The overall performance, as well as the difference, is reduced in experiment 2 when the images are shown upside down.

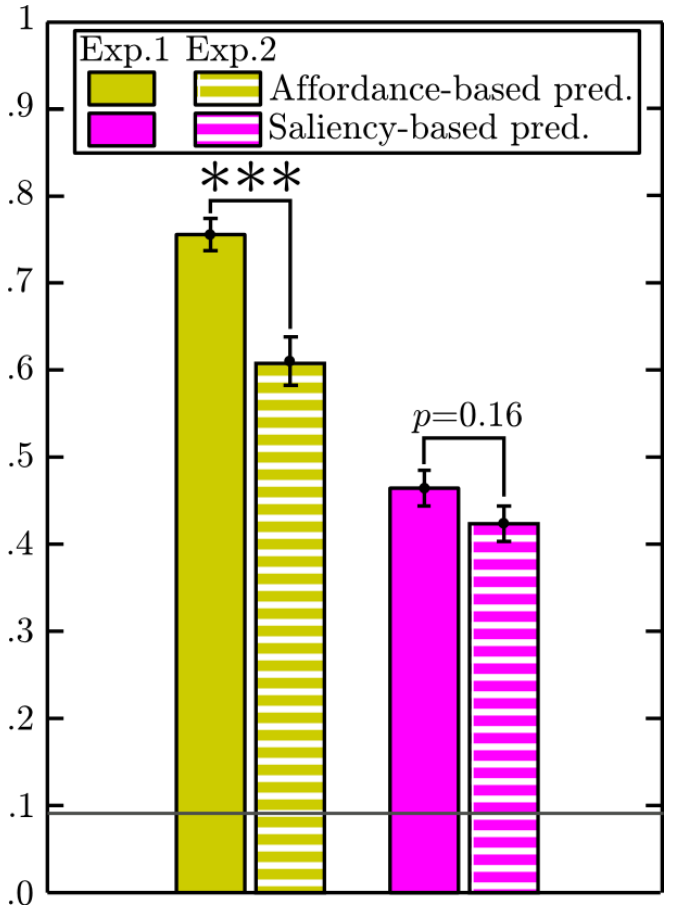


Combining Affordance and Saliency

Results of Experiment 1 & 2

- Affordance-based predictions have a significantly better detection performance.
- The overall performance, as well as the difference, is reduced in experiment 2 when the images are shown upside down.
- The affordance-based predictions differ significantly in experiment 1 and 2, whereas the saliency-based predictions are not significantly different.

→ Affordance produces an attentional benefit.

