

---

# *Computational Visual Attention Systems*

Simone Frintrop

Rheinische Friedrich-Wilhelms-Universität Bonn,  
Germany

Aug 27th, 2011

Based on: "Computational Visual Attention Systems and their Cognitive Foundations: A Survey",  
Simone Frintrop, Erich Rome, Henrik Christensen, ACM Transactions on Applied Perception, 7 (1), 2010

---

# *Content of Survey*

---

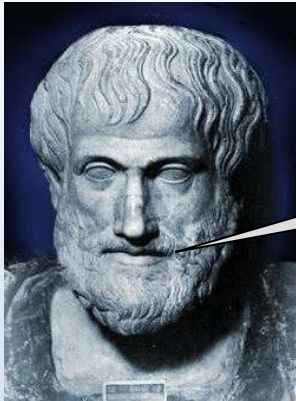
1. What is Visual Attention?
2. Human Visual Attention
3. **Computational Attention Systems (with technical objective)**
4. Important Computational Attention Systems (Review)
5. **Applications in Computer Vision and Robotics**

**Focus in this talk**

# *What is Attention?*

---

Aristotle



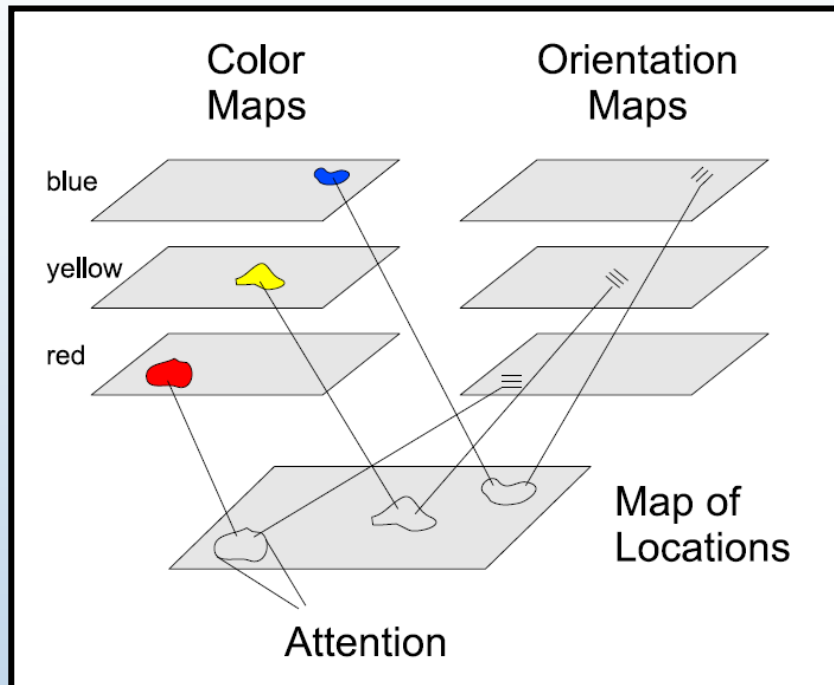
“it is impossible to perceive two objects  
coinstantaneously in the same sensory act”  
(384 – 322 v. Chr.)

**Attention** is the cognitive process  
of selectively concentrating on one  
aspect of the environment

(Wikipedia)

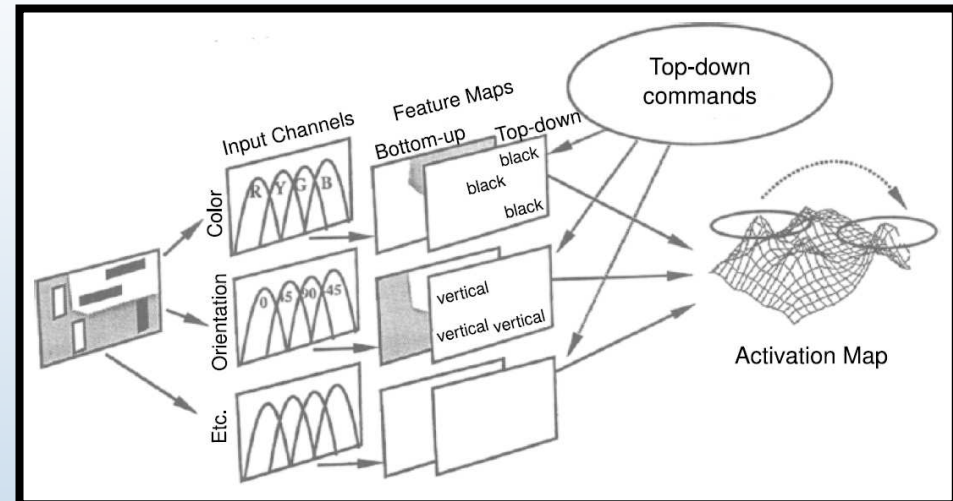
# Psychophysical Attention Models

The Feature Integration Theory:



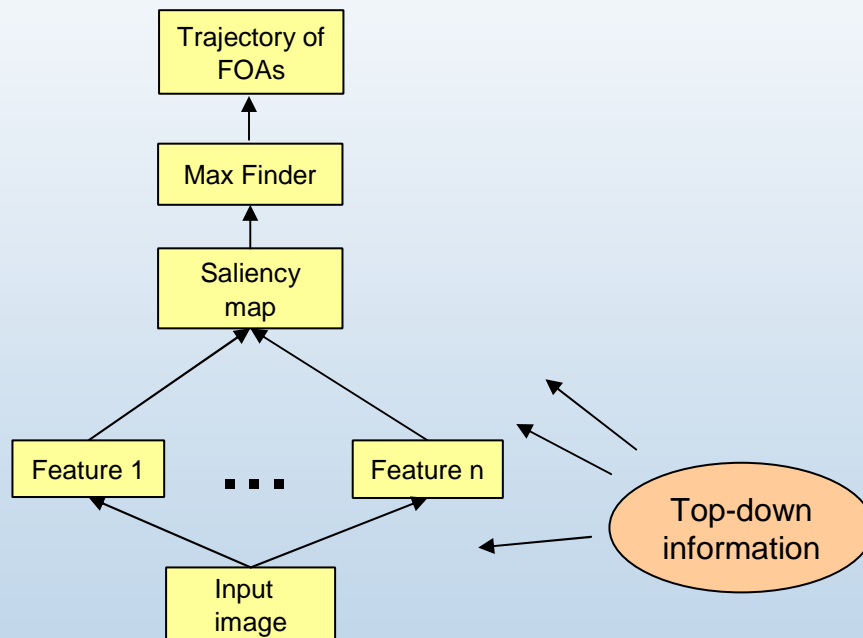
[Treisman and Gelade 1980]  
[Treisman 1993]

