Analysis of Local Appearance-based Face Recognition on FRGC 2.0 Database

HAZIM KEMAL EKENEL

Interactive Systems Labs (ISL), Universität Karlsruhe (TH), Germany

e- mail: ekenel@ira.uka.de

web: http://isl.ira.uka.de/~ekenel/
http://isl.ira.uka.de/face_recognition/

Face Recognition Grand Challenge Workshop
March 22, 2006
Arlington, VA, USA

Outline

- Interactive Systems Laboratories (ISL)
- Local Appearance-based Face Recognition
- Feature Sets
- Feature Normalization
- Experiments
- Conclusions

Interactive Systems Labs (ISL)

- Founded by Prof. Dr. Alexander Waibel at 1991.
- Research on:
  - translation, speech, language, vision technologies, multimodal, and cross-modal perceptual interfaces, smart rooms
- Located at Carnegie Mellon University, USA and University of Karlsruhe (TH), Germany.

ISL Vision Group

- Directed by Assist. Prof. Rainer Stiefelhagen
- Focus on visual perception of people in smart rooms
- Areas of interest
  - Person Identification
  - Person Tracking
  - Head Pose / Focus of Attention Estimation
  - Activity Analysis

ISL Face Recognition Group

- Objective:
  - Face recognition for smart environments.
  - Developing and deploying fully automatic face recognition systems with the research focus on to build simple, fast & robust face recognition algorithms.
- Implementation areas:
  - Recognizing individuals entering a room with a zoom camera
  - Recognizing individuals in a room using fixed camera
  - Human Robot Interaction –Humanoid Robots

ISL Door Face Database

- Ten thousands pictures of more than 100 individuals have been collected during 86 recording days (Feb. 2005, August-Dec. 2005)
- ~30000 images of 30 individuals will become public.

Face Recognition @ CLEAR Evals (FG2006)

- CHIL Project (EU IP FP6)
  http://chil.server.de/
- **NIST**
- **Goal:** To recognize the lecturer/seminar participant by using video and multi-view data acquired by four cameras mounted to the corners of the room.
- **Problem Conditions:** Low resolution faces with improper lighting, especially because of the projector's beam.