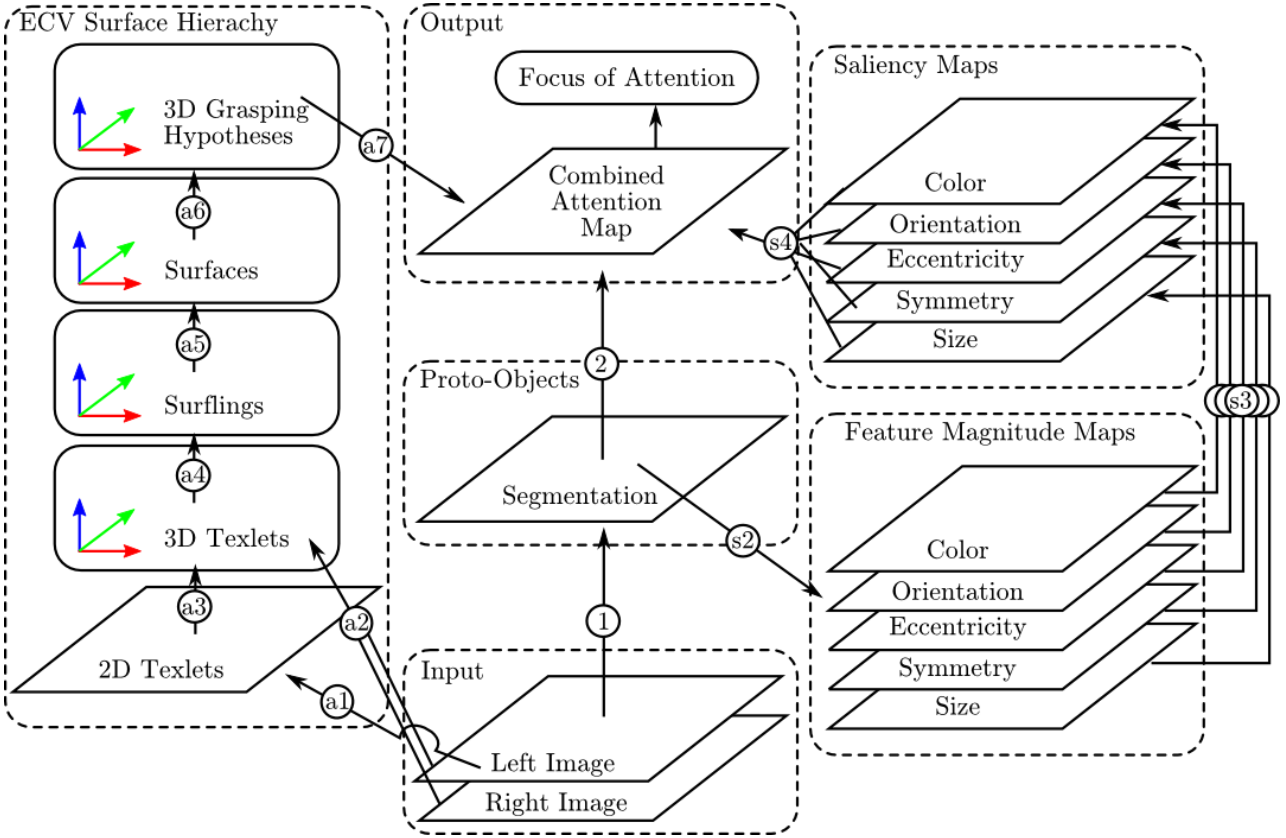


Outline

- Introduction
- Measuring Human Attention in Natural Scenes
- Affordance-based Attention
- Combining Affordance and Saliency
- Discussion and Outlook

Combining Affordance and Saliency

Region-based affordance can be combined with the region-based saliency:



Does this further improve the model with regard to human performance in the change detection task?

Affordance-based Attention

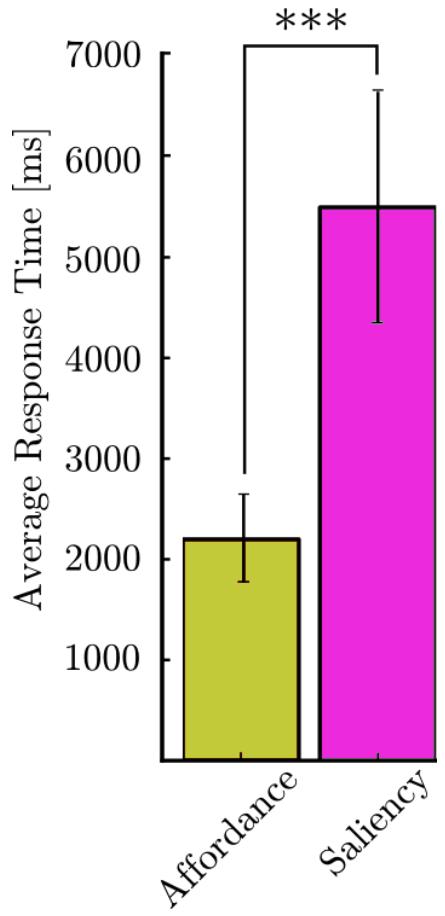
Design of Experiment 3 & 4

- The flicker change-blindness paradigm was used.
- 10 participants in experiment 3 (pilot to test the flicker paradigm).
  - Stimulus material from experiment 1 & 2 with the region-based affordance and saliency (Itti) predictions.
- 40 participants in experiment 4
  - A pure region-based affordance model is compared to pure region-based saliency and their combination.

