User-Independent Gaze Estimation by Exploiting Similarity Measures in the Eye Pair Appearance Eigenspace

Nanxiang Li and Carlos Busso

Multimodal Signal Processing (MSP) Laboratory

Department of Electrical Engineering,

The University of Texas at Dallas,

Richardson TX-75080, USA

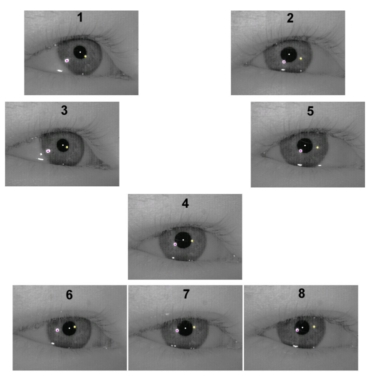
nxl056000@utdallas.edu, busso@utdallas.edu

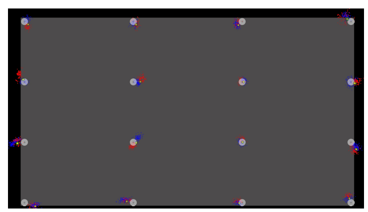
Motivation

- Gaze indicates visual awareness
  - Human-computer interaction (HCI)
  - Market analysis
  - Social behavior
  - Driver distraction

Common Approach

- User and setting specific parameters are acquired by calibration
  - Repetitive and tedious calibration process
  - Constraints imposed on the user





Field of Interest

IR light

Our Vision

- User-independent gaze estimation that do not require calibration, or put any constraints on the users
  - Appearance-based approach (simple system setting)
  - Focus on eye pair image (implied head pose)
  - Large database (capture variance among the users)

MSP-GAZE Database

It considers:

- Individual differences
  - 46 subjects
  - Gender balanced
  - Diverse ethnic groups (Caucasian - 16, Asian - 10, Indian - 10, and Hispanic - 10)
- Head movement
  - Balanced between with and without head movement
- Inter-session variability
  - Two sessions for each subject on two different days
- Distance between user and screen
  - User defined
  - Near - 0.4m
  - Medium - 0.5m
  - Far - 0.6m

MSP-GAZE Database

System setup

MSP-GAZE Database

Recording	Head Movement	Distance	Pattern
1	Yes	User-defined	Testing
2	Yes	User-defined	Training
3	Yes	Near	Training
4,5	Yes	Medium	Training
6,7	Yes	Far	Training
8	No	User-defined	Testing
9	No	User-defined	Training
10	No	Near	Training
11,12	No	Medium	Training
13,14	No	Far	Training

Highlights of the Study

- **Appearance based model**
  - Eigenspace representation (PCA)
- **Reduce the gap between user-dependent (UD) and user-independent (UI) gaze estimation**
  - Similar frame
  - Similar subject

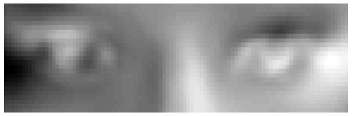
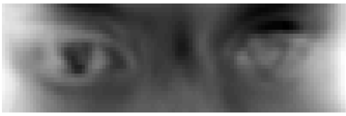
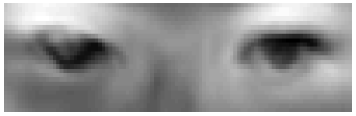
Appearance Based Model

- Eigenspace representation (PCA)
  - PCA construct the orthogonal basis - eigenvectors to the larger eigenvalues of the covariance matrix Σ.

