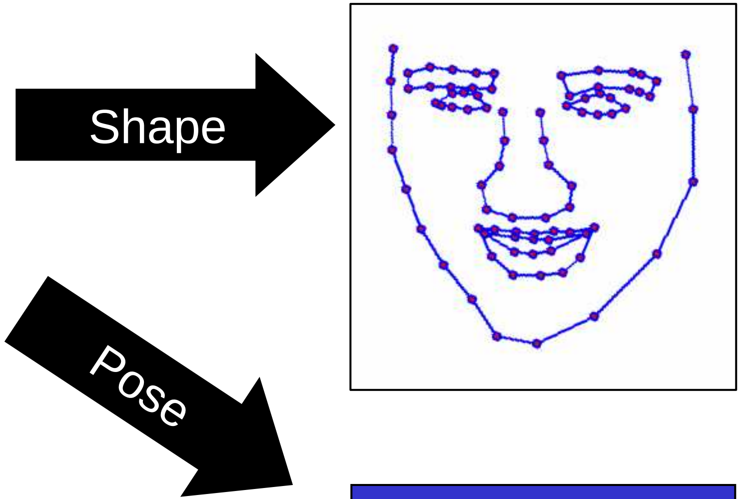
Bayesian Tangent Shape Model For Face Alignment

Leon Gu
Carnegie Mellon University

Presented at CVPR 2003, Wisconsin

Face Alignment

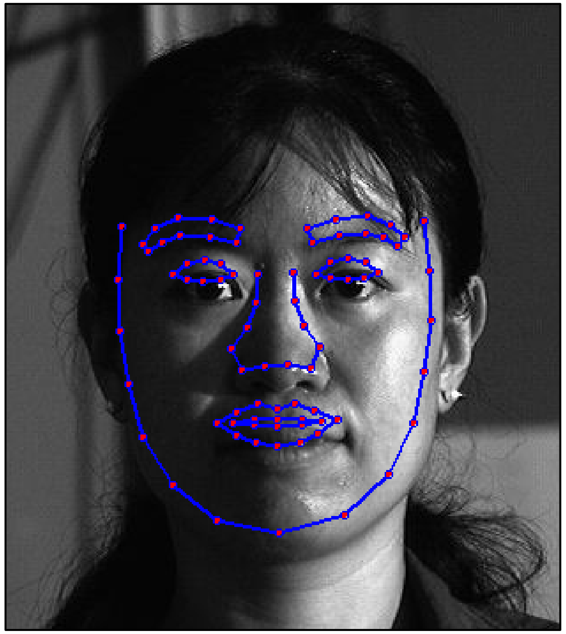
Locate shape structure



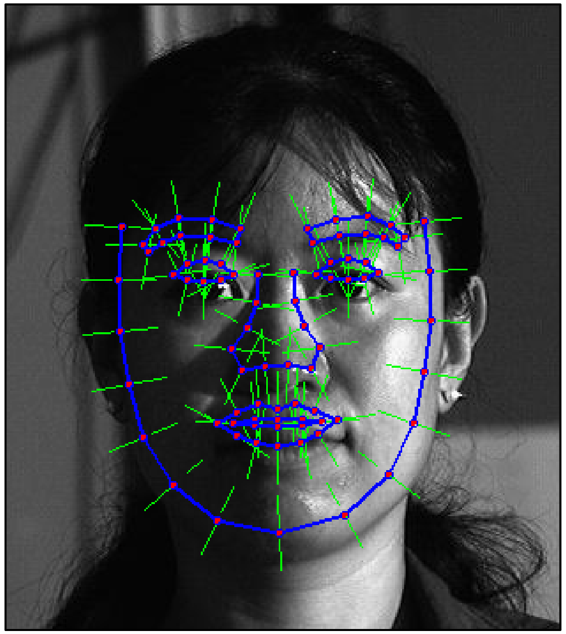
Geometrical Transform

- **Rotation**
- **Translation**
- **Scale**
- **……**

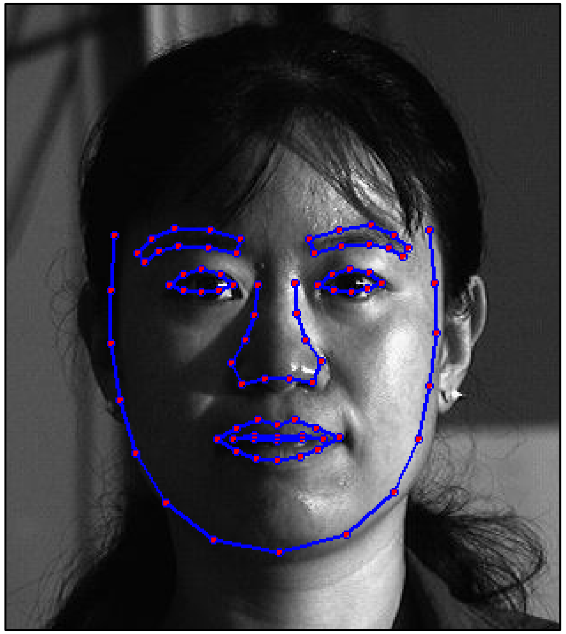
Observation and Regularization



Starting Approximation



Observation



Regularization

Related Work

Active Contour or Snakes