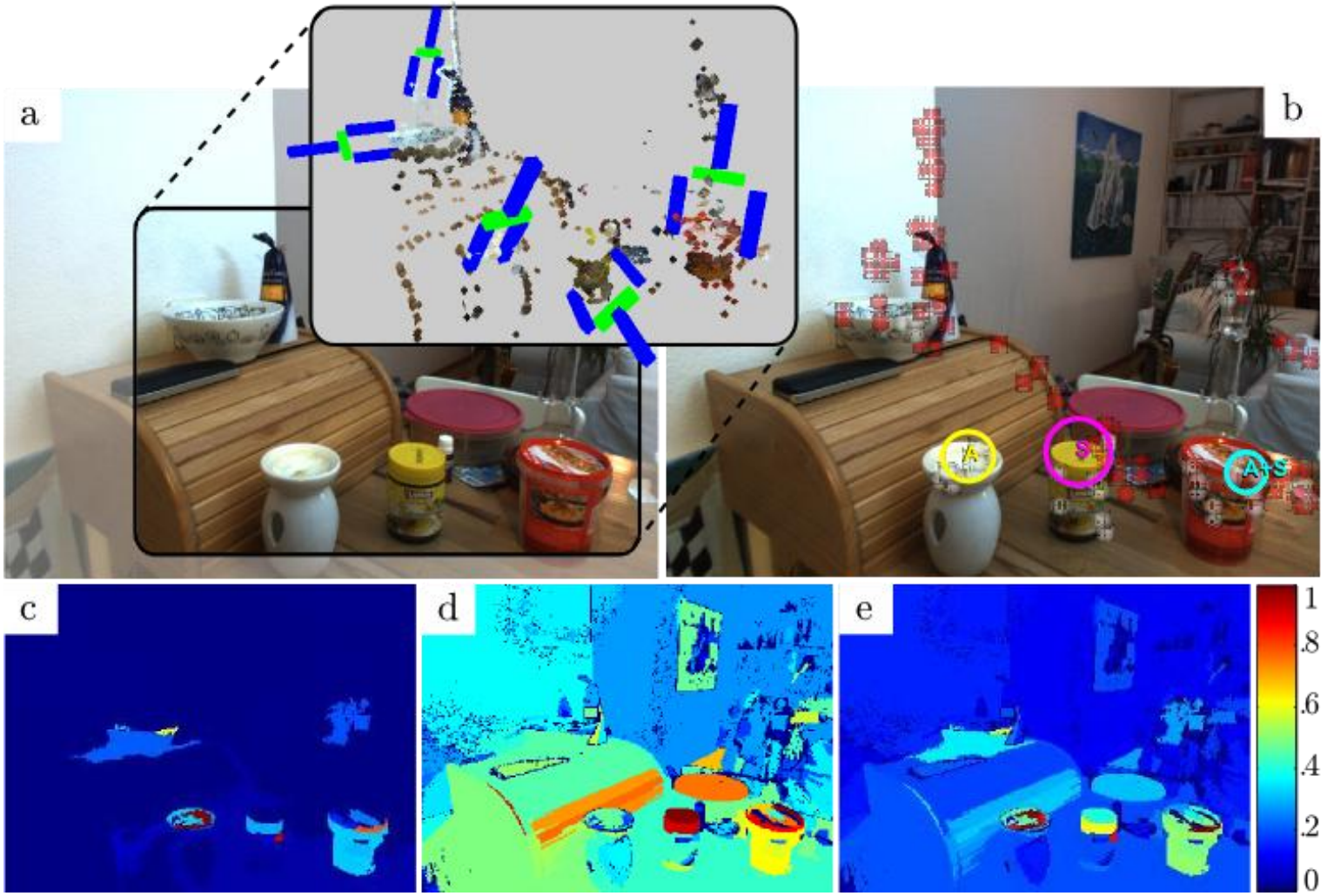
Affordance-based Attention

Results of Experiment 3

- In the flicker paradigm the images alternate until the response is given.
- The results of the flicker paradigm pilot replicate the earlier findings of the single-shot paradigm.



