LUVSS @ FRI @ UL

Visual Cognitive Systems Laboratory
Faculty of Computer and Information Science
University of Ljubljana
Ljubljana, Slovenia

http://vicos.fri.uni-lj.si

Head: **Professor Dr. Aleš Leonardis**

Staff:

Prof. **dr. Aleš Leonardis**
Assist. prof. **dr. Danijel Skočaj**
Researcher **dr. Marko Boben**
Researcher **dr. Matej Kristan**
Researcher **dr. Jaka Krivic**
Researcher **Luka Čehovin, M.Sc**
Researcher **Barry Ridge, M.Sc.**
Researcher **Peter Uršič, M.Sc.**
Researcher **Marko Mahnič**
Researcher **Alen Vrečko**
Researcher **Domen Tabernik**

ViCoS Areas of Research

- Research of intelligent (artificial cognitive) systems
- Focus is on visual learning and recognition:
  - recognition and categorisation of objects, places and actions
  - supervised/unsupervised learning, learning in interaction with a tutor and a scene
- Other activities:
  - panoramic imaging for mobile robotics
  - range image modeling and interpretation
- Target platforms: mobile robots, intelligent environments, personal devices



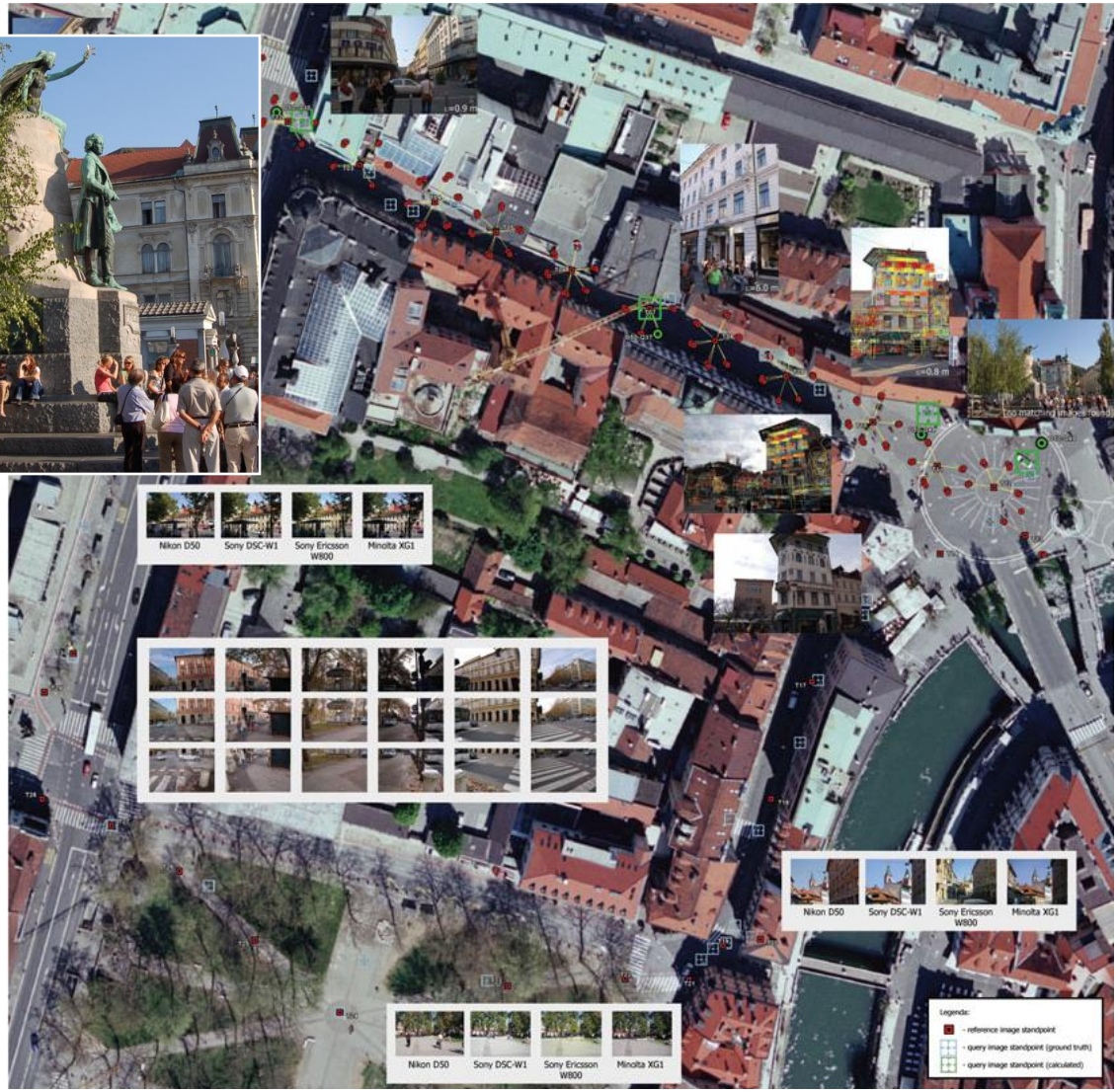


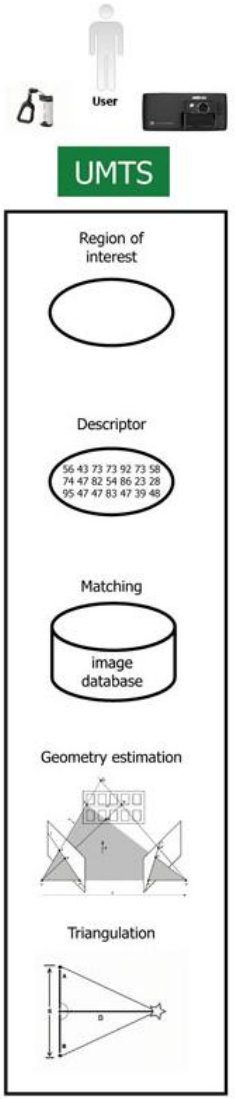

EU projets @ ViCoS

- 5FP:
  - **CogVis** - Cognitive Vision Systems, 2001-2004. (4M EUR/107.000 EUR)
  - **ECVision** - European Research Network for Cognitive Computer Vision Systems, 2002-2005
- 6FP:
  - **CoSy** - Cognitive Systems for Cognitive Assistants, IP, 2004-2008 (7.5M EUR/600.000 EUR)
