Unconstrained Face Recognition and Analysis

S. Kevin Zhou

Siemens Corporate Research, Inc.

kzhou@scr.siemens.com

Roadmap to Unconstrained Face Recognition and Analysis

- Introduction

- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

Roadmap to Unconstrained Face Recognition and Analysis

- Introduction

- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

Why Face Recognition and Analysis?

- Application.
  - Non-intrusive biometric.
  - Homeland security, law enforcement, surveillance.
  - Virtual reality, HCI, multimedia.

- Theory.
  - Interdisciplinary: Image/video processing, mathematics, physics, vision, statistics and learning, psychophysics, neuroscience, etc.

SIEMENS

State-Of-The-Art

- Current FR systems work well ONLY under controlled situations.
  - Neutral expression, no makeup (Intrinsic).
  - Frontal illumination, frontal view (Extrinsic).
  - Mugshot of good quality.
- Apply pattern recognition techs. to face image.
  - Appearance-based: Subspace methods
    - PCA [Turk & Pentland, 91], LDA [Belhumeur et al., 97].
    - Local feature analysis (LFA) [Penev & Atick 96], ICA
    - Neural network, evolutionary computing, genetic algorithm
  - Feature-based:
    - Elastic graph matching [Lades et al., ’93].

Unconstrained Face Recognition and Analysis

- Motivation: deal with unconstrained conditions
  - Intrinsic variations: expression, makeup, aging.
  - Extrinsic variations: illumination and pose.
  - Surveillance video.
  - Age-related: Aging process, age estimation.
  - Expression and animation.
- Feasible approaches
  - Combine pattern recognition with variation modeling
  - Face modeling and animation
  - Utilized video characteristics
  - Statistical learning

Roadmap to Unconstrained Face Recognition and Analysis

- Introduction
- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

* **S. Zhou**, R. Chellappa, and D. Jacobs, *"Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints," European Conf. on Computer Vision, 2004.*

Illumination affects appearance

* Courtesy of Prof. David Jacobs.

Approach

- Generalized photometric stereo.
  - Describes all possible human face images under all possible illumination conditions.
  - Combines a physical illumination model with statistical regularity in the human class.
  - Derive an illumination-invariant signature for robust face recognition under illumination variation.

Key Derivations of Generalized Photometric Stereo

$$\mathbf{h}_{d\times n} = \mathbf{T}\mathbf{s} = (f_1\mathbf{T}_1 + f_2\mathbf{T}_2 + \ldots + f_m\mathbf{T}_m)\mathbf{s}$$

$$= [\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_m](\mathbf{f} \otimes \mathbf{s})$$

$$= \mathbf{W}_{d\times 3m}(\mathbf{f}_{m\times 1} \otimes \mathbf{s}_{3\times 1})$$

Statistical regularity in identity

Lambertian illumination model

FR Across Illumination: Recognition Results

Training set	Yale	Yale ($m=10$)	Vetter ($m=100$)
Method	Eigenface	Generalized Photometric Stereo	Generalized Photometric Stereo
Average Recognition Rate	35%	**67%**	**93%**