Affordance-based Attention

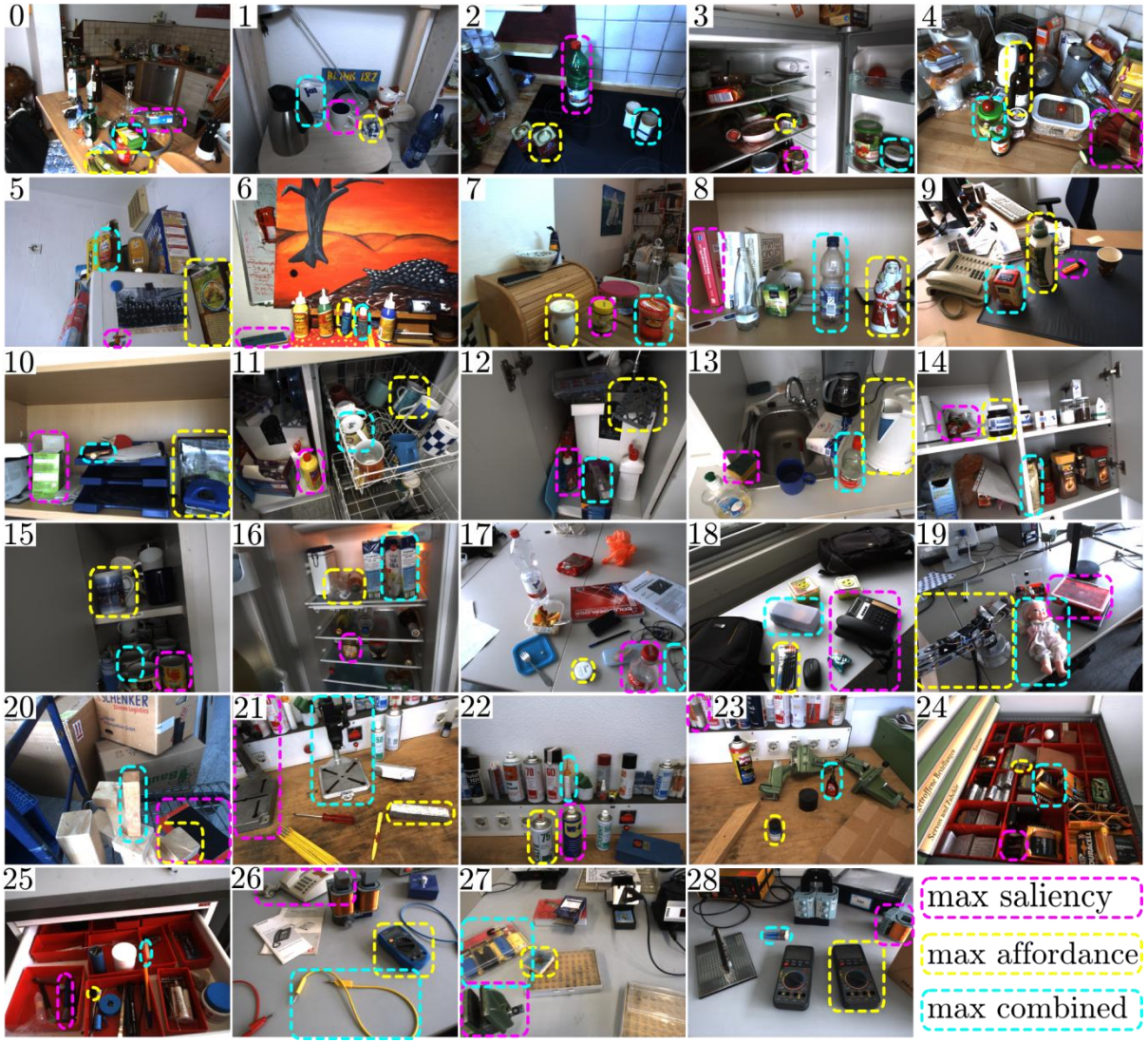


(c) Affordance, (d) Saliency, (e) Combination of both.

Affordance-based Attention

Stimulus Material of Experiment 4

- “Foreground scenes”



