PROBALISTIC VISUAL LEARNING OBJECT DETECTION
A SUMMARY

ADVANCED METHODS APPLIED STATISTICS – ARTICLE SUMMARY

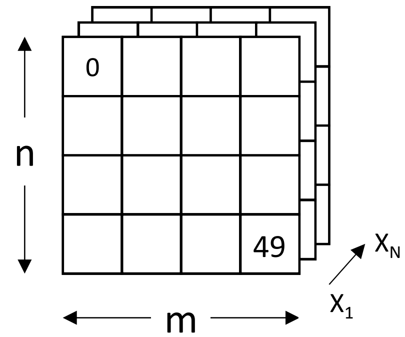
By *Tai Thorsen, Daniel Ramyar & Andrew Khoudi*

March 7th 2019

UNIVERSITY OF COPENHAGEN

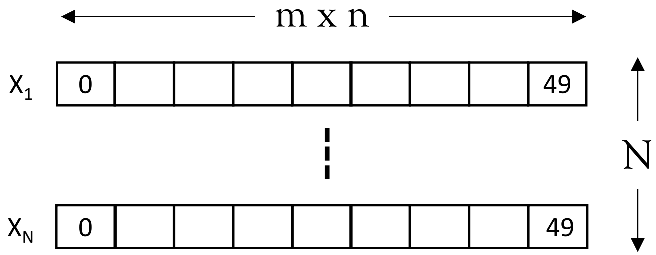
Transforming the training data

Training images



lexicographic ordering

Data matrix

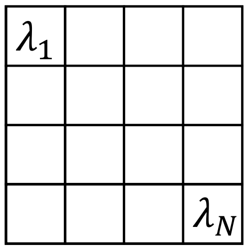


KL transformation

$$\Lambda = \Phi^T \Sigma \Phi$$

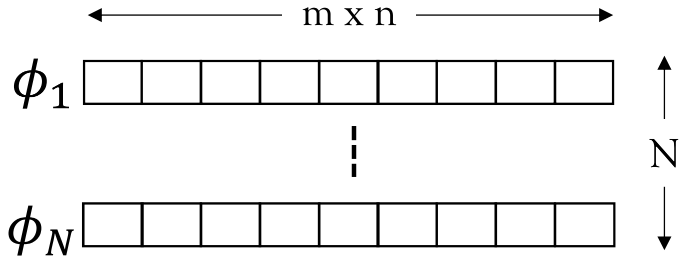
Eigenvalues

$\Lambda =$



Eigenmatrix w Eigenvectors

$\Phi =$

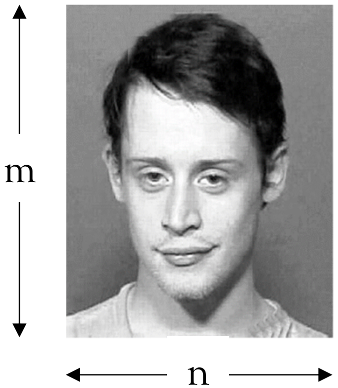


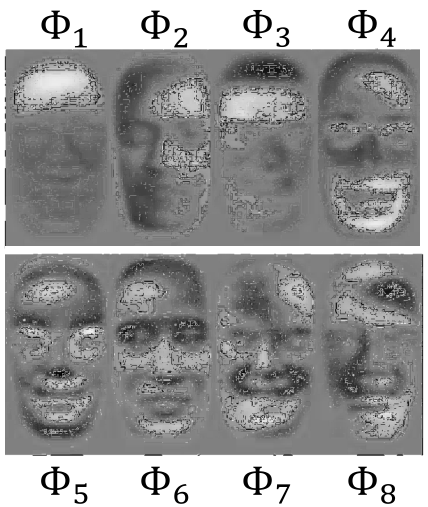
Covariance Matrix

Σ is the covariance matrix of the Data matrix

Analyzing a new image

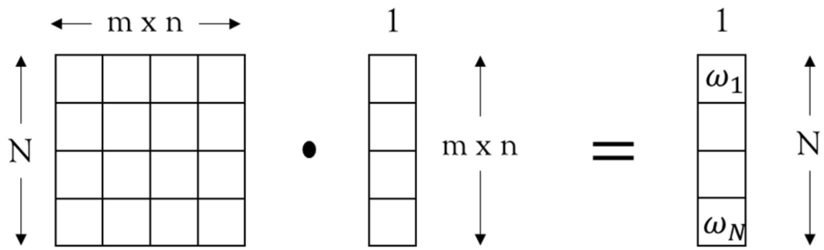
New (random) image, $\boldsymbol{x}$





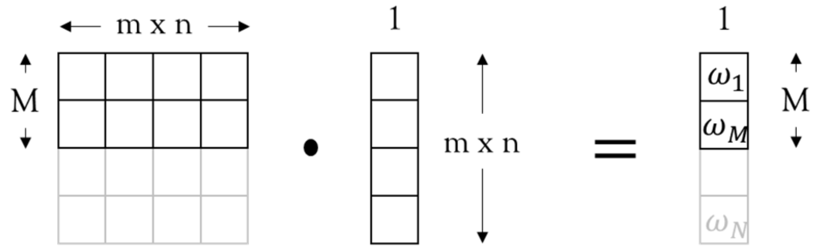
$$\tilde{x} = x - \overline{x}$$

$$y = \Phi^T \tilde{x} =$$



PCA, chosing
M highest eigenvalues

$$y_M = \Phi_M^T \tilde{x} =$$

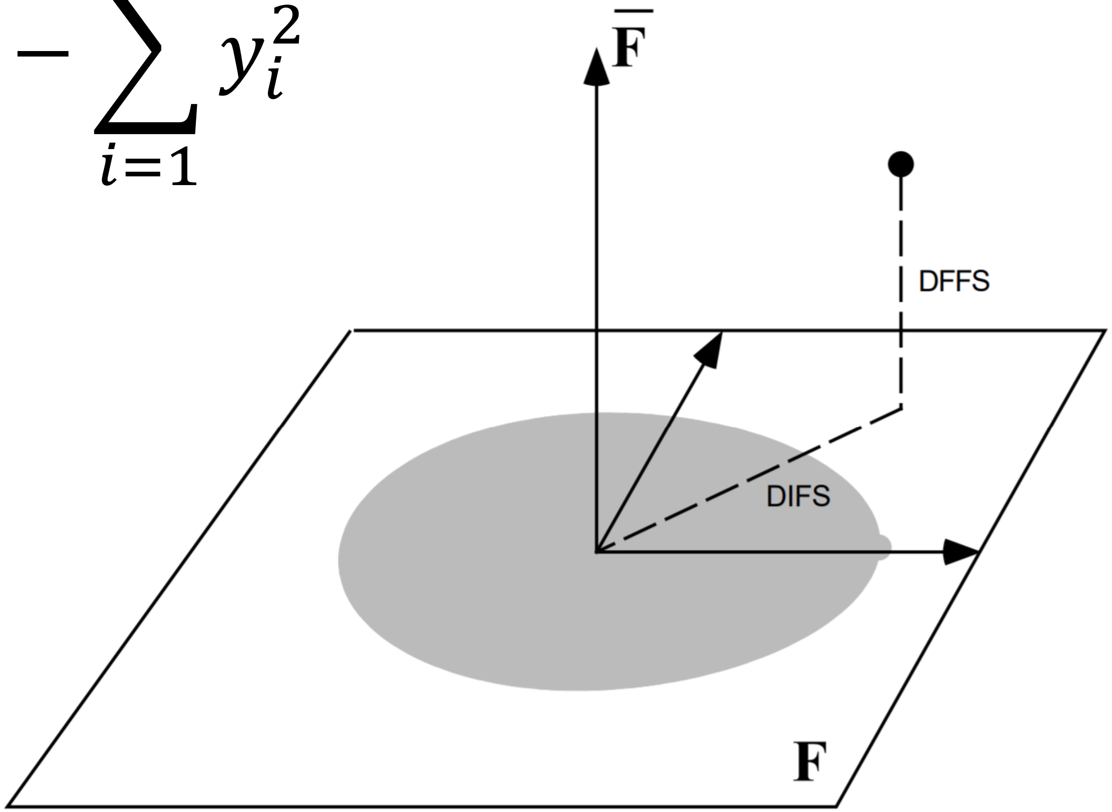


Distance From Feature Space (DFFS)

$$\epsilon^2(\boldsymbol{x}) = \sum_{i=M+1}^{N} y_i^2 = \|\widetilde{\boldsymbol{x}}\|^2 - \sum_{i=1}^{M} y_i^2$$