Face Recognition for Humanoid Robots

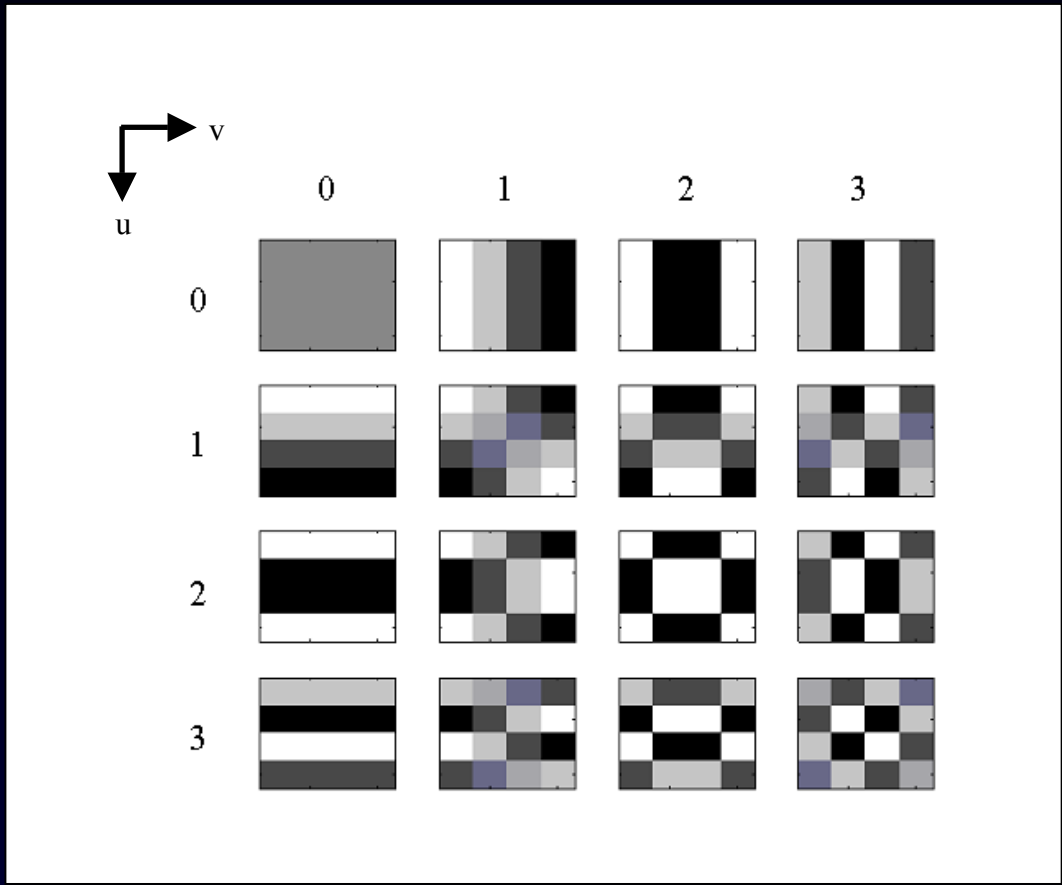
Local Appearance-based Face Recognition

- Merits:
  - Robust against local variations
  - Facilitates weighting/selection of the “important“ local regions for face recognition
- Previous approaches:
  - Salient region based
    - Modular Eigenfaces (Pentland et al., 1994)
    - FR with SVMs: Global vs. Component-based Approach (Heisele et al. 2003)
    - FR using Component-based DCT/LDA (Lee et al. 2005, MPEG7)
  - **Generic**
    - Modular PCA (Gottumukkal & Asari, 2004)

Face Representation

- Local appearance modelling: 8x8 pixels blocks
  - Less sensitiveness to illumination and local variations
- Data-independent basis (Discrete Cosine Transform)
  - Fast
  - Compact representation
- Feature/Decision fusion to perform classification

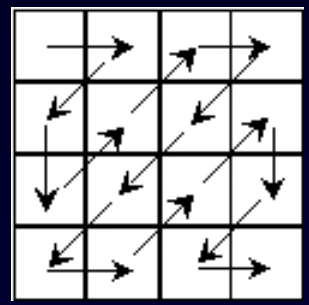
DCT Basis



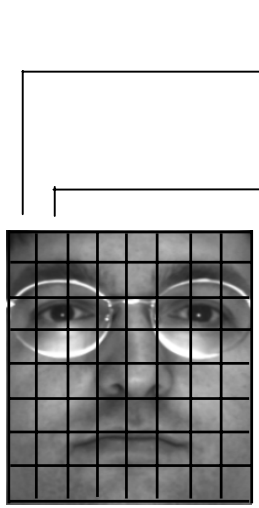
Sample DCT Output & Scan Pattern

415	60	59	26	61	48	10	11
146	9	69	54	8	5	4	5
77	70	28	4	13	1	3	1
33	11	19	9	8	0	2	3
14	21	11	5	4	6	4	2
1	10	0	15	8	1	4	2
1	3	1	4	2	5	4	2
0	2	1	0	2	1	1	1

Zig-Zag Scan



Feature Extraction



Feature Selection

- Selecting the first M DCT coefficients **(DCT-all)**
- Removing the first coefficient, and selecting the first M DCT coefficients from the remaining ones **(DCT-0)**
- Removing the first three coefficients, and selecting the first M DCT coefficients from the remaining ones **(DCT-3)**

Feature Normalization

- The blocks with different brightness levels lead to DCT coefficients with different value levels.