$$\Sigma=\sum_{i=1}^{N}\Phi_i^T\Phi_i$$

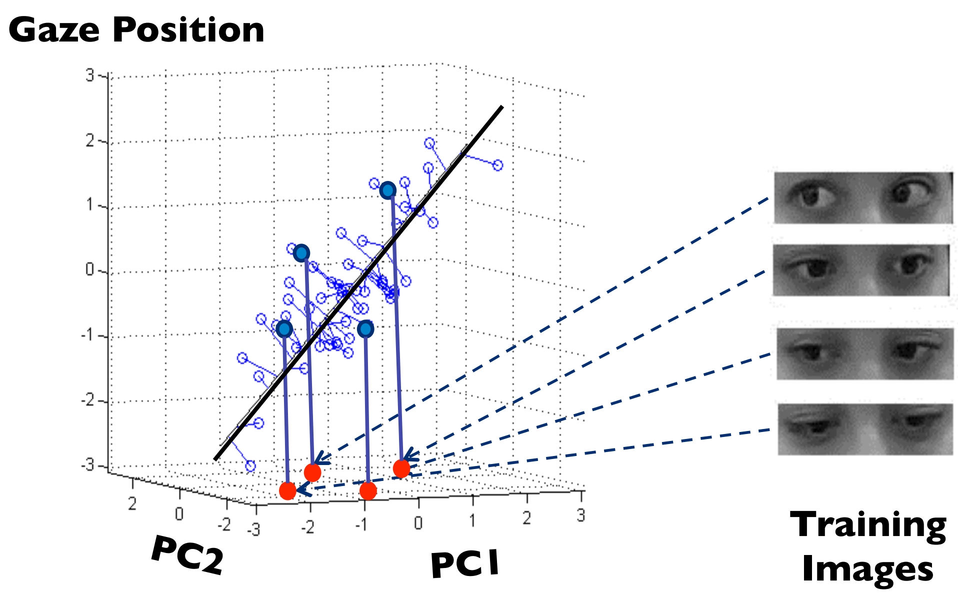
Φ_i - mean removed vectorized eye pair images

  - These bases capture the main variance among the eye appearance
  - We use the projections to these bases as independent variables to build regression model for gaze estimation

Appearance Based Model

- PCA and Regression model for gaze estimation



y = β{ β p2 + β + β p 30 + β p2 @ 1280, ..., β > 1080 @ 10

Difference between UD and UI

- User Dependent Results

	Without head motion			With head motion		
Distance	ρ_x	ρ_y	Θ_{error}	ρ_x	ρ_y	Θ_{error}
Near	0.90	0.85	4.7	0.91	0.84	4.5
Medium	0.89	0.84	3.8	0.91	0.83	3.9
Far	0.88	0.83	3.5	0.90	0.83	3.4
User-Defined	0.89	0.82	3.9	0.88	0.82	3.9

2.0

2.1

- User Independent Results

	Without head motion			With head motion		
Distance	ρ_x	ρ_y	Θ_{error}	ρ_x	ρ_y	Θ_{error}
Near	0.85	0.76	7.0	0.87	0.75	6.8
Medium	0.86	0.75	6.0	0.85	0.74	5.9
Far	0.85	0.68	5.3	0.85	0.73	5.2
User-Defined	0.85	0.78	5.9	0.86	0.70	6.0