SIEMENS

Roadmap to Unconstrained Face Recognition and Analysis

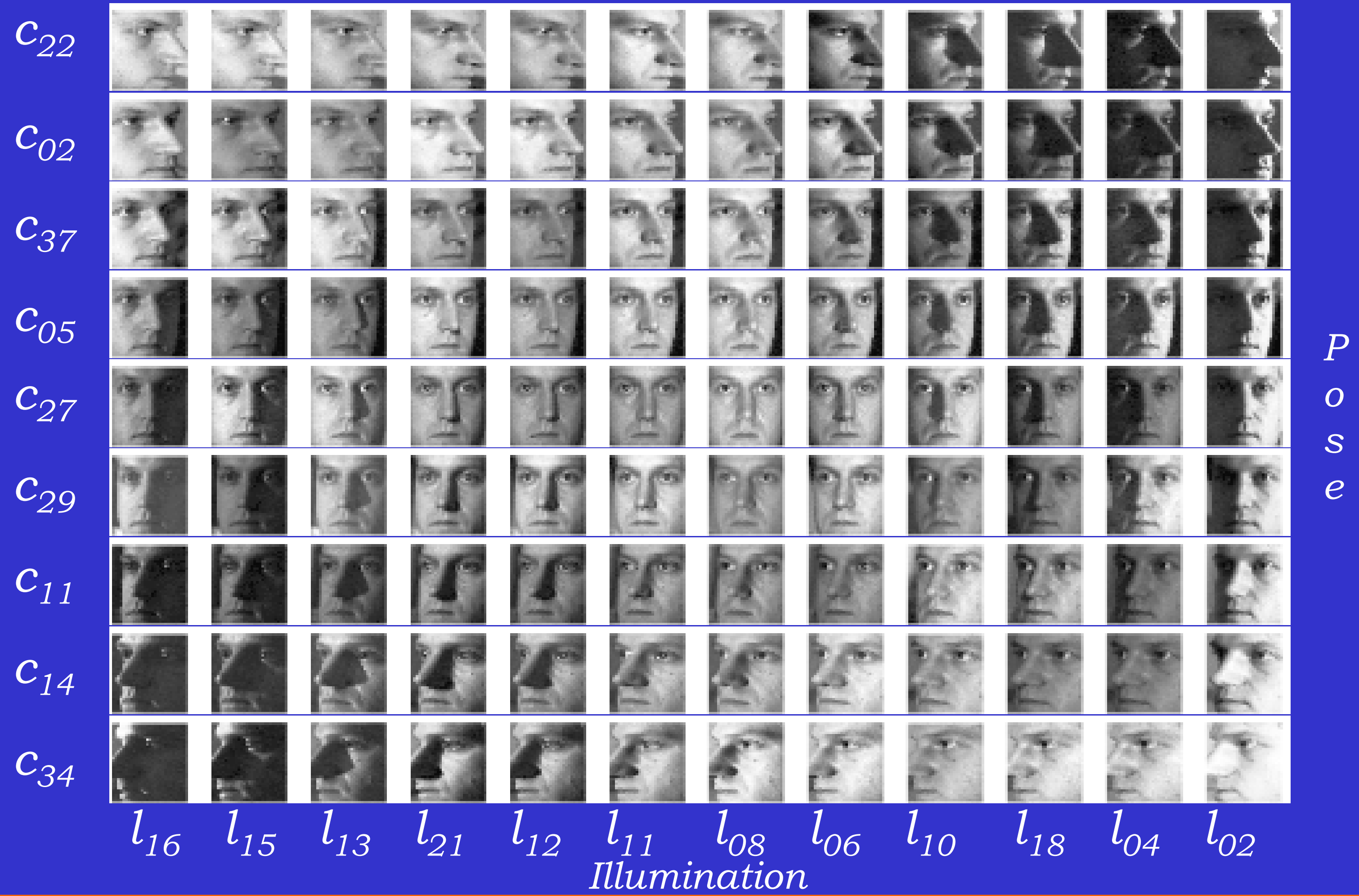
- Introduction to unconstrained face recognition.

- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

* **S. Zhou** and R. Chellappa, *"Image-based face recognition under illumination and pose variations,"* Journal of Optical Society of America (JOSA), Feb., 2005.

Appearances under illumination and pose variation

- 68 objects, 12 lights, 9 poses.



SIEMENS

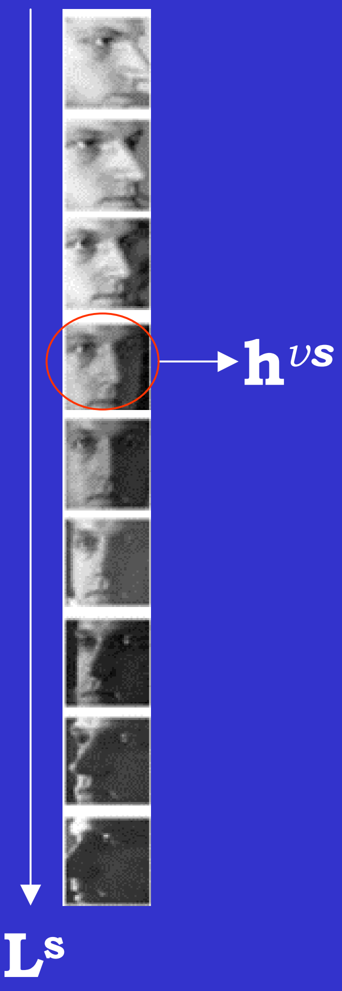
Approach

- Illuminating light field
  - Describes all possible human face images under all possible illumination conditions and at all possible poses.
  - Extends generalized photometric stereo to handle pose variation.
  - Derives an illumination- and pose-invariant signature for robust face recognition under illumination and pose variations.