Kass, Witkin and Terzopoulos, 1987

- Spline curves
- Stretchness and smoothness
- General amorphous objects

Active Shape Model
Cootes and Taylor, 1995

- Landmark points
- Prior Density
- Object with specific structure

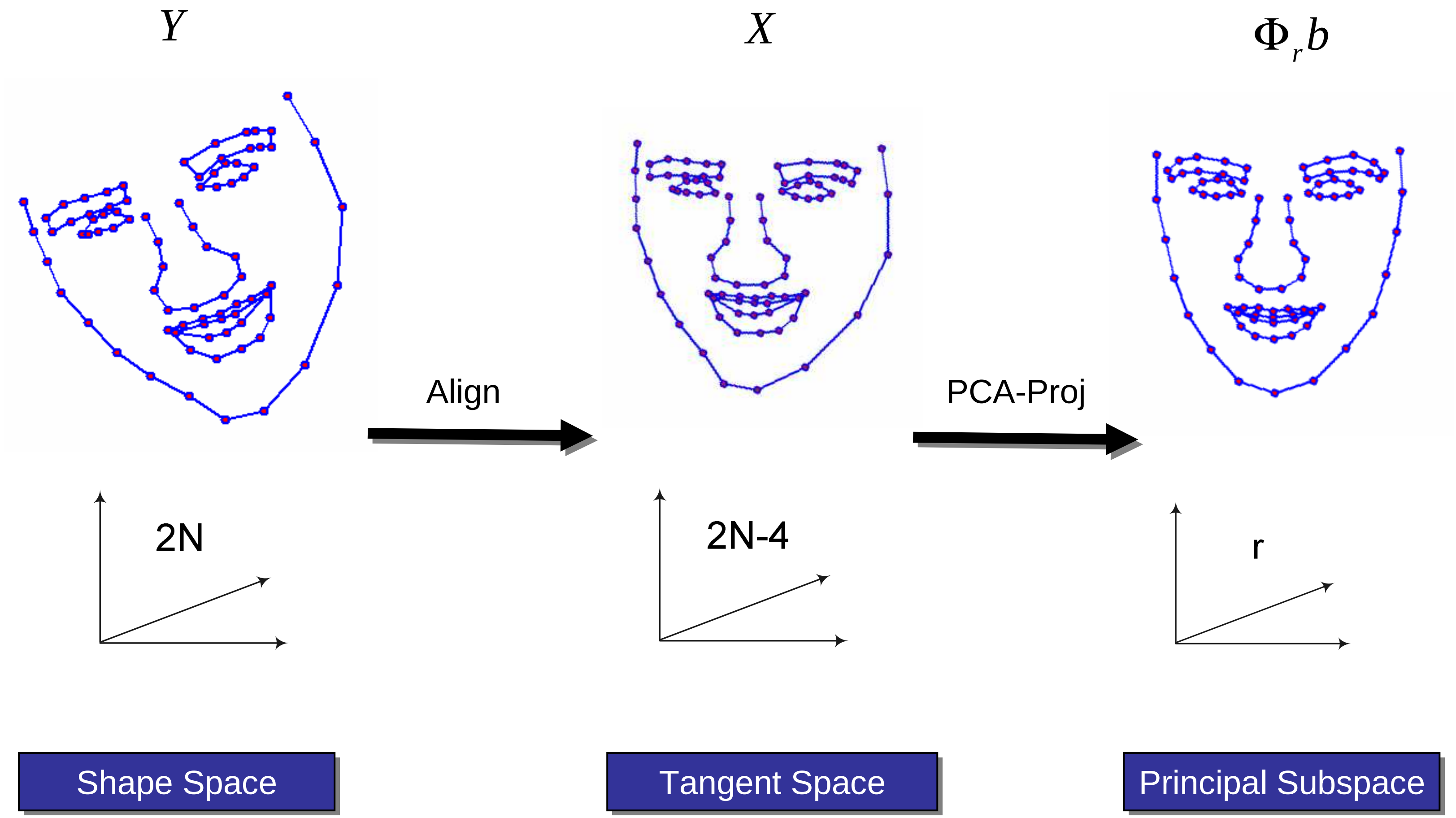
Our Method

- Prior knowledge: same as ASM
- Regularization rules
  - Shape: a weighted average of regularized shape and observed shape.
  - Shape parameters: continuously regularized by multiplying a shrinking factor.
  - Pose parameters: constrained by observation noise.
- Convergence guaranteed by EM

