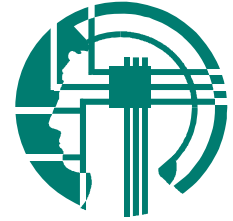




MAX-PLANCK-GESELLSCHAFT

**Heinrich H. Bülthoff  
Christian Wallraven**



MPI FOR BIOLOGICAL CYBERNETICS

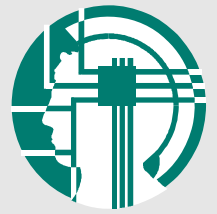
# **Multimodal Recognition and Categorization**

**Max Planck Institute for Biological Cybernetics,  
Tübingen, Germany**

**[www.kyb.mpg.de](http://www.kyb.mpg.de)**

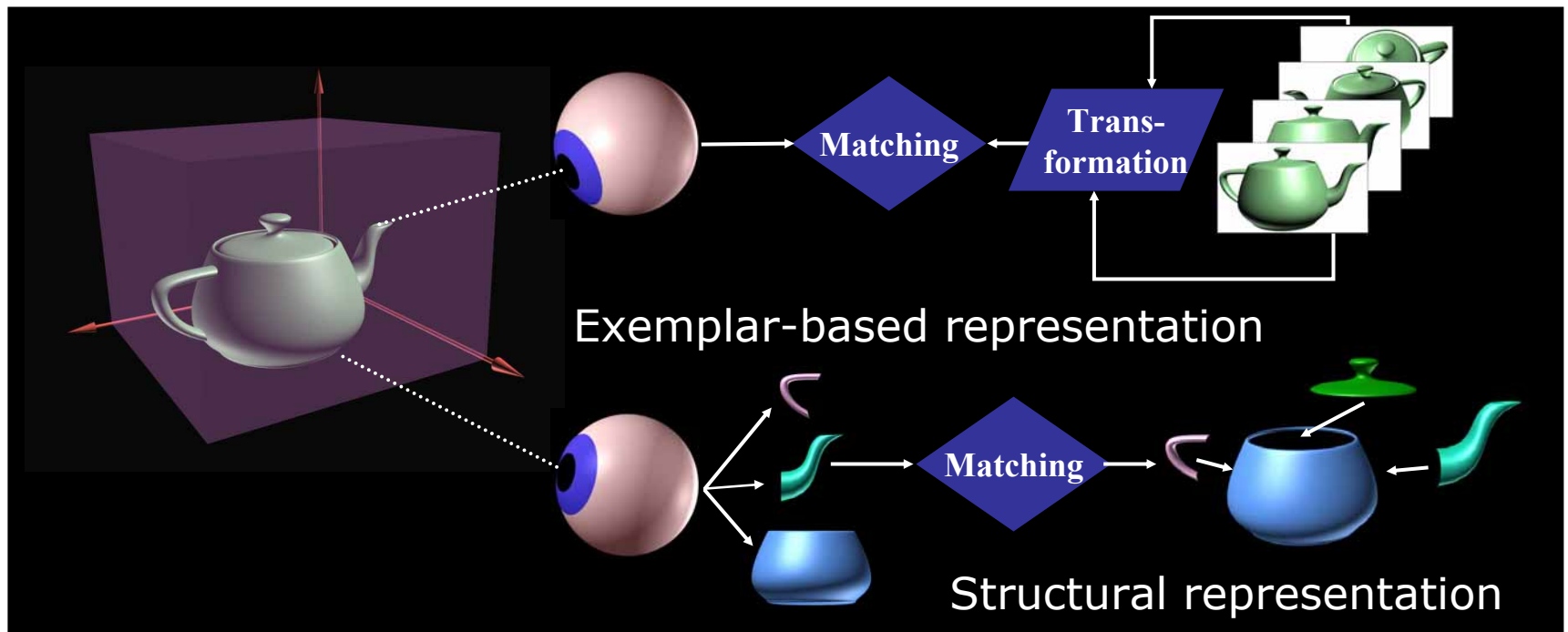
# Scientific perspective

## How did it all start?

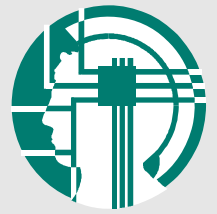


MPI for Biological Cybernetics

- How does recognition work?
  - view-based or exemplar-based representations [Bülthoff, Edelman, 1992,...]
  - or structural representations [Biederman, 1987, ...]



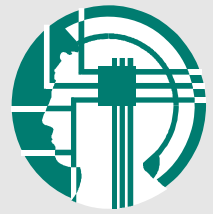
# Scientific perspective



MPI for Biological Cybernetics

- But recognition is of course not as simple as these two frameworks might suggest
- Aspects one needs to look at:
  1. transformation-costs of recognition processes
  2. spatio-temporal representations
  3. multi-modal/multi-sensory aspects
  4. recognition/categorization in the perception/action loop
- Before I address these points I will first report what I have learned from...