  - **MOBVIS** - Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios, FET STREP, 2005-2008. (2.3M EUR/295.000 EUR)
  - **VISIONTRAIN** - Computational and Cognitive Vision Systems: A Training European Network, MRTN, 2005-2009. (3.5 M EUR/240.000 EUR)
  - **EuCognition** - The European Network for the Advancement of Artificial Cognitive Systems, 2006-2009.
- 7FP:
  - **Poeticon** – The "poetics" of everyday life: Grounding resources & mechanisms for artificial agents, STREP, 2008-2010, (4.3M/430.000 EUR)
  - **CogX** - Cognitive Systems that Self-Understand and Self-Extend, IP, (8.8M/888.000 EUR)
  - **EUCogII** - 2nd European Network for the Advancement of Artificial Cognitive Systems, Interaction and Robotics (2009-2012)

Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios

Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios



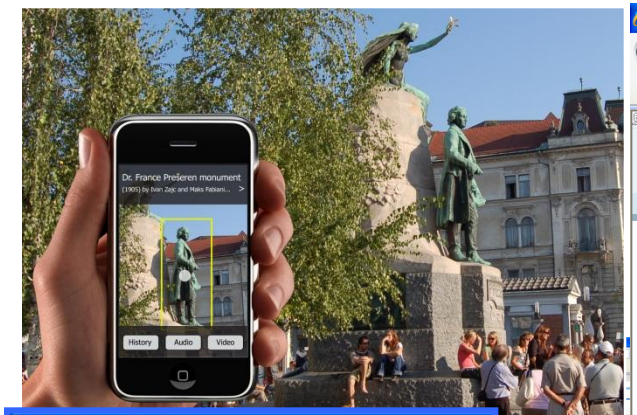


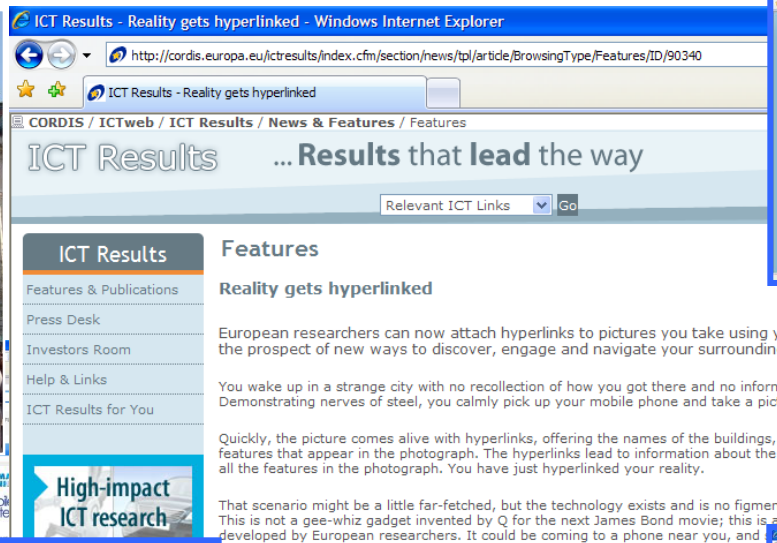
D. Omerčević, O. Drbohlav, and A. Leonardis. *High-dimensional feature matching: Employing the concept of meaningful nearest neighbors.* Accepted to ICCV 2007.

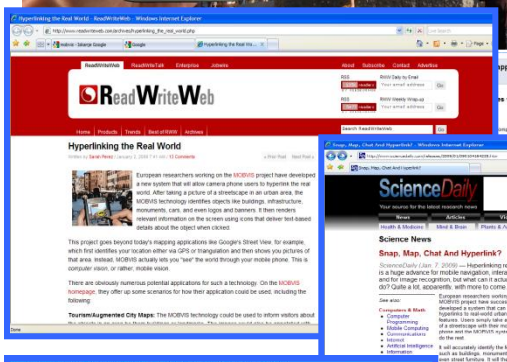
MOBVIS

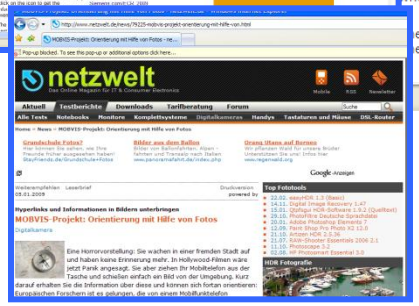
Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios

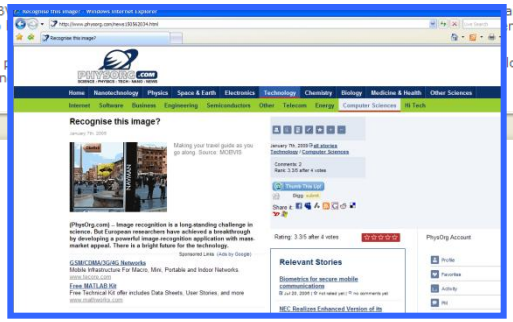
ViCoS Visual Cognitive Systems Laboratory





