JNHR 2010

Designing Human Inspired SLAM for Humanoid Robots

Laboratoire Images, Signaux et Systèmes Intelligents (LISSI / EA 3956),
Université Paris-Est,
Institut Universitaire de Technologie de Sénart

Dominik Maximilián Ramík, Christophe Sabourin, Kurosh Madani

Overview

Presentation Outline:

1. Nao Humanoid Platform
2. State of the Art in Robotic SLAM, Semantic SLAM
3. Learning and Detecting Objects, Semantic Mapping
4. Conclusion, Perspectives

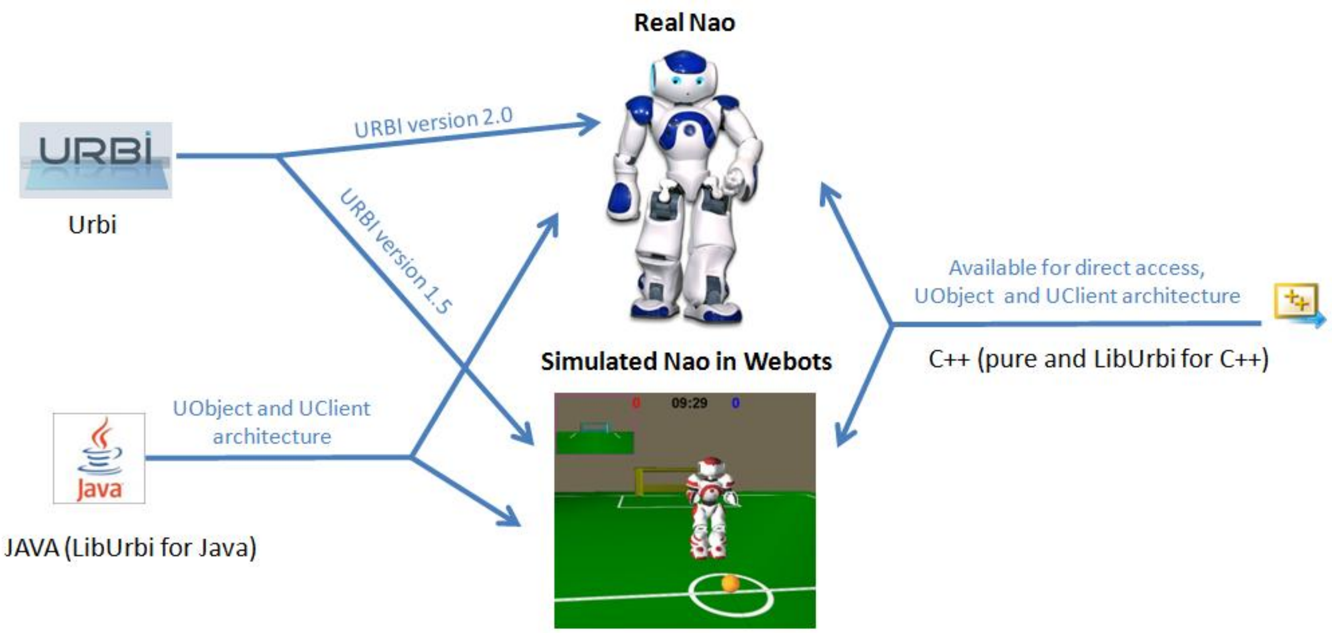
Nao Humanoid Robot

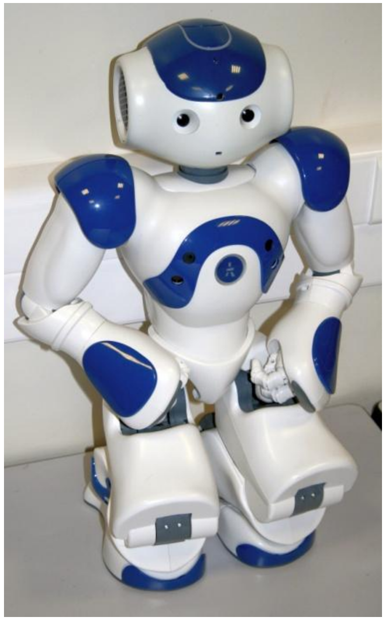
- ***Aldebaran's Nao*** humanoid robotic platform for experiments