The Guided Search model:



Guided Search 1.0: [Wolfe et al. 1989]  
Guided Search 2.0: [Wolfe 1994]  
Guided Search 3.0: [Wolfe/Gancarz 1996]  
Guided Search 4.0: [Wolfe 2001; 2007]

# Computational Attention Systems – General Structure



Underlying structure can be

- connectionist model

(neural networks)

[Olshausen et al 1993]

[Postma 1994]

[Tsotsos et al 1995]

[Cave 1999]

- filter model

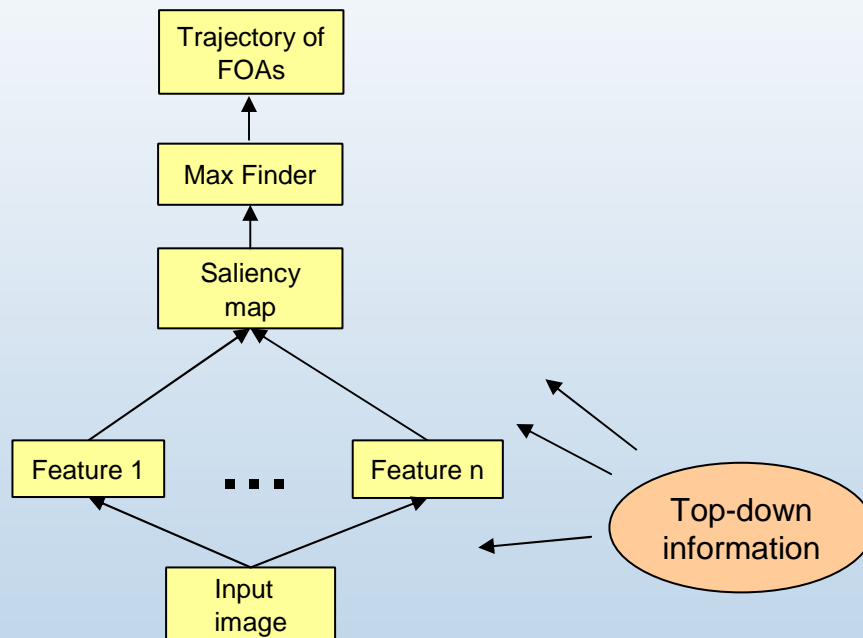
[Milanese 1993]

[Itti, Koch, Niebur 1998]

[Hamker 2005]

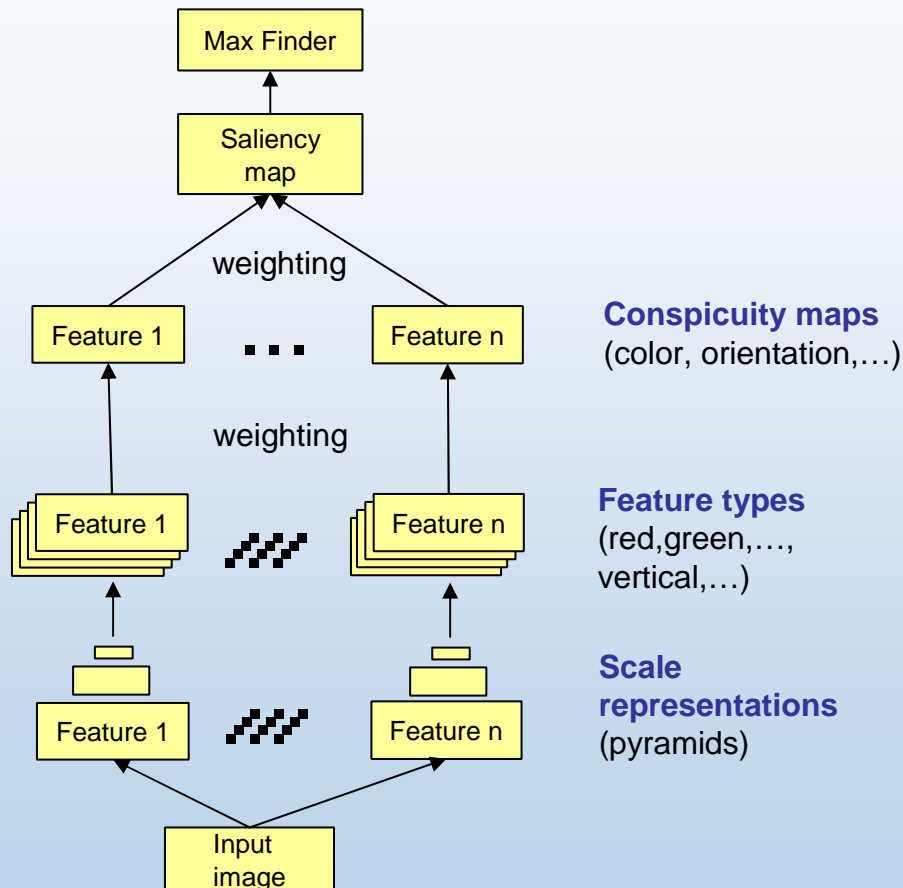
[Frintrop 2006; 2011]