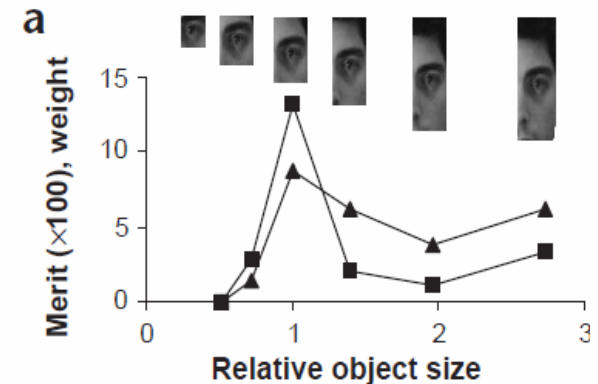
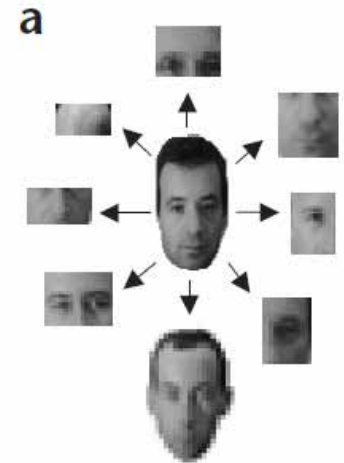
# What have we learned?



MPI for Biological Cybernetics

- **Shimon Ullman - Fragments**

- „simple“ fragments are very informative if applied properly
- the research on informative features is popular in computer vision motivating and validating many successful new approaches [Fergus et al., 2003, Fritz et al., 2005, Lowe, 2004; Ullman et al., 2002, Viola et al., 2004, Wallraven et al., 2001, 2003, ... ]
- it also provides an interesting approach to the problem of creating efficient intermediate, structural representations





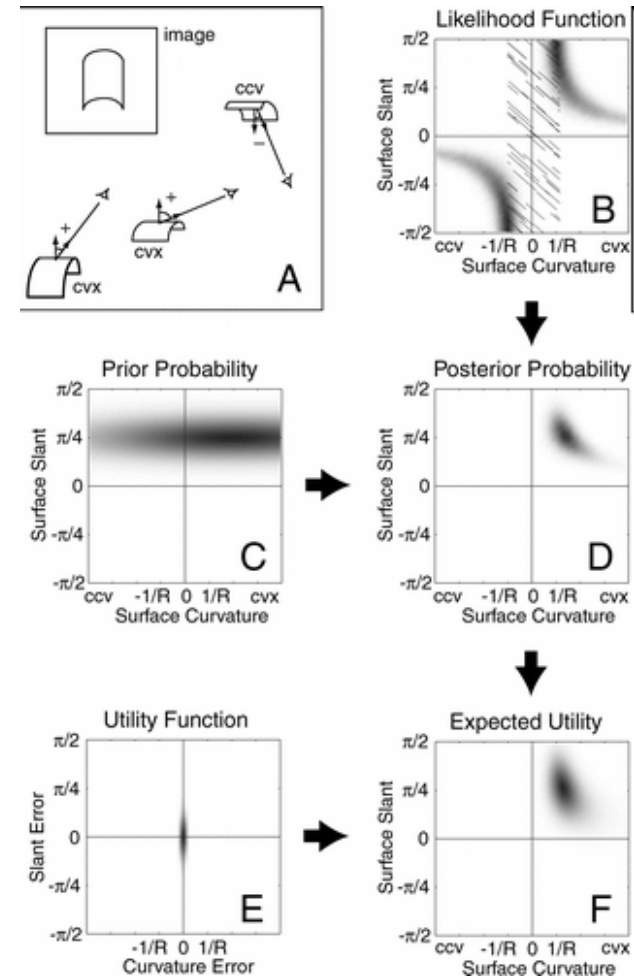
# What have we learned?

## • Dan Kersten - Shape from X

- we share the viewpoint that Bayesian integration as a fundamental way to think about sensor fusion [e.g., Yuille, Bülthoff, 1996, Kersten et al., 2004,...]

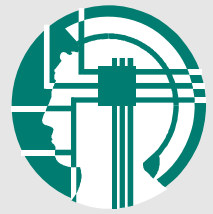


Kersten, Knill, Mamassian & I. Bülthoff  
Nature **379**, 31 (1996)