ρ_x - x correlation; ρ_y - y correlation; Θ_{error} – angular error

Difference between UD and UI

Challenges in UI gaze estimation

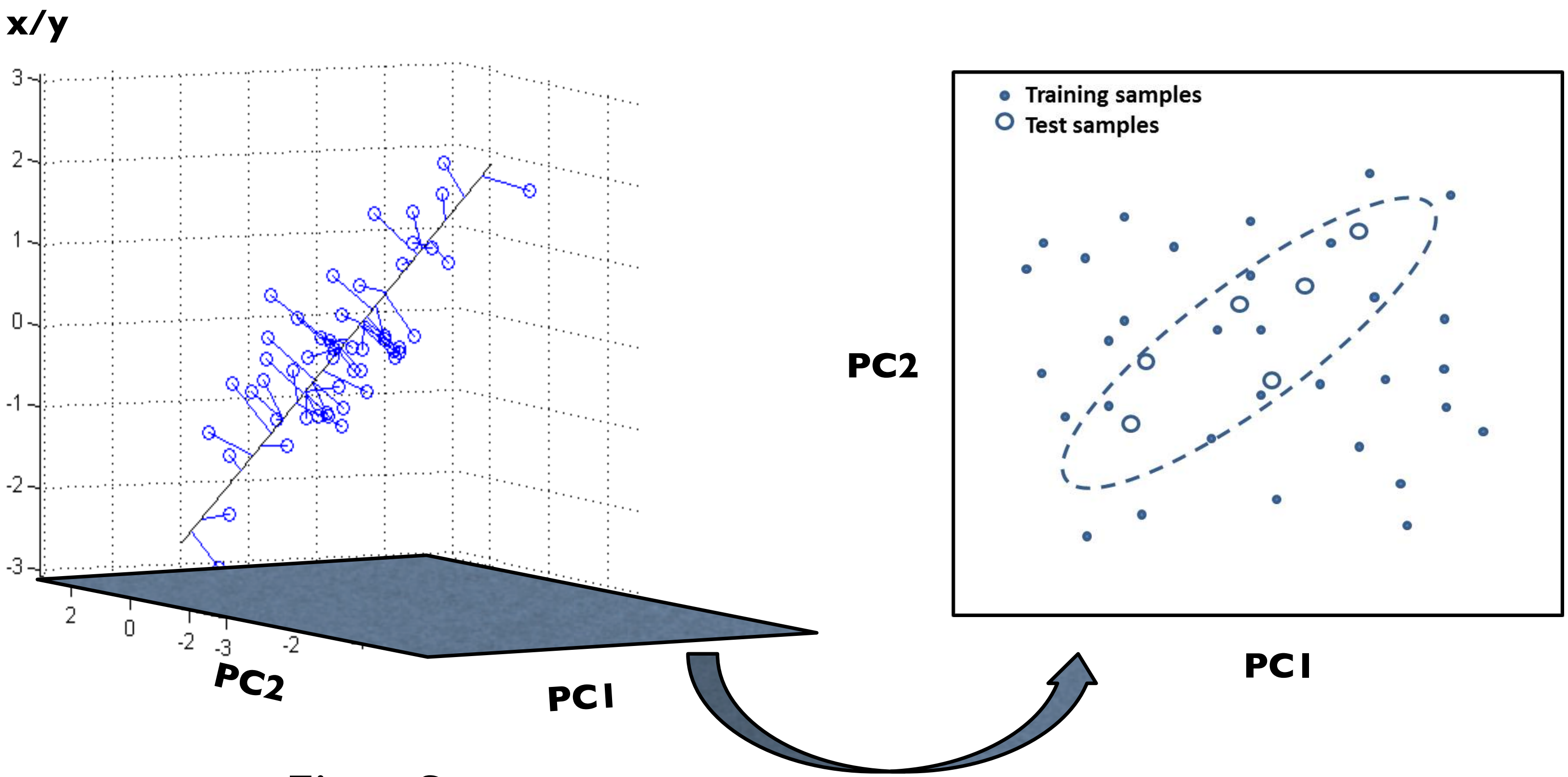
- Individual differences affect the orthogonal basis to represent the target subject

Based on the premise that "*not all data is good data*", we use similarity measures in the eigenspace to find "good data"

- Similar frames
- Similar subjects

$$\Sigma_{All} = \sum_{i=1}^{N} \Phi_i^T \Phi_i \rightarrow \Sigma_S = \sum_{i \in S} \Phi_i^T \Phi_i$$

