The Role of Statistics in Biometric Authentication Based on Facial Images

Sinjini Mitra
Department of Statistics & CyLab
Carnegie Mellon University
smitra@stat.cmu.edu

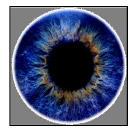
Collaborators: Stephen E. Fienberg
Anthony Brockwell
B.V.K. Vijaya Kumar (ECE)
Marios Savvides (ECE)
Yanxi Liu (Robotics Institute, CS)

What is Biometric Authentication?

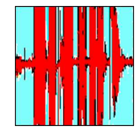
- Biometrics refer to the unique biological traits (physical or behavioral) of individuals that can be used for identification
  - Physical: retinal/iris scan, fingerprint, face, palm-print
  - Behavioral: voice-print, gait, gesture


face


fingerprint


iris scan


palm-print


voiceprint

Biometric Authentication

- Technology for verification of a person's identity based on his/her biometrics
- "Something you are" versus "something you know" (passwords) or "something you possess" (ID card)
- Better security and reliability: cannot be stolen or forgotten and less prone to fraud

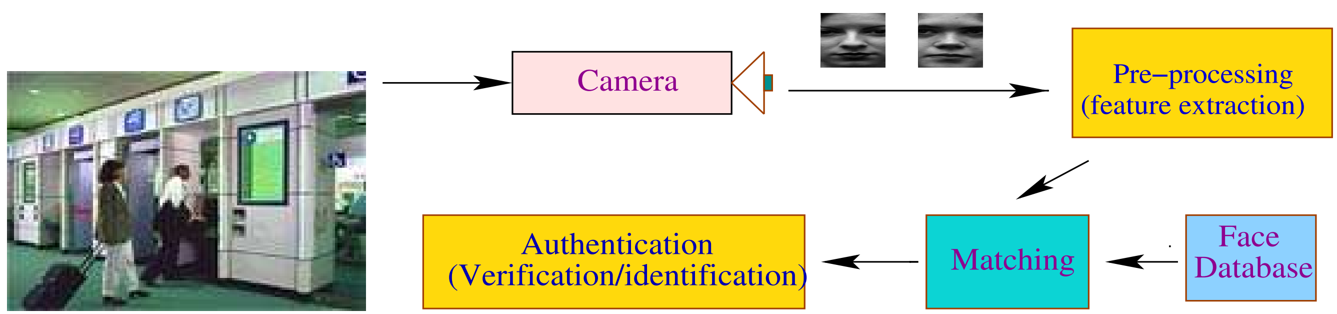
Applications

Forensics, homeland security, access to ATMs and computer networks

Facial Biometrics

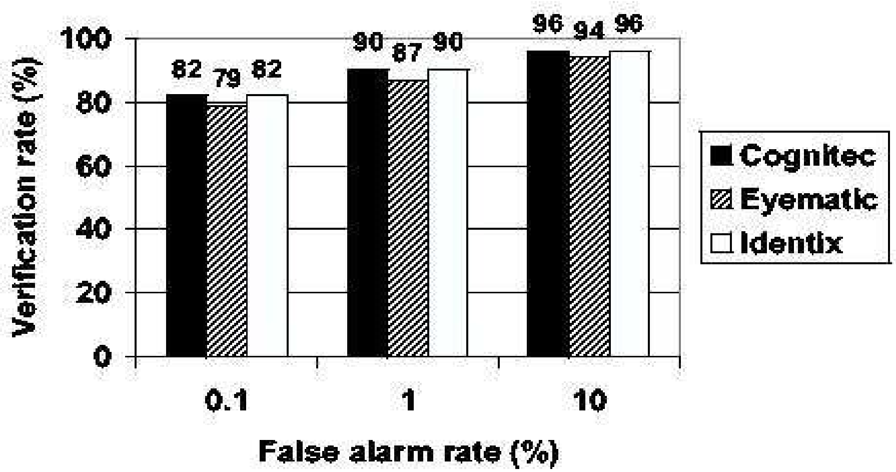
- Fairly accurate, non-intrusive and user friendly
- Analyzes facial characteristics from an image
  Examples: Position relationships between eyes, nose, mouth and chin
- Very challenging - sensitive to external factors

- A face authentication system has 3 components: (i) Enrollment, (ii) Identification, (iii) Decision: authentic or impostor?



