Towards Attentive Robots

Simone Frintrop
Cognitive Vision Group
Institute of Computer Science III
University of Bonn, Germany

Human Attention

Why are there attentional mechanisms in the human brain?

- Deal with the computational complexity:
  - Not enough neurons to process everything (~10^8 bits/second)
  - Many problems in vision are NP-hard (Tsotsos: Analyzing vision at the complexity level, Behavioral and brain sciences, 1990)
- We have to act, so we must decide (direct attention on some things and not on others)

Human Attention

Why are there attentional mechanisms in the human brain?

Itti + Koch, 2001:

“It may **seem ironic that brains employ serial processing**, since one usually thinks of them as paradigmatic ‘massively parallel’ computational structures. However, in any physical computational system, **processing resources are limited**, which leads to bottlenecks similar to those faced by the von Neumann architecture on conventional digital machines.

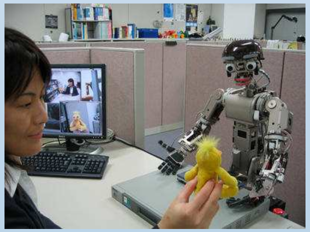
Nowhere is this more evident than in the **primate’s visual system**, where the amount of **information coming down the optic nerve**— estimated to be on the order of 10^8 bits per second **— far exceeds what the brain is capable of fully processing** and assimilating into conscious experience.

The **strategy nature has devised** for dealing with this bottleneck is to **select certain portions of the input to be processed preferentially**, shifting the processing focus from one location to another in a serial fashion

L. Itti, C. Koch, **Computational modelling of visual attention**, *Nature Reviews Neuroscience 2, 194-203, 2001*

Do we need Attentive Machines?

In robotics, attention is especially useful since robots have similar requirements as humans:



- They can process only a fraction of the perceptual input in reasonable time (real-time processing wanted)
  => attention prioritizes
- They have physical constraints (one/few cameras for zooming and pan/tilt, one/few arms,…) and have to decide what to do next
  => attention supports decision making
- Many robots act in the same environments as humans and shall interact with them
  => joint attention useful

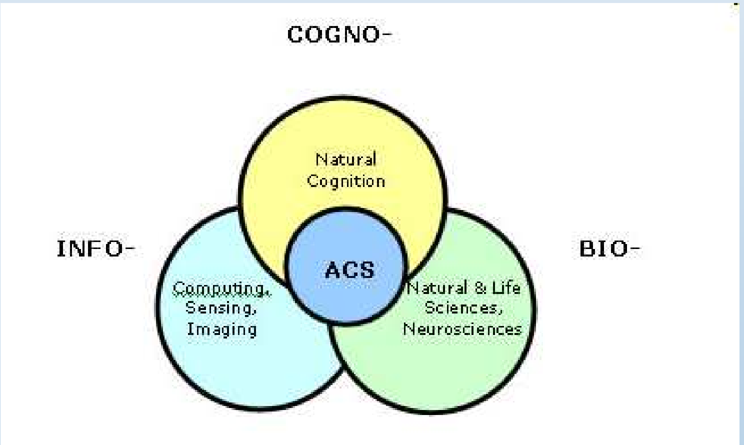
Cognitive Systems in the EU

- Since 2001: Cognitive Systems intensely funded by the EC

"Robots need to be more robust, context-aware and easy-to-use. Endowing them with advanced learning, cognitive and reasoning capabilities will help them adapt to changing situations, and to carry out tasks intelligently with people"

[Challenge 2: Cognitive Systems, Interaction, Robotics]

- More than 100 projects on Cognitive Systems funded; many have an attention module (e.g. MACS, Paco-plus, CogX)



http://cordis.europa.eu/fp7/ict/cognition/home_en.html

Robot as Social Partners

- Humans treat robots like people:

T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4):143–166, 2003

C. Nass and Y. Moon. Machines and mindle
Social Issues, 56(1):81–103, 2000

- The more a robot interact
and intelligent it is perceiv
are

"As we look ahead, it seems clear that social robots will play an ever larger role in our world, working for and in cooperation with humans. Social robots will assist in health care, rehabilitation, and therapy. Social robots will work in close proximity to humans, serving as tour guides, office assistants, and household staff.[…]
Central to the success of social robots will be **close and effective interaction between humans and robots.** Thus, although it is important to continue enhancing autonomous capabilities, we must not neglect **improving the human-robot relationship**."
[Fong et al 2003]

Shared attention

Sociological study of Muhl et al (university of Bielefeld):

- Humans interact with a robot face on a screen.
- The robot's attention is guided by an attention system and usually the robot looks at the toys the human is interacting with.
- If the robot was artificially diverted and directed its gaze away from the object, humans tried to re-obtain the robots attention by waving hands, making noise, or approaching to the robot.
- This shows that people established a communicative space with the robot and accepted it as a social partner.

http://www.claudia-muhl.de/

Shared attention

Attentive Robots

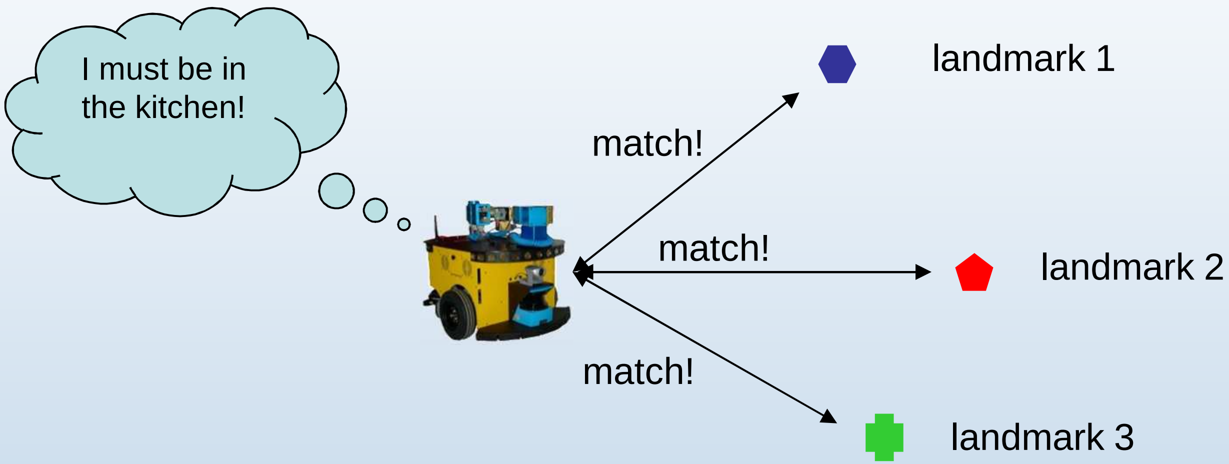
Main application areas in robotics:

- Attention as salient interest point detector
- Attention as front-end for object recognition
- Attention to guide robot action
- Attention to establish joint focus of attention of human and machine (human-computer interaction, human-robot interaction)

More:
Simone Frintrop: **Towards Attentive Robots**, *PALADYN Journal of Behavioral Robotics, Springer, Vol. 2, Issue 2, 2011*

Attentive Robot Localization

Visual localization: use landmarks to determine robot position



ARK project

- An early project (~1998): the ARK project (Autonomous Robot for a Known environment)
- Task: navigate in industrial environments
  (navigation in walled areas: laser
  navigation in open areas: visual landmarks detected by an attention system)

S. B. Nickerson, P. Jasiobedzki, D.Wilkes, M. Jenkin, E. Milios, J. K. Tsotsos, A. Jepson, and O. N. Bains. The ARK project: Autonomous mobile robots for known industrial environments. *Robotics and Autonomous Systems*, 25(1-2):83–104, 1998

Attentive Robot Localization

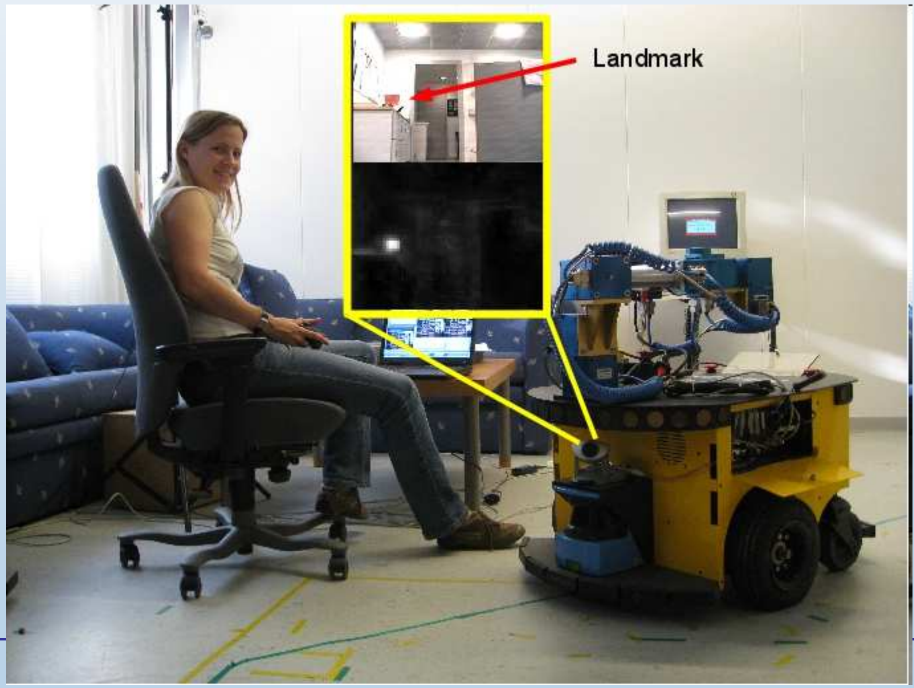
- Beobot 2.0: a robot at the University of Southern California (USC) (group of Laurent Itti) that can operate in large scale unconstrained environments
- The robot uses an attention system to localize the robot
- More on: http://ilab.usc.edu/beobot2

Attentive SLAM

Simultaneous robot localization and mapping (SLAM):

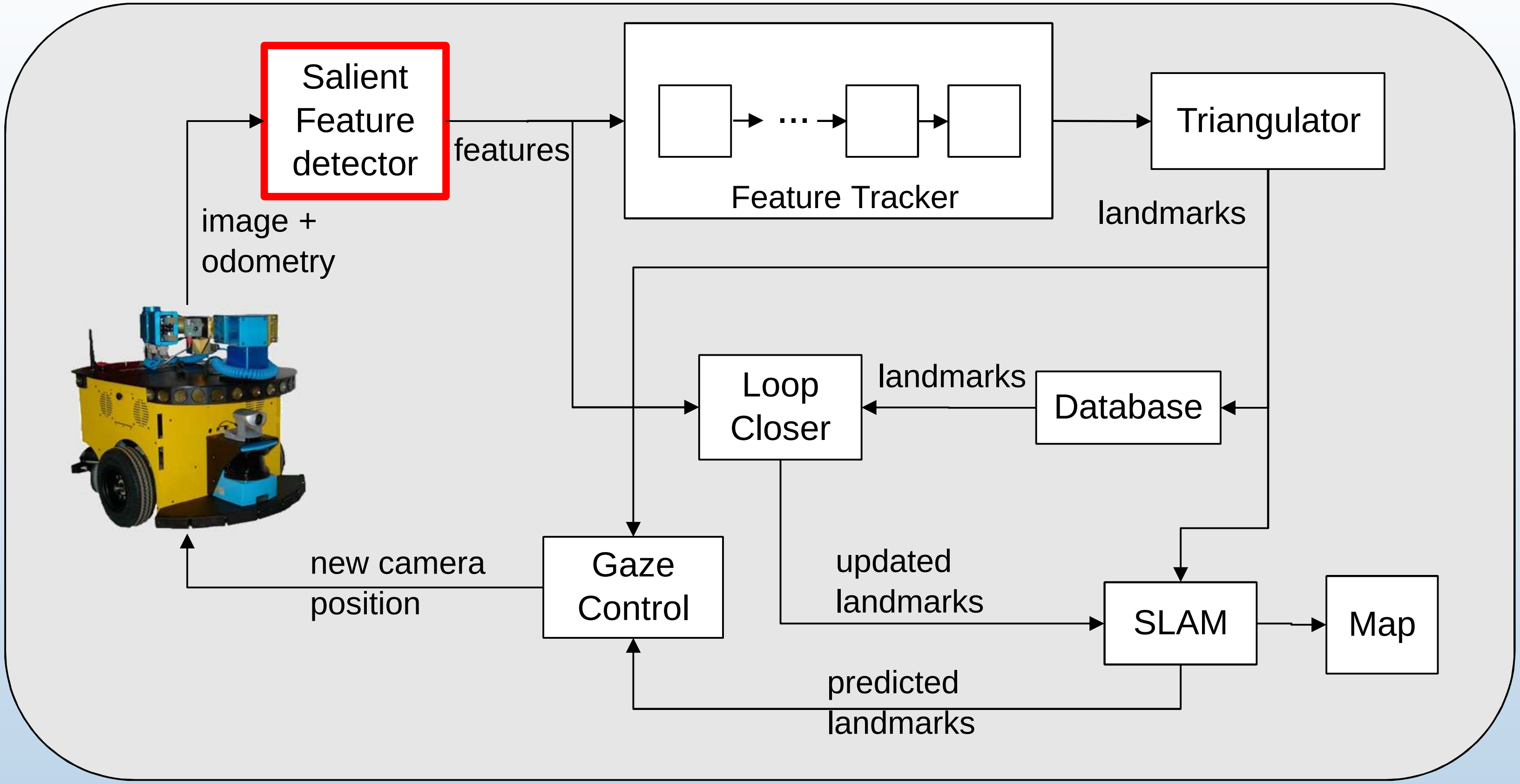
EU project NEUROBOTICS, work at KTH 2005/2006:

Robot Dumbo detects salient landmarks and uses them to build and maintain an internal representation of the environment (mapping) and localizes itself within the map



[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

The Attentional SLAM System



[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

Data acquisition

[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

Landmarks

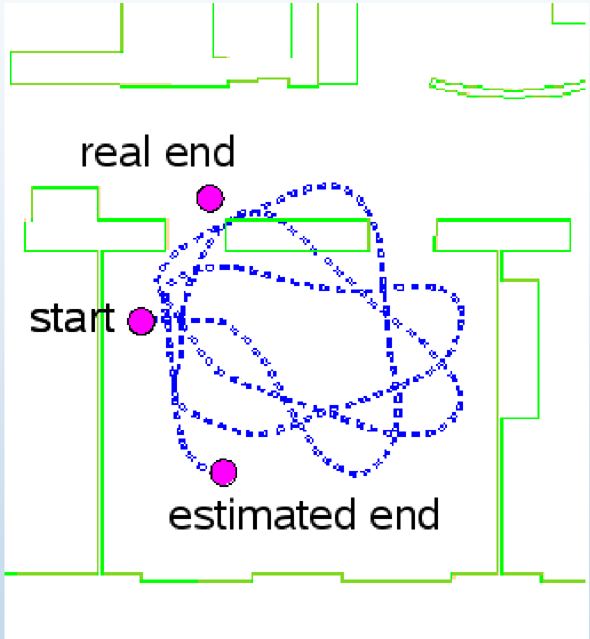
Discriminative landmarks are easy to redetect:

current view:

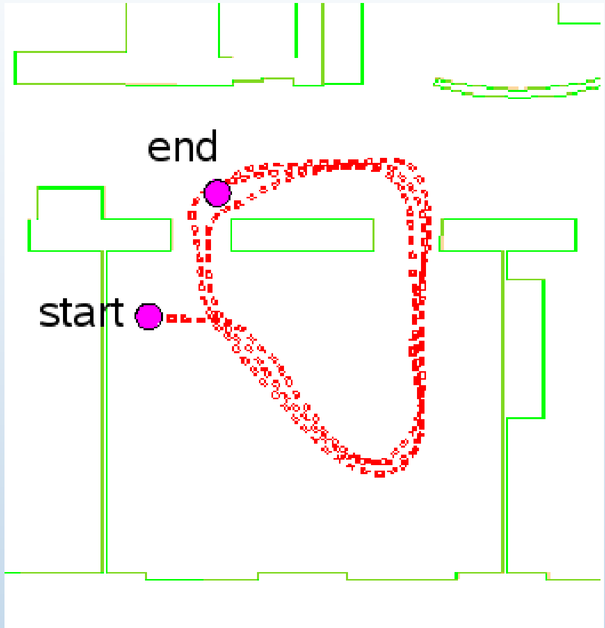
database:

[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

Experiment



estimated path,
only odometry

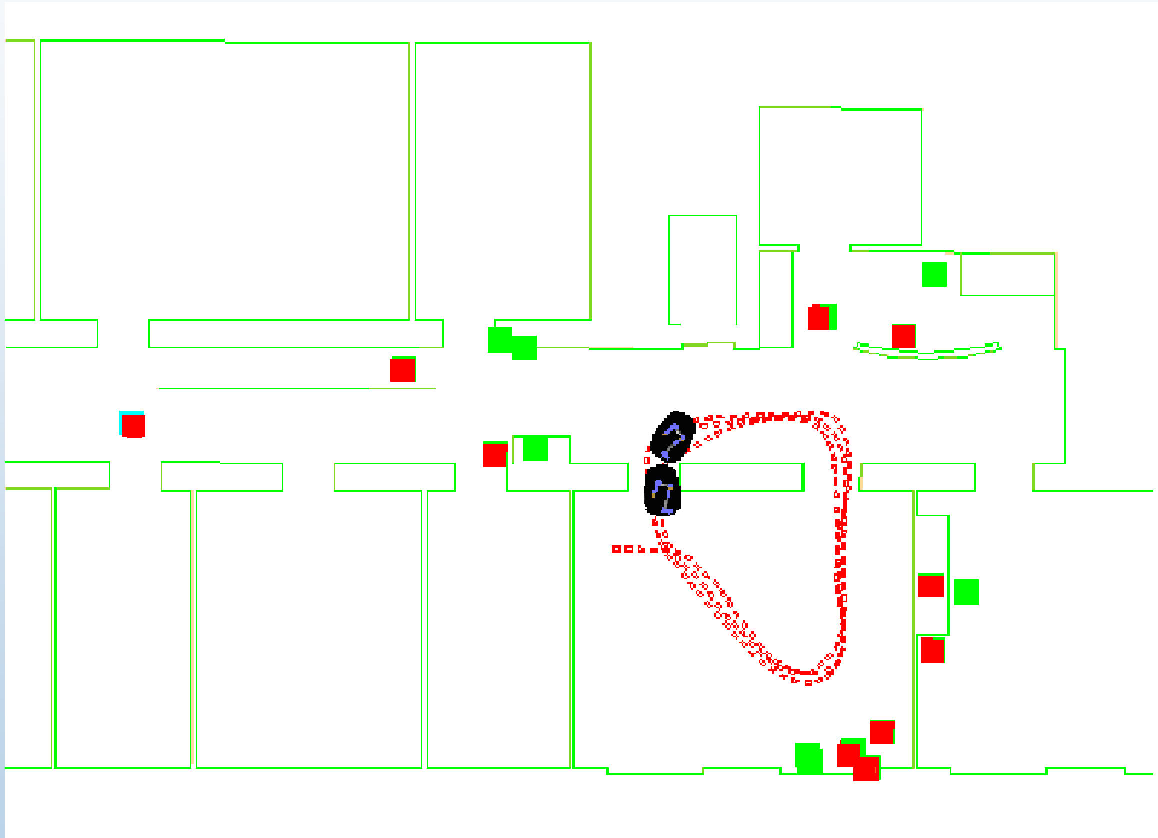


estimated path,
attentional visual SLAM

[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

Experiment

Obtain a sparse landmark representation:



3 runs,
17 landmarks,
21 matches

More than half of the landmarks were redetected

[Frintrop,Jensfelt: IEEE Trans. on Robotics 2008]

Attention for Object Detection

- An early project (~1999) at University of Hamburg, Germany
- A Pioneer1 robot uses the active vision system NAVIS (with an attention module) to play at dominoes

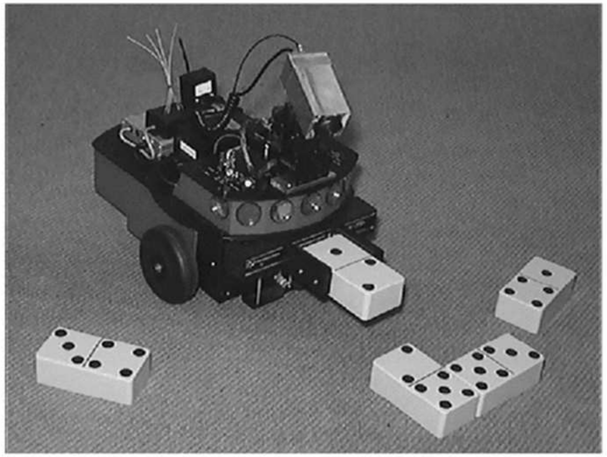
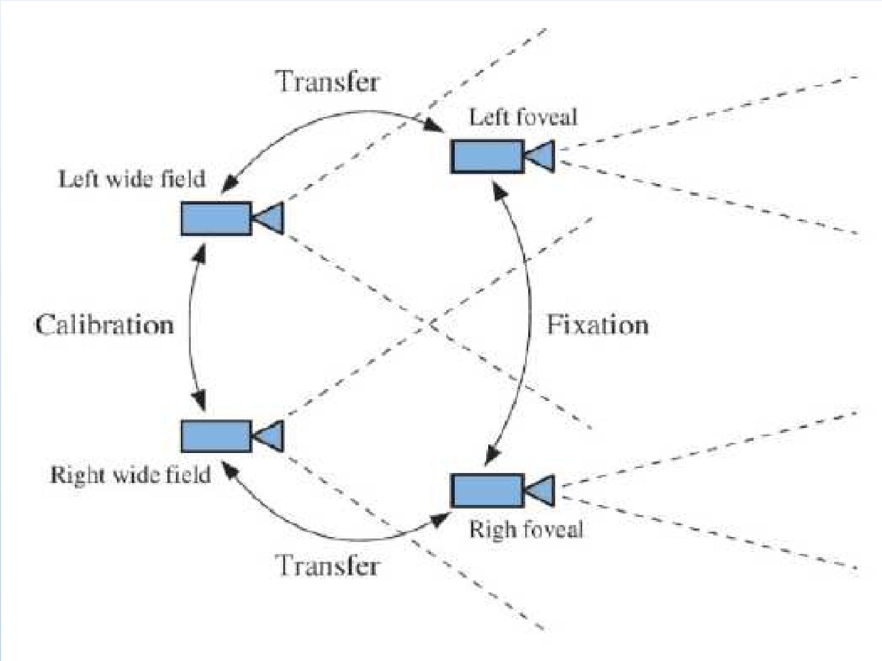


Fig. 1. A Pioneer1 robot serves as the mobile experimental platform. The basic robot consists of a differential wheel drive, seven sonars, an onboard controller, and a one DoF gripper. Additionally, the robot is equipped with a pan-tilt unit and a CCD camera

M. Bollmann, R. Hoischen, M. Jesikiewicz, C. Justkowski, and B. Mertsching. Playing domino: A case study for an active vision system. In H. Christensen, editor, *Computer Vision Systems*, pages 392–411. Springer, 1999

Attention for Object Detection

- Often, two cameras are used to simulate the different resolutions of the human eye:
  – A wide-angle camera for peripheral vision
  – A narrow-angle, high-resolution camera for foveal vision



B. Rasolzadeh, M. Björkman, K. Huebner, and D. Kragic. An active vision system for detecting, fixating and manipulating objects in real world. *International Journal of Robotics Research*, 29(2-3), 2010

Attention for Object Detection

- The STAIR platform (Stanford Artificial Intelligence Robot)
- uses an attention system to determine regions of interest in the peripheral view which are then attended by the foveal camera and used for classification

Figure 1: STAIR platform (left) includes a low-resolution peripheral camera and high-resolution PTZ camera (top-right), and alternative setup with HDV camera replacing the PTZ (bottom-right). In our experiments we mount the cameras on a standard tripod instead of using the robotic platform.

S. Gould, J. Arfvidsson, A. Kaehler, B. Sapp, M. Messner, G. Bradski, P. Baumstarck, S. Chung, and A. Y. Ng. Peripheral-foveal vision for real-time object recognition and tracking in video. In *Proc. of the 20th Int. Joint Conference on Artifical intelligence (IJCAI)*, 2007