Kersten, Mamassian & Yuille, 2004

# What have we learned?

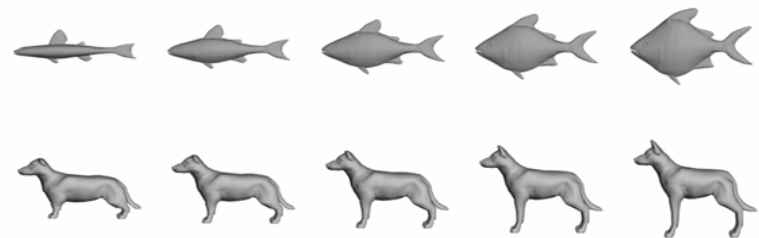
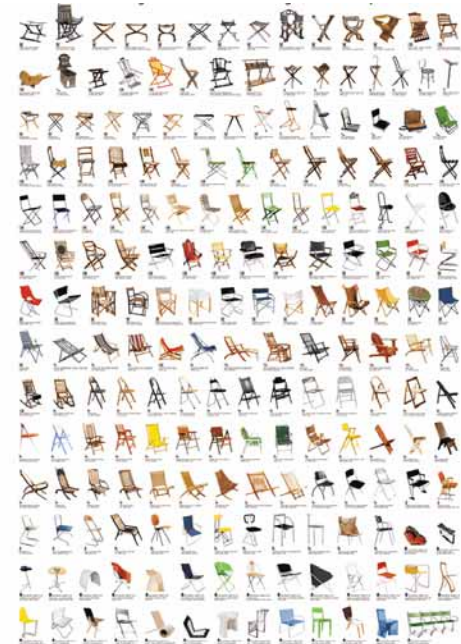


MPI for Biological Cybernetics

- **Irv Biederman** -  
challenging question:
  - how can we recognize something we have never seen before in a view-based framework?

**My answer:**

- view-based framework originally described only recognition at the subordinate/exemplar level
- difficult problem also for RBC
- more interesting is perhaps the question how to integrate recognition and categorization processes

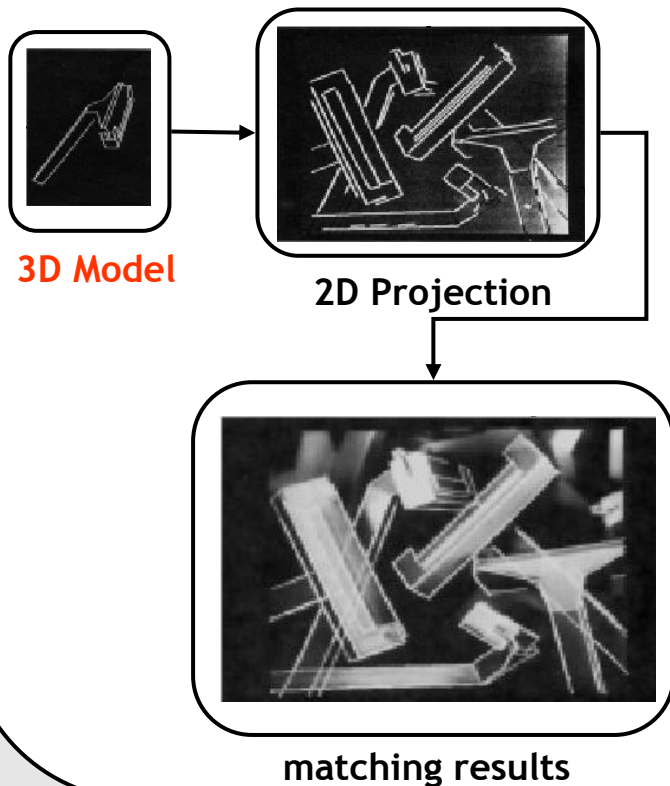


[Graf 2006]

