Lecture 31
High-Level Vision: Overview
ECEN 5283 Computer Vision

Dr. Guoliang Fan
School of Electrical and Computer Engineering
Oklahoma State University

Goals

- To introduce some basic issues about high-level vision.
- To present the major topics in this chapter.

Pathways in Human Vision System

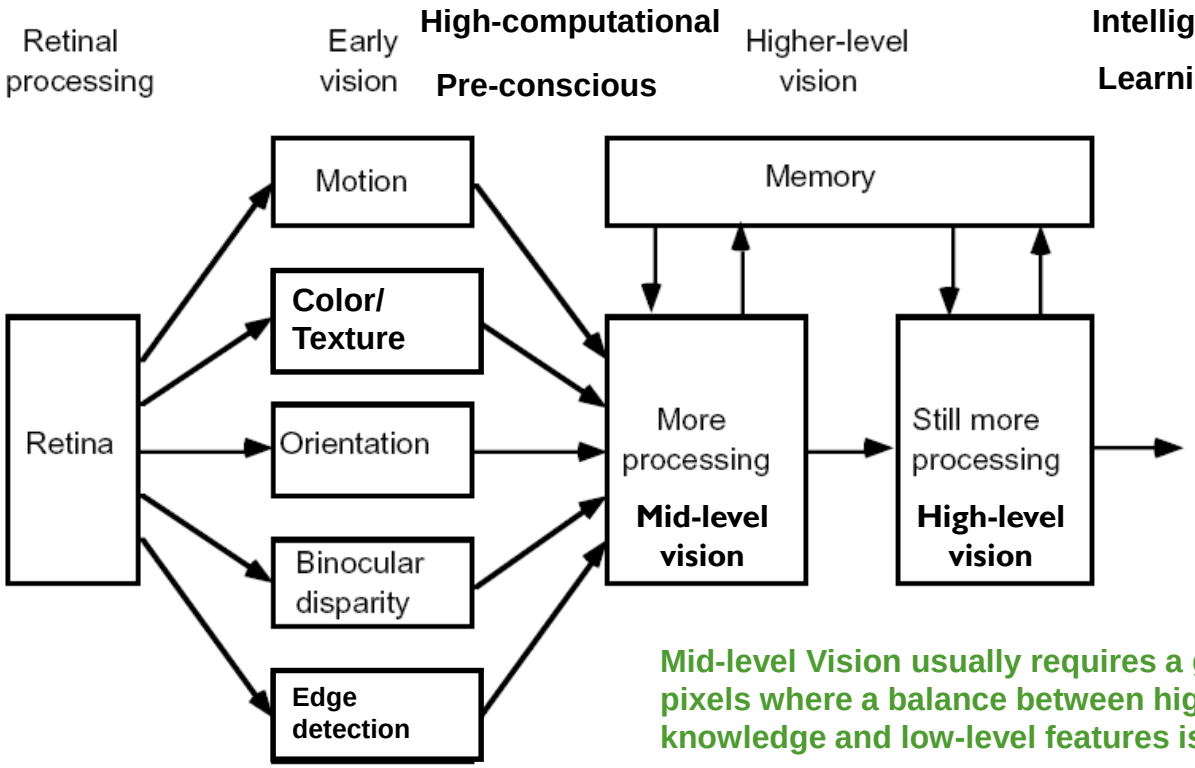


Fig. 1.1
A generic diagram for visual processing. In this approach, early vision consists of a set of parallel pathways, each analyzing some particular aspect of the visual stimulus.