# Computational Attention Systems – General Structure

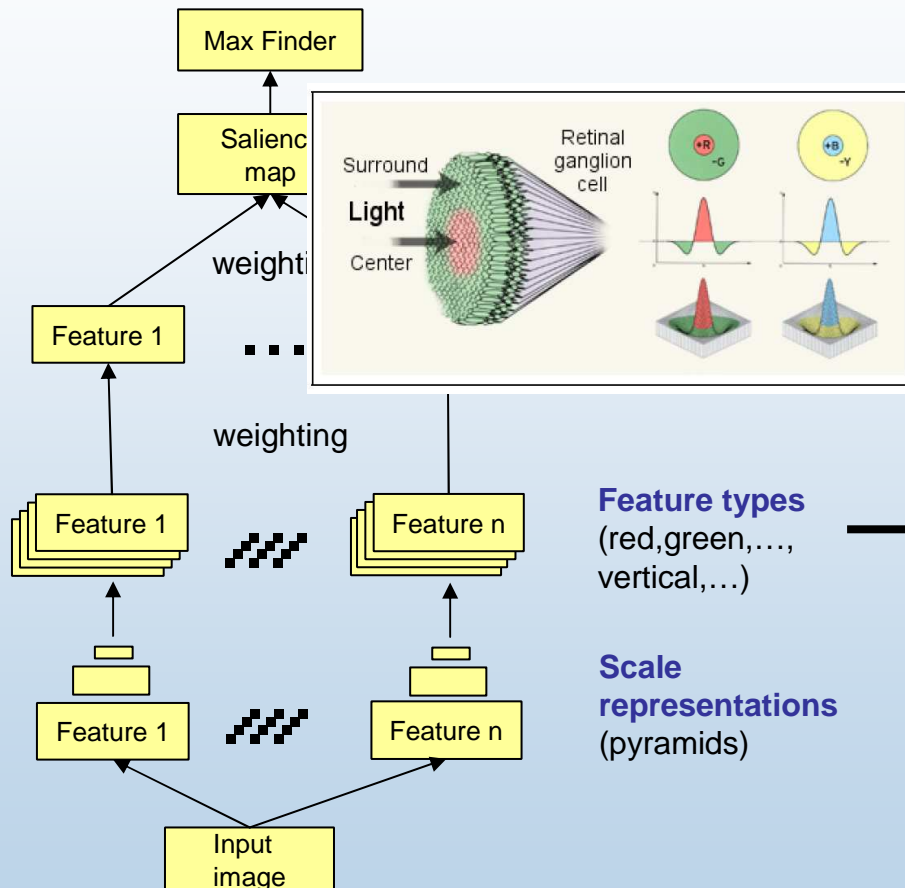


- Most common features:  
**intensity, color, orientation**  
[e.g. Itti,Koch,Niebur 1998]
- Motion [e.g. Tsotsos et al. 2005]
- Depth (with stereo or 3D sensors (e.g. Laser or Kinect))  
[e.g. Frintrop et al. 2005]
- Flicker [e.g. Itti et al. 2003]
- Symmetry [e.g. Backer et al. 2001]
- Corners [e.g. Heidemann et al. 2004]
- Entropy [e.g. Kadir,Brady 2001]

# Computational Attention Systems – General Structure



# Computational Attention Systems – General Structure



## Central concept:

center-surround contrasts

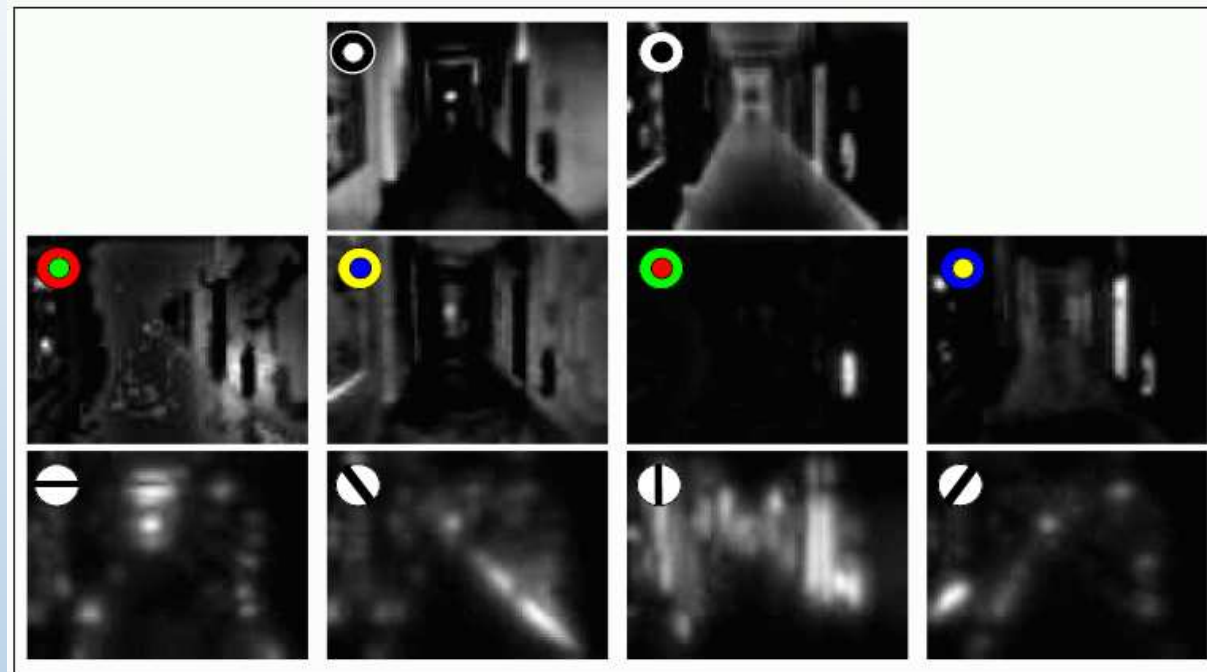
- Classical with Difference-of-Gaussian or Gabor filters  
[Itti, Koch, Niebur 1998]
- Computationally efficient with integral images  
[Frintrop et al. 2007]
- With information-theoretic measures  
[Bruce/Tsotsos 2009; Klein/Frintrop 2011]