  → normalize the local feature vector's, *f*'s, magnitude to unit norm:

$$f_n = f / \|f\|$$

- The first DCT coefficients have higher magnitudes than the later ones, thus having more influence on the classification results.

  → divide each coefficient to its standard deviation that is learned from the training set:

$$f_{n,i} = f_i / \sigma(f_i)$$

Face Recognition Experiments

- Experimental Data derived from FRGC ver. 2 Experiments 1 & 4
- Individuals that have at least 10 images (target, query) are selected
- 120 individuals (10 images for training & testing)
- Controlled vs. Controlled (Fall 2003 recordings for training & Spring 2004 recordings for testing)
- Uncontrolled vs. Uncontrolled (Fall 2003 recordings for training & Spring 2004 recordings for testing)

Distance Metrics

L1:
$$d = \sum_{m=1}^{M} \left| f_{training,m} - f_{test,m} \right|$$

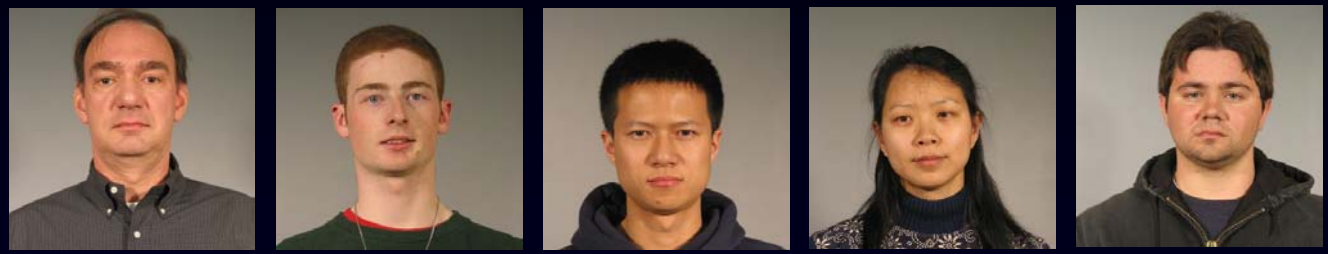
L2:
$$d = \left(\sum_{m=1}^{M} \left| f_{training,m} - f_{test,m} \right|^2 \right)^{1/2}$$

cor:
$$d = \frac{f_{training} * f_{test}}{\left\| f_{training} \right\| * \left\| f_{test} \right\|}$$

cov:
$$d = \frac{\left(f_{training} - m_{training} \right) * \left(f_{test} - m_{test} \right)}{\left\| f_{training} - m_{training} \right\| * \left\| f_{test} - m_{test} \right\|}$$

Sample Images

Controlled Samples:



Uncontrolled Samples:

FRGC Controlled vs. Controlled -DCT & PCA scores @ 320 –no normalization

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	74.8%	**94.3%**	92.8%	89.2%	88.1%
L2	62.2%	86.7%	82.2%	81.8%	85.8%
Cos	78.8%	87.4%	86.3%	80.8%	88.6%
Cov	79.0%	87.8%	86.1%	80.6%	88.6%