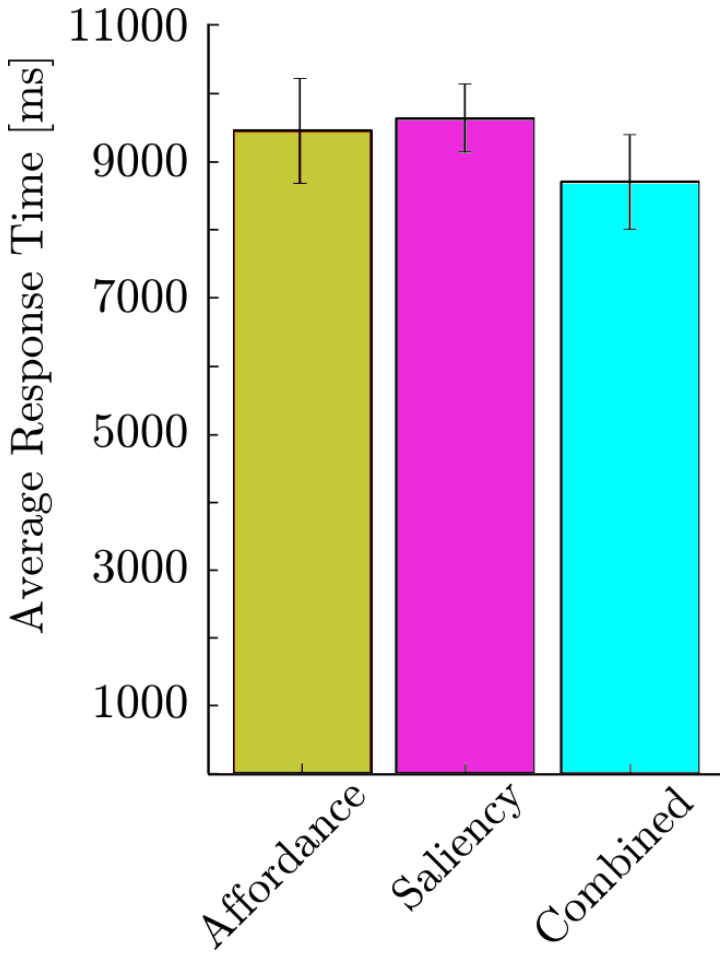
Affordance-based Attention

Evaluation & Results (experiment 4)

- We fail to find a significant difference between the region-based affordance region-based saliency predictions.
- We find no advantage for the combined approach.

What could explain this dissociation with our earlier results?

- No enhancement due to saliency is in line with other studies [Stirk / Underwood 1997].
- Different saliency model.
- Scene type (no real background objects).
- More difficult images.