# Computational Attention Systems – General Structure

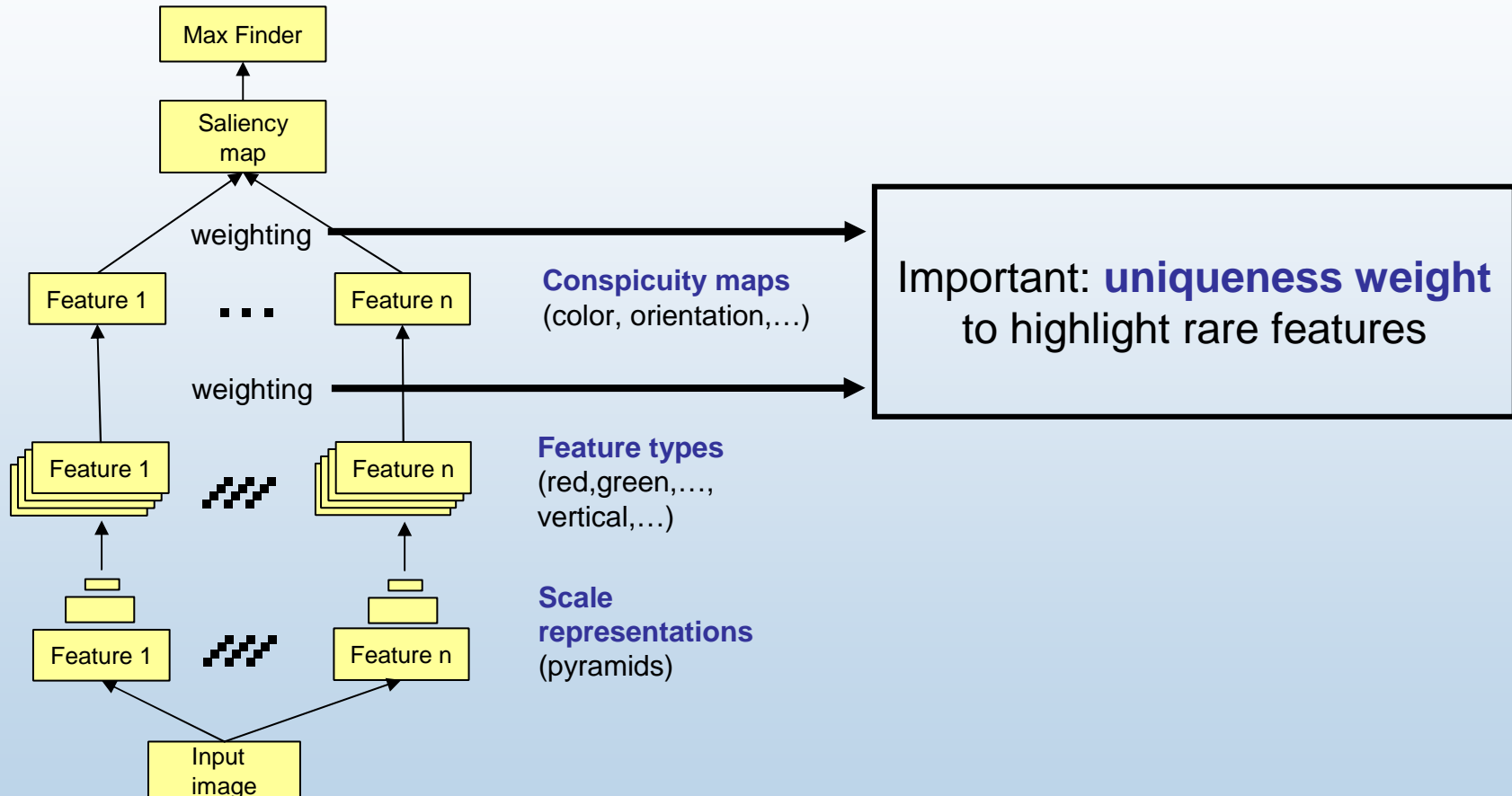


Feature types



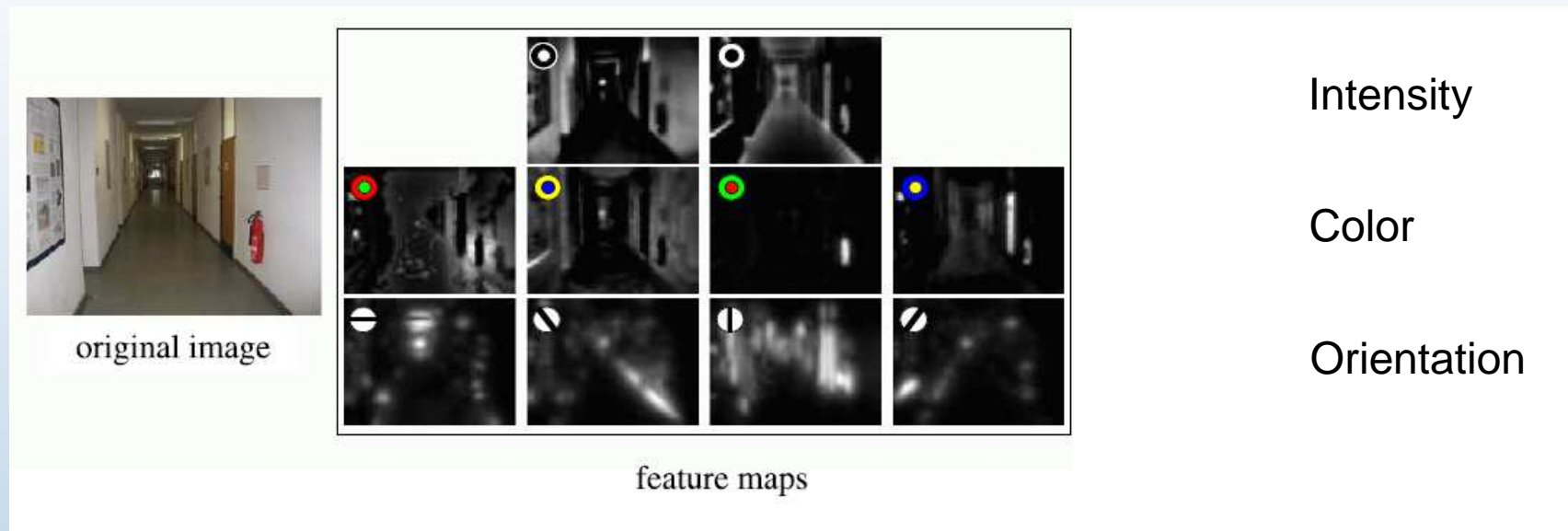
Computed by system VOCUS [Frintrop 2006]