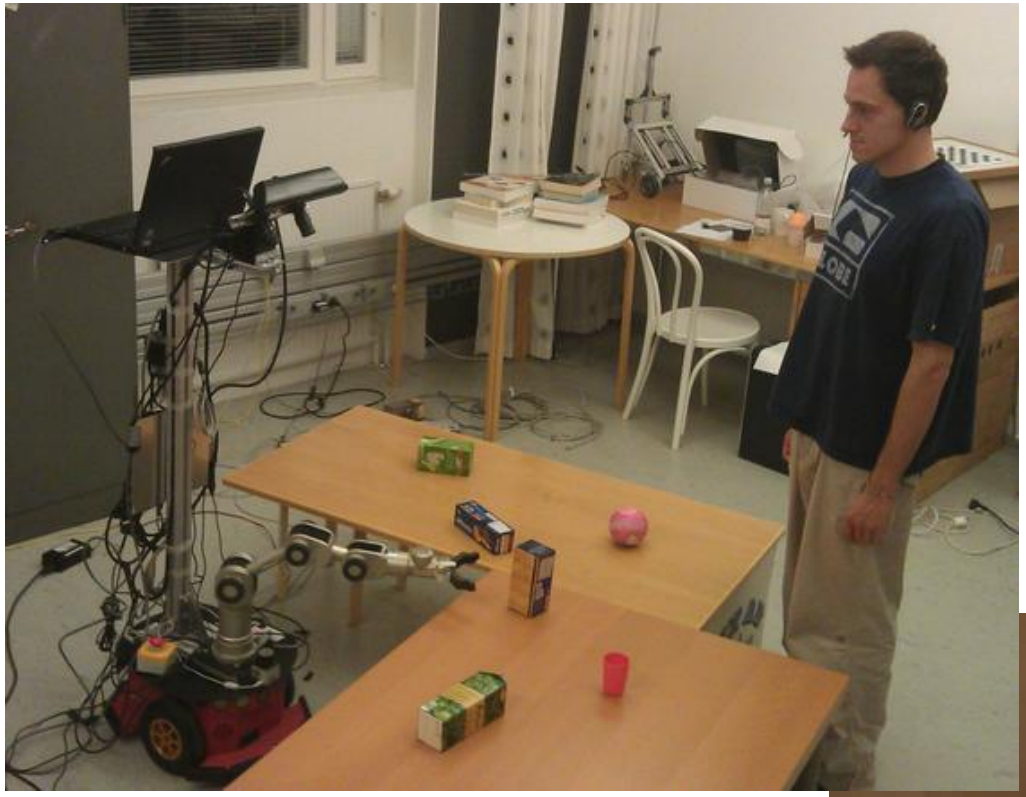




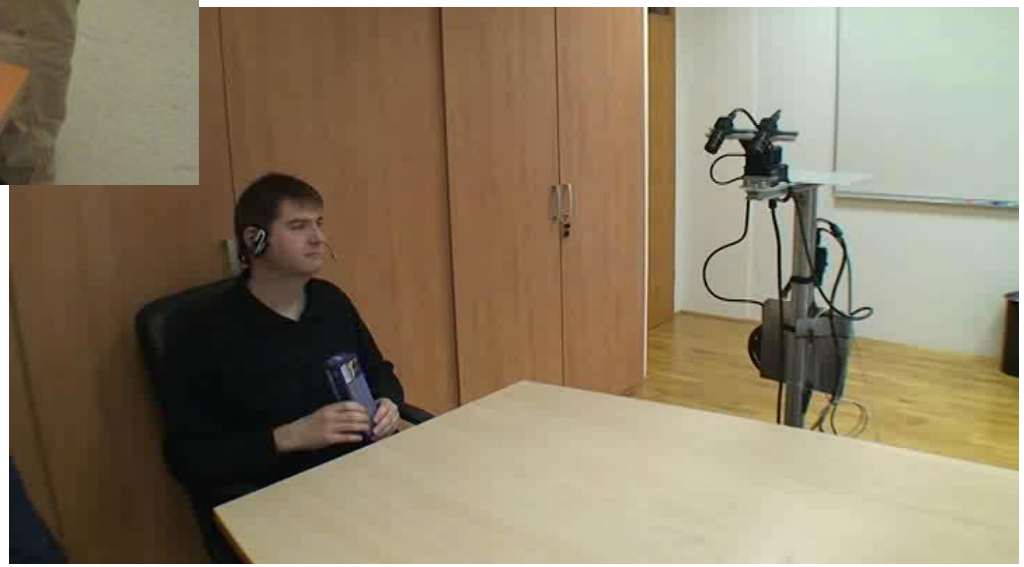


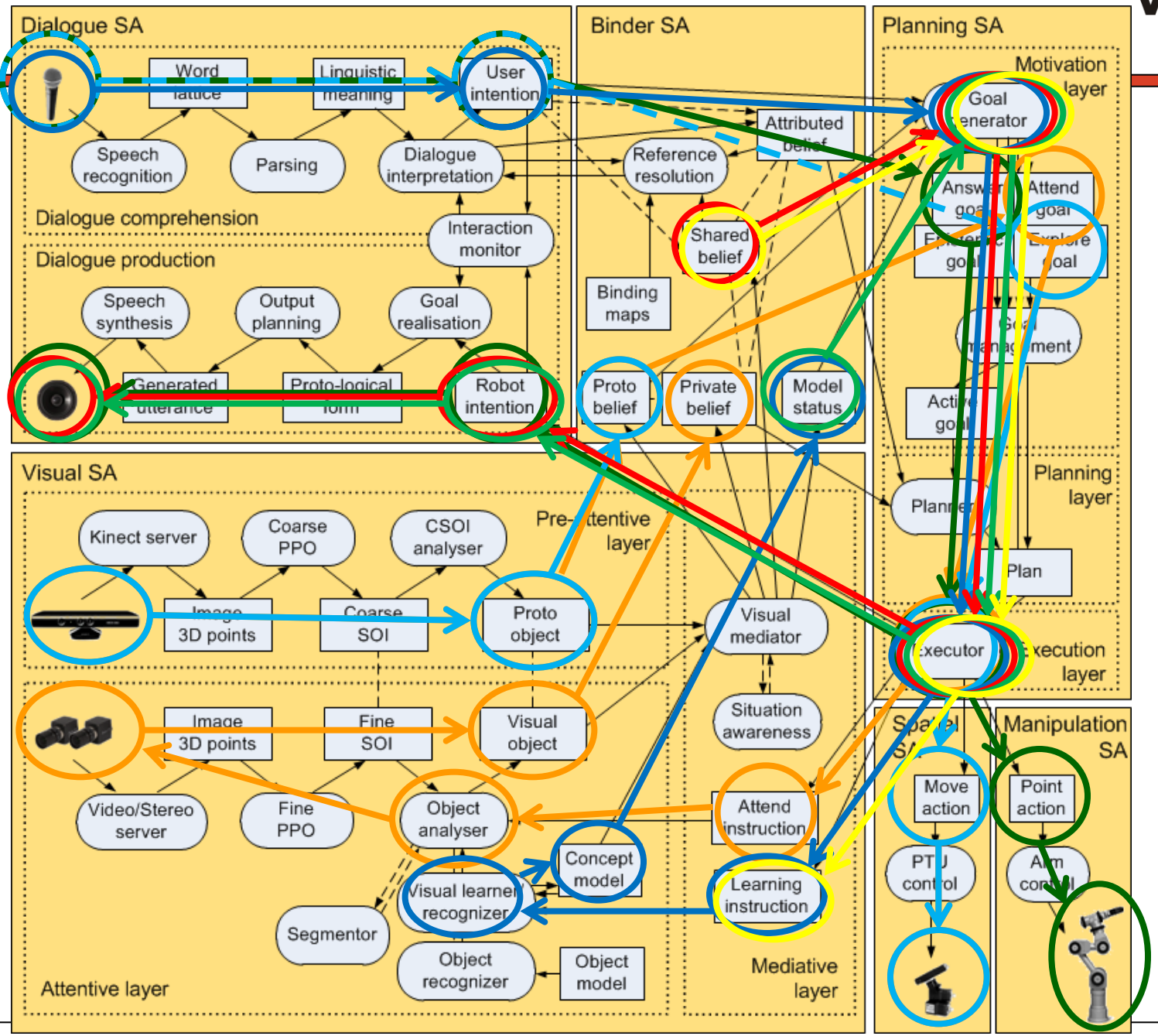
Pronounced as a success story by EU

CogX



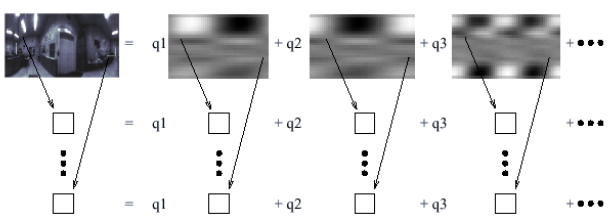
http://cogx.eu/results/george/



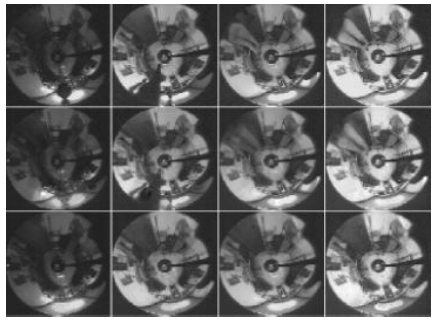

Dialogue SA
Word lattice
Linguistic meaning
User intention
Speech recognition
Parsing
Dialogue interpretation
Dialogue comprehension
Interaction monitor
Dialogue production
Speech synthesis
Output planning
Goal realisation
Generated utterance
Proto-logical form
Robot intention
Binder SA
Attributed belief
Reference resolution
Shared belief
Binding maps
Proto belief
Private belief
Model status
Planning SA
Motivation layer
Goal generator
Answer goal
Attend goal
Explore goal
Goal management
Active goal
Planning layer
Planner
Plan
Execution layer
Executor
Visual SA
Pre-attentive layer
Kinect server
Coarse PPO
CSOI analyser
Image 3D points
Coarse SOI
Proto object
Visual mediator
Situation awareness
Image 3D points
Fine SOI
Visual object
Video/Stereo server
Fine PPO
Object analyser
Attend instruction
Concept model
Visual learner recognizer
Learning instruction
Segmentor
Object recognizer
Object model
Attentive layer
Mediative layer
Spatial SA
Move action
PTU control
Manipulation SA
Point action
Arm control

Subspace methods

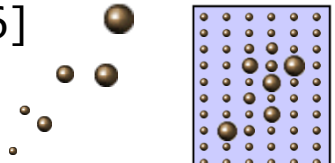
- Recognition
  - Robust PCA recognition [CVIU00]
  - Illumination insensitive recognition [CVIU04]