Attention for Object Detection

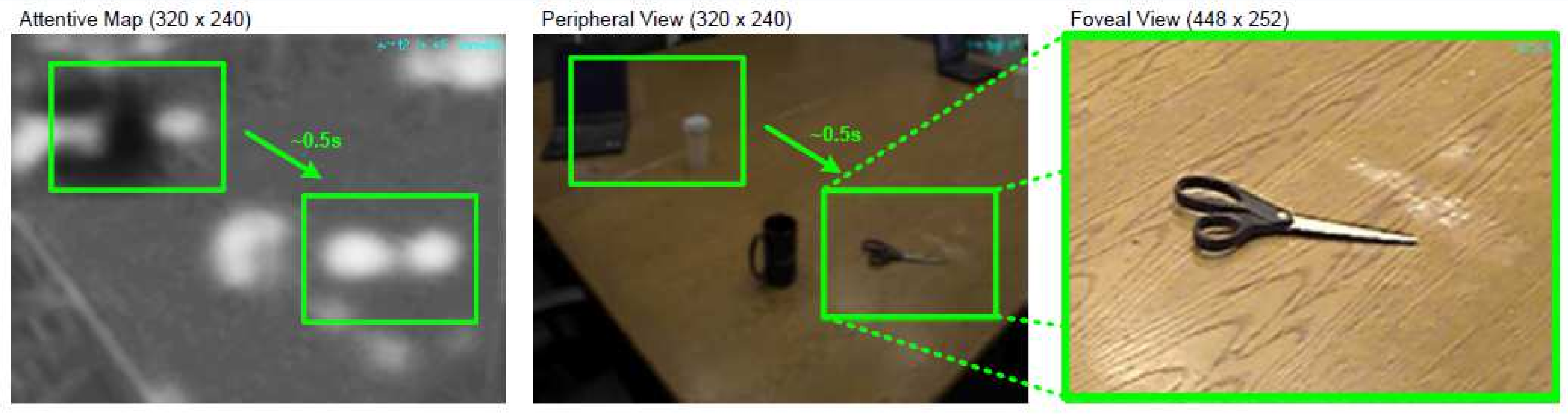
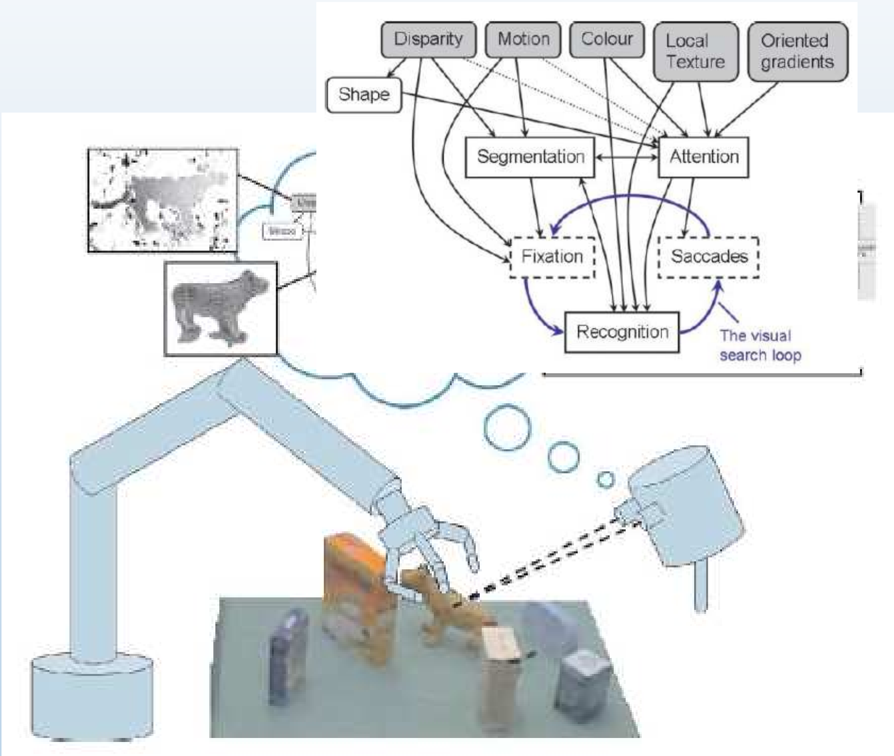


Figure 2: Illustration of the peripheral (middle) and foveal (right) views of a scene in our system with attentive map showing regions of high interest (left). In our system it takes approximately 0.5 seconds for the PTZ camera to move to a new location and acquire an image.

S. Gould, J. Arfvidsson, A. Kaehler, B. Sapp, M. Messner, G. Bradski, P. Baumstarck, S. Chung, and A. Y. Ng. Peripheral-foveal vision for real-time object recognition and tracking in video. In *Proc. of the 20th Int. Joint Conference on Artifical intelligence (IJCAI)*, 2007

Attention for Object Detection

KTH, Stockholm: a KUKA arm is controlled by bottom-up and top-down attention for detecting, recognizing, and grasping objects



B. Rasolzadeh, M. Björkman, K. Hübner, and D. Kragic. An active vision system for detecting, fixating and manipulating objects in real world. *International Journal of Robotics Research*, 29(2-3), 2010

Object detection

Curious George (EU project CogX - Cognitive Systems that Self-Understand and Self-Extend, 2008 - 2012)

- A peripheral vision system identifies object candidates with help of a visual attention system
- A foveal vision system takes close-up views and investigates the candidates with a recognition module