- the human-oriented nature of our approach encourages use of a humanoid

- 25 DOF, 58cm high, among other sensors two 640x480 CMOS cameras (non-stereo)

- programmable in C, C++, Python, Urbi, connectors for Urbi-Java and URBI-Matlab available, newly C# connectors available too

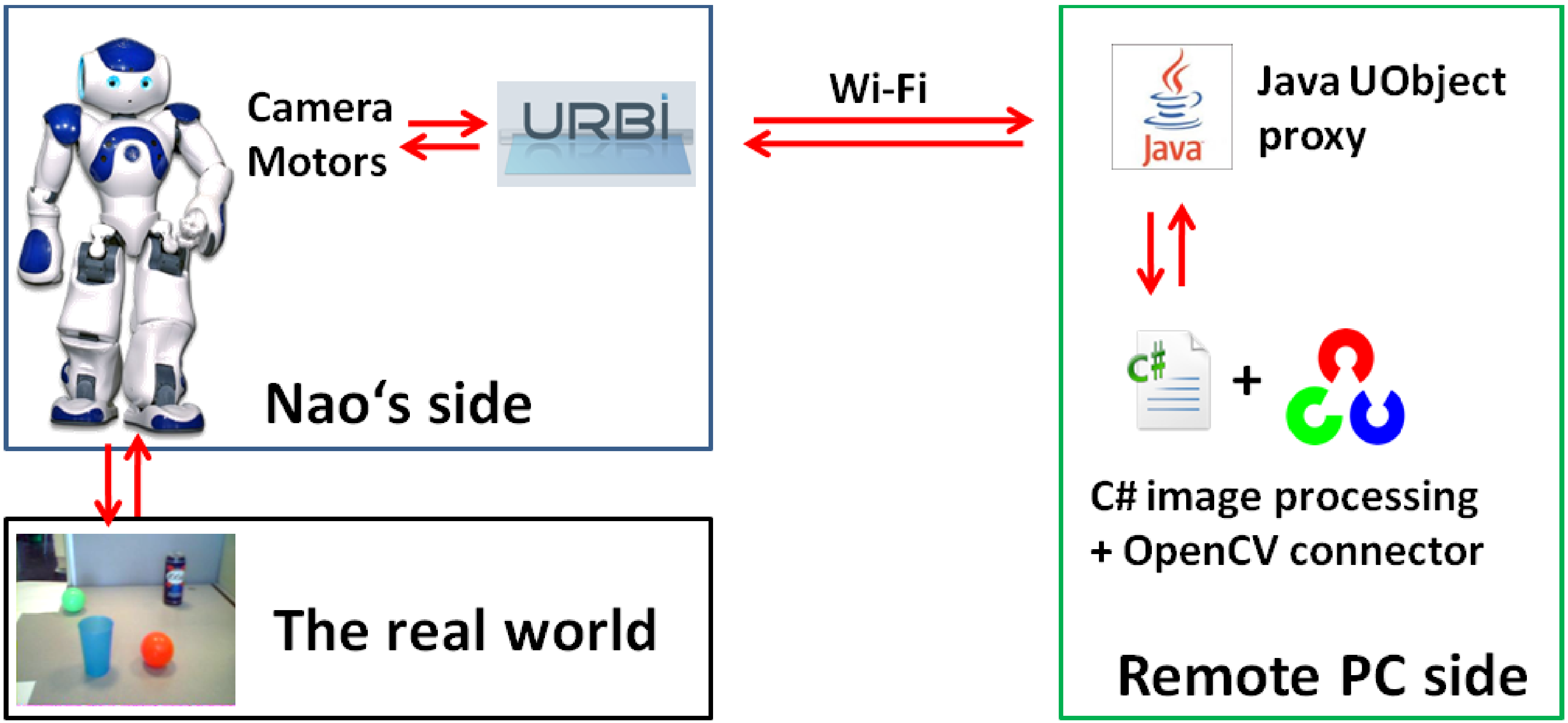




Our robotic platform

- Platform for real-world experiments:
  - for behaviors like walking and image acquisition ***URBI*** custom ***UObjects*** had been used
  - computationally heavy image processing deferred to a remote PC (WiFi connected), a C# program using ***Emgu OpenCV*** wrapper
  - URBI was interfaced with C# by a Java proxy component
  - Solely C# connector without URBI in the future