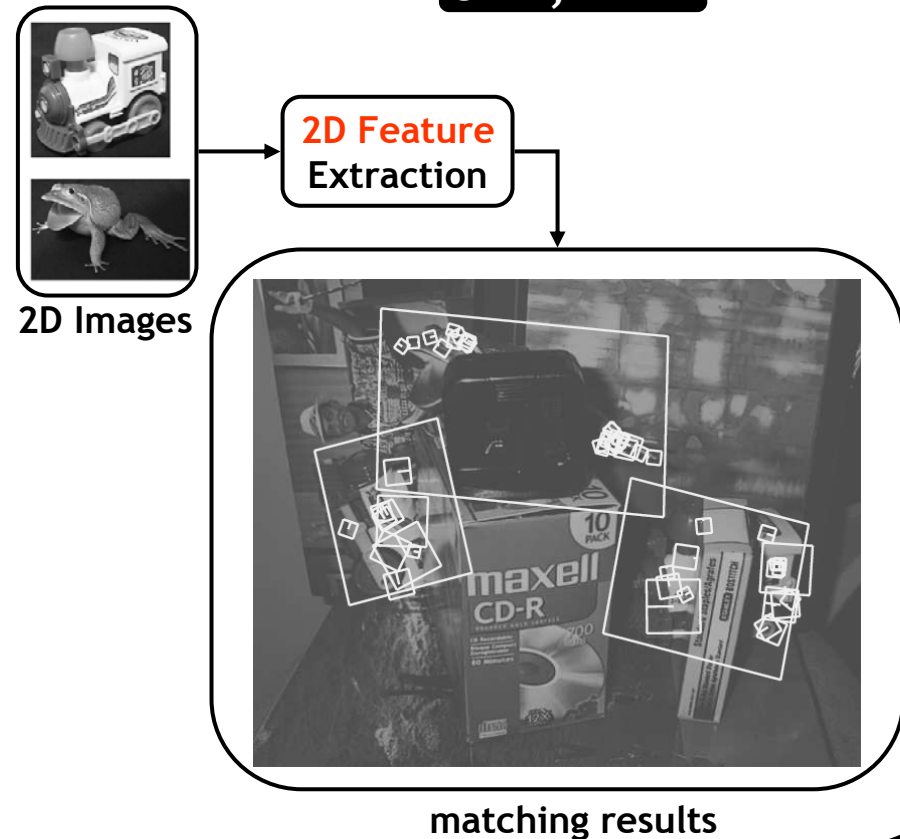
# What have we learned

- David Lowe - the fate of non-accidental features

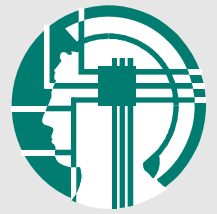
**SCERPO, ~1986**



**SIFT, 2004**



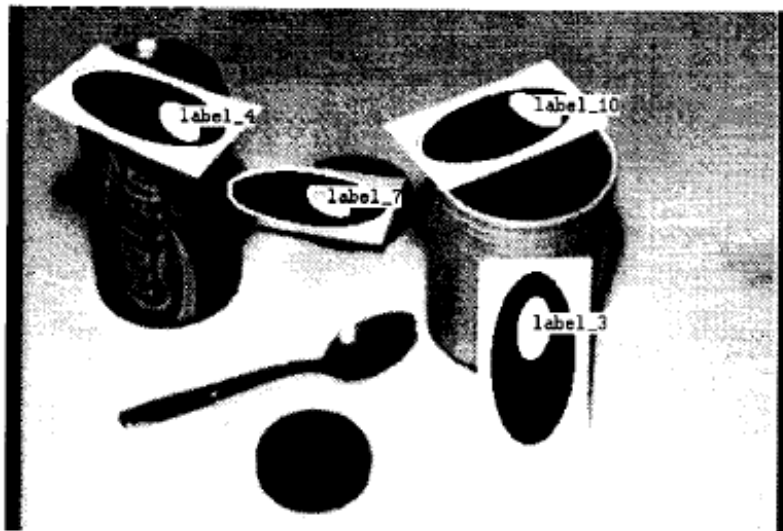
# What have we learned



MPI for Biological Cybernetics

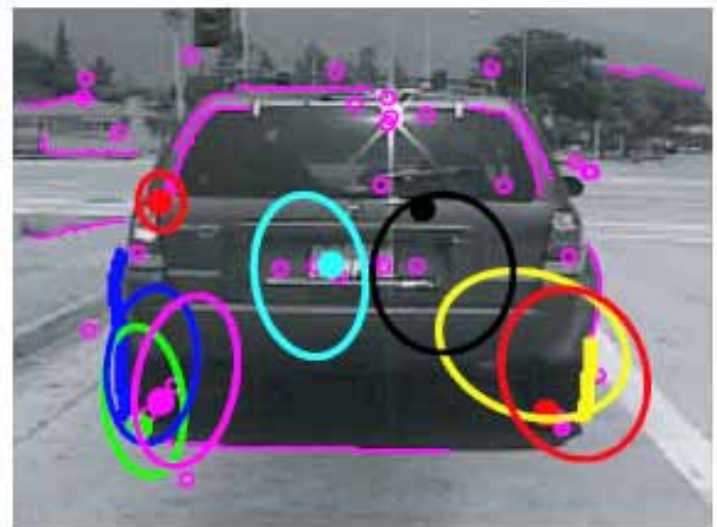
- Andrew Zisserman - the fate of geometry?

Forsyth et al., 1991



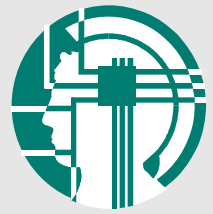
3D projective invariants

Fergus et al., 2005



2D appearance-based parts

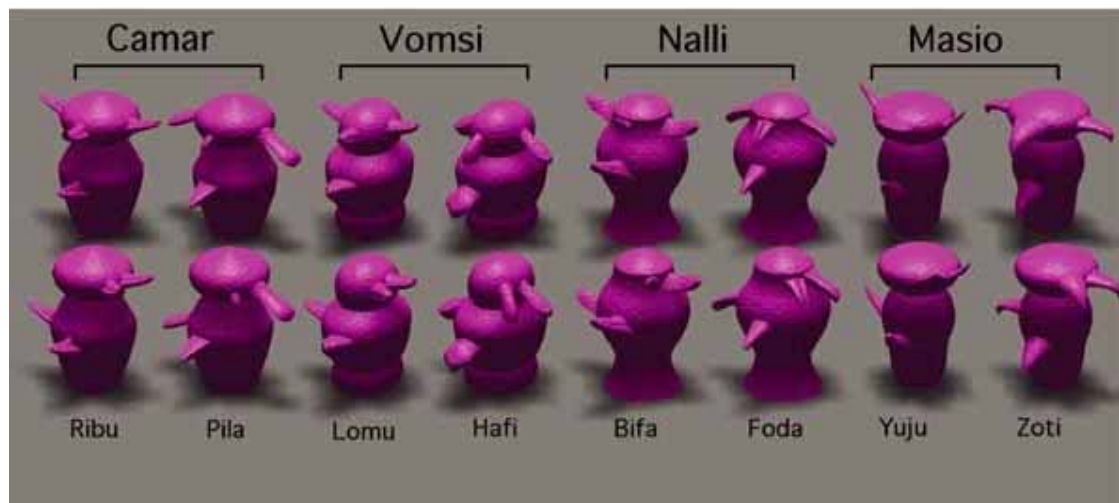
# What have we learned?