- Learning
  - Incremental building of eigenspaces [IMAVIS06]
  - Weighted PCA and PCA on missing data [PR06]
  - Robust building of eigenspaces [PR06]

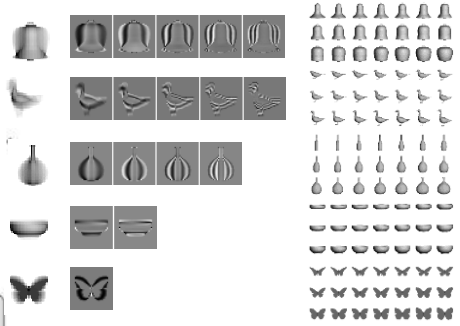


- Multiple eigenspaces [PR02]

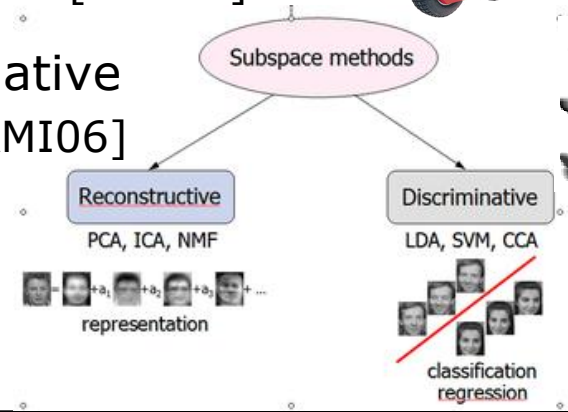


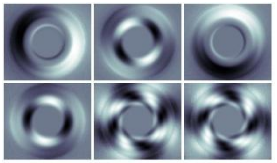
- Eigenspace of spinning images [IP03]

- Appearance-based localization [RAS03]

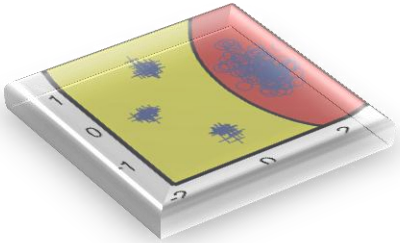


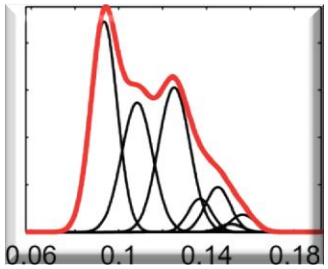
- Reconstructive vs. Discriminative
  - robust LDA, robust CCA [PAMI06]
  - incremental LDA





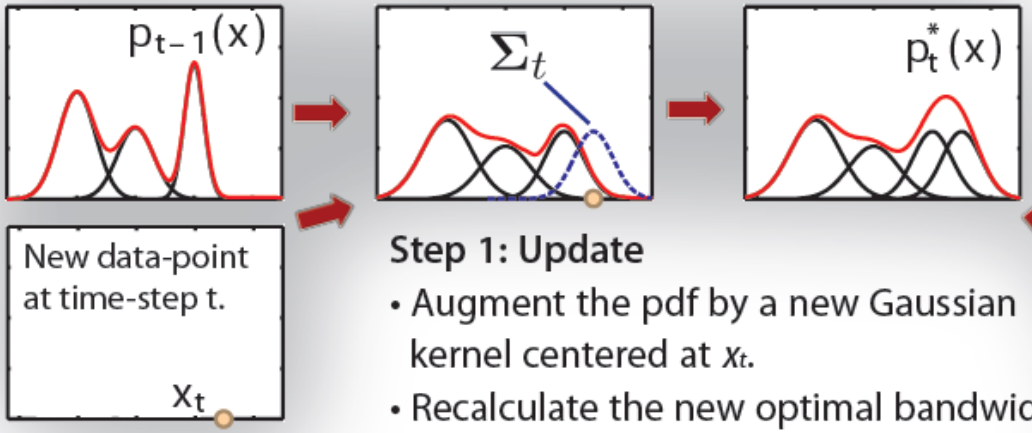
oKDE





The oKDE [1] produces a generative model from streaming data in form of a Gaussian mixture model. When a new data-point arrives, the oKDE proceeds in **two main steps**:

$$p(\mathbf{x})_{t-1} = \sum_{i=1}^{N} w_i \phi_{\Sigma_i}(\mathbf{x} - \mathbf{x}_i)$$

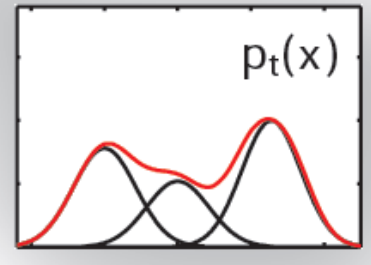


Step 1: Update

- Augment the pdf by a new Gaussian kernel centered at x_t.
- Recalculate the new optimal bandwidth (covariance) H_t and readjust the kernels.