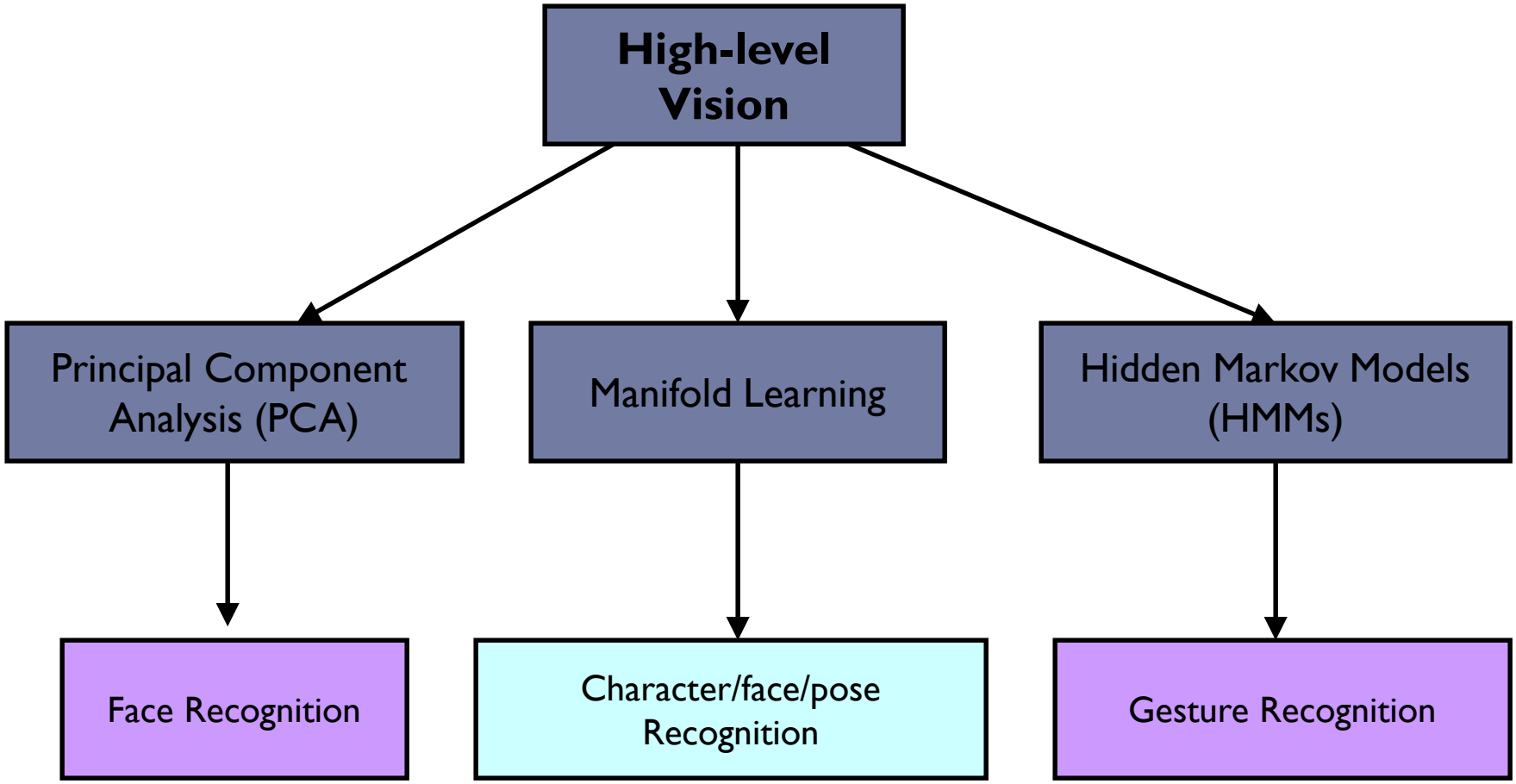
The Plenoptic Function and the Elements of Early Vision web.mit.edu/persci/people/adelson/pub_pdfs/elements91.pdf

What is high-level vision?

- High-level vision is the final stage of visual perception that deals with the interpretation and use of what is seen in the image or from a scene.

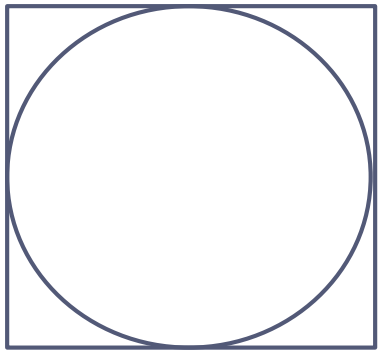
	Low-level vision	Mid-level vision	High-level vision
Purpose	To extract a set of visual primitives for further processing at the pixel-level.	To bridge low-level vision and high-level vision with some inference of the structure by involving multiple pixels.	To interpret or understand the content of an image, such as object recognition.
Computational model	Pixel-level linear filtering (convolution)	Region-level statistical Inference (clustering and state estimation)	Inference based on region-based features or visual cues from the whole image
High-level knowledge involved	It is a pre-conscious process.	It is a process that requires some intelligence.	It is a process that requires more intelligence and prior knowledge.
Bottom-up and Top-down flow	It is mainly a data-driven bottom-up process.	Both data-driven bottom-up and knowledge-driven top-down are involved.	The knowledge-driven top-down plays a more important role than the data-driven bottom-up flow.