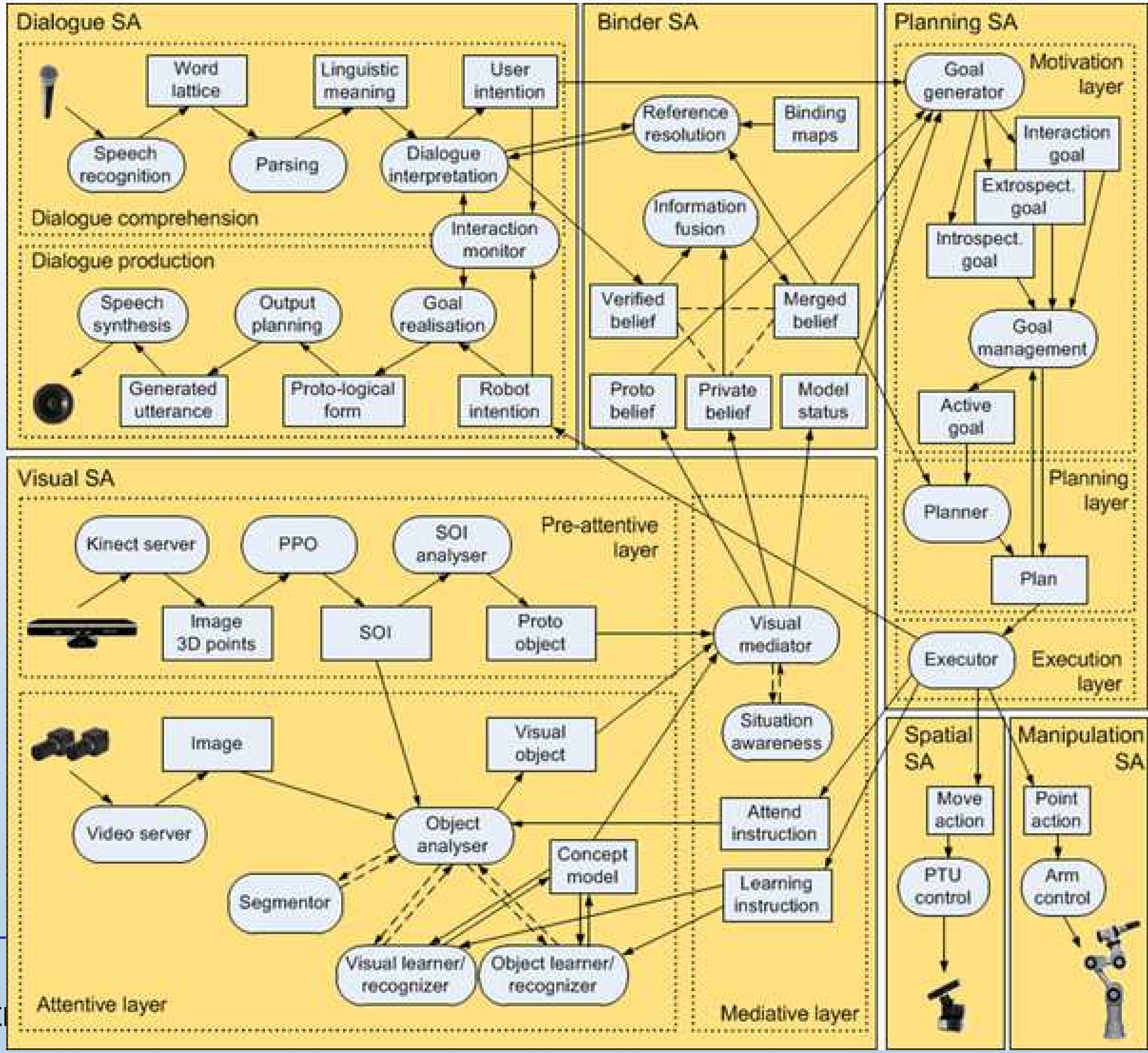
"Curious George" won the robot league of the Semantic Robot Vision Challenge in 2007 and 2008

[Forssén 2008]

universitätbonn

IVS intelligent vision systems

Curious George



http://cogx.eu/results/george/

Curious George

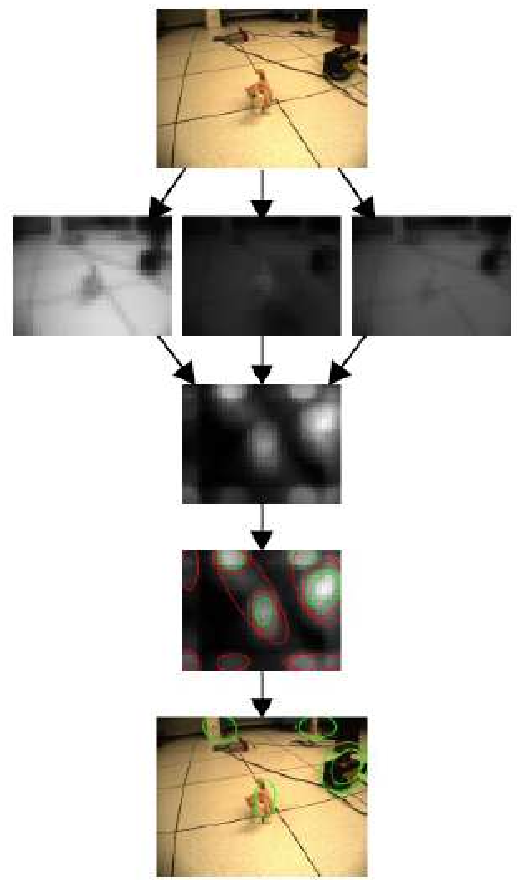
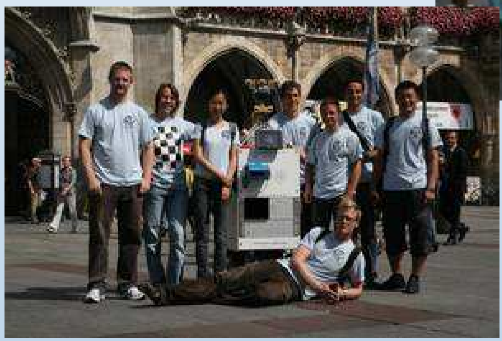
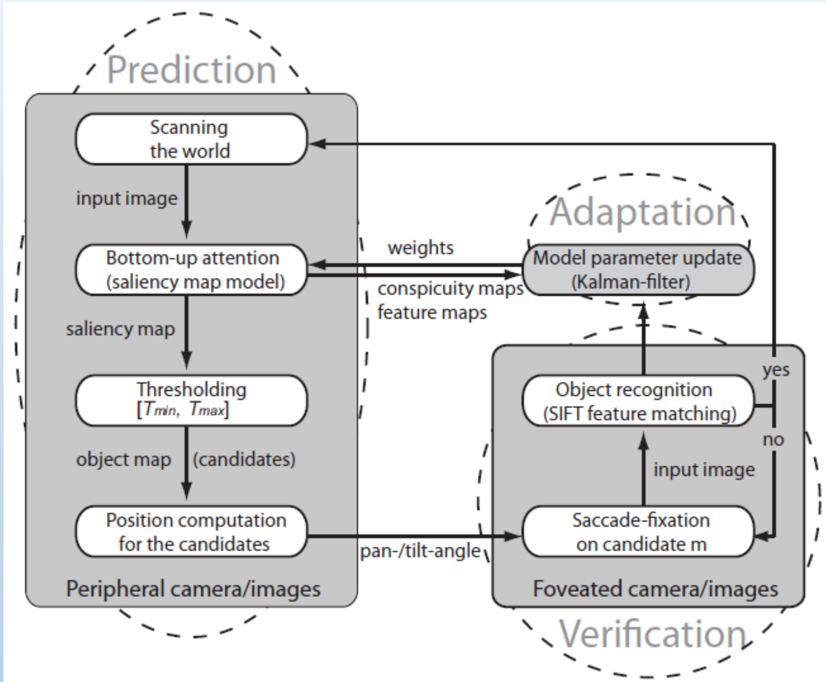


Fig. 13 Saliency computation on curious George. Top to bottom: input image, colour opponency channels (intensity, red-green, yellow-blue), spectral saliency map, detected regions, regions superimposed on input image (image from [43]).