The Face Recognition Vendor Test (FRVT) 2002

- Provide performance measures for assessing the ability of 10 commercially used automatic face recognition systems to meet real-world requirements
  - Participants tested on large data not previously seen - 121, 589 images of 37, 437 people



  - Effect of demographics (sex, age), image properties (location, resolution, pose, illumination), time difference between enrollment and testing
  - Performance degraded with increasing database and "watch-list" size
- **Drawback:** Impressive results but based on observational studies and are empirical in nature - no statistical basis (modeling, ROC curves) and scope for valid inference

My Research Goals

I. Statistical analysis and evaluation of existing authentication systems

II. Explore new approaches to building statistical model-based authentication systems

III. Explore other ways to develop distortion-tolerant authentication systems

Motivation: Minimum Average Correlation Energy (MACE) Filter

- Introduced by Kumar, et al. (2002)
- Easily detected features for distinguishing authentics and impostors
- A linear filter and reports impressive results

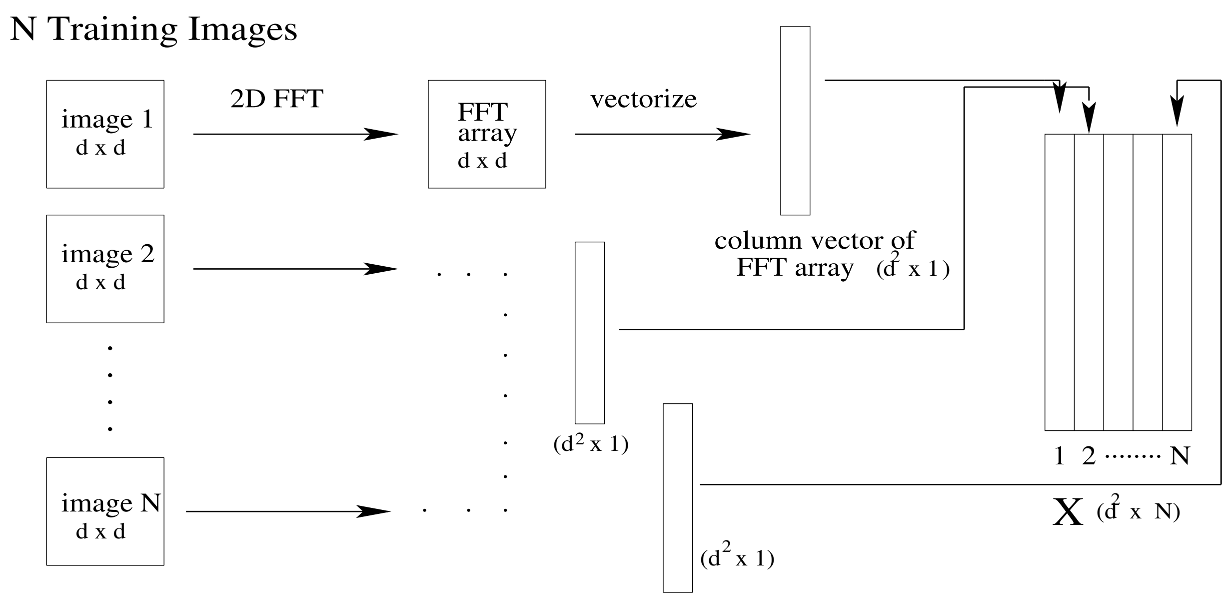
Objective:

- Use MACE as a baseline for developing statistical methods of analysis and evaluation of face authentication systems, in order to make them more rigorous and useful in practice