$$F = \{\Phi_i\}_{i=1}^{M}$$

$$\overline{F} = \{\Phi_i\}_{i=M+1}^{N}$$



Maximum Likelihood (ML)

$$\mathrm{g}(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$$

Rewriting the gaussian to a Likelihood:

$$\mu - x = \phi_i^T \tilde{x} = y_i \qquad \sigma^2 = \lambda_i$$

$$\mathrm{P}(\boldsymbol{x}|\boldsymbol{\Omega}) = \frac{1}{(2\pi)^{N/2} \prod_{i=1}^{N} \lambda_i^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N} \frac{y_i^2}{\lambda_i}\right)$$

Maximum Likelihood (ML)

$$P(\boldsymbol{x}|\boldsymbol{\Omega}) = \frac{1}{(2\pi)^{N/2}\prod_{i=1}^{N}\lambda_i^{1/2}}\exp\left(-\frac{1}{2}\sum_{i=1}^{N}\frac{y_i^2}{\lambda_i}\right)$$

$$\hat{P}(\boldsymbol{x}|\boldsymbol{\Omega}) = P_F(x|\Omega)\cdot\hat{P}_F(x|\Omega)$$

$$\hat{P}(\boldsymbol{x}|\boldsymbol{\Omega}) = \frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{M}\frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{M}\prod_{i=1}^{M}\lambda_i^{1/2}}\cdot\frac{\exp\left(-\frac{\epsilon^2(x)}{2\rho}\right)}{(2\pi\rho)^{(N-M)/2}} \quad , \quad \rho^* = \frac{1}{N-M}\sum_{i=M+1}^{N}\lambda_i$$