Object Detection

- The Autonomous City Explorer (ACE) at TUM (Munich, Germany):
- Task: autonomously navigate in an unstructured urban environment and find its way through interaction with humans (no map or GPS)
- Attention system supports object detection





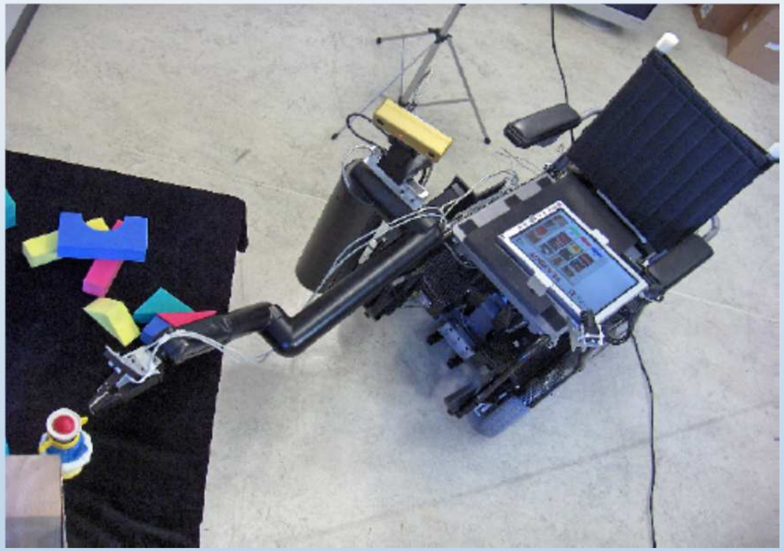
- http://www.ace-robot.de

T. Xu, T. Zhang, K. Kühnlenz, and M. Buss. Attentional object detection of an active multifocal vision system. *International Journal of Humanoid Robotics*, 7(2), 2010

Object detection

PlayBot: project at the university of Toronto, group of John Tsotsos.

- Goal of the project: a robotic wheelchair for disabled children that detects toys with help of a visual attention system



[http://www.cse.yorku.ca/~playbot]

A. Rotenstein et al., **Towards the dream of intelligent, visually-guided wheelchairs**, *Proc. 2nd Int'l Conf. On Technology and Aging, Toronto, Canada 2007*

Applications

- Attention to guide robot action
  - active vision (control the camera)
  - visual tracking of objects & people
  - object manipulation
  - robot navigation
  - human-robot interaction

Robot iCup

[Image: http://www.robotcub.org/]

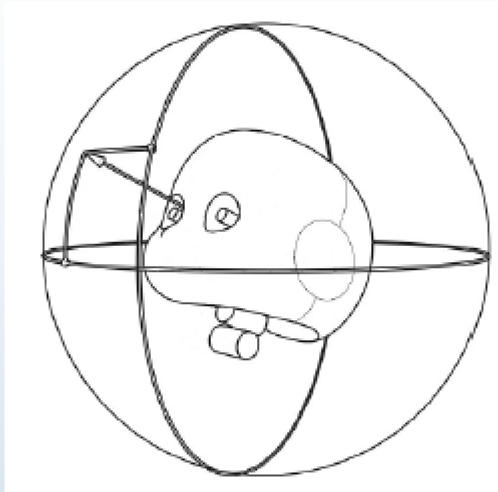
iCub

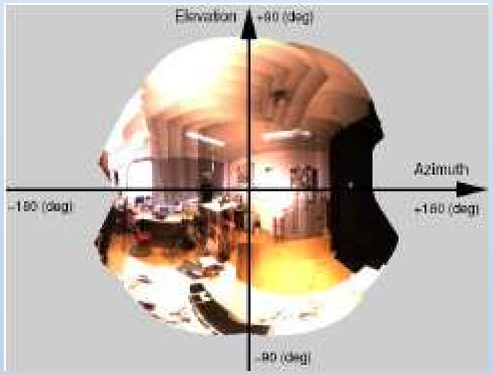
- iCub was developed in the EU project RobotCub
- Main purpose: study cognition
- Base decisions to move eyes and neck on visual and acoustic saliency maps

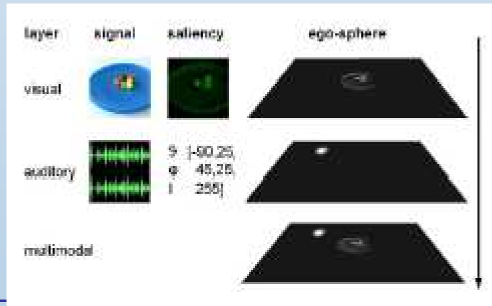
- http://www.icub.org/

iCub

- Saliencies from different modes (visual + audio) are integrated into a topologically organized ego-spere
- Ego-spere: a continuous spherical surface with infinite radius that is centered at the robot‘s head
- Can be used to control the attention of the robot in order to explore the environment
  - Attent to most salient location on spere
  - Inhibit this region in the ego-sphere



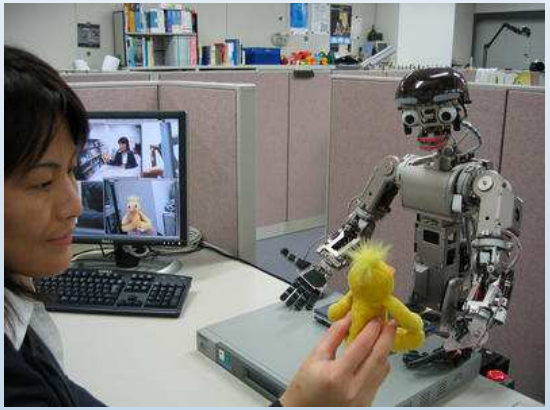




J. Ruesch, M. Lopes, A. Bernardino, J. Hörnstein, J. Santos-Victor, and R. Pfeifer. Multimodal saliency-based bottom-up attention: A framework for the humanoid robot icub. In *Proc. of the Int’l Conf. on Robotics and Automation (ICRA 2008)*, 2008

Shared attention

- Shared attention (also joint attention) mean that an "agent" (human or artificial agent) shares the attention with another agent. I.e. they look at the same object, talk about the same object
- Shared attention is an important aspect of human-robot interaction



Human-robot interaction at the university of Bielefeld

[http://cnr.ams.eng.osaka-u.ac.jp/~yukie/research-en.html#jointattention]