Results




Shape Spaces



BTSM Formulation

b: shape parameter

ε : shape noise

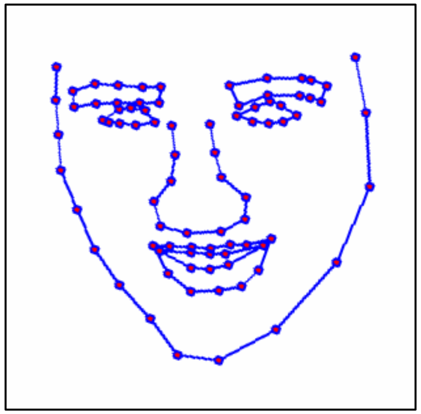
X : tangent shape

Y: observed shape

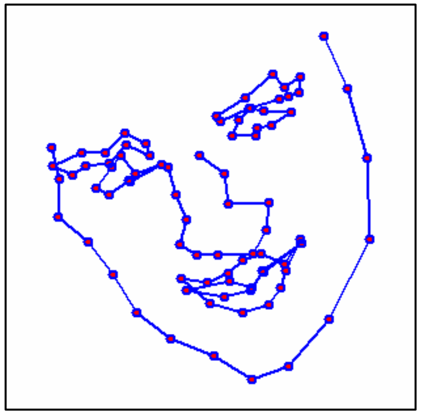
θ : pose parameter

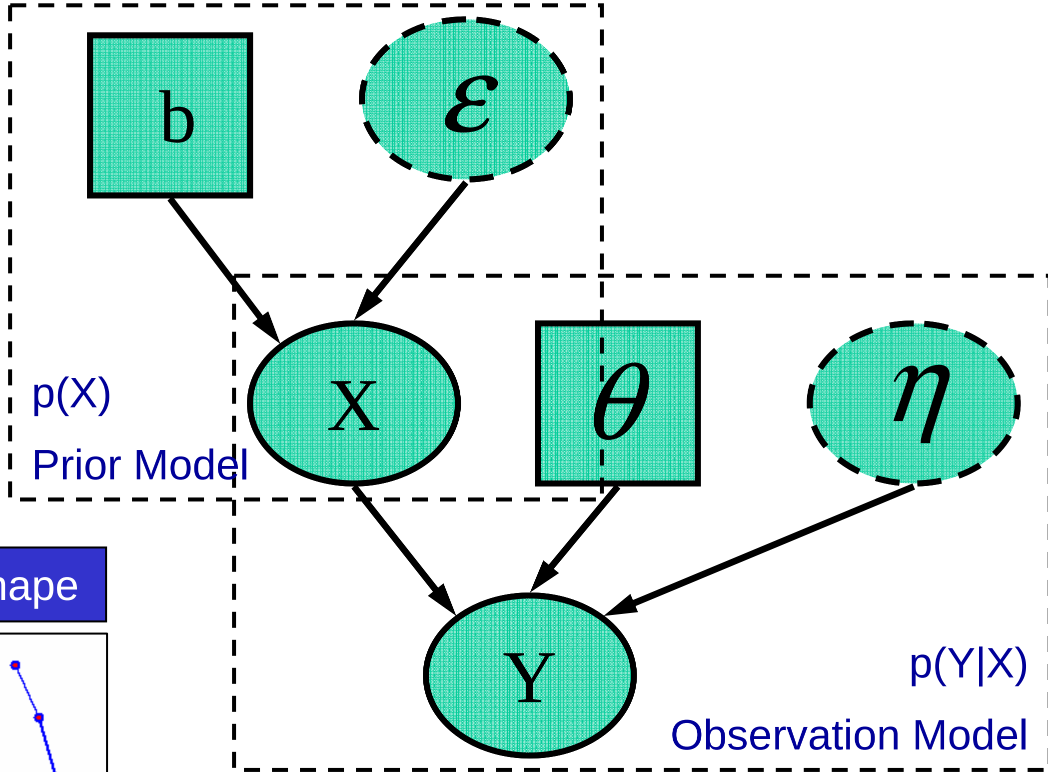
η : observation noise

X: Tangent Shape



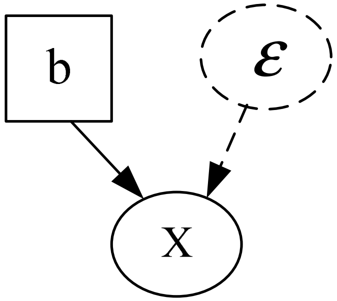
Y:Observed Shape





Prior Model P(X)