# Computational Attention Systems – General Structure

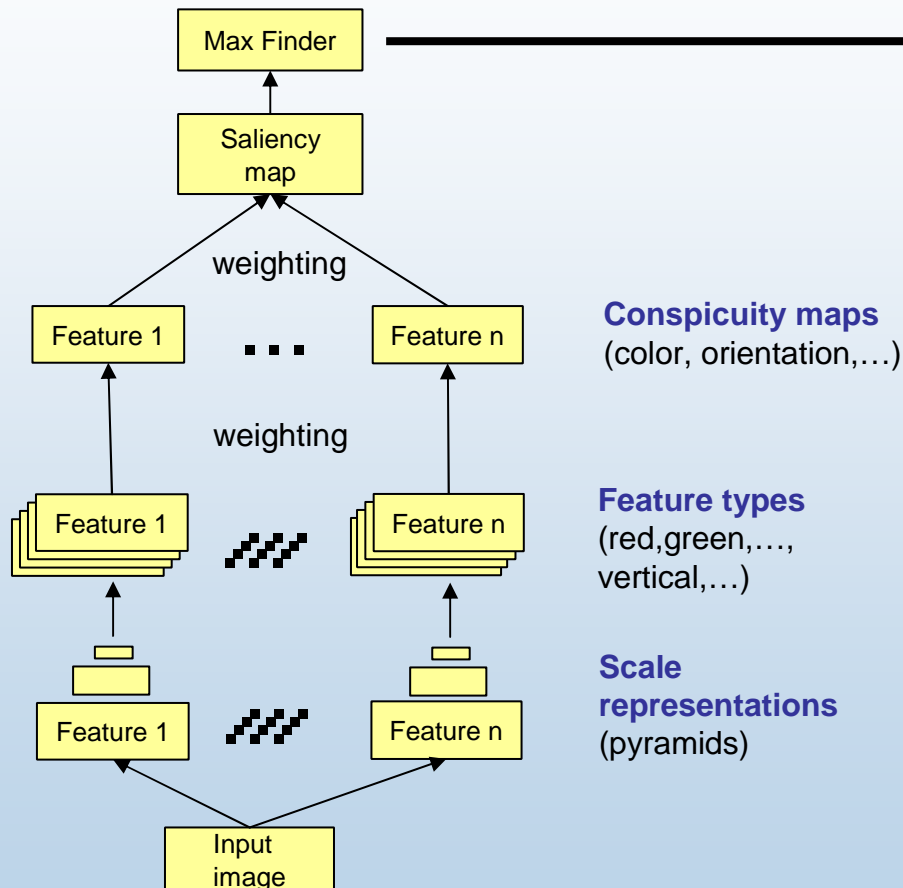


# Computational Attention Systems – General Structure

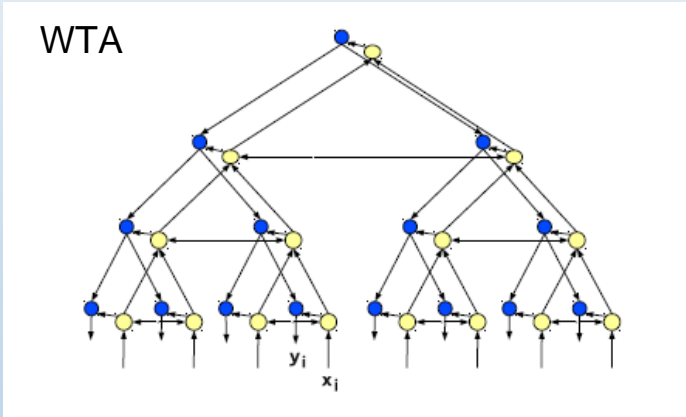
---



# Computational Attention Systems – General Structure



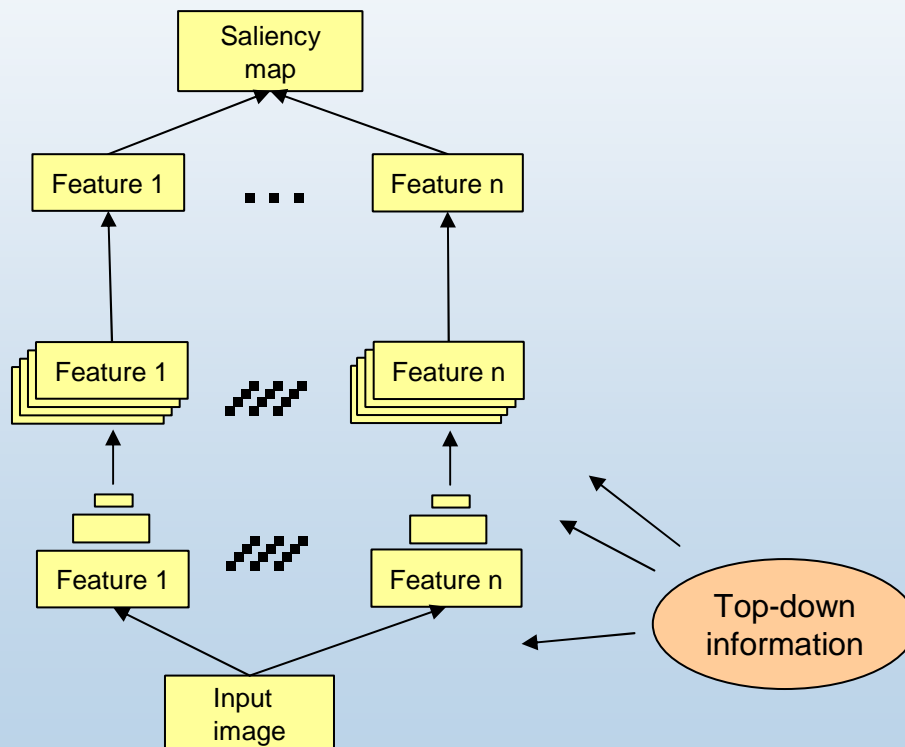
E.g. Winner-Take-All (WTA) Network  
 [Koch, Ullman 1985]  
 or simple maximum finder  
 [Frintrop 2006]