I. The MACE Filter

$$\mathbf{h}_{MACE} = D^{-1}X(X^T D^{-1} X)^{-1}\mathbf{c},$$

D : a diagonal matrix (ave. power spectrum), $\mathbf{c}$: a column vector of ones



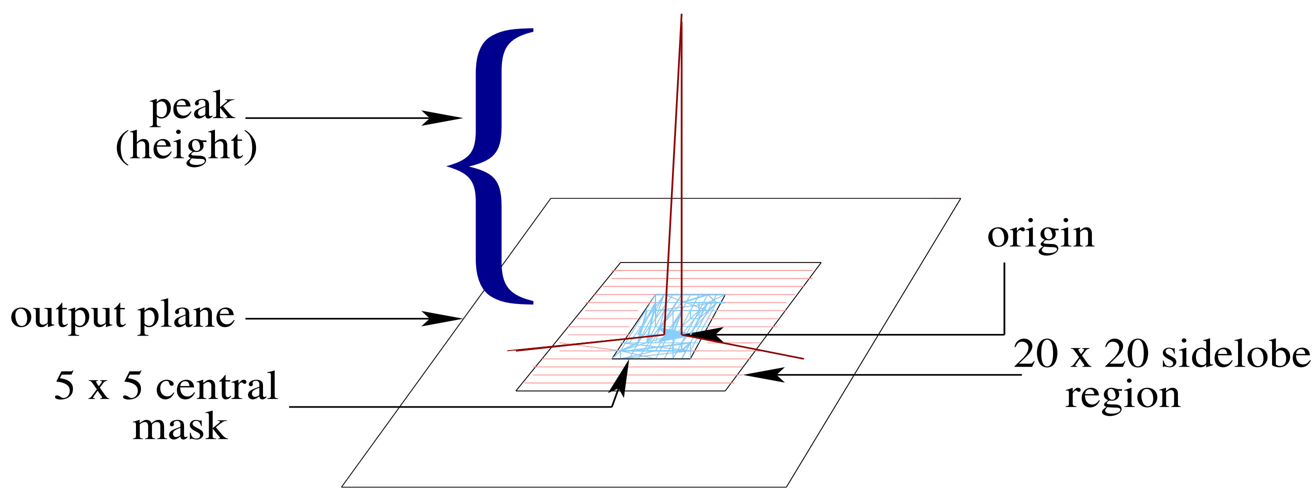
- Obtained by minimizing the average correlation plane energy $E_{ave} = h' D h$ while satisfying $X^+ h = \mathbf{c}$ (constraint at the origin)
- Such a design forces the output plane to have low values everywhere except at the origin – facilitates easy distinction

Filter-based Authentication

- One MACE filter is synthesized per person
- Filter applied to each test image via convolution (frequency domain)
- Inverse Fourier transform yields final spatial output

Peak-to-Sidelobe Ratio (PSR):

- Quantitative measure for authentication



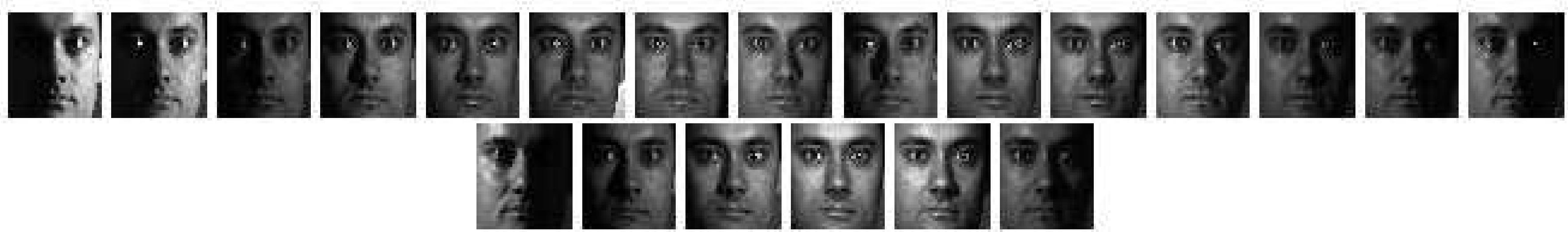
$$\text{PSR} = \frac{\text{peak-}\bar{x}}{s}$$