$$x = \mu + \Phi_r b + \varepsilon$$



Prior model

b: shape parameter
ε: isotropic noise
X: tangent shape

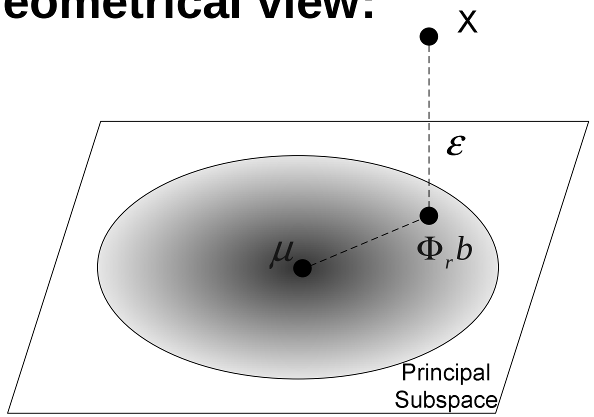
- **Isotropic shape noise:**

$$\varepsilon \sim N(0, \sigma^2 I)$$

- **Variance:**

$$\sigma^2 = \frac{1}{n} \sum_{i=r+1}^{n} \lambda_i$$

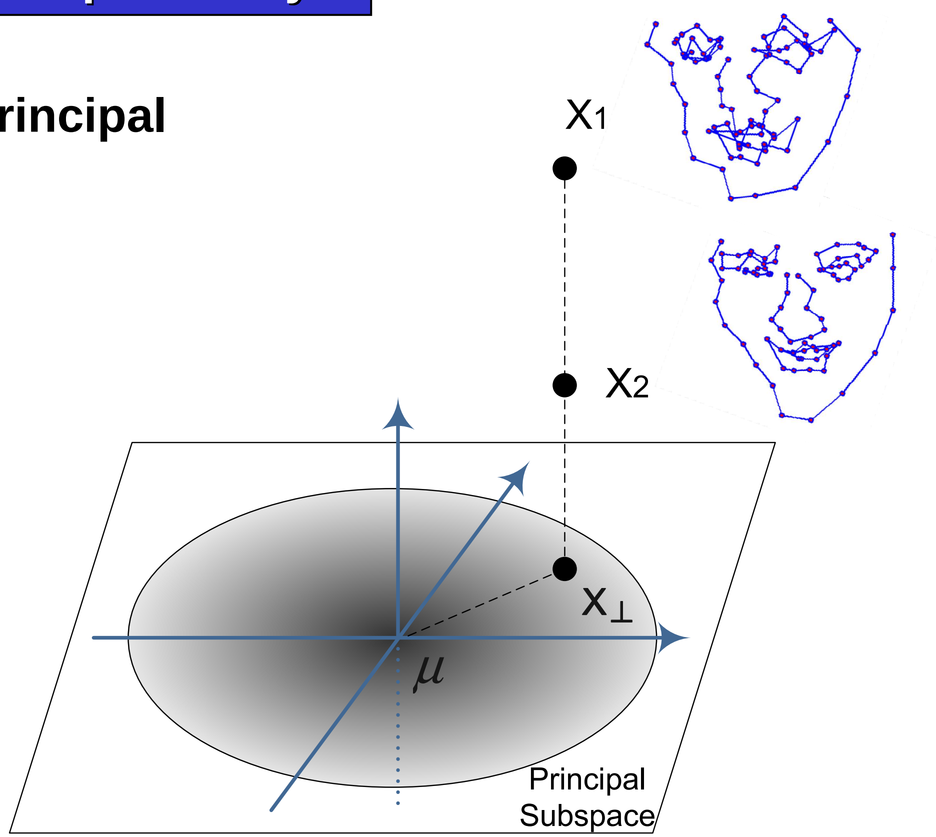
- **Geometrical view:**



Why Model ε

Work with Principal Subspace Only?

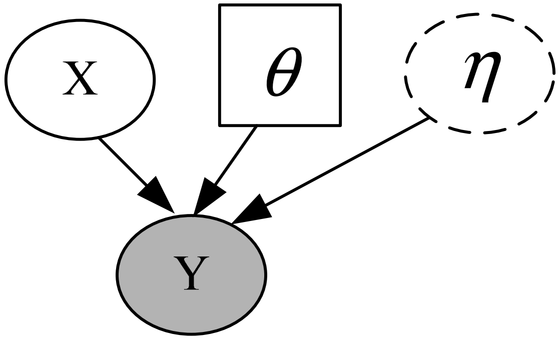
- **Distance outside Principal Subspace**
- **Compensate noise**



Observation Model P(Y|X)

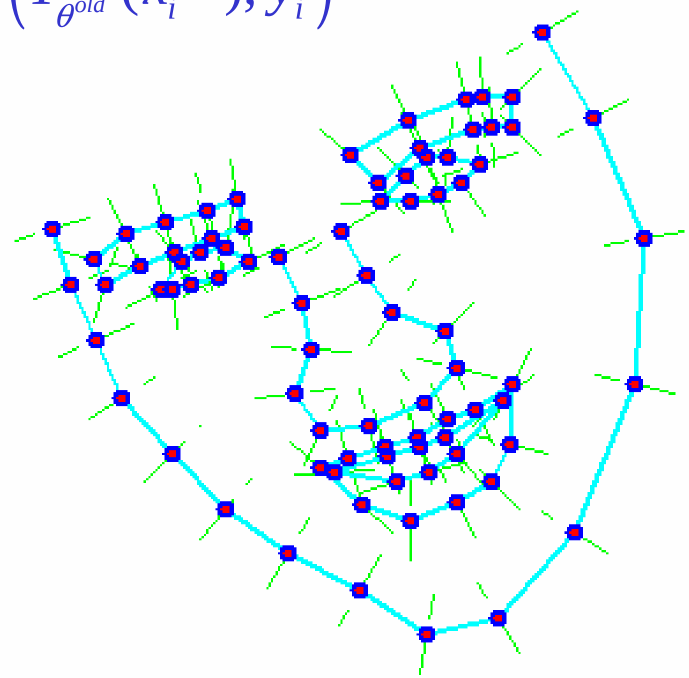
$$y = T_\theta(x) + \eta$$

$$\eta \sim N(0,\ \Sigma),$$
$$\Sigma = diag(\rho_1^2, \ldots, \rho_N^2) \otimes I_2 \qquad \rho_i = D\left(T_{\theta^{old}}(x_i^{old}), y_i\right)$$