# Computational Attention Systems – General Structure

## Types of top-down cues:

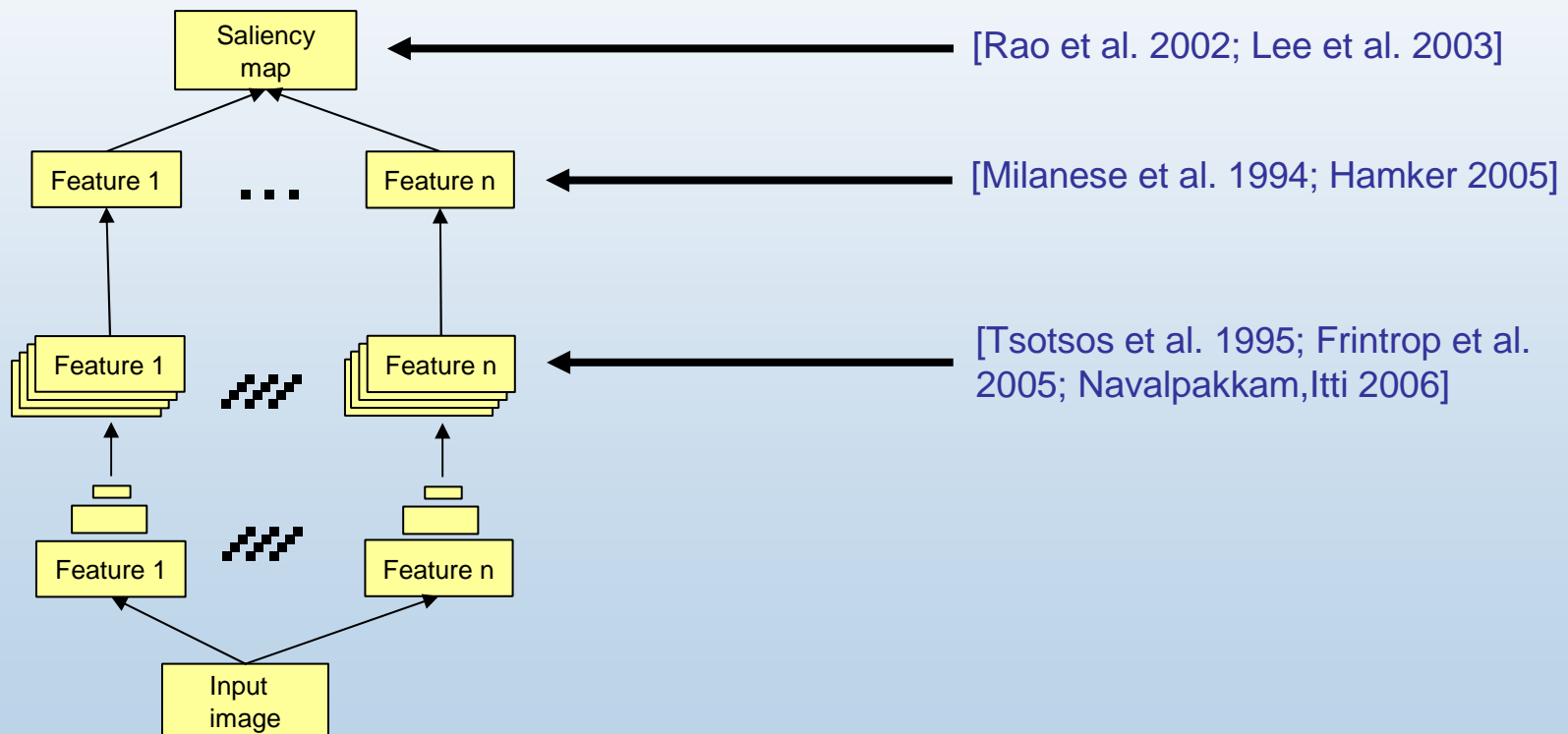
- target information (visual search) [Milanese et al. 1994; Tsotsos et al. 1995; Frintrop 2005; Navalpakkam, Itti 2006; Vincent et al. 2007; Xu et al. 2009]
- scene/context information (e.g. gist) [Oliva et al. 2003; Torralba 2003; Siagian, Itti 2009]
- emotions, desires, motivations (psychological review: [Wells, Matthews 1994], psychological model: [Balkenius 2000], neurobiological model: [Fragopanagos, Taylor 2006]. No computer simulations yet)



# Computational Attention Systems – Top-down

Visual Search:

- **Where** is the target information integrated?



# *Computational Attention Systems – Evaluation*

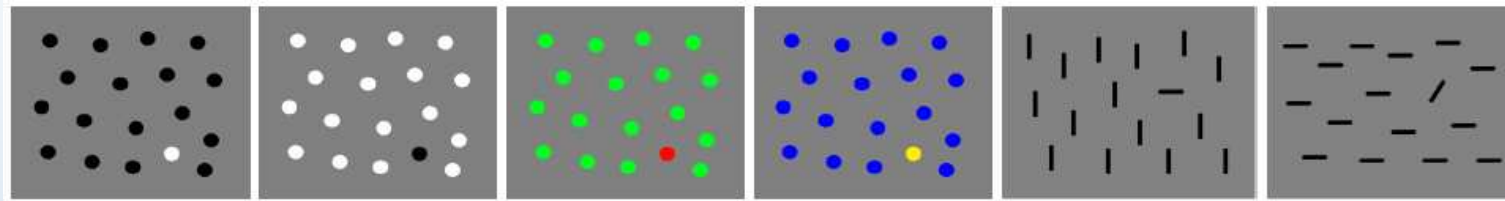
---

Two options to evaluate computational attention systems:

- Compare to human behavior
- Evaluate system in application (compare to other algorithms)