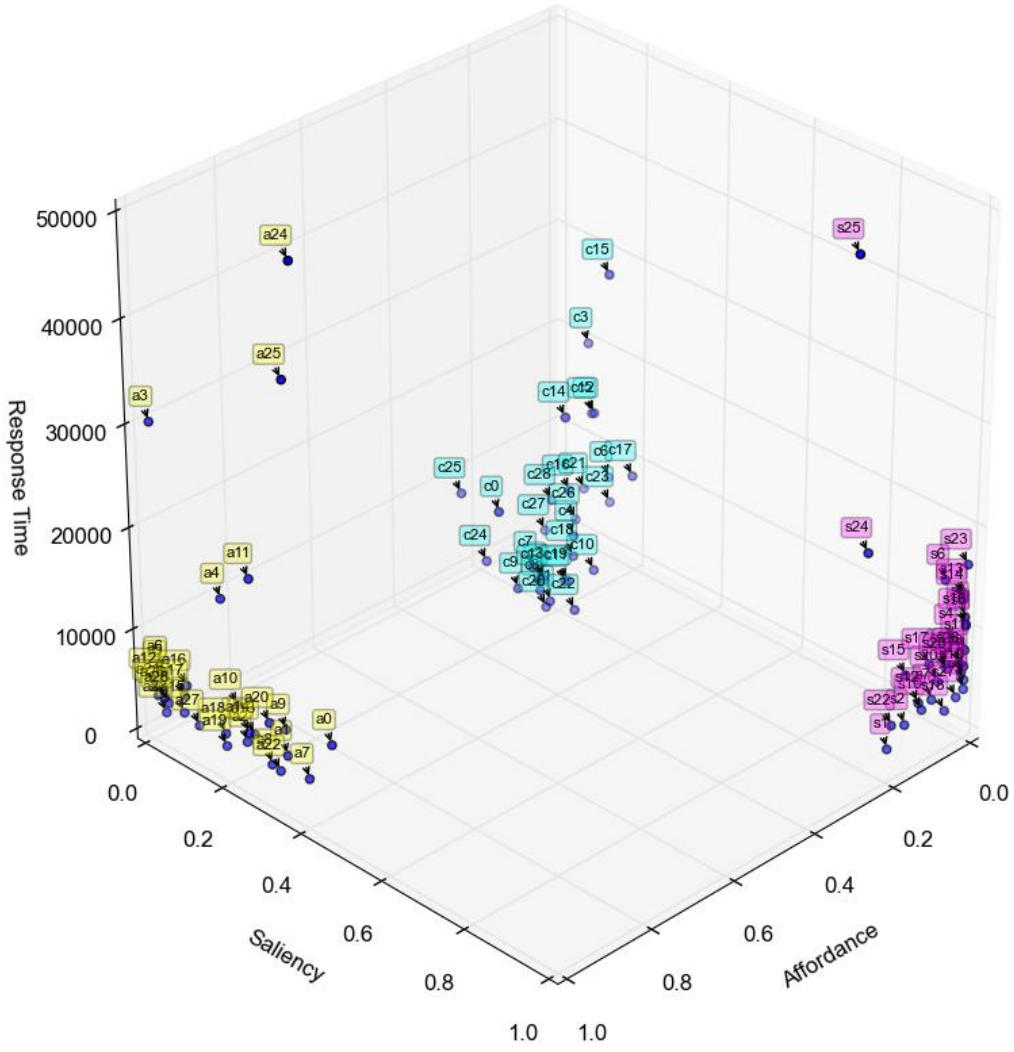


Affordance-based Attention

Affordance-based Attention

Results of Experiment 4 (contd.)

- Distribution of response times for every image mapped in the estimated affordance—saliency space.
- Difficult images have long response times.
- For the saliency-based predictions there may be an effect of affordance on response times.



Outline

- Introduction
- Measuring Human Attention in Natural Scenes
- Affordance-based Attention
- Combining Affordance and Saliency
- Discussion and Outlook

Discussion and Outlook

Affordances are important in the guidance of attention

- Support from experiments 1, 2 and 3 where they outrank saliency.
- Support from the literature:
  - Affordances influence attention in cueing experiment [Roberts / Humphreys 2011], [Garrido-Vasquez / Schubo, 2014].
  - EEG and brain imaging studies [Handy et al. 2003].

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

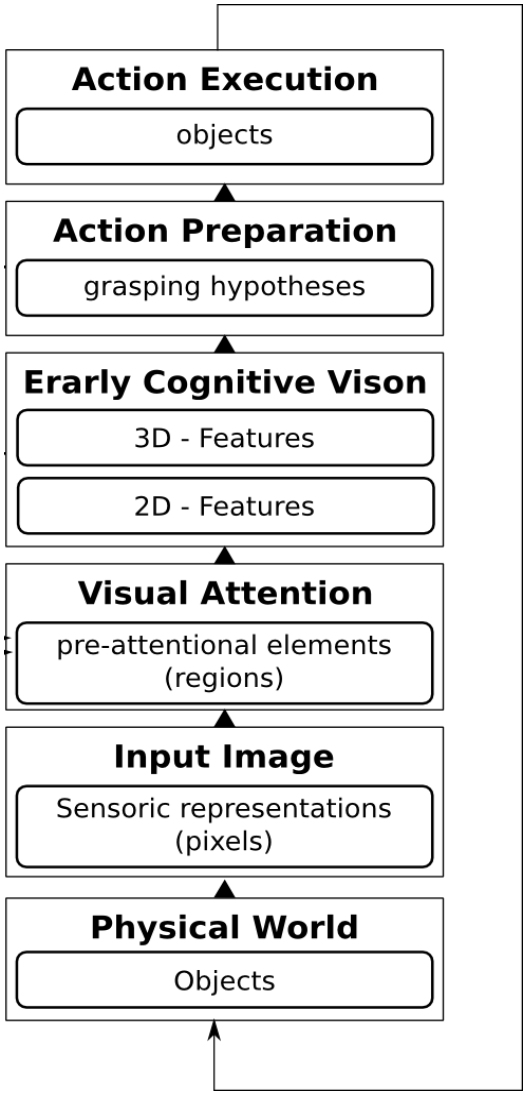
- Combination strategies?
- Prefer affordances if available?

Discussion and Outlook

Goal:

Deeper integration of affordance and attention.

- Attention may influence the scene representation and thereby the grasp estimation.
- Affordances may in turn feed back to attention.