Our robotic platform



SLAM basics

- ***SLAM*** for Simultaneous Localization And Mapping
- informal definition:

  „A process, in which a mobile robot explores an unknown environment, creates a map of it and uses it simultaneously to infer its own position on the map“
- a ***crucial ability*** for modern mobile robots
- state of the art methods and areas of interest
  - Extended Kalman Filter
  - FastSLAM, Particle filters
  - SLAM in dynamic environments
  - ***semantic SLAM***

Semantic SLAM

- motivation:
    - most current SLAM approaches treat the environment without „understanding“ it

    BUT

    - the semantic information about the environment is valuable and may significantly improve the performance in terms of SLAM
- current robots perceive environment in a very different way if compared to humans
- incorporating semantics into maps and the mapping process should
    - improve robot‘s mapping capability by understanding the ***nature of encountered objects***
    - allow better place recognition and description
    - improve ***sharing the knowledge*** about the environment between humans and robots

Human inspired semantic SLAM

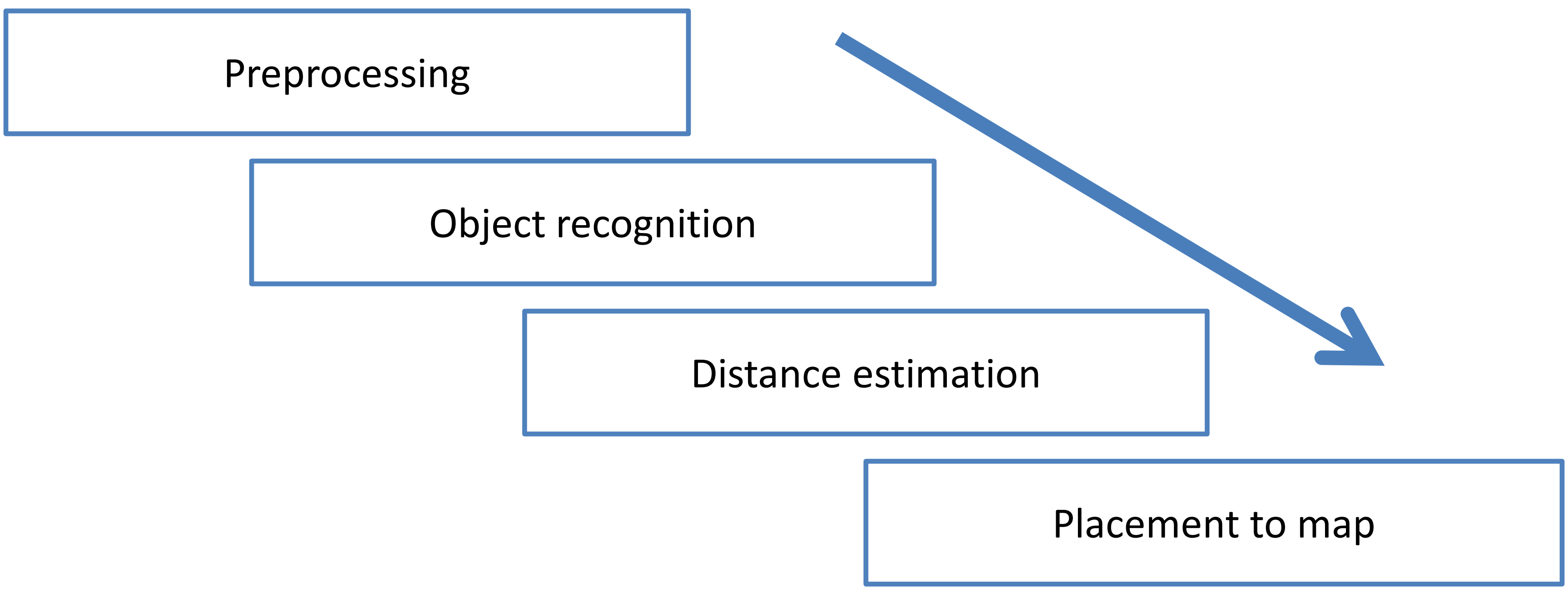
- observation:
    - humans perform SLAM implicitly without any significant effort while using completely different means of place description and orientation than contemporaneous robots
    - use of fuzzy language terms and description of places by means of objects and their relations
    - use of exact metric information about objects is minimal
- goal:
    - for prospective adoption: determine interesting strategies that humans use in context of SLAM
    - explore feasibility of semantic SLAM using human inspired place and object description