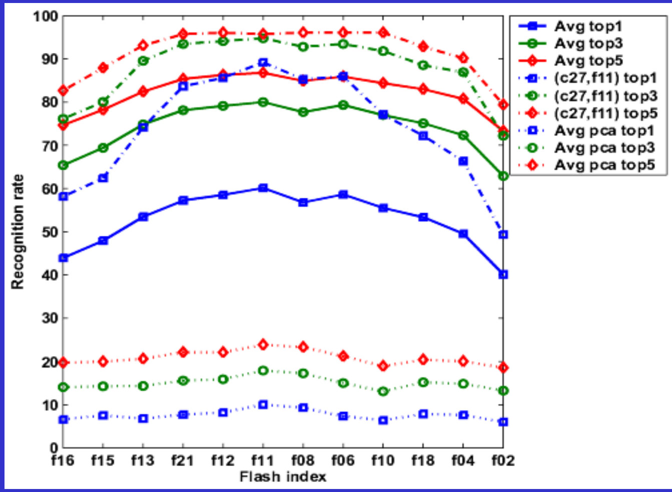
Illuminating Light Field (ILF)

[Zhou & Chellappa JOSA'05]

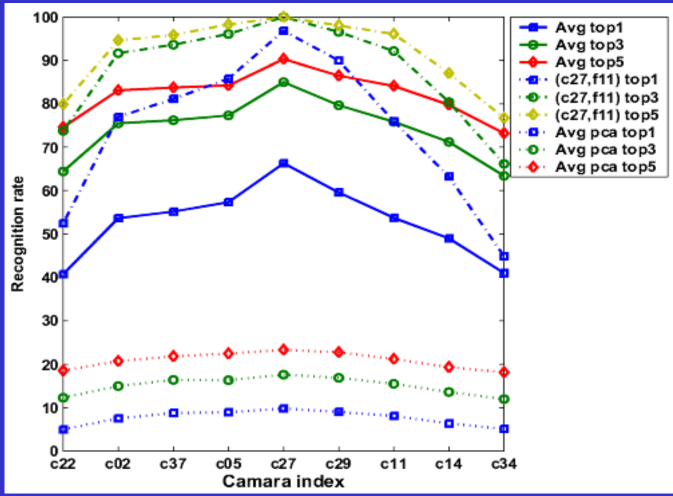
- The concept of light field (LF).
  - $\mathrm{L}^{\mathbf{s}}_{Vd\times 1} = \mathbf{W}_{Vd\times 3m}(\mathbf{f}_{m\times 1} \otimes \mathbf{s}_{3\times 1})$
  - $\mathbf{h}^{v\mathbf{s}}_{d\times 1} = \mathbf{W}^{v}(\mathbf{f} \otimes \mathbf{s})$
  - $\mathbf{f}$: illumination- and pose-invariant.



FR Across Illumination and Pose: Recognition Results



Across illuminations



Across poses

Illumination variation is easier to handle than pose variation!!

Roadmap to Unconstrained Face Recognition and Analysis

- Introduction

- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

* **S. Zhou**, V. Krueger, and R. Chellappa, *"Probabilistic recognition of human faces from video,"* Computer Vision and Image Understanding (special issue on Face Recognition), Vol. 91, pp. 214-245, August 2003.