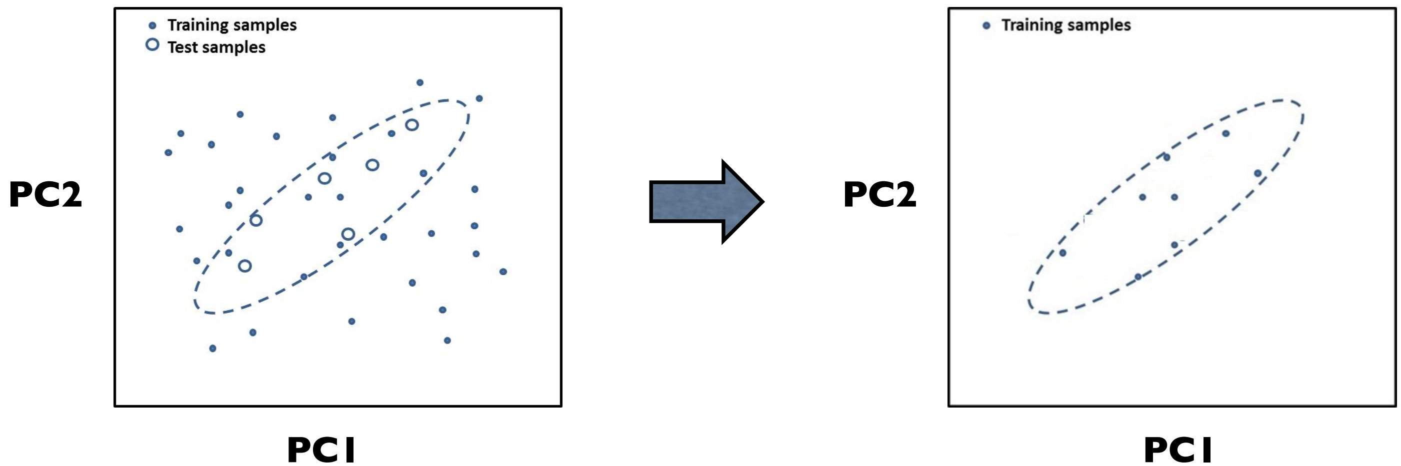
Finding Similar Data



Eigen Space

Finding Similar Data

- Similar frames



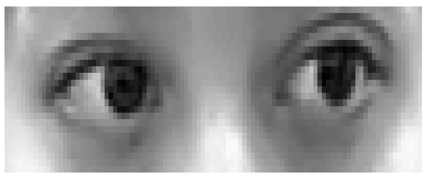
- Similar subjects
  - The most frequently appeared subject in the similar frames

Finding Similar Data

Test image

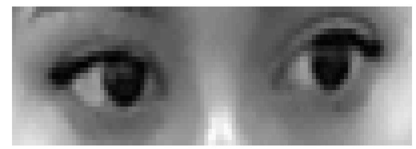
Identified Similar frames

Sub 36



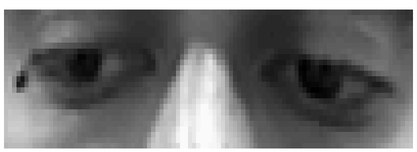
Gaze positions (68, 160)

Sub 44



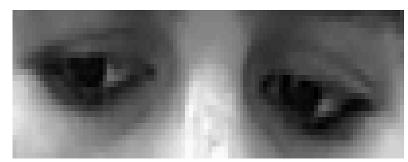
Gaze positions (75, 99)

Sub 1



Gaze positions (1600, 967)

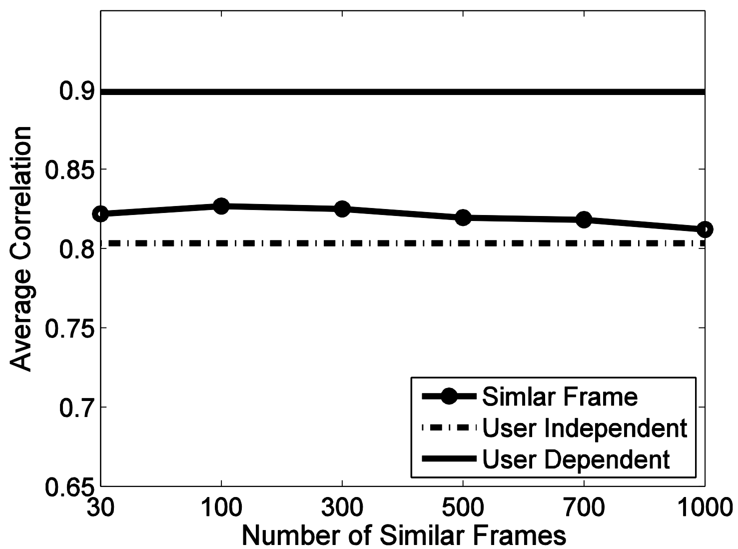
Sub 24



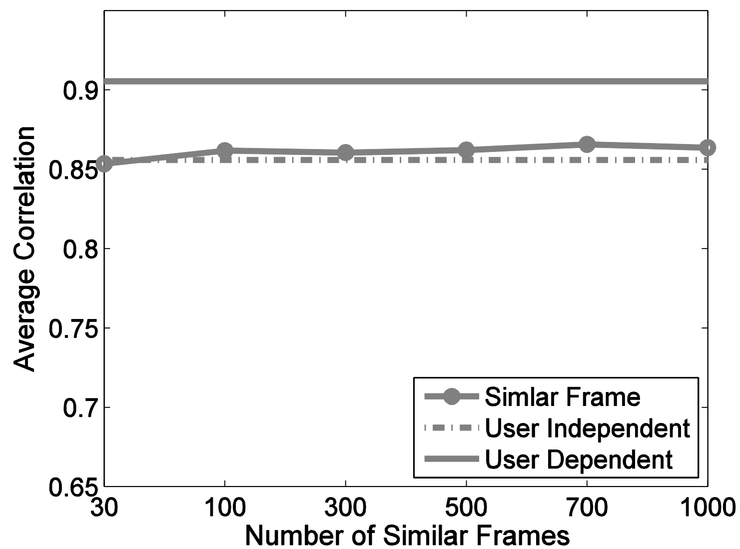
Gaze positions (1526, 931)