„Proto-map“ construction

- apart of stereo vision, humans are able to ***infer the distance*** of an object from its apparent size and en experience of its real size

- recognizing objects in robots environment enables creation of a proto-map with ***approximated distances*** of surrounding objects

- a monocular color camera is used as a source sensor



Object detection and recognition

- fast object recognition needed for real-time operation

-idea: present the robot arbitrary objects in different conditions and let it to learn them

- steps:

    - **training samples acquisition**: robot records images of different places with objects of interest present (but not a-priori marked)

    - ***segmentation***: „object level“ segmentation using salient region detection

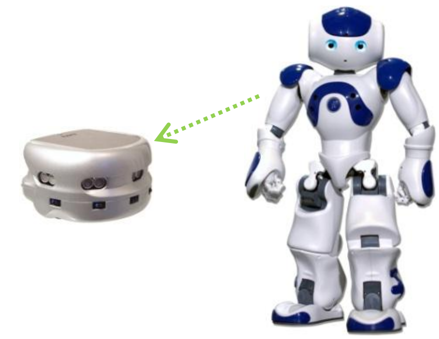
    - **learning object determination**: by statistical analysis of object presence through the training sequence

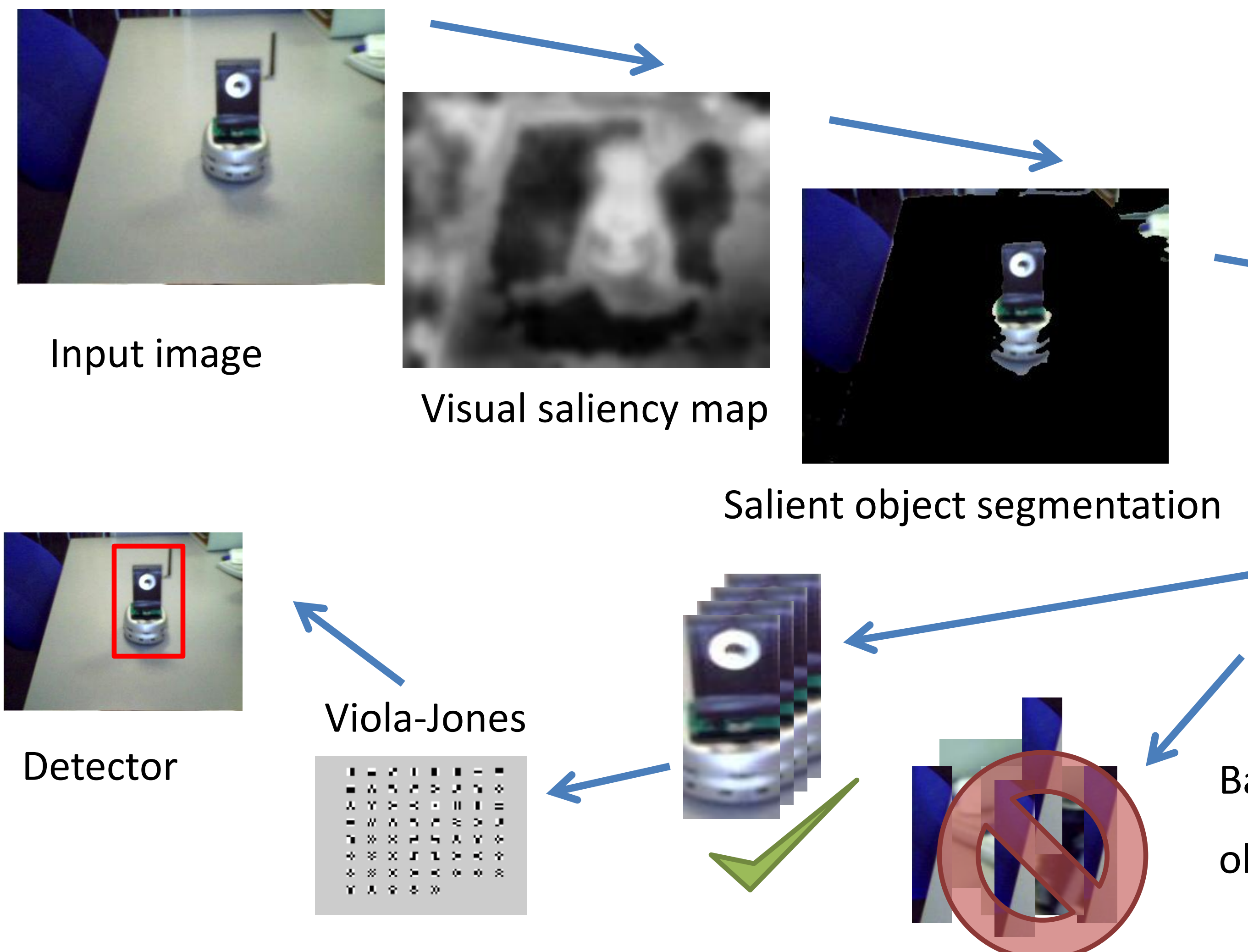
    - **training set preprocessing**: automatic ROI marking around the found objects