Observation Model

X: tangent shape
θ: pose parameter
η: observation noise
Y: observed shape



Distance Map: $T_{\theta^{old}}(x_i^{old})$ Blue: ρ_N)

Posterior

Posterior

$$p(b,c,s,\theta \mid y) \propto \exp\{-\frac{1}{2}[(\sigma^2 + s^{-2}\rho^2)^{-1}(\| \Phi_r^T T_\theta^{-1}(y) - b \|^2 + \| \Phi_{-r}^T T_\theta^{-1}(y) \|^2) + s^2\rho^{-2} \| A^T T_\theta^{-1}(y) \|^2 + b^T \Lambda^{-1} b]\} \cdot \frac{const}{(\sigma^2 + s^{-2}\rho^2)^{(N-2)} s^{-4} \rho^4}$$

Q-Function

$$\log p(b,s,c,\theta \mid x,y) = \log p(b \mid x) + \log p(\gamma \mid x,y)$$
$$= -\frac{1}{2}\{b^T \Lambda^{-1} b + \sigma^{-2} \| x - \mu - \Phi_r b \|^2\} - \frac{1}{2}\rho^{-2} \| y - X\gamma \|^2 + const$$
$$where\ \gamma = (c_1, c_2, s \times \cos\theta, s \times \sin\theta)^T \ and\ X = (x, x^*, e, e^*).$$

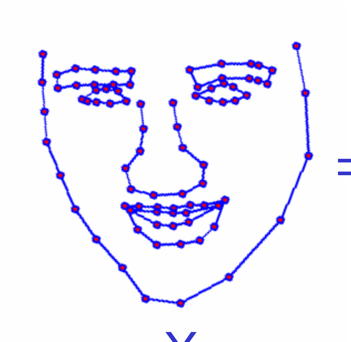
Expectation-Maximization

E-Step: computing $\langle x \rangle, \langle \|x\|^2 \rangle$

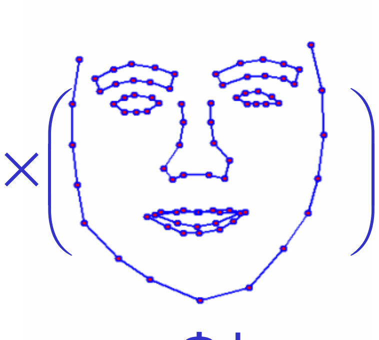
M-Step: estimate b and θ

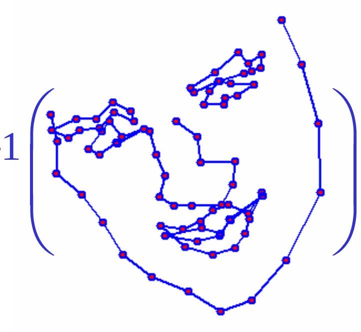
E-Step

Updating Tangent Shape



$$= (1-p)\times\Big(\quad\Big) + p\times\Phi\Phi^T T_\theta^{-1}\Big(\quad\Big)$$