# Computational Attention Systems – Evaluation

- Compare to human behavior
  - check behavior on psychophysical patterns



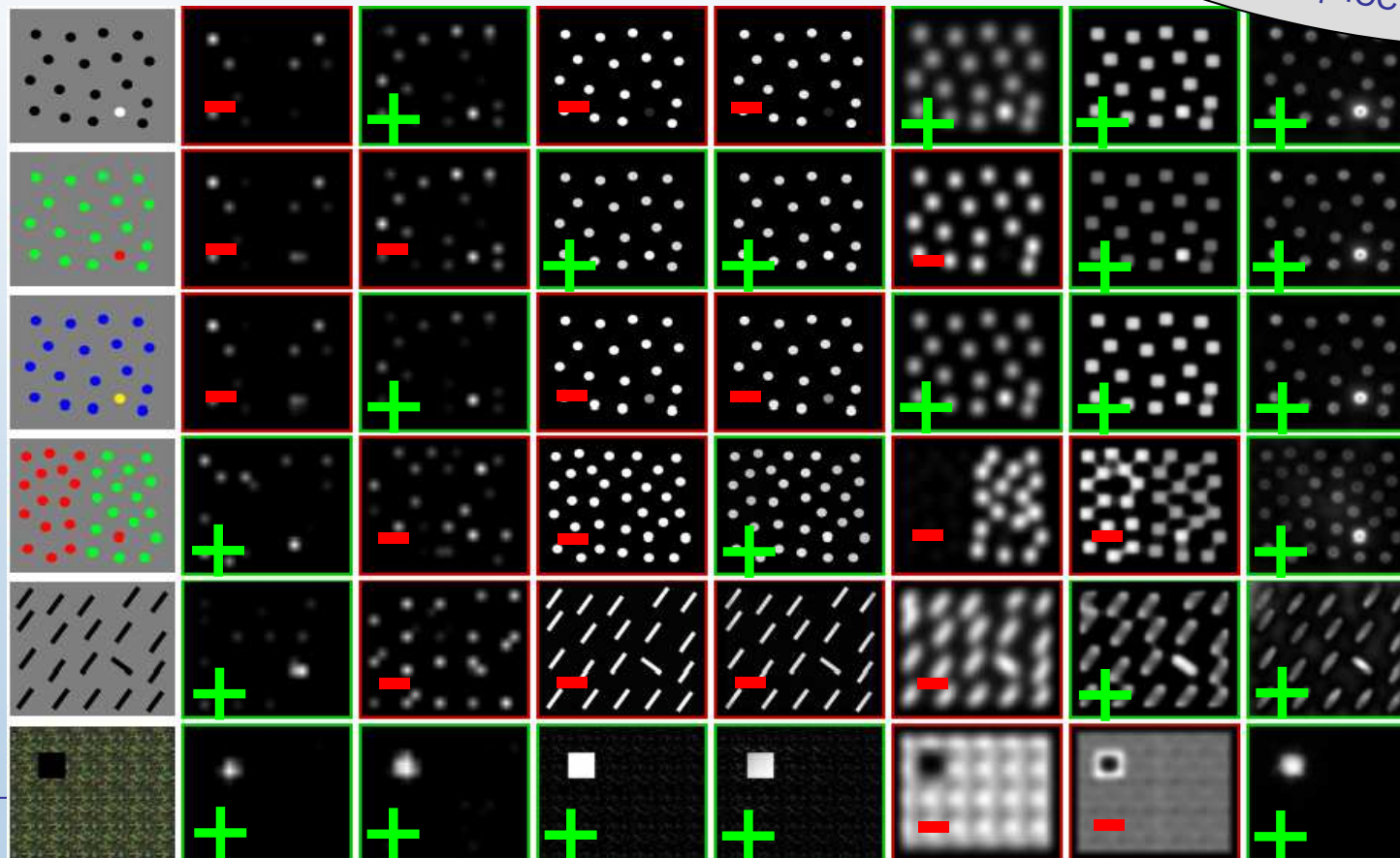
"What is the outlier?"



# Computational Attention Systems – Evaluation

- Compare to human behavior
  - check behavior on psychophysical patterns

Results from:  
Klein/Frintrop ICCV 2011



Simone Frintrop    iNVT [Itti et al]    Sal.Toolbox [Walther]    [Achanta 2009]    [Achanta 2010]    [Hou/Zhang 2008]    AIM [Bruce/Tsotsos 09]    BITS [Klein/Frintrop 11]

# Computational Attention Systems – Evaluation

---

- Compare to human behavior
  - check behavior on psychophysical patterns
  - compare with eye movement data [Parkhurst et al. 2002; Ouerhani et al. 2004; Tatler et al. 2005; Elazary/Itti 2008; Peters/Itti 2008; Bruce/Tsotsos 2009]

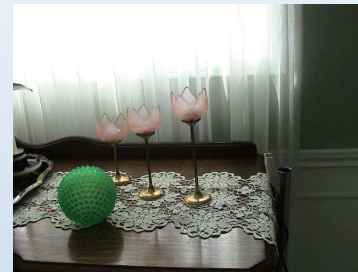
There are several online databases available, e.g.:

- Eye tracking data of Neil Bruce (20 test persons, 120 images):  
<http://www-sop.inria.fr/members/Neil.Bruce>
- Eye tracking data of Gert Kootstra (31 test persons, 101 images):  
<http://www.ai.rug.nl/~gert>
- Eye-tracking data from humans watching videos on the CRCNR website: <http://crcns.org/data-sets/eye>

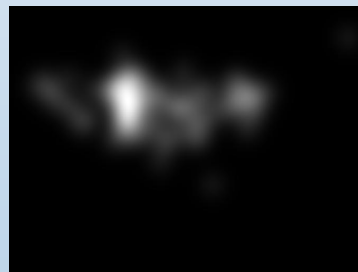
# Computational Attention Systems – Evaluation

- Compare to human behavior
  - check behavior on psychophysical patterns
  - compare with eye movement data [Parkhurst et al. 2002; Ouerhani et al. 2004; Tatler et al. 2005; Elazary/Itti 2008; Peters/Itti 2008; Bruce/Tsotsos 2009]

original images



human fixation density map



Images from <http://www-sop.inria.fr/members/Neil.Bruce>.