Video presents challenges and chances

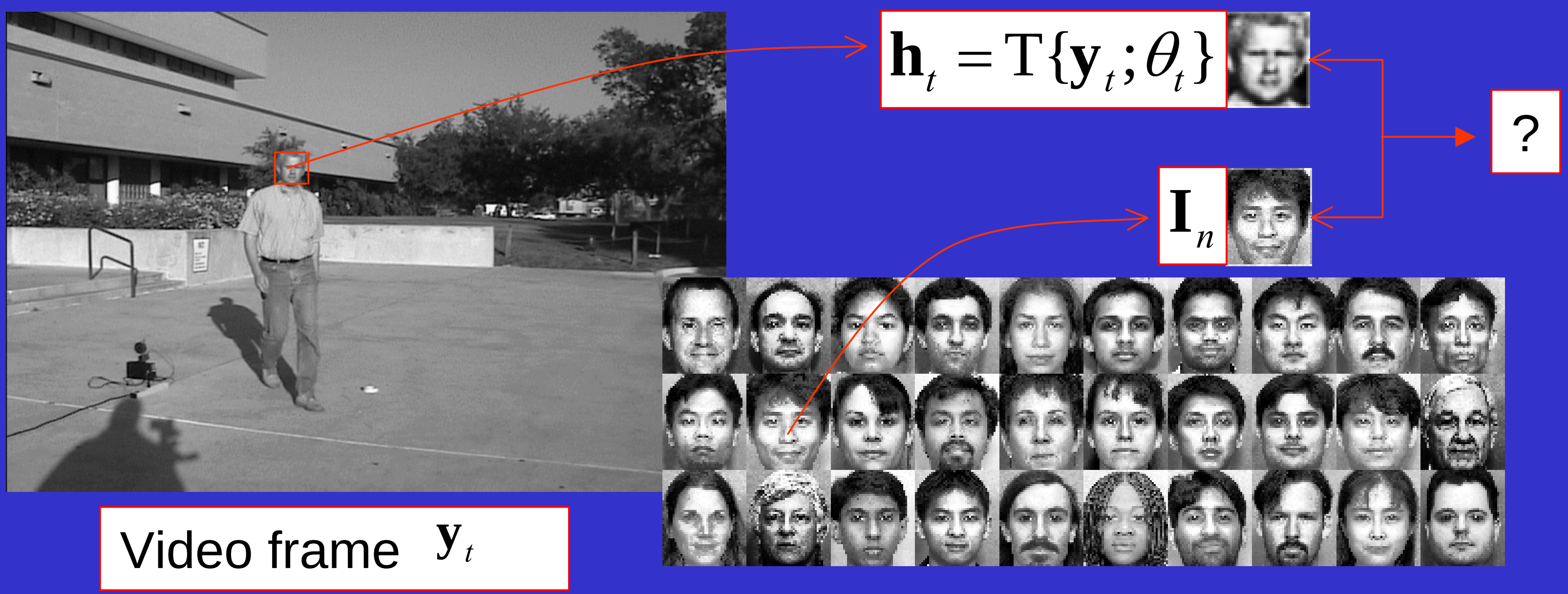
- Requires solving both tracking and recognition.
- Appearance variation.
- Poor image quality.
- Multiple frames with temporal continuity.

Tracking-then-Recognition v.s. Tracking-and-Recognition Approaches

Tracking-then-recognition	Tracking-and-recognition
Essentially still-image-based face recognition	Simultaneous tracking-and-recognition
Utilize temporal information for tracking only	Utilize temporal information for tracking and recognition
Recognition performance relies on tracking accuracy	Improves tracking accuracy and recognition performance
	Probabilistic, interpretable