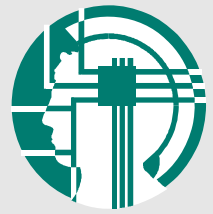
MPI for Biological Cybernetics

- **Mike Tarr**

- share very similar view [e.g., Tarr, Bülthoff, 1998]  
but Greebles are more fun than paperclips and amoebae
- the **role of expertise** in recognition
  - birds, cars, chicken, cows, dogs [Tarr, Cheng, 2003 - but, see also Tsao et al., 2006]

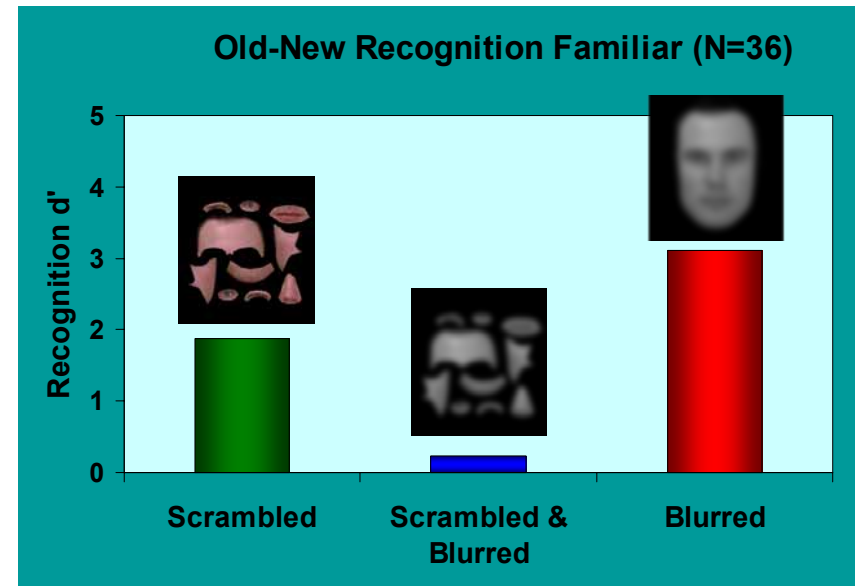


# What have we learned?



MPI for Biological Cybernetics

- **Mary Peterson** - holistic/analytic processes
  - inspired the perceptual work and computational modeling in our face recognition lab
  - the work stresses the importance of mid-level visual processes (figure ground segmentation, holistic representations)
  - need for integration of low-level and high-level processing

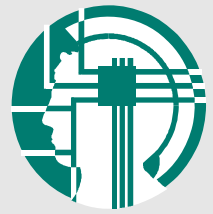


Schwaninger, Lobmaier & Collishaw, BMCV, 2002  
Wallraven, Schwaninger & Bülthoff, Network, 2006