X $\qquad \Phi_r b \qquad$ Y

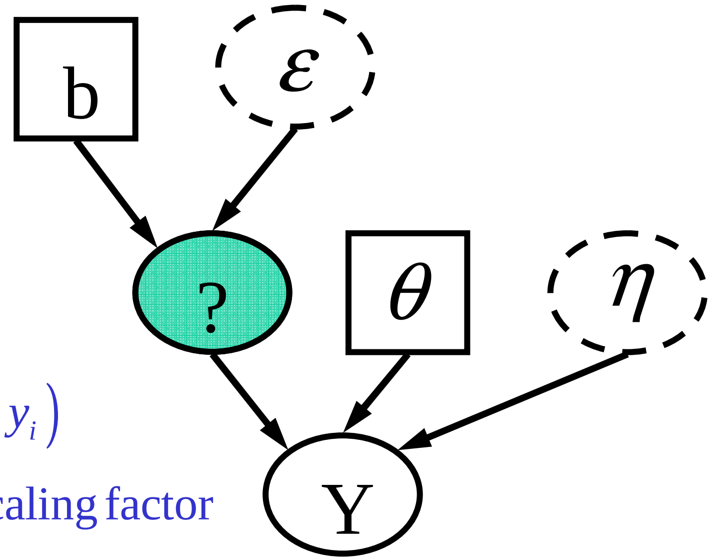
Weight p

$$p = \frac{E_{shp\ noise}}{E_{shp\ noise} + s^{-2}E_{obs\ noise}}$$

$$= \frac{\sigma^2}{(\sigma^2 + s^{-2}\rho^2)}$$

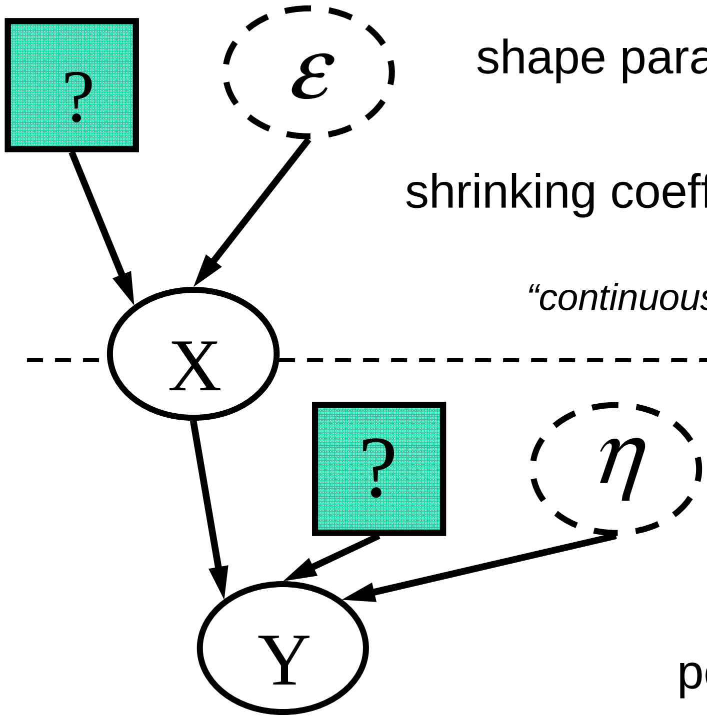
$$\rho_i = D\left(T_{\theta^{old}}(x_i^{old}) - y_i\right)$$

$$\sigma^2 = \frac{1}{n}\sum_{i=r+1}^{n}\lambda_i \qquad s: \text{scaling factor}$$



M-Step

Update Shape Parameter



shape para: $b_i = \alpha_i \left(\Phi_r^T x \right)_i$

shrinking coeff: $\alpha_i = \dfrac{\lambda_i}{\lambda_i + \sigma^2}, \quad \sigma^2 = \dfrac{1}{n} \sum_{i=r+1}^{n} \lambda_i$

"continuous regularization by SNR"

Update Pose Parameter

pose para: $\theta = \arg\min \sum_i w_i \left(T_\theta(x_i) - y_i \right)^2$

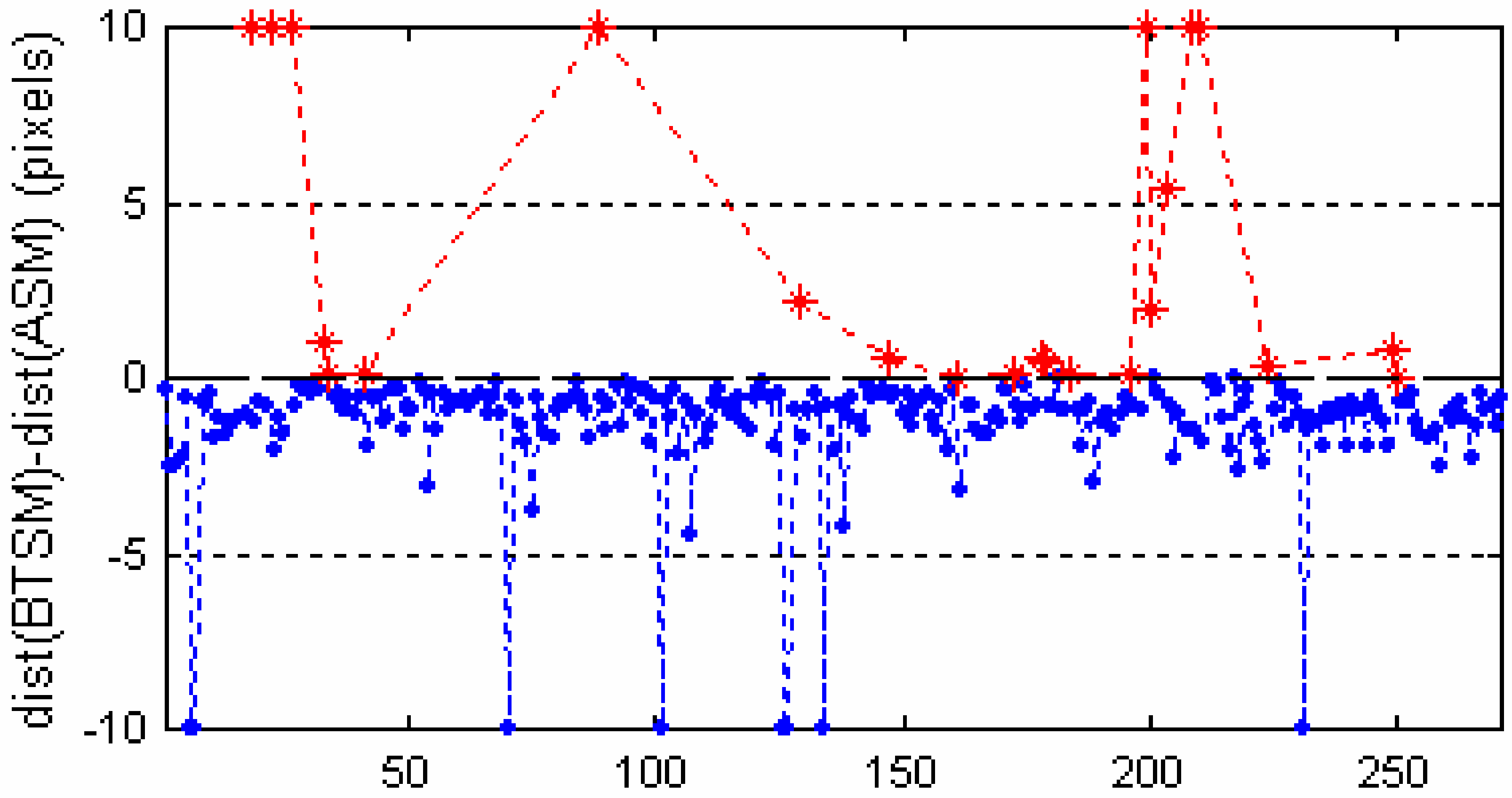
"weighted procrustes analysis"