Experiment results

- Similar frames



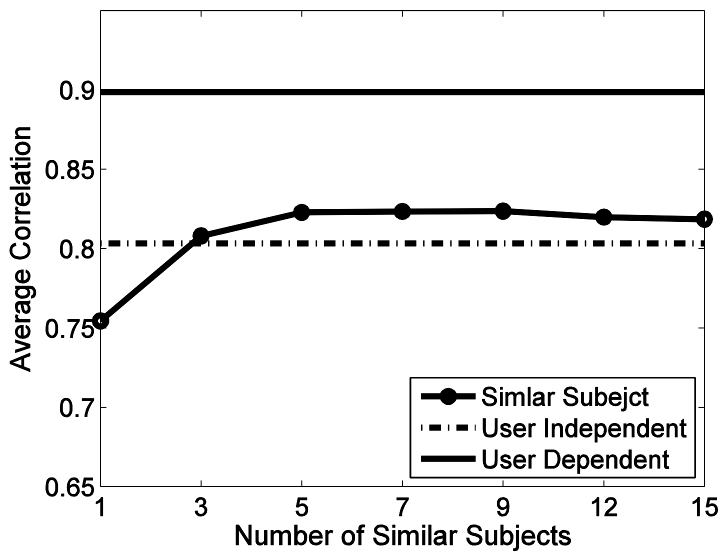
With Head Movement



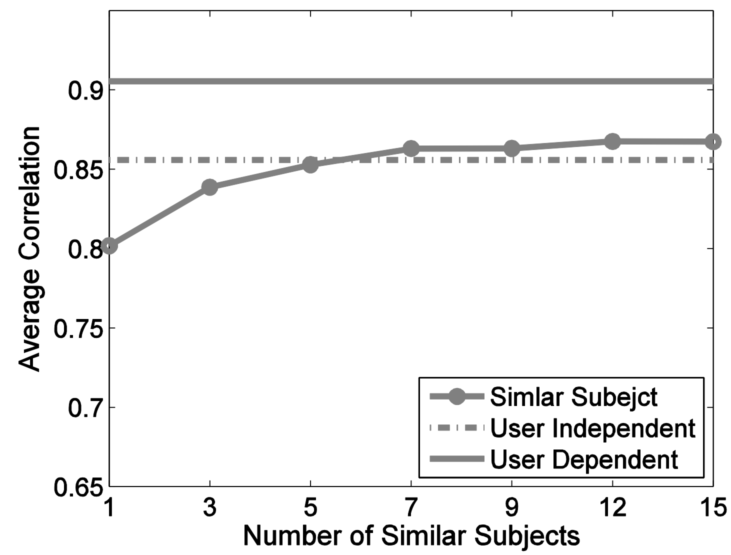
Without Head Movement

Experiment results

- Similar subjects



With Head Movement



Without Head Movement

Experiment results

		Without head movement		With head movement	
	Distance	ρ_x	ρ_y	ρ_x	ρ_y
7 Similar subjects	**Far**	0.90	0.79	0.88	0.72
	Medium	0.93	0.82	0.90	0.74
	Near	0.93	0.87	0.93	0.78
	User-Defined	0.91	0.78	0.88	0.74
100 Similar frames	**Far**	0.90	0.81	0.88	0.72
	Medium	0.92	0.80	0.90	0.74
	Near	0.93	0.82	0.93	0.77
	User-Defined	0.92	0.81	0.90	0.76

Conclusion

- We identified similar training data samples in the eigenspace for user-independent gaze estimation
- We used the reduced training set to build the eye appearance eigenspace
- By user-independent model, the proposed approach eliminates the requirement for calibration
- For real application, start with user-independent models, as more test samples become available, create and update the reduced set S

We would like to thank our sponsor SAMSUNG

Thank you!

Questions?