# Current Work

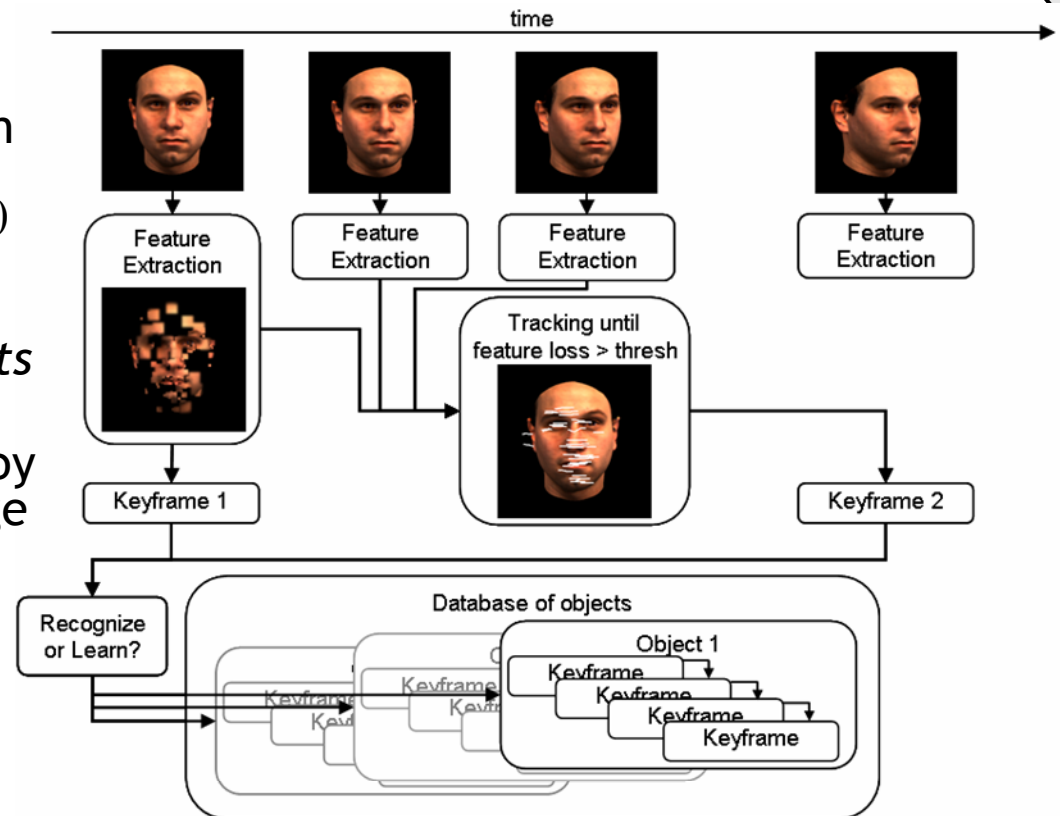
[Christian Wallraven, 2001-2005]



MPI for Biological Cybernetics

- Keyframe Model

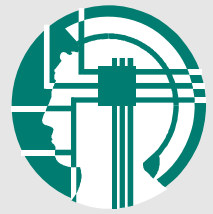
- motivated by temporal association advantage in face recognition (Wallis & Bülthoff, PNAS, 2001)
- image sequences are represented as trajectories of *fragments* between *keyframes*
- keyframes are defined by smoothly changing image information
- this model has been applied to:
  - face, car recognition
  - perceptual modeling
  - perceptual learning
  - machine learning



Wallraven & Bülthoff, CVPR, 2001

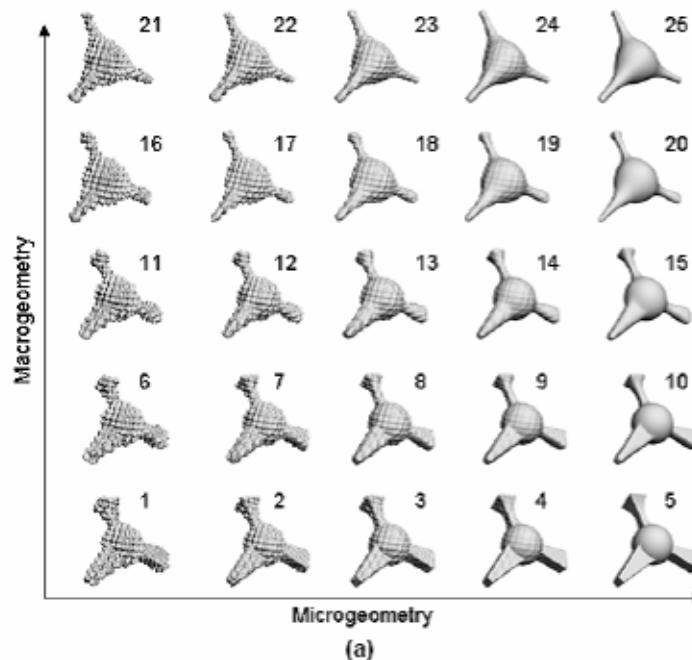
# Multi-modal/multi-sensory

[Theresa Cooke, 2003-2006]



MPI for Biological Cybernetics

- Multi-modal representations
  - controlled space of visual/haptic stimuli
  - micro/macro geometry
  - use MDS to find perceptual space for haptic, visual and bimodal exploration
  - compare similarity and categorization performance across modalities
  - we could show that representations are shared across modalities



(b)

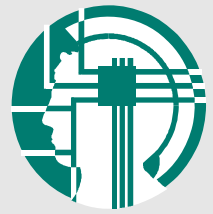
Figure 1: Variations of macrogeometry and microgeometry create a two-dimensional stimulus space (a). The stimuli are printed out into 3D objects (b).