Weight: $w_i = \dfrac{1}{\rho_i^2}, \quad \rho_i = D\left(T_{\theta^{old}}(x_i^{old}) - y_i \right)$

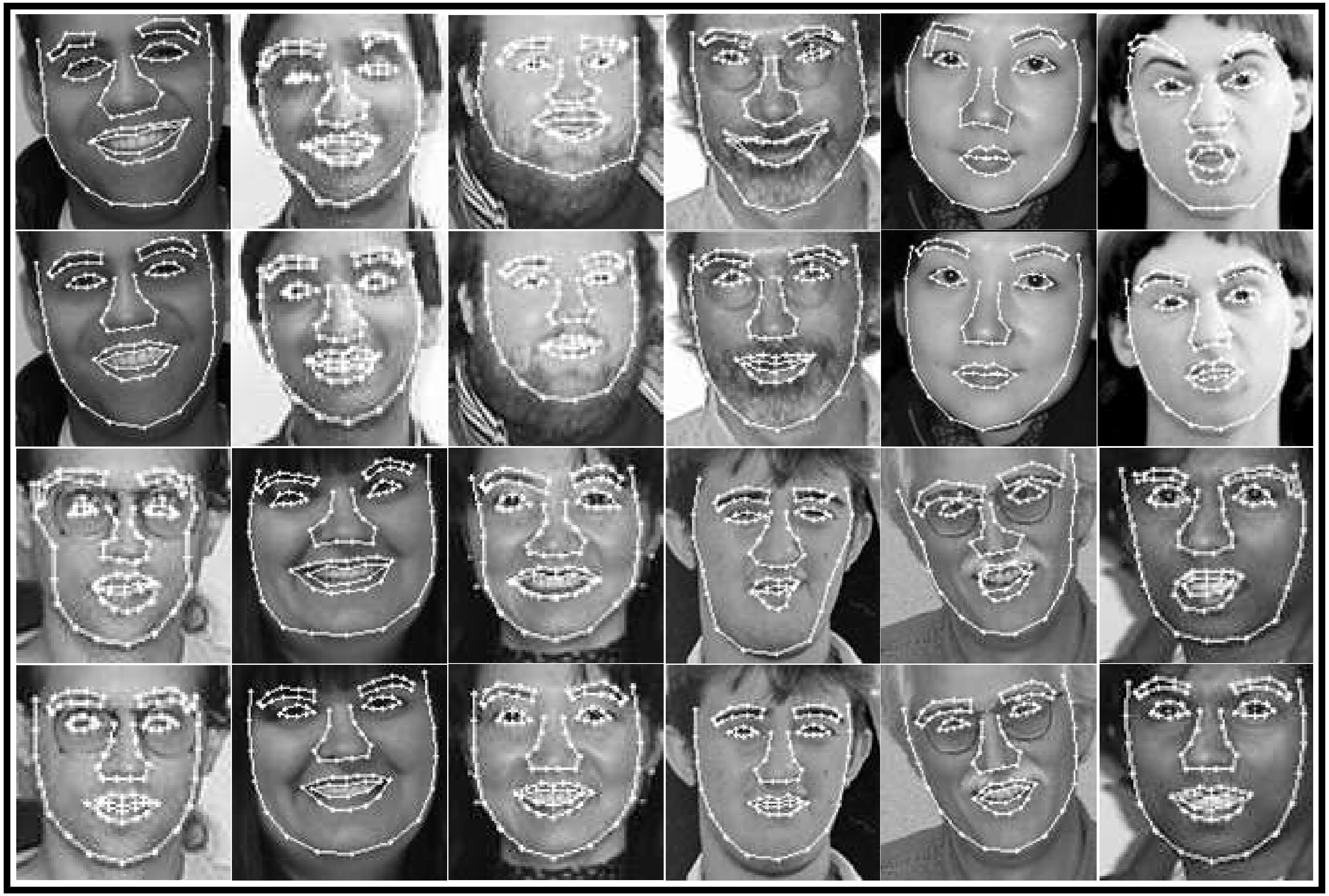
More Results

Precision: BTSM vs ASM

- 870 manually labeled face (training: 599, testing: 271)
- X: the index of testing faces; Y: dist(BTSM) – dist(ASM)