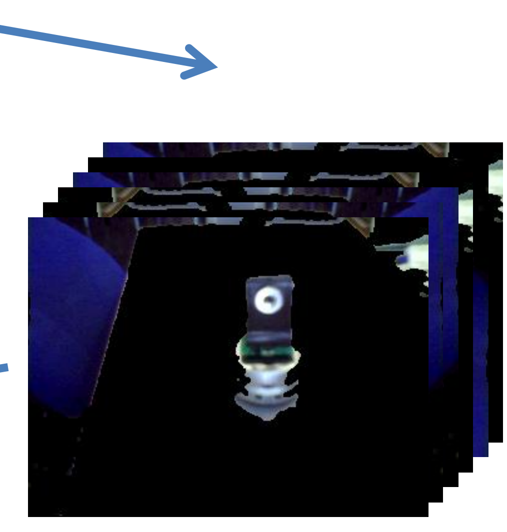
    - ***training using Viola-Jones object detection framework:*** for fast object detection

Object detection and recognition

- Learning to recognize a friend: Nao vs. Khepera

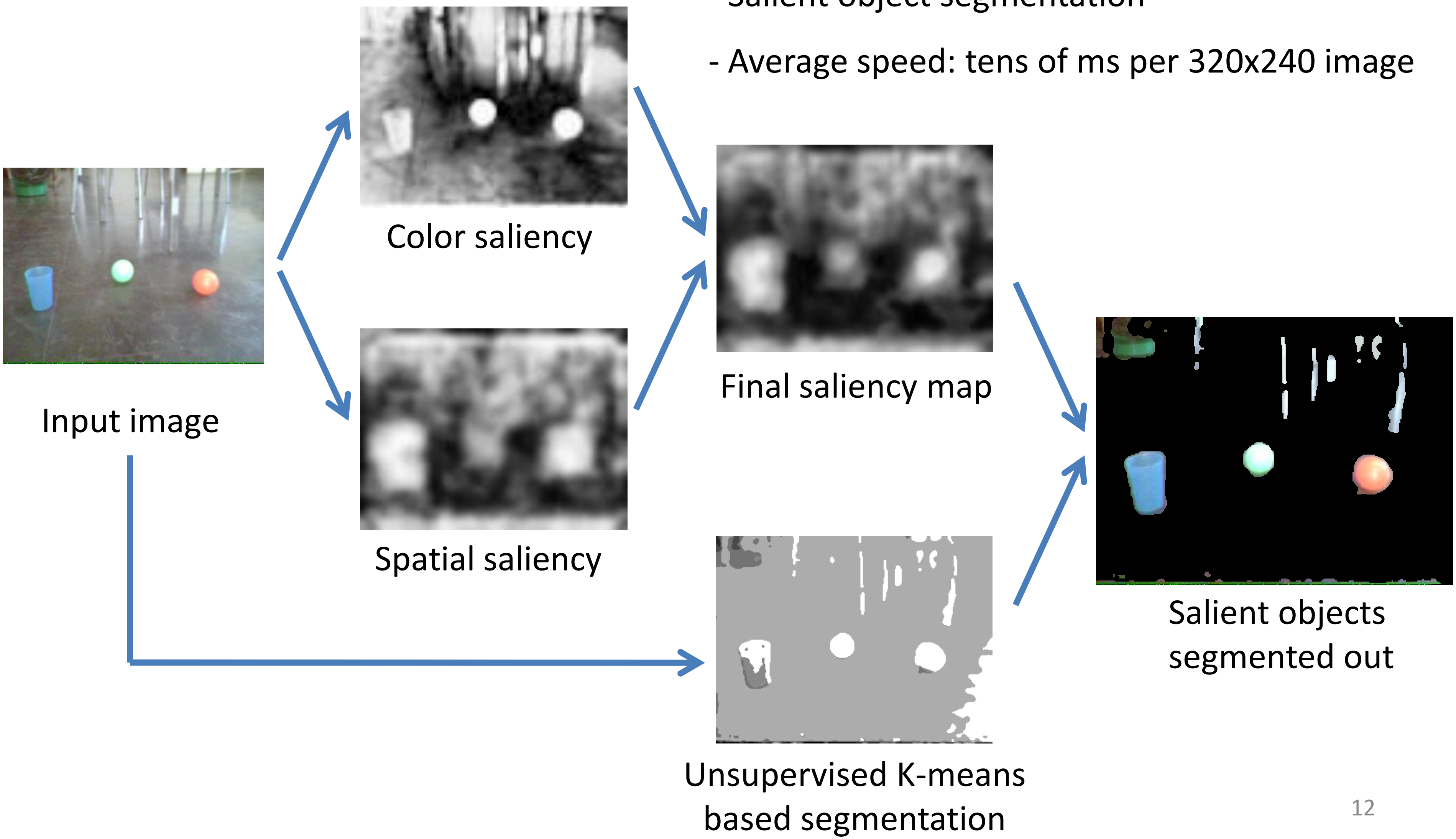






Object detection and recognition

- Salient object segmentation
- Average speed: tens of ms per 320x240 image



Distance estimation