Step 2: Compress

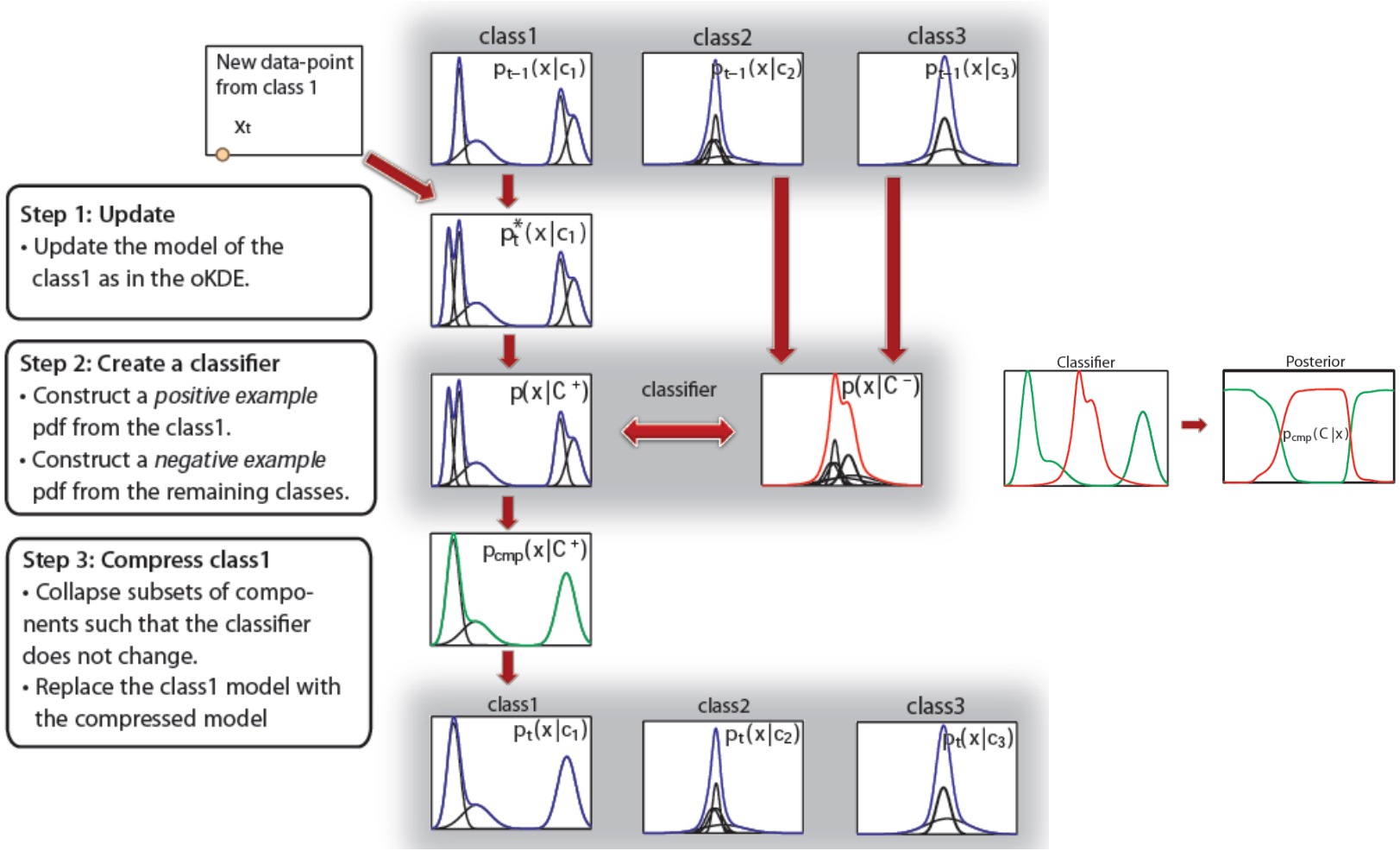
- Collapse subsests of components such that the distribution remains approximately the same.
- Prevent loss of reconstruction in the compression.



$$\phi_{\Sigma}(\mathbf{x} - \mu) = (2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} e^{(-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu))}$$

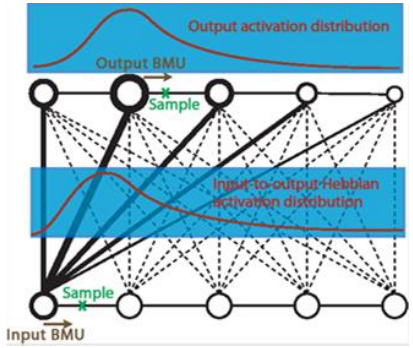
odKDE

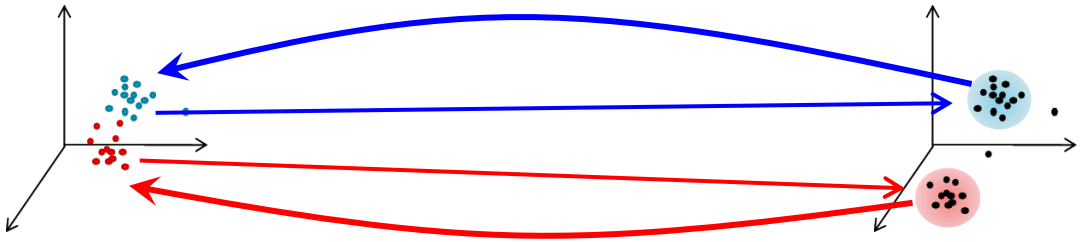
- Updates a model for class1 and compresses it such that classification properties of the Bayes classifier dc

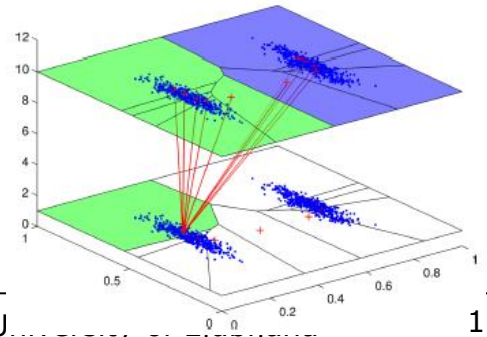


Learning of Object Affordances

- **Low-level cross-modal object affordance learning.**
- **Two distinct modalities:**
  - Object shape properties.
  - Object effects under action.
- **Broad Aims:**
  - Form percepts within each modality.
  - Form cross-modal associations between them.