Comparison



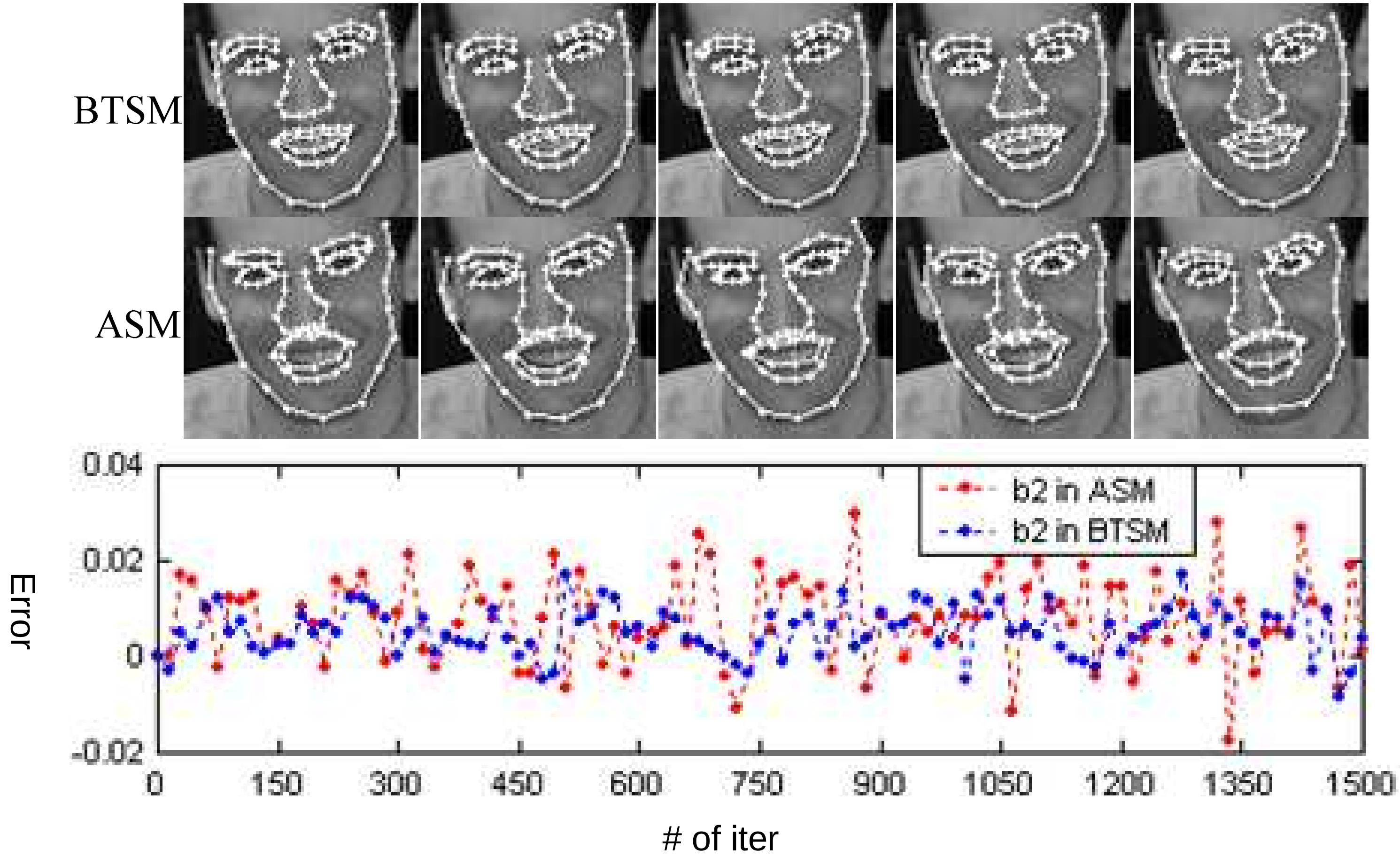
ASM

BTSM

ASM

BTSM

Numerical Stability



Summary

- Two Simple Ideas
  - De-Noising by Shrinkage
    - Suppress noise
    - Preserve major shape deformations
  - Penalize outliers by iterative re-weighting
- Pro's
  - Well generalized to novel, unseen faces
  - Robust to image noise
  - Fast (from 30ms to 170ms); fully automatic
- Con's
  - Relies on face detector for good initialization
  - Limited to frontal faces