# *Content of Survey*

---

1. What is Visual Attention?
2. Human Visual Attention
3. Computational Attention Systems (with technical objective)
4. Important Computational Attention Systems (Review)
5. **Applications in Computer Vision and Robotics**

# Applications in Computer Vision and Robotics

---

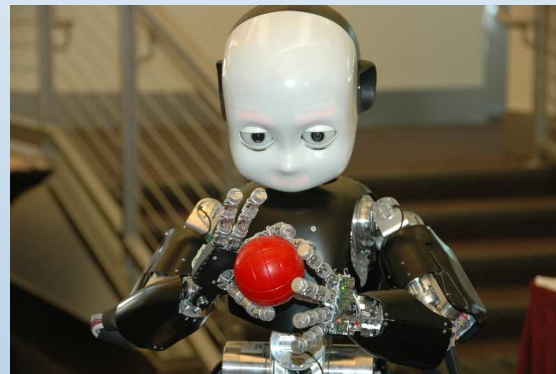
Do we need "Attentive Machines" and why?

Robots have similar requirements as humans:

- They can process only a fraction of the perceptual input in reasonable time
- They have physical constraints (one/few cameras for zooming and pan/tilt, one/few arms,...)
- They act in the same environments as humans and interact with humans



Romeo



iCub



Rhino

# *Applications in Computer Vision and Robotics*

---

Main application areas:

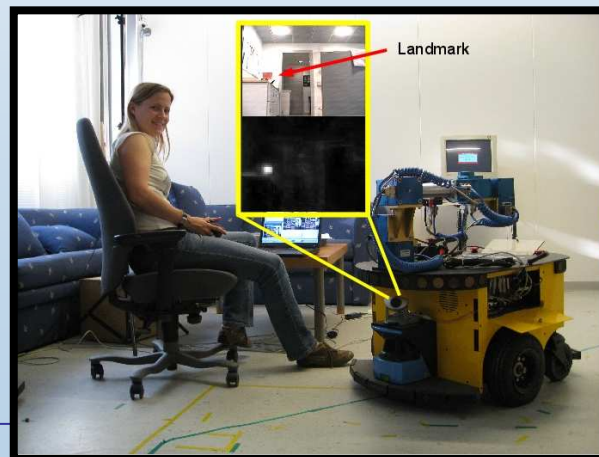
1. Attention as salient interest point detector
2. Attention as front-end for object recognition
3. Attention to guide robot action

# Applications in Computer Vision and Robotics

---

## 1. Attention as salient interest point detector

- image segmentation [Ouerhani 2003]
- image/video compression [Ouerhani 2003; Itti 2004]
- image matching [Walther 2006; Siagian/Itti 2009]
- robot localization [Nickerson et al. 1998; Ouerhani et al. 2005]
- scene classification [Siagian/Itti 2009]
- simultaneous localization and mapping (SLAM) [Frintrop/Jensfelt 2008]



Robot Dumbo detects salient landmarks for visual SLAM [Frintrop/Jensfelt 2008]

# Applications in Computer Vision and Robotics

---

## 2. Attention as front-end for object recognition

- object detection/localization [Rybak et al. 1998; Mitri et al. 2005; Hamker 2005; Gould et al. 2007; Meger et al. 2008; Xu et al. 2010]
- object recognition/classification [Miau et al. 2001; Walther 2006; Walther/Koch 2007; Vogel/de Freitas 2008]



"Curious George" won the robot league of the Semantic Robot Vision Challenge in 2007 and 2008 [Forssén 2008]

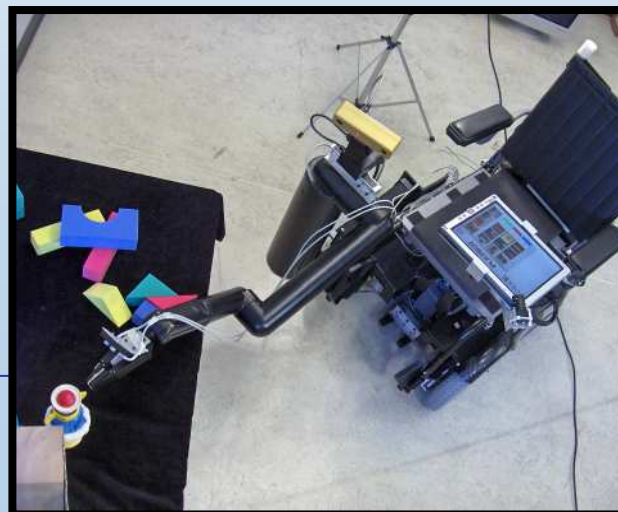


# Applications in Computer Vision and Robotics

---

## 3. Attention to guide robot action

- active vision (control the camera) [Clark/Ferrier 1988; Bollmann 1999; Vijayakumar et al. 2001; Frintrop/Jensfelt 2008]
- visual tracking of objects & people [Walther et al. 2004; Frintrop et al. 2010]
- object manipulation [Bollmann et al. 1999; Rae 2000; Tsotsos et al. 1998; Rotenstein et al. 2007]
- robot navigation [Clark/Ferrier 1992; Scheier/Egner 1997; Baluja/Pommerleau 1997]
- human-robot interaction [Breazeal 1999; Heidemann 2004; Rae 2000; Belardinelli 2008; Nagai 2009; Muhl 2007]



PlayBot: a robotic wheelchair for disabled children (image from <http://www.cse.yorku.ca/~playbot>)

# Conclusion

---

1. What is Visual Attention?
2. Human Visual Attention
3. **Computational Attention Systems**
4. Important Computational Attention Systems
5. **Applications in Computer Vision and Robotics**

For references and details see:

"Computational Visual Attention Systems and their Cognitive Foundations: A Survey",  
Simone Frintrop, Erich Rome, Henrik Christensen, ACM Transactions on Applied Perception, 7 (1), 2010

"Towards Attentive Robots", Simone Frintrop,  
to appear in the PALADYN Journal of Behavioral Robotics (2011)

"Computational Visual Attention", Simone Frintrop, to appear in A. A. Salah and T. Gevers, editors,  
Computer Analysis of Human Behavior, Advances in Pattern Recognition. Springer, 2011

---