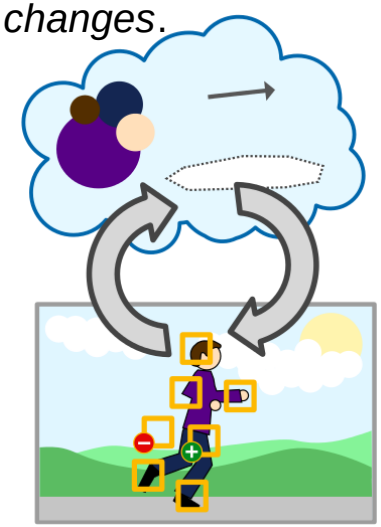
Visual tracking

Our goal: Visual tracking of objects undergoing *rapid structural appearance changes*.

Our contribution: Visual model structured into two layers:

Local layer – Set of patches used for short-term tracking.
Patches can be removed or added to the set.

Global layer – Describes the appearance of the entire object.
Updated on-line.
Used to add new patches.

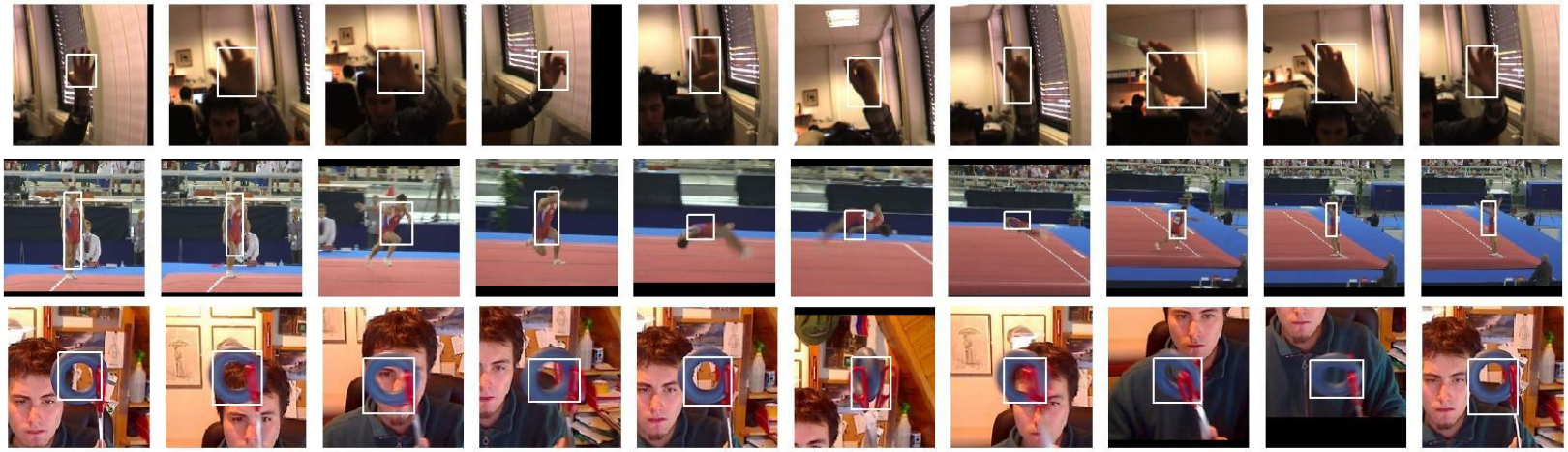


Theory: Bayesian formulation of structural model, updating and tracking.

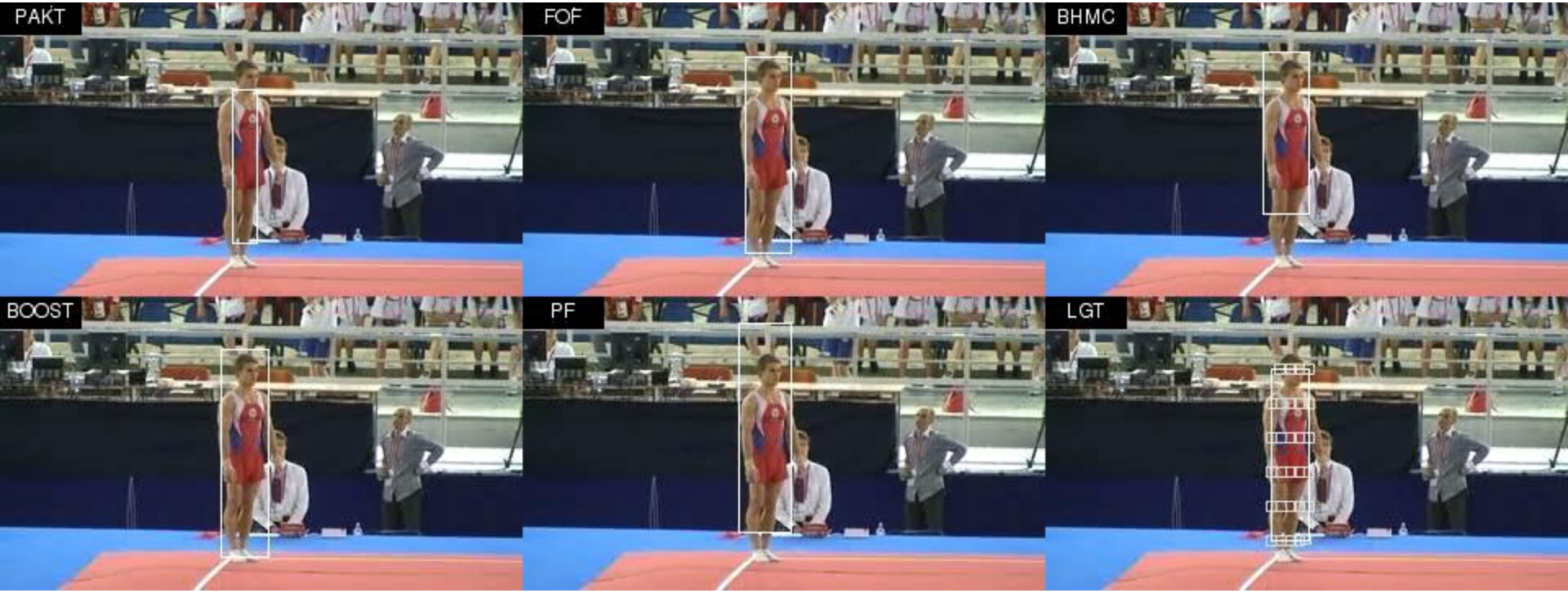
Experiments: 5 state-of-the art reference trackers, 6 challenging sequences