Discussion and Outlook

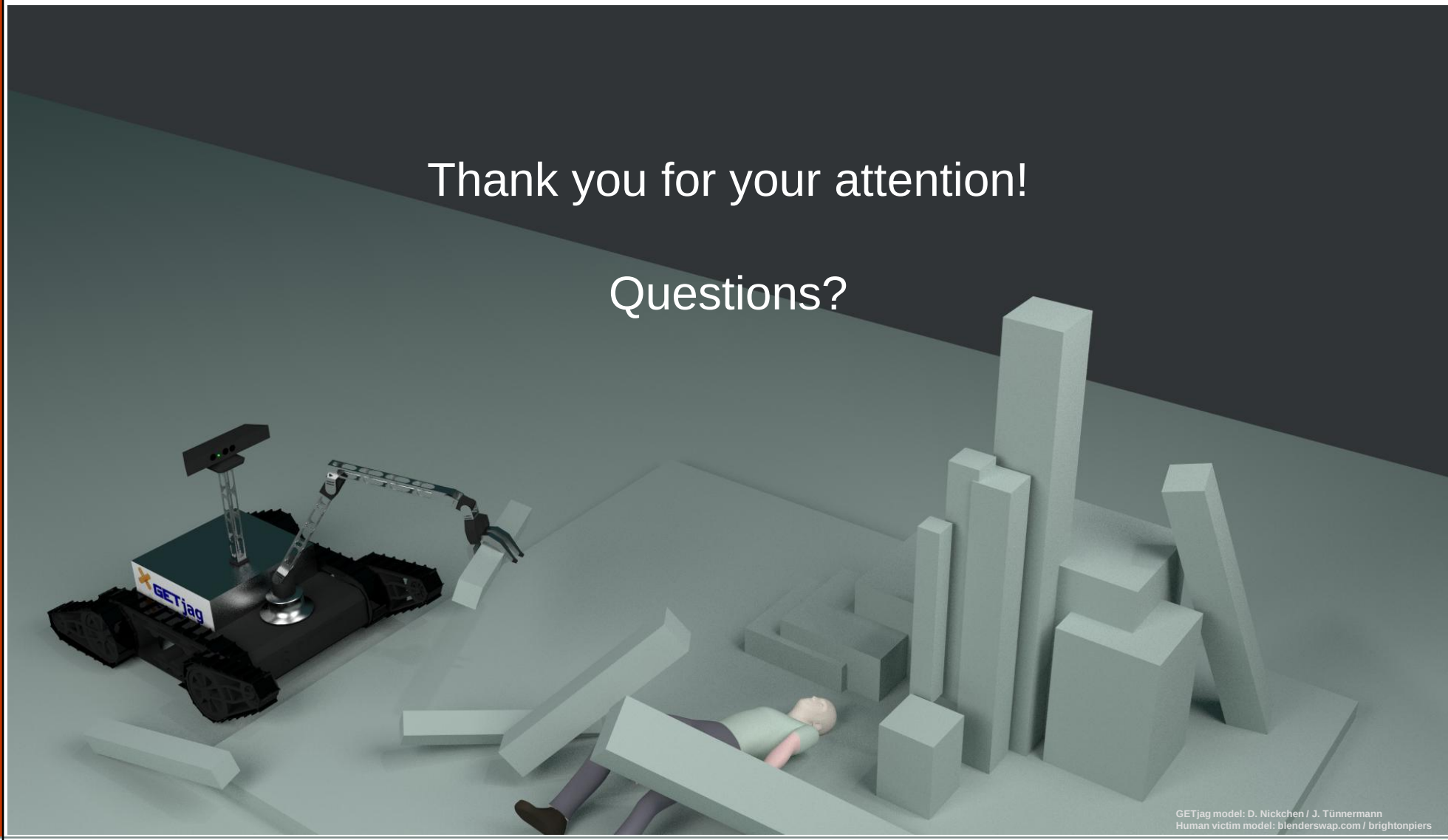
Goal:

Real time implementation in a robot.

- Practical evaluation.
- Integration with action execution.

Thank you for your attention!

Questions?



GETjag model: D. Nickchen / J. Tünnermann
Human victim model: blenderswap.com / brightonpiers