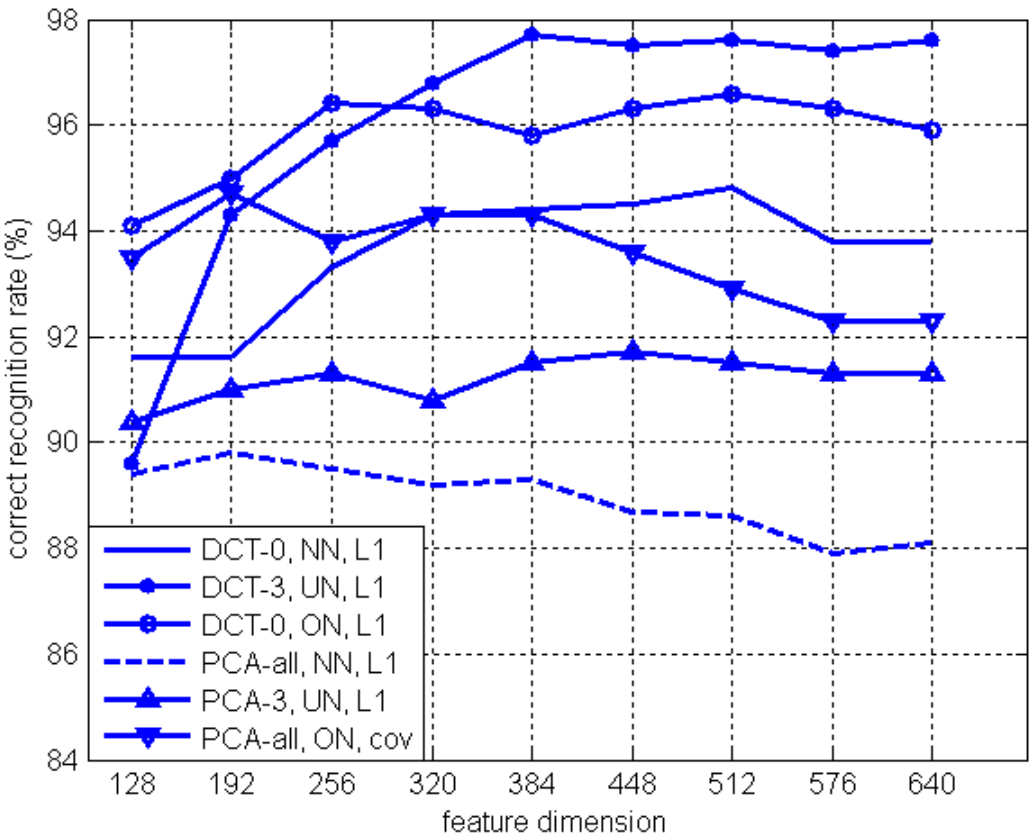
FRGC Controlled vs. Controlled -DCT & PCA scores @ 320 –unit norm

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	90.6%	**96.8%**	**96.8%**	87.3%	90.8%
L2	79.9%	93.6%	94.3%	81.0%	89.0%
Cos	79.9%	93.6%	94.3%	80.8%	88.6%
Cov	80.0%	93.6%	94.4%	80.6%	88.6%

FRGC Controlled vs. Controlled -DCT & PCA scores @ 320 –over dim. norm

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	91.3%	**96.3%**	95.7%	80.8%	79.9%
L2	89.4%	93.1%	91.2%	80.9%	79.6%
Cos	92.7%	93.8%	94.4%	94.2%	94.0%
Cov	93.2%	93.9%	94.4%	94.3%	94.0%

FRGC Controlled vs. Controlled Overview



FRGC Uncontrolled vs. Uncontrolled -DCT & PCA scores @ 320 –no normalization

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	43.6%	61.4%	60.6%	49.0%	44.1%
L2	36.9%	56.1%	55.8%	40.3%	39.1%
Cos	39.4%	69.3%	65.8%	38.4%	37.8%
Cov	39.3%	**69.6%**	66.1%	38.4%	37.8%