Our approach outperfoms state-of-the-art in accuracy and lower failure rate

Examples of tracking

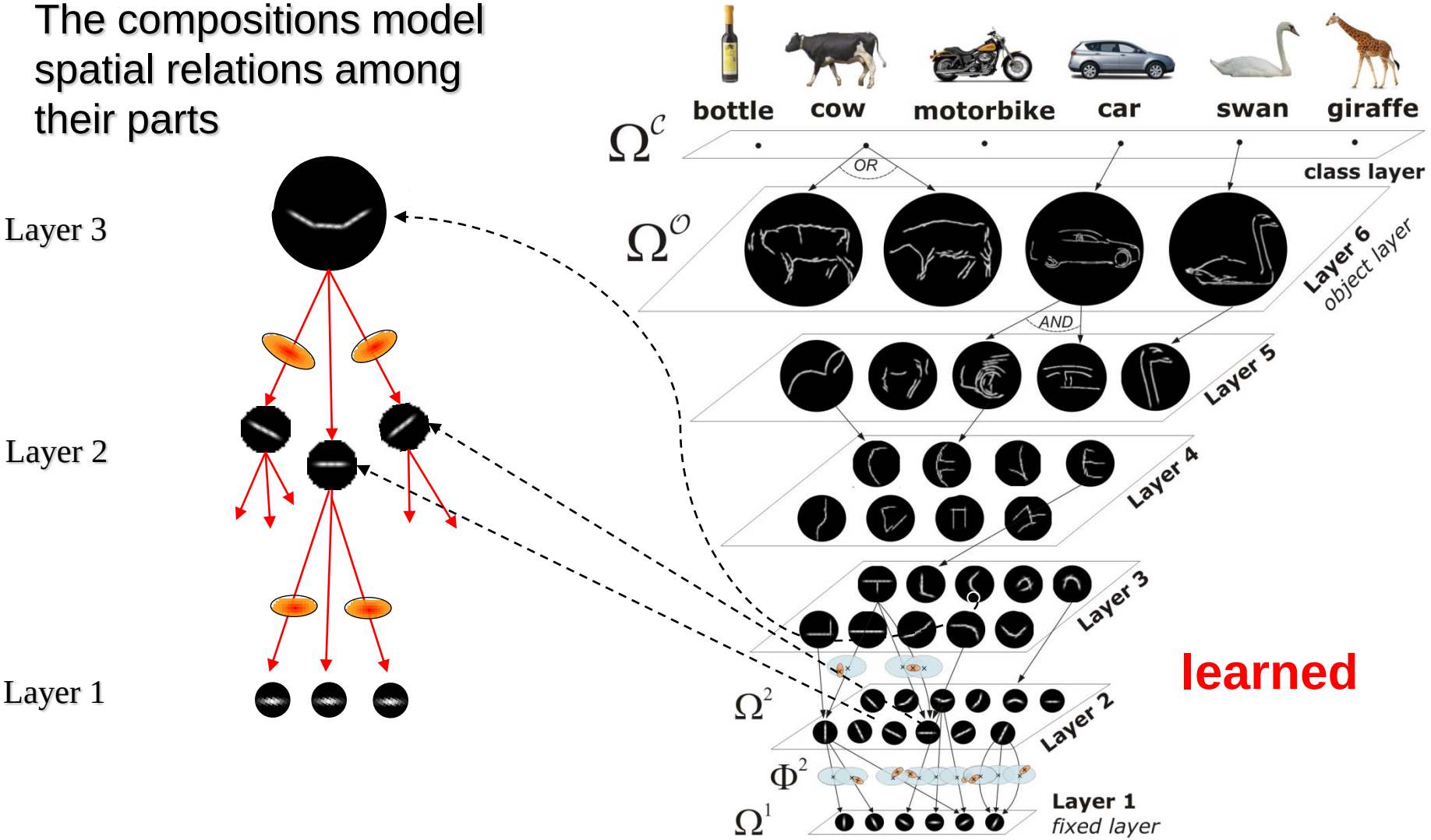


Visual tracking



Representation – the big picture

- Object representation: A **hierarchical compositional shape vocabulary**
- The compositions model spatial relations among their parts



Detection of object classes