- knowing the nature of objects in the scene, we can recall their typical sizes (given as a prior knowledge or as a result of learning)

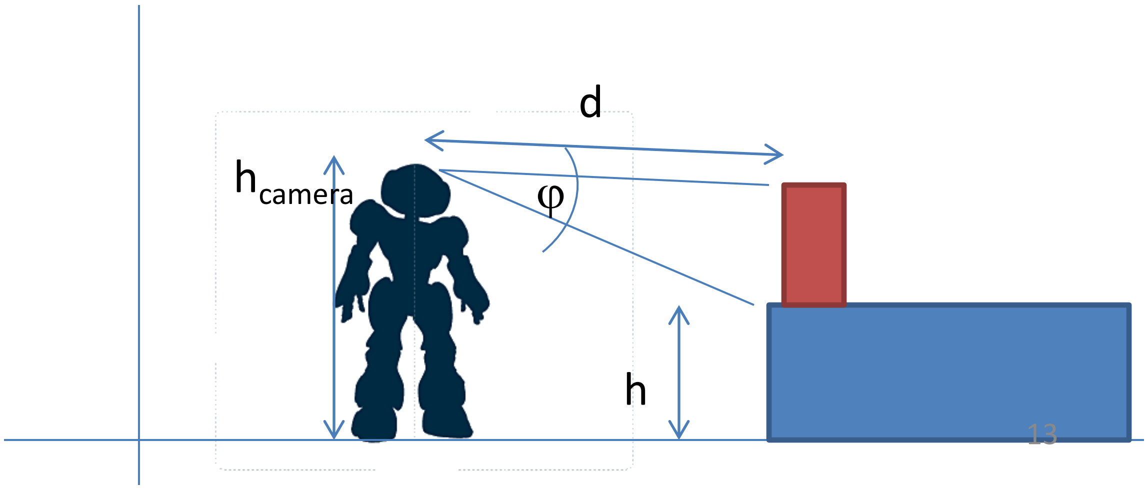
- using its typical size, we can approximate the distance *d* of the object form camera

$$d = w_{real} * \tan (w_{px} * \varphi / w_{image})$$

- for height of the object above the robot‘s ground, we use the following equation