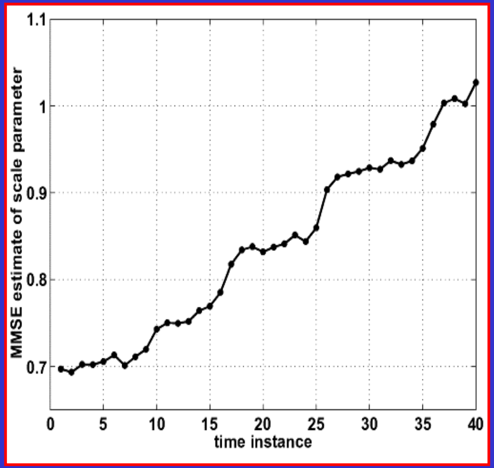
Time Series State Space Model

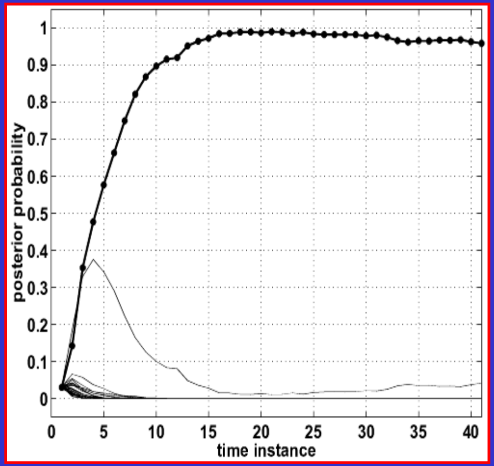
- Motion equation: $\theta_t = g(\theta_{t-1}) + \mathbf{u}_t$
- Identity equation: $n_t = n_{t-1}$
- Observation equation: $\mathbf{h}_t = \mathrm{T}\{\mathbf{y}_t; \theta_t\} = \mathrm{I}_{n_t} + \mathbf{v}_t$



Model Solution

- Posterior distribution: $p(n_t, \theta_t \mid \mathbf{y}_{0:t})$
  - $p(n_t \mid \mathbf{y}_{0:t})$: posterior recognition density.
  - $p(\theta_t \mid \mathbf{y}_{0:t})$: posterior tracking density.
- Particle filter with efficient computation.





Tracking Accuracy and Recognition Result

- NIST database
  - Case 1: Pure tracking using a Laplacian density.
  - Case 2: Tracking-then-recognition using an IPS density.
  - Case 3: Tracking-and-recognition using a combined density.

Case	Case 1	Case 2	Case 3
Tracking Accuracy	87%	NA	100%
Recognition within top 1	NA	57%	93%
Recognition within top 3	NA	83%	100%

* Courtesy of the HumanID project

Roadmap to Unconstrained Face Recognition and Analysis

- Introduction to unconstrained face recognition.

- Selected Approaches
  - Face recognition across illumination.
  - Face recognition across illumination and pose.
  - Video-based face recognition.
  - Age Estimation.

* **S. Zhou** *et al., "Image based regression using boosting method,"* Submitted.

What is Image Based Regression?

- Regression or function approximation
  - Given an input image $\mathbf{x}$, infer or approximate an output $\mathbf{y}(\mathbf{x})$ that is associated with the image $\mathbf{x}$.
- Age estimation: $\mathbf{y}(\mathbf{x}) = age$

SIEMENS

State-Of-The-Art: Data-Driven Approach

- **Nonparametric regression (NPR)** $\mathbf{g}(\mathbf{x}) = \sum_{n=1}^{N} w_n(\mathbf{x})\mathbf{y}(\mathbf{x}_n);\ w_n(\mathbf{x}) \propto h(\mathbf{x}, \mathbf{x}_n);$
  - Smoothed k-NN regressor
- **Kernel ridge regression (KRR)** $\mathbf{g}(\mathbf{x}) = \sum_{n=1}^{N} \mathbf{w}_n k(\mathbf{x}, \mathbf{x}_n) = \sum_{n=1}^{N} \mathbf{w}_n \phi(\mathbf{x}_n)^{\mathrm{T}} \phi(\mathbf{x})$
  - Hyperplane in RKHS
- **Support vector regression (SVR)** $g(\mathbf{x}) = \sum_{i=1}^{I<N} w_{n_i} k(\mathbf{x}, \mathbf{x}_{n_i})$
  - Single output, ε-insensitive loss function
- **Boosting regression** $\mathbf{g}(\mathbf{x}) = \sum_m \alpha_m \mathbf{h}_m(\mathbf{x});\quad \mathbf{h}_m(\mathbf{x}) \in \mathbf{H}$
  - Using boosting method
  - Not data-driven

Challenges: Appearance Variation

- Appearance variation
  - Inter-object variation.
  - Extrinsic variations: camera, geometry, lighting, etc.
  - Alignment/background.