Authentication: High for authentics and low for impostors

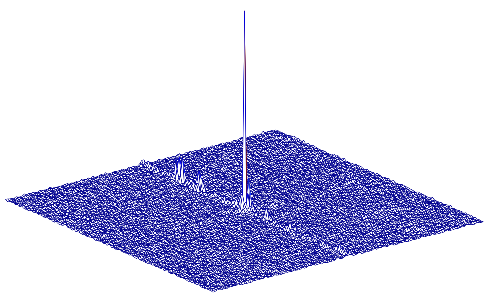
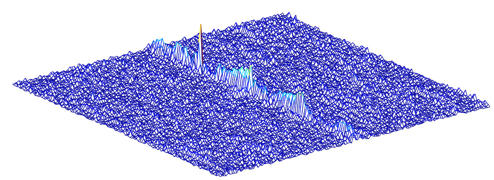
The Databases

I. Cohn-Kanade AU-Coded Facial Expression Database: 55 subjects expressing neutral, joy, anger and disgust

II. CMU-PIE Database: 65 subjects under different illumination conditions



My Authentication Results

Authentic (Subject 1)	Impostor (Subject 2)
Test Image	Test Image
	
Filter 1 Output (Subject 1)	Filter 1 Output (Subject 2)
	
PSR: 30.6360	PSR: 8.8697

Properties and Distortion Tolerance

- Easy to implement and has attractive features
- Shift-invariance, tolerance to illumination and partial occlusion (Savvides and Kumar, 2003), but sensitive to other distortions like noise, expressions, pose, etc.
- Many heuristics involved: (i) training images, (ii) PSR threshold

Distortion-tolerant MACE

- Obtained by minimizing the compromise criterion (Kumar, 1992): $E_{ave} + \alpha\sigma^2 = h'Dh + \alpha h'h = h'(D + \alpha I_d)h$, where σ^2: noise variance, α: tuning parameter
- Replace D in $\mathbf{h}_{MACE}$ by $D + \alpha I$, so that

$$\mathbf{h}_{\mathbf{noise}} = (D + \alpha I_d)^{-1} X [X^T (D + \alpha I_d)^{-1} X]^{-1} \mathbf{c}.$$

- Reasonable performance under distortions - lower false alarms
- Drawback: No deterministic way of choosing α: “brute-force” and ad-hoc, and no fixed optimal value

Statistical Analysis of MACE

- MACE is not a model-based technique
- Model variation in PSR values with changes in image properties (noise, resolution), filter design parameters (sidelobe dimension, α) and demographics (age, sex)
- Performance evaluation and inference:
  - Confidence intervals and hypothesis tests for PSR and error rates
  - Predict PSR values and error rates for unobserved large new data
- Model performance statistics as a function of database and "watch-list" sizes

Exploratory Analysis

- More training images required in presence of distortions
- Often better authentication with (i) fewer training images, (ii) lower resolution images, (iii) smaller sidelobe dimension

Performance Evaluation: The Literature

- A decision-theoretic framework: a match score T and a threshold τ
  - $T > \tau$: match (authentic), $T \leq \tau$: mismatch (impostor)
  - Type I error: FNR $= P(T \leq \tau | T \in \text{Authentic}) = \int_{-\infty}^{\tau} f_A(x)dx$
    Type II error: FPR $= P(T > \tau | T \in \text{Impostor}) = \int_{\tau}^{\infty} g_I(y)dy$
  - Trade-off between FPR, FNR and their behavior with τ can be represented by a Receiver Operating Characteristic (ROC) curve
- Error rates estimated empirically by sample proportions
- Confidence intervals and hypothesis tests for error rate estimates: binomial distribution and bootstrapping (Bolle et al, 2000), beta-binomial distribution (Schuckers, 2003)
- Drawback: Based on many assumptions - independence, equality of variances, which seldom hold in practice for real image data
- ROC curves help in evaluation of score distributions - Ishwaran and Gatsonis (2000) used hierarchical models for clustered data