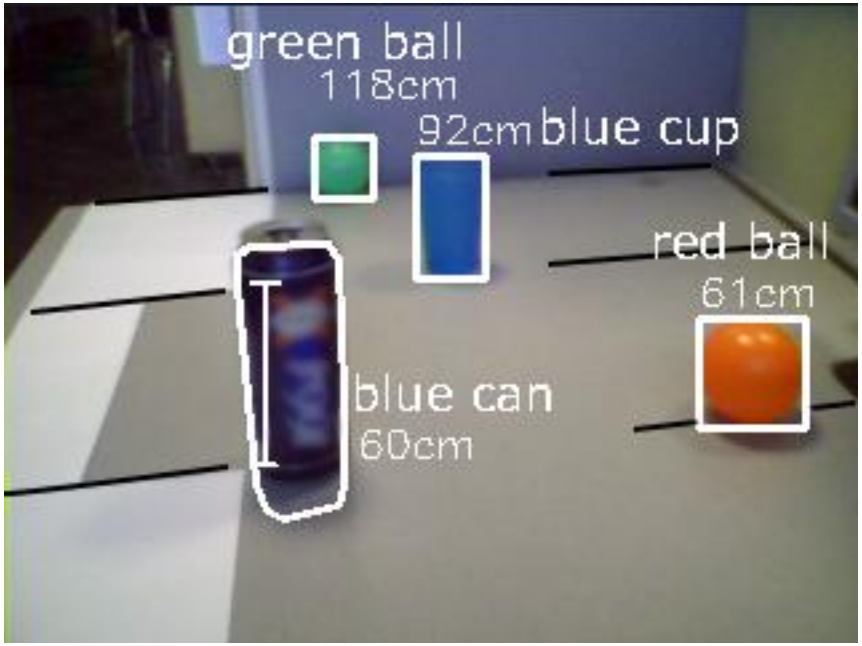
$$h = h_{camera} - d * \cos(\pi/2 – \alpha)$$

- the result is a rough estimation of the object‘s position with respect to the robot, the aim is to enable fast inference of spatial relations of objects (behind, above etc...)



Results

- in course of experiments, Nao walked through the real environment, identifying surrounding objects and inferring their distances
- the distance estimation precision turns to be accurate enough to allow inference of spatial relations of perceived objects (the error doesn‘t exceed 5-10% depending on objects distance from the camera)