Overview of High-level Vision



Curse of High Dimensionality

- In recent years, the effect of the curse of high dimensionality has been studied in great detail on several problems such as clustering, nearest neighbor search, and indexing.
  - Data become sparse;
  - The data mass is in the corner of hypercube;
  - Distance or similarity is no longer meaningful;
  - The concept of proximity is no longer valid;

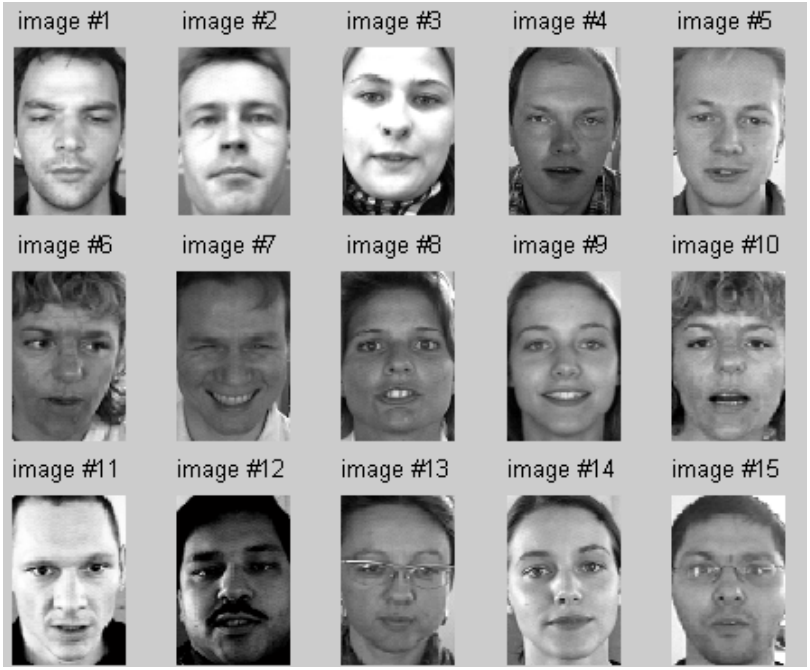


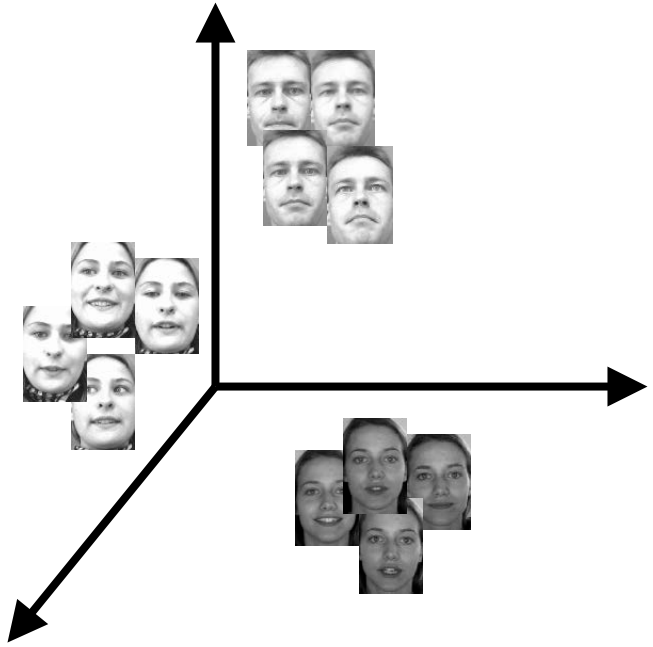
$$\lim_{d\to\infty} \frac{\pi^{d/2}}{d\,2^{d-1}\Gamma(d/2)} = 0 \Leftarrow \lim_{d\to\infty} \frac{\text{volume of hypersphere}}{\text{volume of hypecube}}$$

- Distance-based clustering and classification become a problem in a high-dimensional space, and dimensionality reduction (DR) is needed which supports low-dimensional analysis.