My approach:

- Use ROC curves to study the score distributions and use the robust modeling approach

Inference for New Data

- **Goal:** Predict PSR value for new large face data, estimate the expected error rates and model variation as a function of database and "watch list" size

Random Effects Hierarchical Model (Gelfand, et al. JASA 1990)

- $Y_{ij} \overset{ind.}{\sim} N(\alpha_i + \sum_{m=1}^{M} \beta_i^m x_{ij}^{(m)}, \sigma^2),$
  $\boldsymbol{\theta_i} \equiv (\alpha_i, \beta_i^1, \ldots, \beta_i^M)^T \sim MVN(\boldsymbol{\theta_0} \equiv (\alpha_0, \beta_0^1, \ldots, \beta_0^M)^T, \Sigma),\ \ \sigma^2 \sim IG(a, b)$
- Conjugate hyperpriors for $\boldsymbol{\theta_0}$, Σ and MCMC-based posterior simulation
- Inference based on $\boldsymbol{\theta_0}$ and posterior predictive distributions $p(y_{ij}|\mathbf{y})$
- PSR is the MACE "score", so $Y:\ \log(PSR)$, covariates x_{ij} :

Image properties	Filter parameters	Database properties
Authentic/Impostor (binary)	# training images	size
Distortions (categorical)	α	"watch-list" size
Image resolution	sidelobe dimension	

- Model the odds of false alarm (FPR, FNR) in a logistic regression framework
- Model checks for validity of assumptions: linearity, independence, homoscedasticity

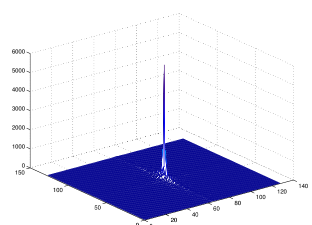
II. Statistical Model-based Systems

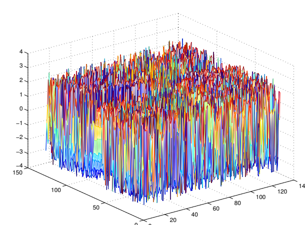
- Spatial models (2D AR, MRF)
  - inadequate for building model-based classification tools
- Spectral models: **My approach**
  - No one has modeled the image spectrum directly
  - The Fourier transform of an image $x(n_1, n_2)$ is defined as:

$$X(j,k) = \frac{1}{N_1^2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_1-1} x(n_1, n_2) e^{-i2\pi(n_1 j/N_1 + n_2 k/N_1)}$$

$$\overset{\text{(polar form)}}{=} \underbrace{|X(j,k)|}_{\text{magnitude}} \, e^{i \overbrace{\theta_x(j,k)}^{\text{phase}}}, \quad j, k = 0, 1, \ldots, N_1 - 1$$

  - We will model magnitude and phase

Magnitude: 

Phase: 

Spectral Modeling

- Phase captures most of a face image identifiability (Hayes, 1982)

Subject 1 | Subject 2 | Mag 1 + Phase 2 | Mag 2 + Phase 1

- **Difficulties in Phase Modeling:**
  - No stationarity assumptions work - “wrapping around” property
  - Hard to isolate location of discriminating information in phase
  - Varies considerably with any kind of distortion
- **Model Selection:**
  - **Idea:** Generate models for an “optimal” number of Fourier coefficients by preserving identifiability - dimension reduction
  - An image of good quality can be reconstructed using few low frequency components (high energy) while higher ones (low energy) represent finer facial details

original | 40% | 16% | 3%

Mixture Models

- Flexible semi-parametric framework for modeling unknown distributional shapes
- Mixtures represent different illumination conditions for each person
- Model log-magnitude and phase for pixels within a 50×50 grid around origin:

$$\mathbf{Y_j} = \begin{pmatrix} L_{s,t}^{k,j} \\ P_{s,t}^{k,j} \end{pmatrix} \sim BVN\left(\boldsymbol{\mu}_{s,t}^{k}, \boldsymbol{\Sigma}_{\mathbf{s,t}}^{\mathbf{k}}\right)$$