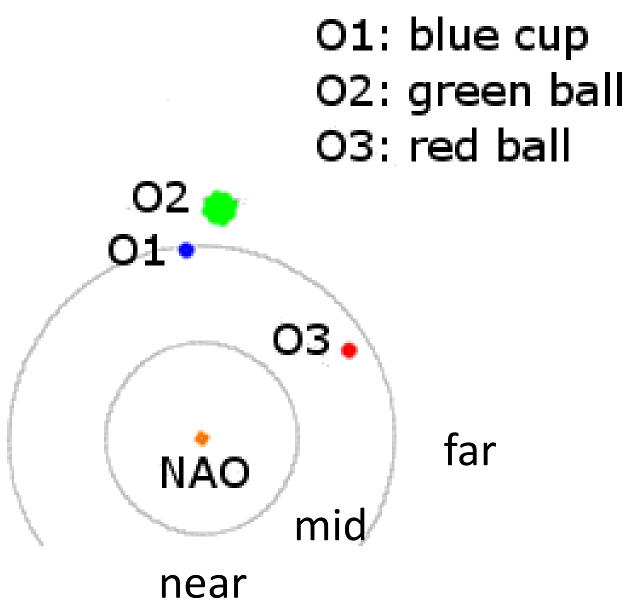


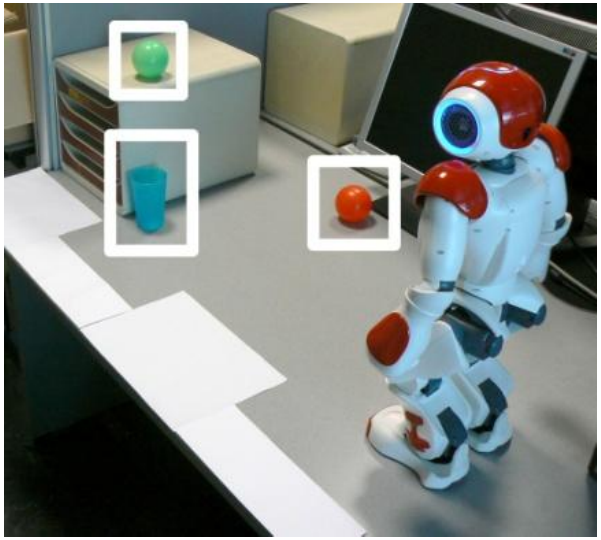
Results

- a ***semantic proto-map*** of surrounding objects has been created during real-world experiments

- looking around (in approx 180deg) Nao detected different objects, estimated their distance and height above the surface and registered them into a proto-map

- whenever Nao moves, new measures are taken and the map is updated





Results

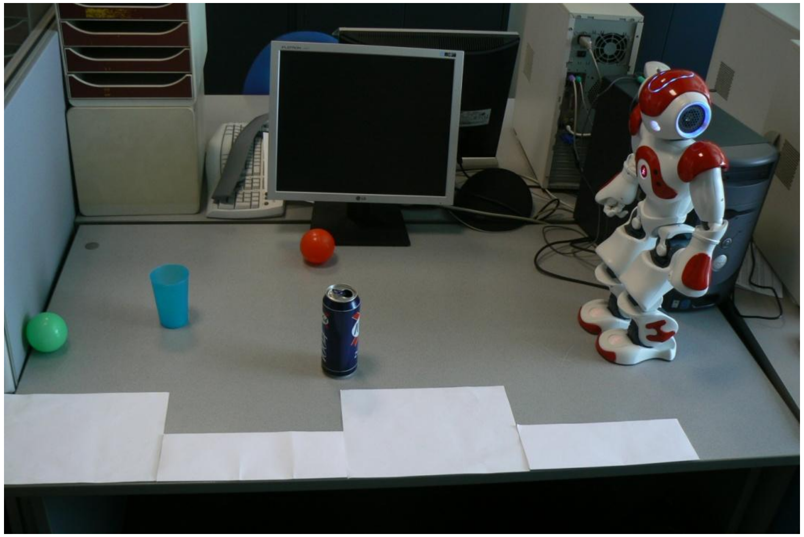
- several test run to evaluate semantic mapping of Nao

- each time, Nao has first created a proto-map of its surroundings

- command has been given to it in natural language (eg. „fetch the red ball“) by voice or by keyboard

- using the generated map, Nao walked to fetch the ball and finally announced successful accomplishment of the task with its voice synthesizer

- Nao could be equally asked to tell the position and distance of certain object, giving the response in human language (eg: „The blue cup is far on the left“)



[VIDEO]

Conclusion

- state of the art in SLAM has been investigated and ***semantic SLAM*** has been identified as an important direction for future mobile robots
- we point out the pertinence of semantic SLAM for autonomous navigation and human-robot interaction with ***a special focus on humanoid robots***
- we present an initial research on our concept of human-inspired semantic SLAM with experiments carried out on the humanoid robot Nao, making use of some of human spatial concepts

Future work

- results of our experiments are encouraging for further development of a human inspired SLAM algorithm
- our future work will be focused on
    - development of an operational algorithm of human inspired semantic SLAM, detached as much as possible from the metric information and relying on linguistic relations between objects and fuzzy logic
    - learning of spatial relations of important objects found in the environment and their representation in a semantic map using human language categories

JNHR 2010

Thank you for your attention

Designing Human Inspired SLAM for Humanoid Robots

Laboratoire Images, Signaux et Systèmes Intelligents (LISSI / EA 3956),
Université Paris-Est,
Institut Universitaire de Technologie de Sénart

Dominik Maximilián Ramík, Christophe Sabourin, Kurosh Madani