- Treatment of appearance variation
  - Data-driven approach: Kernel function $k(\mathbf{x}, \mathbf{x}_n)$ is global and sensitive to appearance variation.
  - Boosting approach: Feature function $\mathbf{h}_m(\mathbf{x})$ is local and robust to appearance variation.

Boosting

- Boosting [Freund & Schapire'95][Friedman *et al.*, AS'00]
  - AdaBoost is the state-of-the-art classification method.
  - Ensemble method: Combines weak learners into a strong learner using an additive form:

$$\mathbf{g}(\mathbf{x}) = \sum_m \alpha_m \mathbf{h}_m(\mathbf{x}); \quad \mathbf{h}_m(\mathbf{x}) \in \mathbf{H}$$

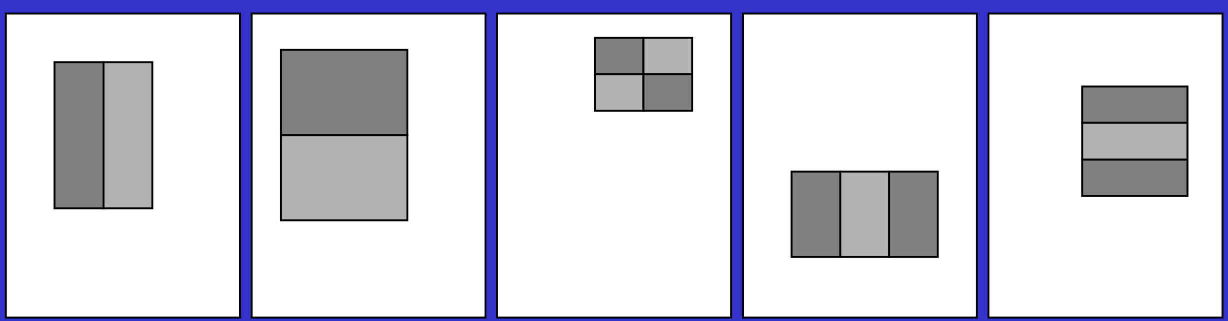
  - Selects weak learners (or features) from the dictionary set.

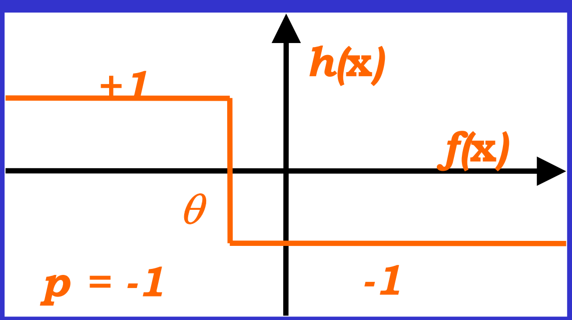
- Three elements of boosting
  - (a) Loss function or error model $L(\mathbf{g}(\mathbf{x}), \mathbf{y}(\mathbf{x}))$
  - (b) Dictionary set $\mathbf{H} = \{\mathbf{h}(\mathbf{x})\}$
  - (c) Selection algorithm

Dictionary Set

- Primitives: 1-D decision stump [Viola & Jones, CVPR'01]

$$h(\mathbf{x}) = \begin{cases} +1; & \text{if } pf(\mathbf{x}) \geq p\theta \\ -1; & \text{otherwise} \end{cases}; \qquad f(\mathbf{x}): \text{ simple feature}$$





Result: Age estimation

- Variations
  - Pose, illumination, expression, beard, moustache, spectacle, etc.

- Performance (1002 images, 800 training/202 testing, 5-fold CV)

	NPR	KRR	SVR	IBR
mean err.	8.44	13.56	6.60	5.81
25% per. err.	2.54	3.99	1.38	1.26
median err.	5.50	10.80	4.39	3.15
75% per. err.	10.87	17.99	9.04	7.79
testing time(s)	3.6s	3.6s	3.3s	0.016s

Visual Tracking

* **S. Zhou**, *et al.*, *"Visual tracking and recognition using appearance-adaptive models in particle filters,"* IEEE Trans. on Image Processing, November 2004.

THANKS for Listening!!!

* Shaohua Kevin Zhou, kzhou@scr.siemens.com

http://www.cfar.umd.edu/~shaohua/