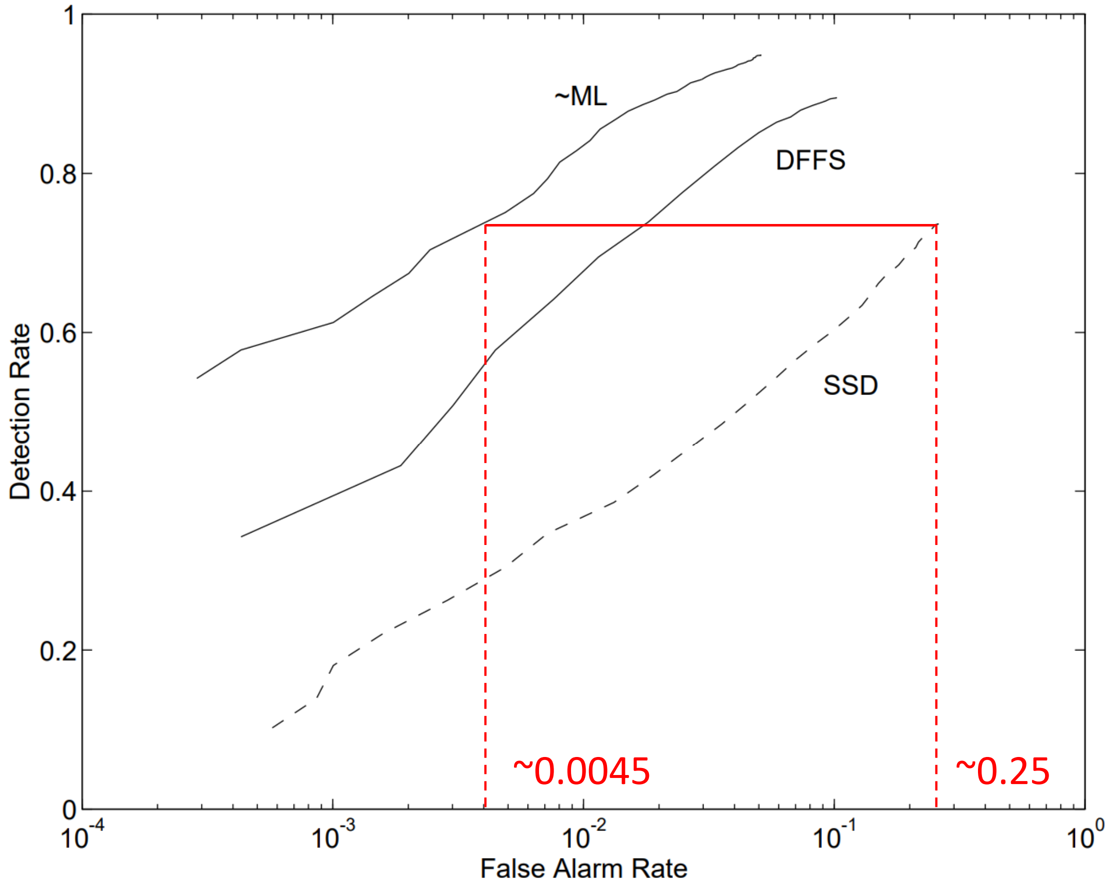
Results I

Eigenvectors based on ~7,600 "mugshots"

Test images analyzed ~7,000

5 dimension principal subspace for DFFS & ML

- ML detector exhibits best detection vs. false-alarm trade-off
- ML detector yields highest detection rate (of 95%)
- ML has a false-alarm rate nearly 2 orders of magnitudes lower than SSD

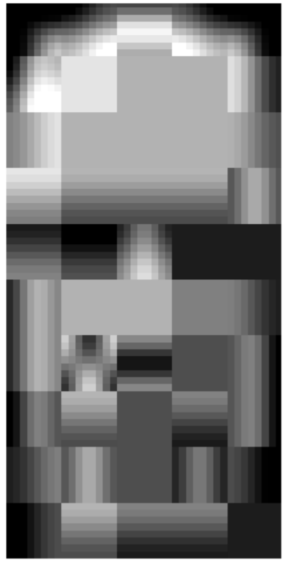


Example w 8 dim. Principal Subspace

ϕ_1 ϕ_2 ϕ_3 ϕ_4 ϕ_5 ϕ_6 ϕ_7 ϕ_8

$$\Phi_8^T \tilde{x} = y_8 = [\omega_1, \ldots, \omega_8]$$

JPEG reconstruction using 5x more data than Eigenvector reconstruction



Summary & what's next…

What they learned back in 1995 @ MIT:

- Eigenspace decomposition & PCA for dimensional reduction
- Estimation of density functions of high-dimensional image space
- Maximum likelihood method for object detection

What they recommended for further studies:

A Bayesian prior to improve the false-alarm rate.

Ω represents different class objects (requires new training data), not only faces. This will improve/reduce false-positives.

$$P(\Omega_i|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|\Omega_i)\cdot P(\Omega_i)}{\sum_{j=i}^{n} P(\boldsymbol{x}|\Omega_j)\cdot P(\Omega_j)}$$