Shared attention

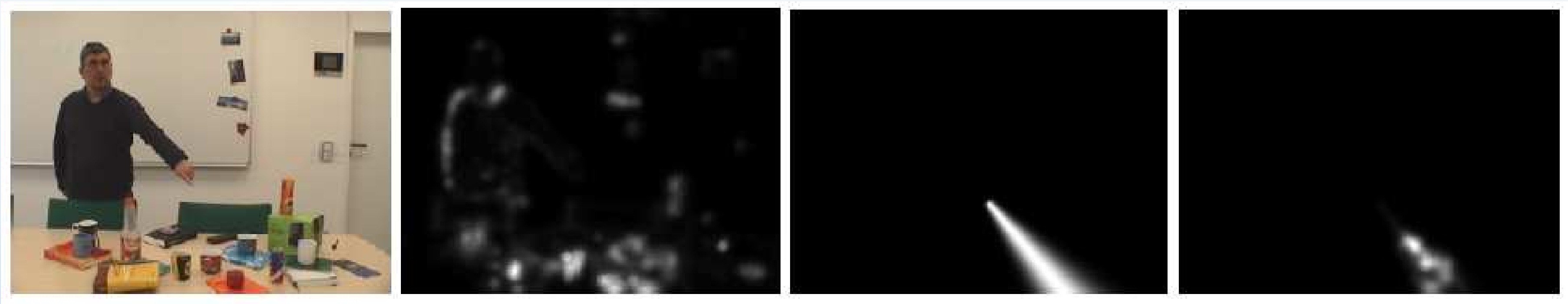
Shared attention can be obtained e.g. by

- bottom-up attention (especially motion cues)
- top-down attention (including pre-knowledge of target of interest)
- following pointing gestures of humans
- including audio information (speech recognition)

and by combinations of the above.

Shared attention

- Example: combining computational visual attention with pointing gestures by Boris Schauerte



[https://cvhci.anthropomatik.kit.edu/~bschauer/]

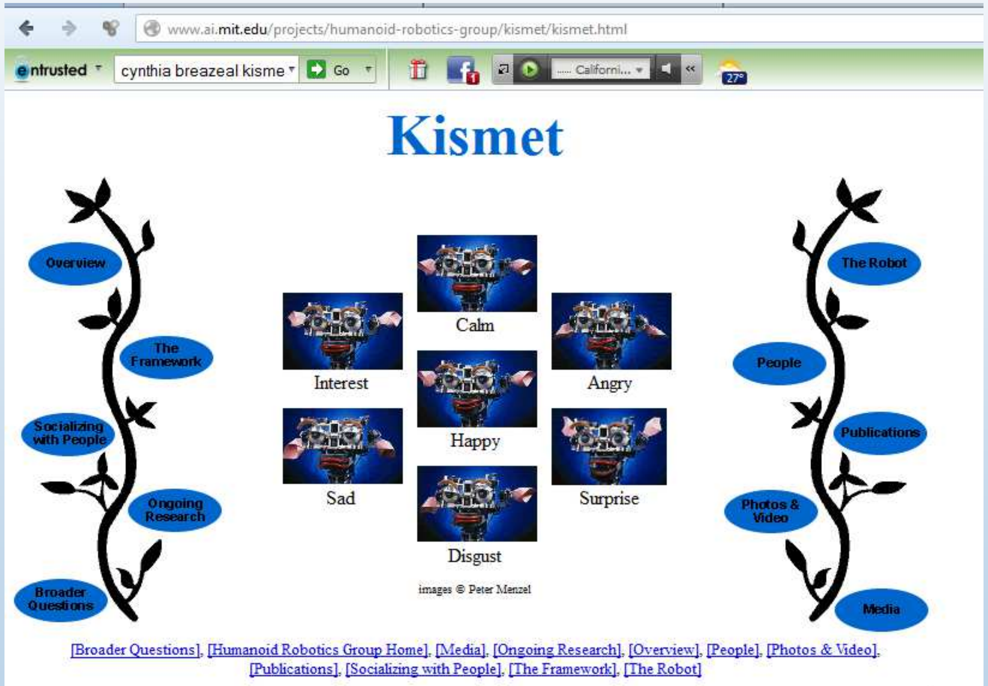
Shared attention

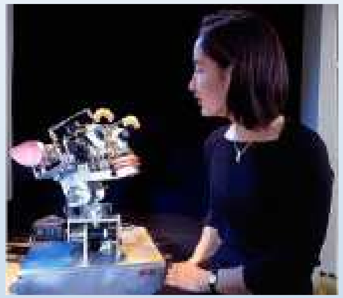
- A Nao robot, equipped with a visual attention system interacts with humans and points at objects

G. Schillaci, S. Bodiroza, and V. V. Hafner. Evaluating the effect of saliency detection and attention manipulation in human-robot interaction. *International Journal of Social Robotics*, 2012

Shared attention

- One of the first robots with shared attention
- Robot Kismet interacts with humans in a natural and intuitive way
- Cynthia Breazeal at MIT




© Sam Ogden)

- http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html

Towards Attentive Robots

Do we have already attentive robots??

We are still in the preliminary phase and a lot has to be done to really achieve human-like capabilities and behavior on robots

But interest in attentive capabilities is growing the more the systems advance!

Literature

- Simone Frintrop: **Towards Attentive Robots**, *PALADYN Journal of Behavioral Robotics, Springer, Vol. 2, Issue 2, 2011*
- Simone Frintrop, Erich Rome and Henrik I. Christensen: **Computational Visual Attention Systems and their Cognitive Foundation: A Survey**, *ACM Transactions on Applied Perception (TAP), Vol. 7, Issue 1, 2010*
- Martín-García/Frintrop ,**A Computational Framework for Attentional 3D Object Detection**, *Annual Meeting of the Cognitive Science Society,* 2013
- Simone Frintrop, Patric Jensfelt: **Attentional Landmarks and Active Gaze Control for Visual SLAM**, *IEEE Transactions on Robotics, Special Issue on Visual SLAM, vol. 24, no. 5, Oct. 2008*
- A. Rotenstein et al., **Towards the dream of intelligent, visually-guided wheelchairs**, *Proc. 2nd Int'l Conf. On Technology and Aging, Toronto, Canada 2007*
- P.-E. Forssén et al., ***Curious George: An Attentive Semantic Robot***, *Robotics and Autonomous Systems Journal, Volume 56, Number 6, 503-511, 2008*