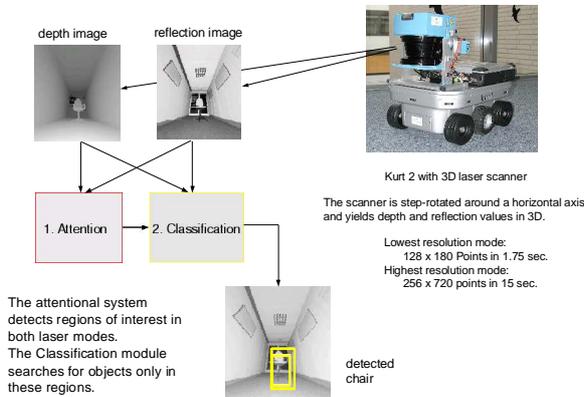# Focussing Object Recognition on Regions of Interest
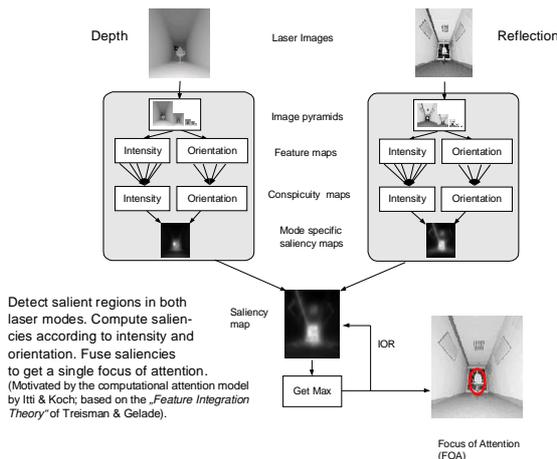
Simone Frintrop,
Andreas Nüchter, Hartmut Surmann
Team Robot Control Architectures (ARC)
Contact: simone.frintrop@ais.fraunhofer.de

Motivation:

Challenge: Detect objects in laser data supported by attention.

depth image   reflection image

1. Attention   2. Classification

Kurt 2 with 3D laser scanner

The scanner is step-rotated around a horizontal axis and yields depth and reflection values in 3D.

Lowest resolution mode:
128 x 180 Points in 1.75 sec.
Highest resolution mode:
256 x 720 points in 15 sec.

The attentional system detects regions of interest in both laser modes.
The Classification module searches for objects only in these regions.

detected chair

The Bimodal Attention System

Depth   Laser Images   Reflection

Image pyramids

| Intensity | Orientation | Feature maps | Intensity | Orientation |

| Intensity | Orientation | Conspicuity maps | Intensity | Orientation |

Mode specific saliency maps

Detect salient regions in both laser modes. Compute saliencies according to intensity and orientation. Fuse saliencies to get a single focus of attention.
(Motivated by the computational attention model by Itti & Koch; based on the „*Feature Integration Theory*" of Treisman & Gelade).

Saliency map

IOR

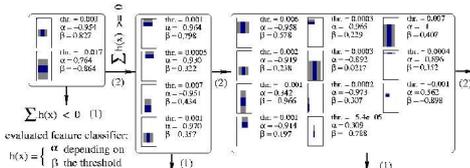Get Max

Focus of Attention (FOA)

Classification

Classification of objects according to Viola & Jones:
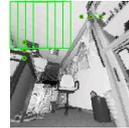- Compute 6 Basis features with *Integral Images* in constant time:

- Train with *Ada Boost Algorithm* on 200 images of office chairs.
- A cascade achieves a high detection rate and a low error rate.

The Chair Cascade:

$$\sum h(x) < 0 \quad (1)$$

evaluated feature classifier:

$$h(x) = \begin{cases} \alpha & \text{depending on} \\ \beta & \text{the threshold} \end{cases}$$

Object detection:

Search for object in specified search window (20 x 40 pixels).

Usual approach (uninformed search):
Scan whole image with search windows.
Time consuming, specially for many object classes!

Our approach (attentive search):
- Detect regions of interest with attention.
- Determine 5 starting points for classification from each focus.
- Search for objects in the neighborhood of these points (30% of the image!).

Focus

Results:

Focus of Attention
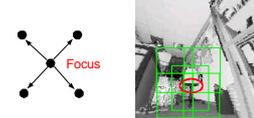
detected chair

Classifier detects almost all focussed chairs: 14 test images with focussed chairs, in 13 images the chairs are detected by the classifier.
(Viola/Jones: Classifier achieves detection and false positive rates equivalent to the best published results).

Classification: 200 ms   for the uninformed search (Pentium IV, 2400 MHz)
60 ms   for the attentive search (**30%** of uninformed search)
230 ms   for the attention system

Classification of m object classes:

m	uninformed m * 200 ms	attentive 230 + m * 60 ms
1	200	290
2	400	350
3	600	410
4	800	470
5	1000	530

=> for more than one object class the attentive search is useful!

Future work:

1) Extend the model to combine laser and camera data:

2) Integrate *top-down influences* depending on the robot's task,
e.g.: (task = find an object to grasp) => increase depth influence
(task = detect a red poster)   => increase color influence

References:

Frintrop, Rome, Surmann, Nüchter: "*An Attentive Multi-modal Laser „Eye*"", ICVS '03
Itti, Koch, Niebur: „*A model of saliency-based visual attention for rapid scene analysis*", in IEEE Trans. Patt. Anal. Mach. Intell. '98
Surmann et al: „*A 3D laser range finder for autonomous mobile robots*", in Proc. ISR '01
Treisman, Gelade: „*A feature-integration theory of attention*", in Cognitive Psychology '80
Viola, Jones: „*Robust Real-time Object Detection*", in Workshop on STCV '01

AIS

Fraunhofer Institut
Autonome Intelligente
Systeme