Dimensionality Reduction (DR): Why?

- DR is the process of reducing the number of random variables under consideration, so that the intrinsic data structure are preserved for efficient and effective data analysis.

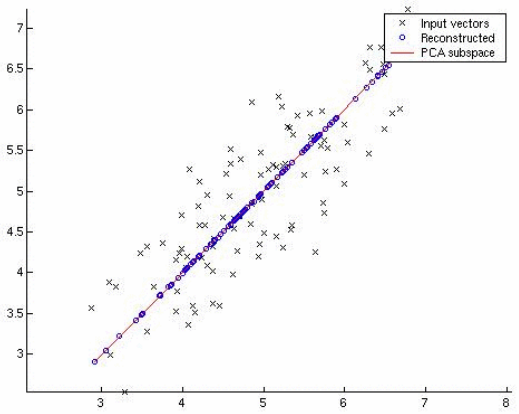




DR: How?

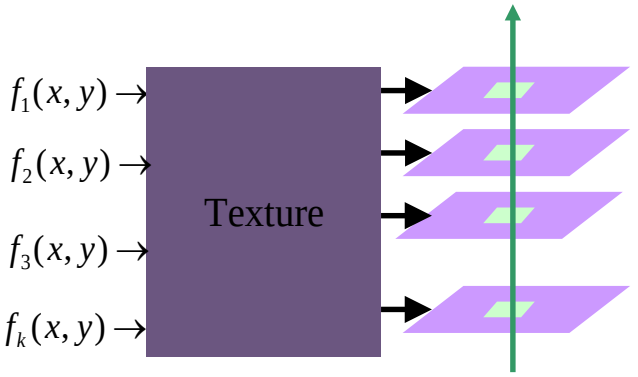
- The DR techniques can be divided into feature selection and feature extraction
  - Feature selection is the DR technique, commonly used in machine learning, of selecting a subset of relevant features for building robust learning models.

    Reconstruction is usually possible.



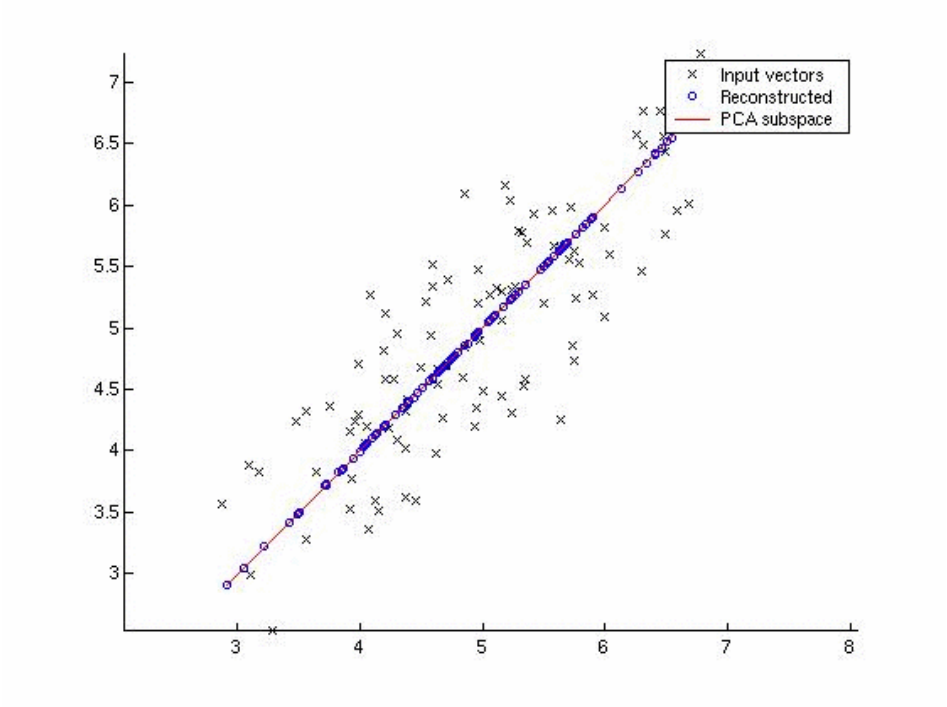
  - Feature extraction transforms the input data that could be notoriously redundant into the set of compact features.