- Mixture model: $\boxed{f(\mathbf{y_j}; \boldsymbol{\Psi}) = \sum_{i=1}^{g} \pi_i \phi(\mathbf{y_j}; \boldsymbol{\mu_i}, \boldsymbol{\Sigma_i})}$
- One mixture model per pixel per person: $f_{s,t}(\mathbf{y_j}; \boldsymbol{\Psi}|k)$
- Gibbs Sampler used for parameter estimation via posterior simulation, using conjugate priors for $\boldsymbol{\pi}$, $\boldsymbol{\mu_i}$ and $\boldsymbol{\Sigma_i}$
- New test image ($\mathbf{x} = (L_{s,t}, P_{s,t})$) classified by MAP estimate based on posterior likelihood:

$$\boxed{C = \arg\max_k g(k|L, P) \equiv \arg\max_k g(L, P|k)p(k)}$$

where $g(L, P|k) = \Pi_s \Pi_t f_{s,t}(\mathbf{x}; \boldsymbol{\Psi}|k), \; p(k) = 1/k$

- Inference based on likelihoods, possibly using similar random effects models as before
- Possible to classify the illumination type of an image of a person

Summary

- Presented a rigorous statistical framework for analysis and evaluation of existing authentication systems which helps in bypassing the need for the empirical system evaluation tools mostly used today
- Shown significance of statistically-based systems:
  - inference for large new data
  - guidelines to users of existing systems, making them more reliable
- Exploiting the key role of phase in face identification for building models in the spectral domain is a promising novel approach with a simple classification scheme
- Current research agenda consists of implementing all these techniques in the context of the MACE filter system and the spectral model-based system (after development)

Future Directions:

- Spectral Models: Model all pixels together using inter-pixel correlations, increase algorithm efficiency
- Other Methods: *Facial Asymmetry* - potential for devising distortion-tolerant authentication systems (Liu et al. 2003)
- Other Biometrics: Fingerprints, Multi-modal systems

References

- Bolle, R.M., Pankanti, S. and Ratha, N.K. (2000). Evaluation Techniques for biometrics-based authentication systems (FRR). *In Proceedings of ICPR 2000*, pages 2831-2837.
- Gelfand, A.E., Hills, S.E., Racine-Poon, A., Smith, A.F.M. (1990). Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *JASA*, 85(412):972-985.
- Hayes, M.H. (1982). The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of its Fourier Transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(2): 140-154.
- Ishwaran, H. and Gatsonis, C. (2000). A general class of hierarchical ordinal regression models with applications to correlated ROC analysis. *The Canadian Journal of Statistics*, 28:731:750.
- Liu, Y., Schmidt, K., Cohn, J., Mitra, S. (2003). Facial Asymmetry Quantification for Expression-invariant Human Identification. *Computer Vision and Image Understanding Journal*, 91(1/2):138-159.
- Savvides, M. and Vijaya Kumar, B.V.K. (2003). Efficient Design of advanced correlation filters for robust distortion-tolerant face identification. *Proceedings of the IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS)*, pages 45-52.
- Schuckers, M.E. (2003). Using the Beta-Binomial Distribution to Assess Performance of a Biometric Identification Device. *International Journal of Image Graphics*, 3(3):523-529.

References (cont.d)

- Vijaya Kumar, B.V.K., Savvides, M., Venkataramani, K., Xie, C. (2002). Spatial Frequency Domain Image Processing For Biometric Recognition. *Proceedings of the International Conference on Image Processing (ICIP)*, Rochester, NY.
- Vijaya Kumar, B.V.K. (1992). Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31(23): 4773-4801.

Acknowledgment:

Funding in part by the Army Research Office contract DAAD19-02-1-3-0389 to CyLab.