**Validation by comparison**

Cooke, T., F. Jäkel, C. Wallraven and H. Bülthoff, *Neuropsychologia*, in press (2006)



# Current work

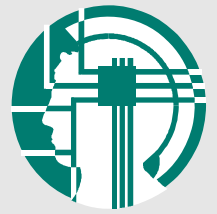


MPI for Biological Cybernetics

- Overall approach:
  - **perception and action in natural settings**
  - **life-like, interactive, multi-sensory environments**
  - **using Computer Graphics and Virtual Reality**
- Benefits:
  - not only more **informative about *real-world* behavior**, but also for developing **new tools for perception research**
    - MPI Face-Database (200 3D head models)
    - MPI Video-Database (40 videos of facial action units (FACS))
    - Morphable Face Models
    - Believable Face Models
    - FACS-driven Face Animation (Avatar)

# Face Database

faces.kyb.tuebingen.mpg.de  
[Troje, Bülthoff, 1996]



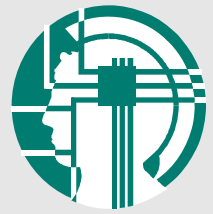
MPI for Biological Cybernetics



Morphing of Average Male Head into Average Female Head [Blanz, Vetter, 1999]

# Face Video Database

vdb.kyb.tuebingen.mpg.de



MPI for Biological Cybernetics

based on Ekman's action units



Action Unit 02 - Right Eyebrow



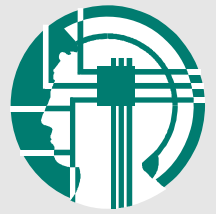
Action Unit 12 - Left Mouth Corner

*Wallraven, Cunningham, Breidt & Bülthoff: APGV - SIGGRAPH (2004)*

*Kleiner, Wallraven & Bülthoff: MPI-Technical Reports (123) (2004)*

# Synthetic Face Animations

[Curio, Breidt, Bülthoff, 2004]



MPI for Biological Cybernetics

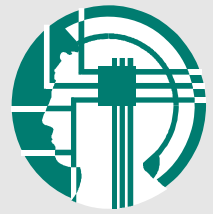
## **"Motion Capture Driven Morph Animation of Realistic 3D Head Models"**

Cristobal Curio, Martin Breidt, Heinrich H. Buelthoff  
Max Planck Institute for Biological Cybernetics

Submitted to Siggraph 2004

# Virtual Human

[Breidt, Wallraven, Cunningham & Bülthoff, 2003]



MPI for Biological Cybernetics

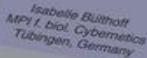
## Animation Test

### "Facial Animation Based on 3D Scans and Motion Capture"

M. Breidt, C. Wallraven, D. W. Cunningham, H. H. Bülthoff

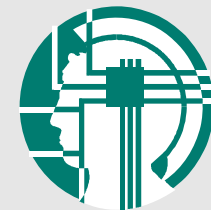
Submitted to SIGGRAPH 2003

- The "Chair" challenge





# Challenges

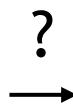


MPI for Biological Cybernetics

- The "Art" challenge: build a computer vision system that learns to interpret art images
  - such a system would need to deal with abstraction

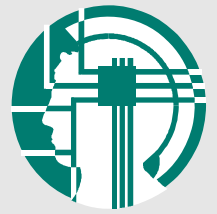


© Robert Pepperell



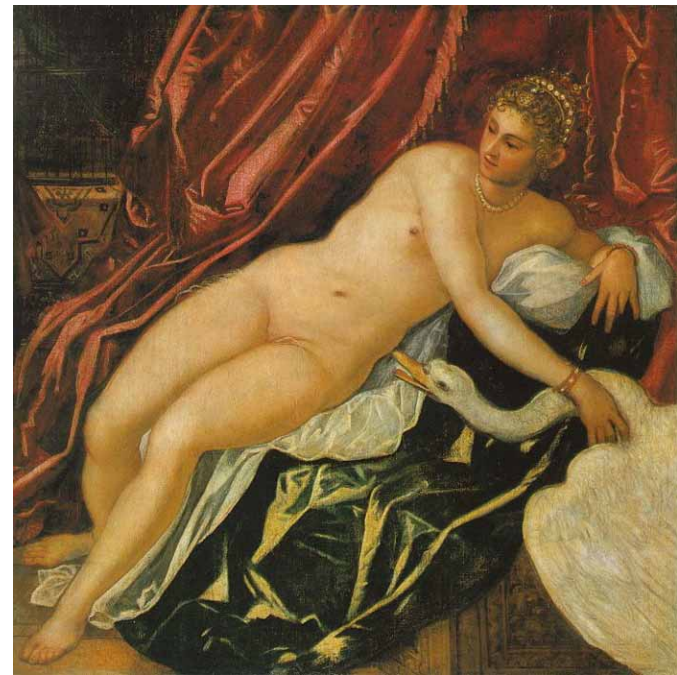
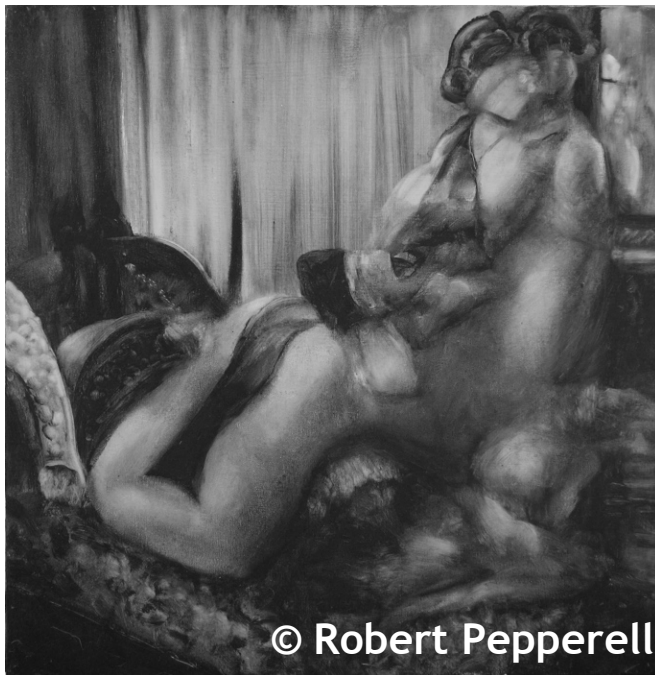
Images © by Robert Pepperell, collaboration with MPI