    Reconstruction is usually NOT possible.



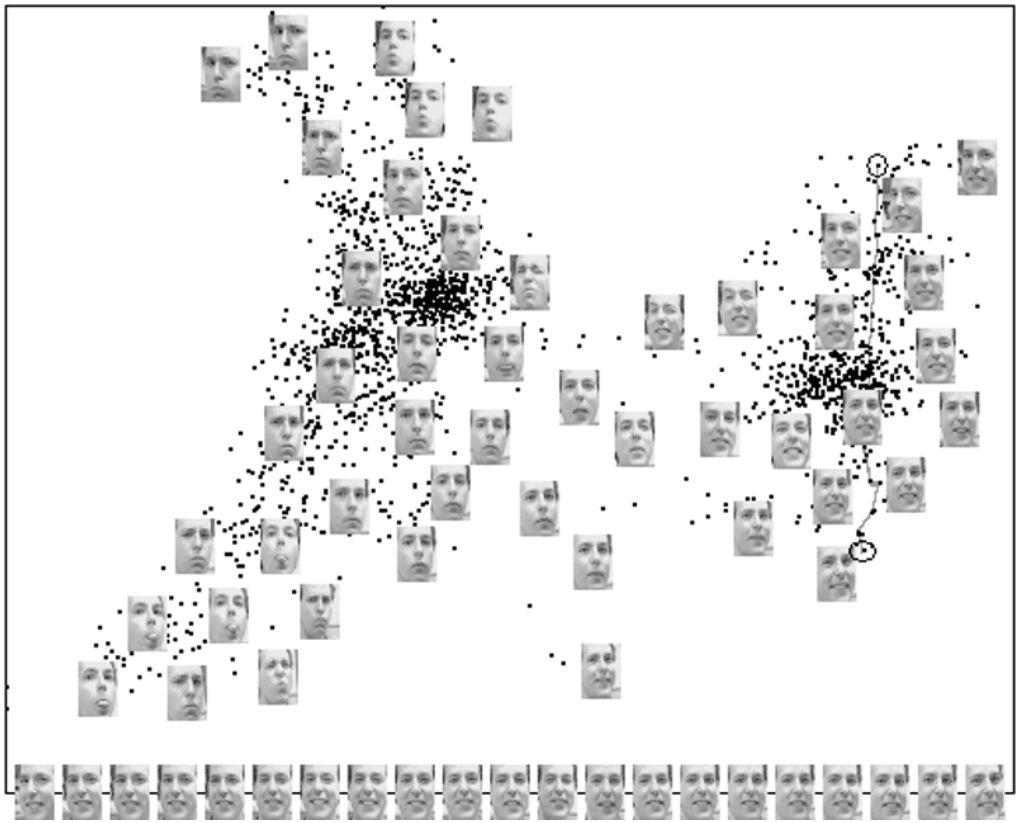
Principal Component Analysis (PCA)

- **PCA** *provides compact data representation*
  - We can construct a lower dimensional linear subspace that "best explains" the variation of these data points from their mean.
  - All data will be represented in this low-dimension feature space.



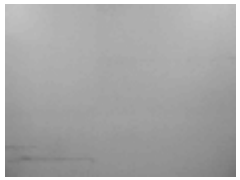
Non-linear Dimensional Reduction (Manifold Learning)

- Manifold learning is the process of exploring a low-dimensional non-linear embedding underlying a set of high-dimensional data.