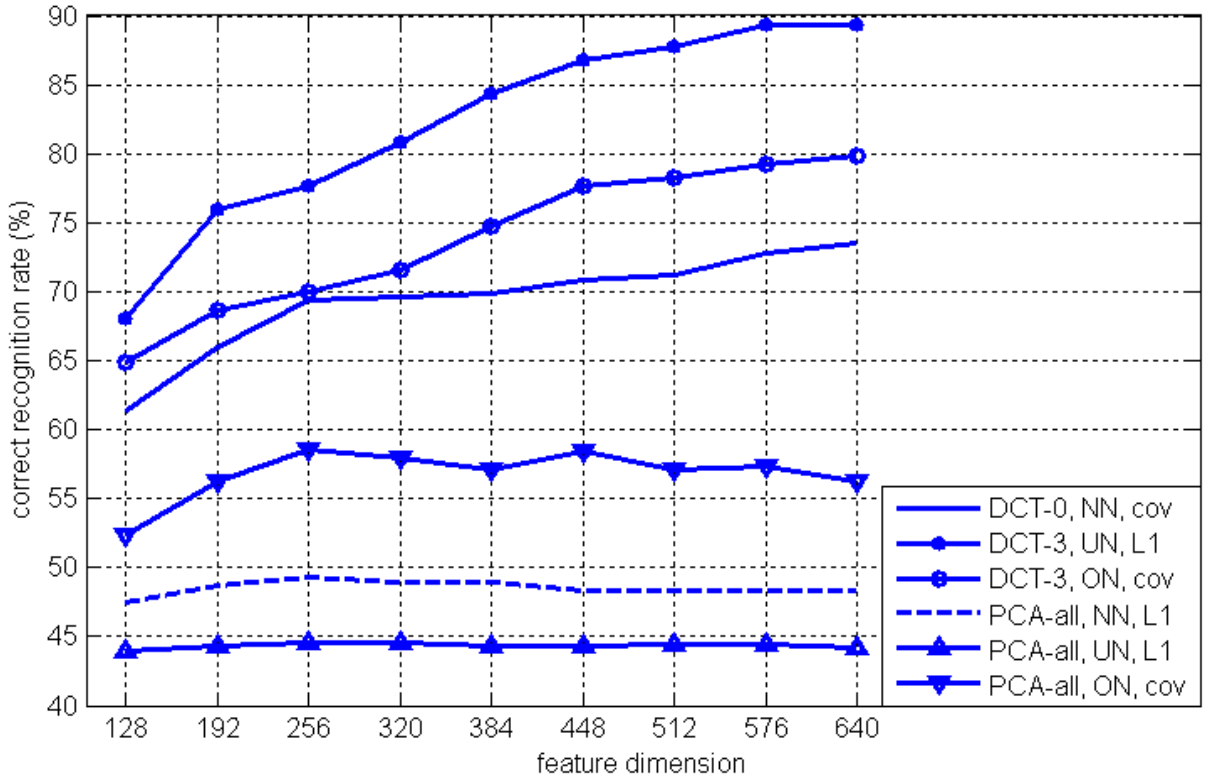
FRGC Uncontrolled vs. Uncontrolled -DCT & PCA scores @ 320 –unit norm

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	70.5%	80.5%	**80.8%**	44.6%	43.5%
L2	63.2%	75.3%	76.8%	38.8%	38.6%
Cos	63.2%	75.3%	76.8%	38.4%	37.8%
Cov	63.6%	75.3%	76.7%	38.4%	37.8%

FRGC Uncontrolled vs. Uncontrolled -DCT & PCA scores @ 320 –over dim. norm

	DCT	DCT – w/o DC	DCT – w/o 3	PCA	PCA - w/o 3
L1	53.3%	63.1%	58.9%	46.5%	45.7%
L2	49.8%	57.8%	56.5%	45.4%	45.1%
Cos	51.2%	67.6%	71.4%	57.4%	57.4%
Cov	50.9%	68.3%	**71.6%**	58.0%	57.8%

FRGC Uncontrolled vs. Uncontrolled Overview



Conclusions

- Using proper local features / normalizing local features contributes face recognition performance (Similar results have been also obtained of the AR and CMU PIE face databases)

- Unit norm DCT-0 / DCT-3 local features perform best

Controlled vs. Controlled @ feature dimension of 320	Uncontrolled vs. Uncontrolled @ feature dimension of 320
PCA, 89.2%	PCA, 49%
DCT ver. 1, 94.3%	DCT ver. 1, 69.3%
DCT ver. 2, 96.8%	DCT ver. 2, 80.5%

Questions?