# Challenges



MPI for Biological Cybernetics

- The "Art" challenge: build a computer vision system that learns to interpret art images
  - such a system would need to deal with abstraction



Images © by Robert Pepperell, collaboration with MPI



# Challenges

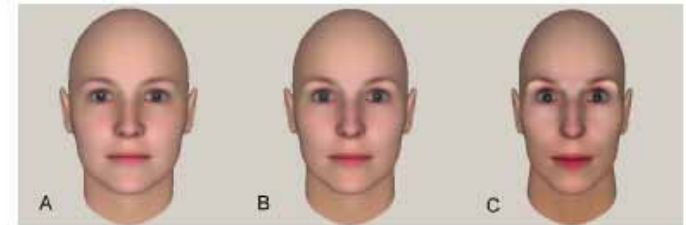
- The "Pawan Sinha" challenge
  - build a computer vision system that integrates the 20 results every CV researcher should know about face recognition [http://web.mit.edu/bcs/sinha/papers/20Results\\_2005.pdf](http://web.mit.edu/bcs/sinha/papers/20Results_2005.pdf)



eyebrows as important features

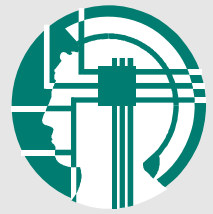


recognition under distortions

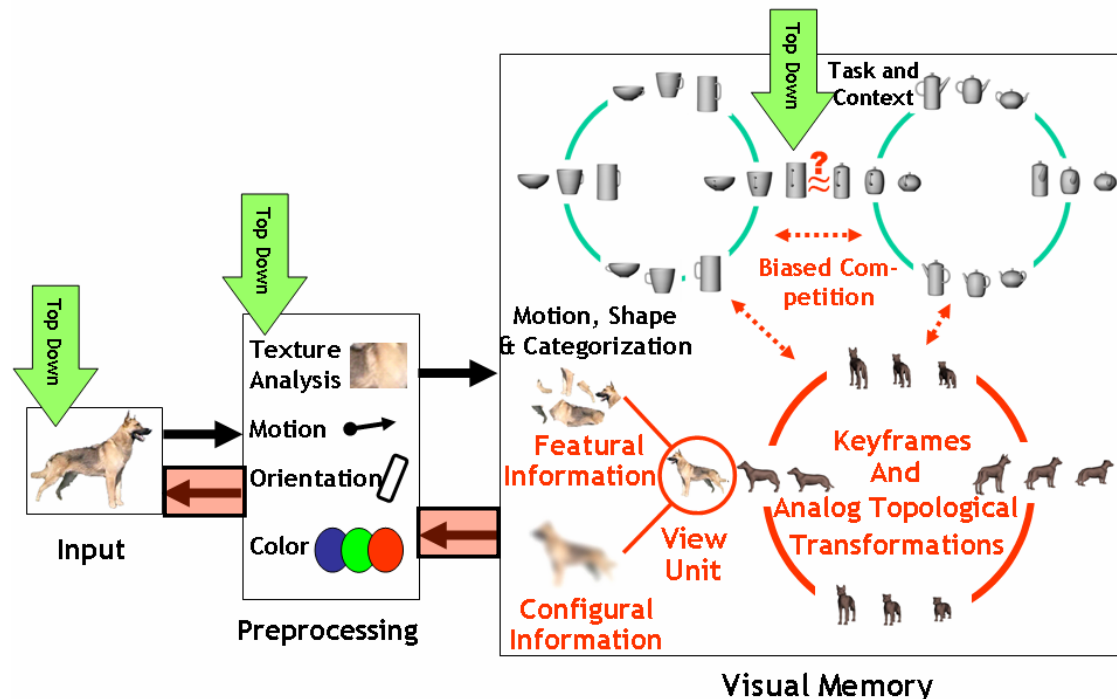


caricature effect for recognition

# Integrative Model?



MPI for Biological Cybernetics



- How can we build an integrated hierarchy of processes going from scene gist to object recognition, and **back** to feature selection?  
[Graf, Schwaninger, Wallraven, Bülthoff, 2002]

- Thank you for your attention!