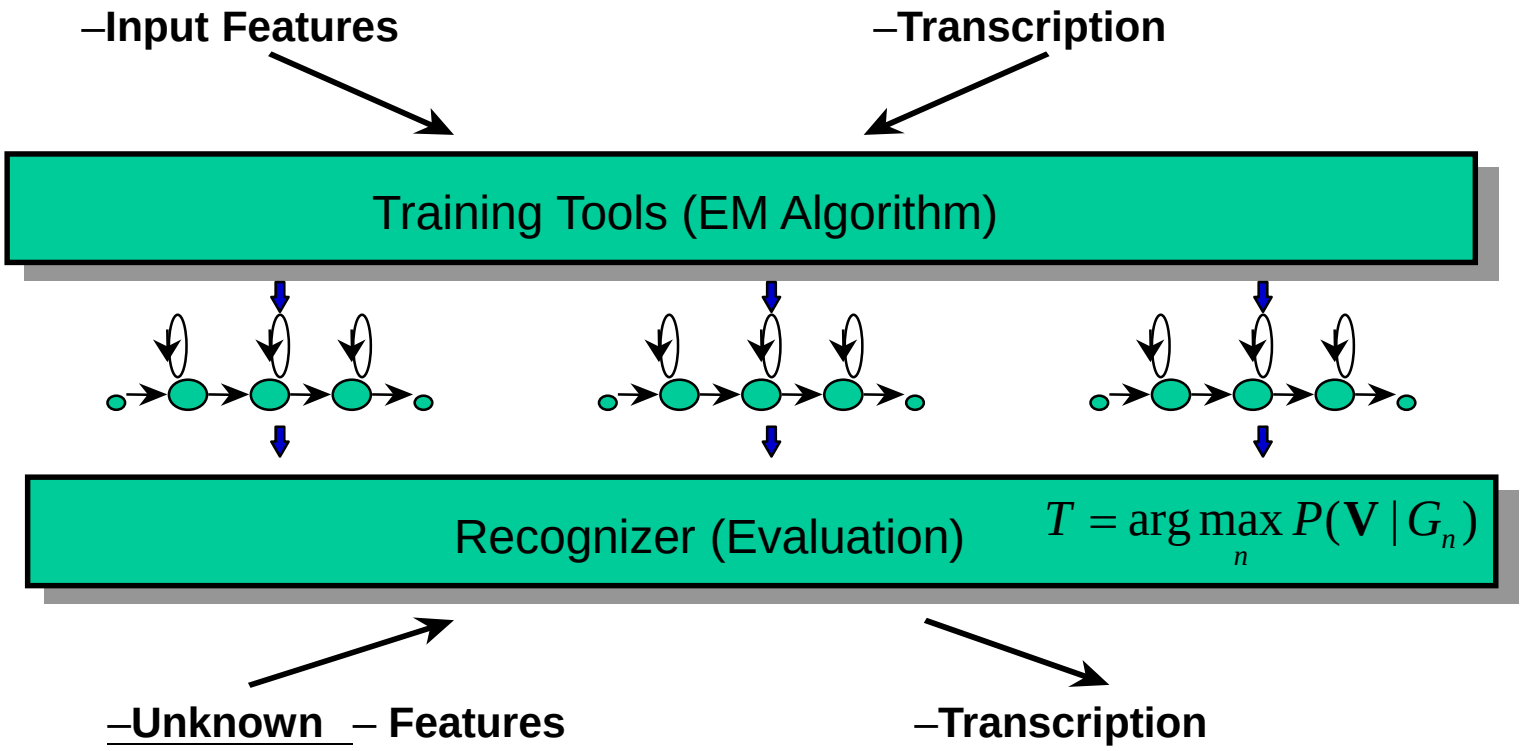
http://sciencewatch.com/ana/st/face/09mayFacRecHe/

Why do we need HMMs?

 Hand-Waving-Both	 Hand-Waving-Right	 Hand-Waving-Left	 Top-Right	 To-Left
 To-Top	 To-Bottom	 Round-Clockwise	 Round-Counterclockwise	 Stop
 Come	 Nod-Yes	 Nod-No	 Clapping	 Kotow
 Spin	 Go-Left	 Go-Right	 Turn-Right	 Turn-Left

http://www.fb9-ti.uni-duisburg.de/projekte/video/video_e.html

HMM Training and Recognition



- Andreas Kosmala and Gerhard Rigoll. On-Line Handwritten Formula Recognition Using Statistical Methods. In *Int. Conference on Pattern Recognition (ICPR)*, pages 1306-1308, Brisbane, August 1998.
- Gerhard Rigoll, Andreas Kosmala, and Stefan Eickeler. High Performance Real-Time Gesture Recognition Using Hidden Markov Models. In *Gesture Workshop*, pages 69-80, Bielefeld, Germany, September 1997.